Facultad de Química   PEDECIBA Química

Tesis de doctorado

Caracterización molecular de PknG, una quinasa de proteínas crucial para la patogenicidad de *Mycobacterium tuberculosis*

Magdalena Gil

Orientadoras   Ana Denicola
                  Rosario Durán

Montevideo, mayo de 2016

Portada: imagen de criomicroscopía electrónica de una cepa salvaje de *Mycobacterium smegmatis* MC^{2}155.

"We dance round in a ring and suppose, but the Secret sits in the middle and knows"

Robert Frost (1874 - 1963)

Resumen

Mycobacterium tuberculosis, agente etiológico de la tuberculosis, representa un importante problema de salud pública a nivel global. El genoma de *M. tuberculosis* codifica 11 Ser/Thr quinasas "de tipo eucariota" denominadas Pkns (PknA a PknL), ocho de las cuales se encontraron expresadas durante la infección, según datos que surgen de análisis proteómicos y transcriptómicos.

En particular una de ellas, PknG, cumple un rol tanto en el metabolismo de las micobacterias como en la patogénesis, y ha cobrado gran relevancia como una de las moléculas centrales para la sobrevida de la bacteria dentro del hospedero. Desde el punto de vista estructural esta es una quinasa singular, ya que además del dominio catalítico posee otros dominios de función aún poco caracterizada (rubredoxina en el N-terminal y un motivo repetido tetratricopéptido en la región C-terminal).

Si bien la importancia de esta quinasa ha sido reconocida hace ya más de una década, al día de hoy se desconocen los mecanismos moleculares subyacentes a su acción, los sustratos fosforilables tanto en la bacteria como en el hospedero, los mecanismos de activación/inhibición de la misma y las señales que los desencadenan.

En el presente trabajo nos propusimos contribuir a la elucidación del rol fisiológico de PknG y su regulación en las micobacterias. Para ello caracterizamos el papel de los distintos dominios de PknG sobre su actividad enzimática, e identificamos nuevos mediadores y sustratos en las vías de señalización mediadas por esta quinasa. Además reportamos un nuevo inhibidor de PknG que actúa sobre un dominio extracatalítico de la quinasa y que representa un nuevo mecanismo para inhibir este importante factor de virulencia. Nuestros resultados indican que el dominio rubredoxina cumple un papel en la regulación de la actividad de esta quinasa. Estudios interactómicos nos permiten además postular la participación de PknG en la regulación del metabolismo bacteriano y la síntesis de la pared celular. En particular, esta enzima estaría encargada de la fina regulación tendiente a mantener el balance entre la síntesis y la degradación de intermediarios nitrogenados en la bacteria.

A partir de esta tesis surge una nueva hipótesis de trabajo que permite integrar las dos funciones previamente reportadas de PknG. La capacidad de *M. tuberculosis* de inhibir la maduración fagolisosomal y de adaptarse al ambiente intracelular del hospedero mediante la reprogramación de vías metabólicas para acceder a nutrientes, minimizar las consecuencias del estrés oxidativo y mantener la infección son elementos centrales en su patogenicidad. Nuestros resultados podrían indicar que PknG está jugando un papel clave en la sobrevida de la bacteria en el hospedero a través de la regulación de la capacidad de la misma de adaptarse a las condiciones nutricionales del ambiente intracelular.

Índice general

Índice de figuras

Abreviaciones

AG	arabinogalactano
AM	ácidos micólicos
AP-MS	purificación por afinidad seguida de espectrometría de masa
BCG	bacilo de Calmette-Guérin
CFP-10	del inglés *culture filtrate protein 10*
DMSO	dimetilsulfóxido
EM	espectrometría de masa
ESAT-6	del inglés *early secreted antigenic target 6 kDa*
ESX-1	del inglés *ESAT-6 secretion system 1*
FHA	del inglés *ForkHead Associated*
GarA	del inglés *glycogen accumulation regulator*
GDH	glutamato deshidrogenasa dependiente de NAD^+
Glc-Nac/Mur-Nac	N-acetilglucosamina ácido N-acetilmurámico
GOGAT	complejo glutamato sintasa
GS	glutamina sintetasa
HK	quinasa de proteínas en histidina
KGD	α-cetoglutarato descarboxilasa
LAM	lipoarabinomanano
LDL	lipoproteína de baja densidad
MAPK	quinasa de proteína activada por mitógenos
mymA	del inglés *mycobacterial monooxygenase*
$OA-NO_2$	ácido 9- y 10-nitro-octadeca-9-*cis*-enoico
O.M.S.	Organización Mundial de la Salud
PDIM	dimicocerosatos de ftiocerol
PG	peptidoglicano
PstP	del inglés *PP2C-family Ser/Thr phosphatase*
PTP	del inglés *protein tyrosine phosphatase*
PykA	piruvato quinasa
RD1	región de diferencia 1
Rbx	dominio rubredoxina
RR	regulador de respuesta
SCID	del inglés *severe combined immunodeficiency*
SPR	resonancia plasmónica de superficie
STPK	quinasa de proteína en serinas y treoninas
TB	tuberculosis
TB-MDR	tuberculosis multidrogorresistente
TNF	factor de necrosis tumoral
TPR	motivo repetido tetratricopéptido
XDR	cepas extremadamente drogorresistentes

I. INTRODUCCIÓN

1. Tuberculosis

En el año 2000 las Organización de las Naciones Unidas estableció los "Objetivos de Desarrollo del Milenio", ocho ambiciosas metas que intentarían alcanzarse para 2015. Entre las mismas figuraba combatir el VIH/SIDA, el paludismo y otras enfermedades. **En este objetivo se incluía detener la propagación y disminuir la incidencia de la tuberculosis (TB) ya que esta es la segunda causa de muerte debida a un agente infeccioso a nivel global.** Los datos de la Organización Mundial de la Salud (O.M.S.) señalan que un tercio de la población mundial está infectada por este bacilo y que en 2014 se registraron 9,6 millones de nuevos casos y 1,5 millones de muertes debidas a esta bacteria [1]. En Uruguay la incidencia de esta enfermedad es de 25,1 casos por cada 100.000 habitantes [2]. A pesar de los esfuerzos realizados, nuestro país no alcanzó en 2015 la meta de reducir a la mitad la incidencia y mortalidad por esta patología. Por el contrario en la última década la incidencia de esta enfermedad en la población uruguaya se ha incrementado [3].

La tuberculosis es una enfermedad infecciosa causada por un grupo de bacterias estrechamente emparentadas que forman parte del complejo *Mycobacterium tuberculosis* y que según datos paleoepidemiológicos ha afectado a la Humanidad los últimos 70.000 años [4]. El efecto que ha tenido esta enfermedad en la historia de nuestra especie ha quedado plasmado en la literatura de todas las épocas. Distintas culturas describen este mal, aparece como *schachepheth* en el Antiguo Testamento, *φϑίσις* (en latín *phthisis*) en la literatura griega del siglo IV a.C., *cunsumptio* en reportes de Cicerón y como *consunción* en obras del siglo XIX tales como "La Dame aux camélias". En 1839 Johann Lukas Schönlein propone llamar tuberculosis a esta afección ya que todas las manifestaciones de la enfermedad incluyen la presencia de tubérculos y hace poco más de 130 años, el 24 de marzo de 1882, Robert Koch comunica a la Sociedad Berlinesa de Fisiología el aislamiento del microorganismo causante de la enfermedad [5].

Los drásticos cambios en el estilo de vida del Hombre durante milenios así como los progresos de la medicina en el último siglo no han logrado detener el avance de esta enfermedad lo que denota la enorme adaptabilidad de este patógeno. Lo anterior ha llevado a algunos autores a considerar a *M. tuberculosis* como el paradigma de la adaptación patógeno-hospedero [6].

1.1. *Mycobacterium tuberculosis: el agente etiológico de la tuberculosis*

M. tuberculosis es un bacilo inmóvil y aerobio estricto de 2 a 4 μm de longitud con un tiempo de generación de entre 15 y 20 horas [7]. De acuerdo al análisis taxonómico, *M. tuberculosis* pertenece al grupo de bacterias Gram positivas, sin embargo, comparte algunas características con bacterias Gram negativas ya que por ejemplo no retiene la tinción de Gram [8, 9] y presenta un

pseudoperiplasma [9]. Por otra parte, la alta densidad de lípidos de superficie vuelve a *M. tuberculosis* resistente a la tinción ácido-alcohol (tinción Ziehl-Neelsen) [10].

Esta bacteria es extremadamente resistente a la deshidratación debido a la compleja organización en estratos de su pared celular. La membrana externa está constituida por el peptidogligano (PG), el arabinogalactano ramificado (AG) y los ácidos micólicos de cadena larga (AM) entre los que se intercalan ceras inertes y glicolípidos [11]. Rodeando a la membrana externa se encuentra una cápsula de proteínas y lipopolisacáridos tales como: lipoarabinomananos (LAM), lipomananos, dimicocerosatos de ftiocerol (PDIM del inglés *phthiocerol dimycoceroserate*), micolato de trealosa y sulfolípidos importantes para la virulencia (Fig. 1) [9].

Los ácidos micólicos son una familia de ácidos grasos complejos que se presenta exclusivamente en la pared celular de bacterias de los géneros *Mycobacterium*, *Nocardia*, *Rhodococcus* y *Corynebacterium*. Estos lípidos, de entre 60 y 90 átomos de carbono compuestos por una cadena α alquílica y una β hidroxilada, son esenciales para la viabilidad celular y contribuyen a la virulencia de *M. tuberculosis* [10-12]. En las micobacterias, esta capa de AM funciona de forma análoga a la membrana externa de las bacterias Gram negativas ya que se genera un espacio entre la membrana interna y la capa de AM equivalente al periplasma [10].
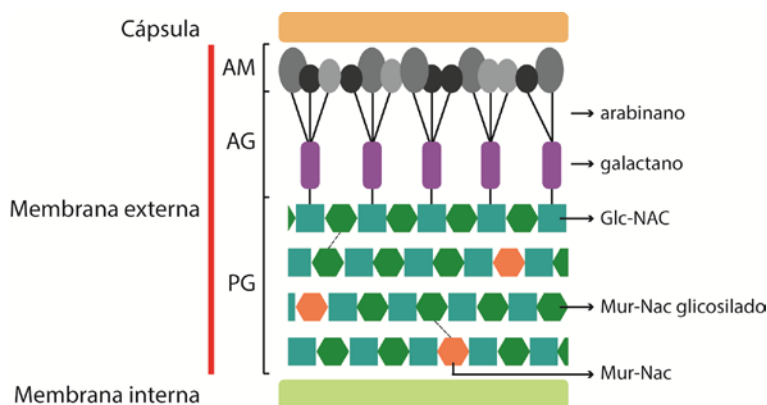
Figura 1. Estructura de la pared celular de *M. tuberculosis*. Se indican los componentes principales de la pared celular micobacteriana. El PG se localiza alrededor de la membrana interna y está compuesto por largos polímeros del disacárido N-acetilglucosamina ácido N-acetilmurámico (Glc-Nac/Mur-Nac) estructurados en hebras. Una capa de AG rodea al PG y está compuesta por galactano, unidad repetida de 6-D-Galfβ1-5-D-Galfβ sintetizada por galactofuranosil transferasas, unido a largos polímeros de arabinano. La mayoría de las moléculas de arabinano se unen a los AM dando origen a la envoltura gruesa y cerosa característica de las micobacterias. Adaptado de [9, 10].

Las enzimas involucradas en la biosíntesis de los ácidos micólicos son esenciales para la viabilidad y virulencia de *M. tuberculosis* lo que las ha convertido en interesantes blancos para el desarrollo de nuevos antibióticos [12].

1.2. *Infección por Mycobacterium tuberculosis y evasión de la respuesta inmune*

M. tuberculosis desarrolló una estrategia completamente diferente a la de otros patógenos para asegurar su persistencia en el hospedero.

En 1964 el trabajo pionero de Max Lurie en conejos distinguió cuatro etapas en la progresión de la tuberculosis pulmonar. La *primera fase* consiste en la inhalación de los bacilos, que son fagocitados y generalmente destruidos por macrófagos alveolares. Las bacterias que escapan a la destrucción intracelular se multiplican y originan la ruptura del macrófago. Cuando esto ocurre, monocitos, linfocitos y neutrófilos son atraídos hacia el pulmón (*segunda fase*) y los monocitos reclutados se diferencian en macrófagos inmaduros que son incapaces de destruir a *M. tuberculosis*. En este estado simbiótico, el crecimiento de las micobacterias se vuelve logarítmico y se acumulan macrófagos que mantienen a las bacterias aisladas en el interior de una estructura denominada granuloma. En las dos o tres semanas posteriores al inicio de la infección comienza a desarrollarse la inmunidad adaptativa, los linfocitos T estimulados por antígenos micobacterianos proliferan y secretan citoquinas que activan a los macrófagos, lo que provoca la interrupción del crecimiento bacteriano (*tercera fase*). La necrosis caseosa de la región central del granuloma inhibe el crecimiento extracelular de las bacterias y de acuerdo a la respuesta inmunitaria del hospedero la infección se detiene en esta etapa o progresa. En individuos inmunocompetentes se contiene la infección dando lugar a una tuberculosis latente o persistente en la que el hospedero es asintomático e incapaz de transmitir el bacilo. Eventualmente los granulomas sanan dejando pequeñas lesiones fibrosas. Sin embargo, si el sujeto no puede controlar la infección inicial o si el sistema inmunitario de la persona con tuberculosis latente se debilita (por drogas inmunosupresoras, infección con el virus de la inmunodeficiencia humana, malnutrición, edad avanzada, etc.) ocurre la licuefacción del granuloma. En este medio rico en nutrientes las bacterias comienzan a replicarse rápidamente y bacilos viables pueden escapar y propagarse dentro de los pulmones (tuberculosis activa) e incluso acceder a otros tejidos a través del sistema linfático (tuberculosis extrapulmonar) (Fig. 2) [6, 13, 14].
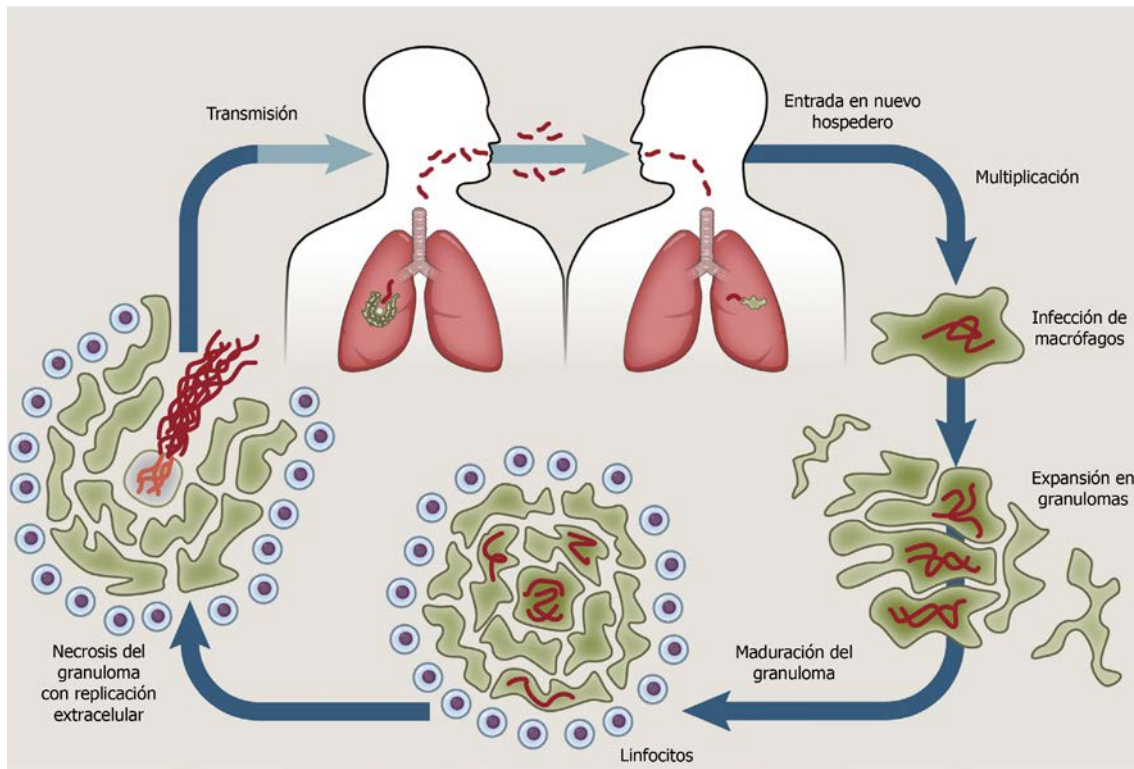
Figura 2. Etapas de la infección por _M. tuberculosis_. Cuando finas partículas conteniendo de una a tres bacterias son expelidas por un individuo con tuberculosis activa (enfermo bacilífero), inhaladas por otro individuo y se depositan en el lóbulo pulmonar inferior se inicia la infección por _M. tuberculosis_. Allí, la bacteria comienza a reclutar macrófagos desde la superficie pulmonar y es fagocitada por estas células especializadas. Una vez dentro de la célula fagocítica, la bacteria es transportada desde el epitelio pulmonar hacia tejidos más profundos donde comienza un nuevo ciclo de reclutamiento de macrófagos dando lugar al granuloma: una estructura organizada que contiene tanto macrófagos diferenciados como otras células inmunes. En una etapa temprana, el granuloma permite que las bacterias se diseminen entre los macrófagos recién llegados, no obstante una vez que comienza a desarrollarse la inmunidad adaptativa, el granuloma restringe el crecimiento bacteriano. Empero, en ciertas circunstancias ocurre la necrosis de los macrófagos infectados contenidos en el granuloma formándose un centro necrótico donde las bacterias vuelven a replicarse y son transmitidas a un nuevo hospedero. Tomado de [6].

El trabajo de Lurie fue fundamental para definir las etapas de la enfermedad. Sin embargo, estudios más recientes proponen una mirada desde los mecanismos bacterianos que permiten el éxito de la infección y es en este marco que se plantean las siguientes etapas:

1.2.1   _Acceso del bacilo a su nicho ecológico_

Para alcanzar su nicho y allí poder replicarse, los patógenos deben eludir las defensas del hospedero; las que a su vez están fuertemente influenciadas por la microbiota normal que puebla las superficies de nuestro organismo [15]. En las mucosas continuamente se están reclutando macrófagos activados con propiedades microbicidas que mantienen controlados a los patógenos oportunistas. La estrategia de evasión de la respuesta inmune de _M. tuberculosis_ se basa en el acceso a una región prácticamente libre de microorganismos comensales: el lóbulo pulmonar inferior. Por lo que, para poder infectar a su hospedero _M. tuberculosis_ evade a los microorganismos comensales [6].

1.2.2. *Multiplicación de la bacteria en el hospedero*

Una vez fagocitadas, las micobacterias tienen la habilidad de modular la maduración fagosomal, proceso por el cual el lumen de los fagosomas adquiere propiedades microbicidas a través de múltiples ciclos de fusión con endosomas tempranos, tardíos y finalmente con lisosomas y que da lugar a un fagolisosoma funcional [16]. En los fagosomas que contienen a *M. tuberculosis* el proceso de maduración fagolisosomal se detiene antes de que el contenido soluble de estas vesículas adquiera las características necesarias para eliminar al bacilo [17]. **La interrupción temprana de la maduración fagosomal permite que la micobacteria perdure en el interior de los macrófagos donde además es capaz de replicarse**. Sin embargo, experimentos en líneas celulares de macrófagos en cultivo han determinado la localización subcelular de *M. tuberculosis*; se halló a la micobacteria tanto en endosomas tempranos levemente ácidos como en fagolisosomas e incluso una pequeña proporción en el citosol proveniente de la ruptura de fagosomas [6].

Los macrófagos infectados se acumulan dando lugar a la formación del granuloma, agregado compacto de macrófagos maduros que surge en respuesta a un estímulo persistente. Hasta hace algunos años se consideraba que este agregado constituía una estructura de protección generada por el hospedero para aislar al bacilo; esta idea se apoyaba en el rol que cumple el factor de necrosis tumoral (TNF del inglés *tumoral necrosis factor*) en la organización del tráfico de macrófagos y leucocitos durante la inflamación [18]. Sin embargo, estudios más recientes sugieren que *M. tuberculosis* promueve la formación del granuloma y que utiliza esta estructura para favorecer su expansión y diseminación. Cuando el número de bacterias que infectan un macrófago supera cierto umbral se disparan los mecanismos que llevan a la muerte celular apoptótica, lo que deja bacterias viables entre los restos celulares. En simultáneo, acceden al granuloma naciente macrófagos no infectados que fagocitan esas bacterias, el proceso de muerte celular y "re"fagocitosis produce una enorme expansión bacteriana [19, 20].

1.2.3. *Salida de Mycobacterium tuberculosis del hospedero*

Para asegurar su supervivencia como especie, *M. tuberculosis* debe salir del granuloma de un individuo infectado y entrar en un nuevo hospedero en el que establecer la infección. Datos epidemiológicos sugieren que la eficiencia de transmisión es mayor cuando la persona infectada presenta granulomas que han sufrido necrosis en la región central. En el árbol bronquial, la ruptura de las áreas necróticas de los granulomas permite el acceso de los bacilos a la vía aérea y cuando el infectado habla o tose se forman finos aerosoles que facilitan la salida de las bacterias [6].

El éxito de *M. tuberculosis* es indudable pero se debe tener en cuenta que el 90% de los infectados logra contener e incluso eliminar la infección a través de mecanismos que implican la respuesta inmune innata o adaptativa. La supresión de la infección puede dar origen a una forma latente de la enfermedad en la cual la bacteria persiste indefinidamente en el hospedero y puede producir tuberculosis activa incluso décadas después [6]. A pesar de que este escenario cuenta con amplia aceptación en la literatura, estudios recientes sugieren que en nuestros días la enfermedad se manifiesta pocos meses después de la infección o es eliminada [21], de hecho, los individuos recientemente infectados son los responsables de transmitir la enfermedad. Aún en la actualidad queda mucho por entender sobre los procesos de reactivación que dan lugar a la tuberculosis activa.

1.3. *Prevención de la tuberculosis*

La única vacuna disponible en la actualidad para prevenir la tuberculosis es el bacilo de Calmette-Guérin (BCG), cepa atenuada de *Mycobacterium bovis* obtenida por Albert Calmette y Camille Guérin a partir de una cepa virulenta luego de 230 pasajes *in vitro* en un período de 13 años [22]. El análisis comparativo de los genomas de las cepas virulentas de *M. tuberculosis* y *M. bovis* y las cepas de BCG mostró que la denominada región de diferencia 1 (RD1) está ausente en estas últimas. En *M. tuberculosis* el *locus* RD1 comprende nueve genes (*rv3871* a *rv3879c*) y codifica, entre otras proteínas, para el sistema de secreción ESX-1 (del inglés *ESAT-6 secretion system 1*). Además, este *locus* codifica para dos pequeñas proteínas altamente inmunogénicas secretadas por ESX-1: ESAT-6 (del inglés *early secreted antigenic target 6 kDa*) y CFP-10 (del inglés *culture filtrate protein 10*) que se sabe son necesarias para la virulencia de *M. tuberculosis*. Múltiples estudios evidencian que la pérdida de RD1 está implicada en la atenuación de la virulencia. No obstante la reintroducción de ESX-1 en el bacilo de Calmette-Guérin no restablece completamente la virulencia y el mutante de deleción de RD1 en *M. tuberculosis* continúa siendo más virulento que BCG en modelos de infección en ratones, lo que sugiere modificaciones adicionales que contribuyen a la atenuación de BCG [23].

Diversos estudios confirman que en niños la eficacia de la vacuna es del 80% en lo que respecta a la protección frente a las formas más severas de la enfermedad (meningitis y tuberculosis diseminada). Sin embargo, para el caso de tuberculosis pulmonar en adultos esta vacuna es de dudosa utilidad, de acuerdo a los estudios se presentan valores de eficacia en el intervalo de 0 a 80% [23]. El objetivo de la O.M.S. de erradicar la tuberculosis para el año 2050 solo podrá ser alcanzado si se desarrollan nuevas vacunas capaces de controlar la expansión de la epidemia limitando la infección inicial, la progresión de la enfermedad y la reactivación de la tuberculosis latente [24].

1.4. *Tratamiento de la tuberculosis*

El advenimiento de fármacos antituberculosos en la década del 40 transformó a la tuberculosis en una enfermedad curable. En 1945 la tasa de mortalidad en nuestro país era de más de 100 por cada 100.000 habitantes y gracias al tratamiento disponible esta cifra descendió a 1.8 muertes por cada 100.000 habitantes (datos de 2004).

El esquema actual de tratamiento de la tuberculosis tiene una duración de seis meses y está dividido en dos etapas: una *fase inicial intensiva* con administración diaria de la medicación seguida de una *fase de consolidación* en la que los fármacos se administran dos veces por semana. Las drogas administradas en la primera fase son isoniazida, rifampicina, pirazinamida y etambutol y en la etapa de administración intermitente se continúa el tratamiento con isoniazida y rifampicina [25]. Debido a que en nuestro país la resistencia inicial a isoniazida es menor al 1% (debajo del 4% tomado como valor de corte por las pautas internacionales) no se incluye etambutol o estreptomicina en la fase inicial del tratamiento [26].

La supervisión estricta del tratamiento por parte del personal de la salud ha demostrado ser la estrategia más efectiva para el control de la tuberculosis ya que en regiones con programas débiles de control de esta enfermedad han aparecido cepas resistentes a las drogas actualmente en uso. La tuberculosis multidrogorresistente (TB-MDR) es una forma de tuberculosis insensible al tratamiento con isoniazida y rifampicina y que causó el 5% de los casos de TB reportados en 2014. Para el tratamiento de la TB-MDR se utilizan algunos de los llamados fármacos de segunda línea: kanamicina, amikacina, capreomicina, fluoroquinolonas (levofloxacina, moxifloxacina, gatifloxacina y ofloxacina), etionamida, protionamida, cicloserina, terizidona y ácido *p*-aminosalicílico; estas drogas provocan más efectos secundarios, son más costosas y debido a su menor eficacia dan lugar a tratamientos más prolongados. Dentro de las cepas MDR se distinguen las extremadamente drogorresistentes (XDR) definidas como resistentes a todos los fármacos de primera línea, las fluoroquinolonas y a al menos una de las drogas inyectables de segunda línea (kanamicina, amikacina o capreomicina). En el año 2014 ya eran 105 los países que habían reportado casos de TB-XDR y se estima que el 9.7% de las muertes debidas a TB-MDR fueron causadas por cepas XDR [27].

Múltiples factores comprometen la eficacia de las terapias actuales, a saber: aparición de cepas multirresistentes; ausencia de nuevos fármacos capaces de lidiar con las cepas que han desarrollado resistencia; y la capacidad del patógeno para permanecer en el hospedero en estado de latencia y comenzar a desarrollarse cuando la integridad del sistema inmune de este se ve disminuida [28]. Para atender la necesidad de nuevas estrategias terapéuticas, se vuelve imperioso el

conocimiento de la fisiología de este microorganismo así como la comprensión de su complejo ciclo de vida capaz de adaptarse a los cambios en el ambiente celular del hospedero. En la última década los elementos de señalización, en especial los que involucran la fosforilación de proteínas en residuos de serinas y treoninas, han emergido como una prometedora nueva clase de blancos terapéuticos [29, 30].

2. Fosforilación reversible de proteínas

La adecuada respuesta de una célula a los cambios en su entorno es esencial para su supervivencia y es por ello que todas las células cuentan con sistemas de detección y procesamiento de estímulos. Las llamadas *cascadas de transducción de señales* son las encargadas de la detección, amplificación e integración de las señales extracelulares que dan lugar a respuestas intracelulares [31].

La fosforilación reversible de proteínas es uno de los mecanismos fundamentales para la transmisión de información en sistemas biológicos. Los procesos de fosforilación y desfosforilación, catalizados por quinasas y fosfatasas de proteínas respectivamente, pueden modular la actividad biológica de las proteínas. La simplicidad, flexibilidad y reversibilidad de la fosforilación sumada a la disponibilidad de ATP como dador de fosforilo explica su selección como mecanismo general de regulación [32].

La fosforilación de proteínas es una modificación muy diversa desde el punto de vista químico ya que involucra distintos grupos funcionales en las cadenas laterales de los nueve aminoácidos susceptibles a esta modificación. Se distinguen cuatro grupos de aminoácidos fosforilables (Fig. 3):

1. aminoácidos de cadena lateral hidroxilada (serina, treonina y tirosina) en los que la formación de un enlace fosfoéster da lugar al aminoácido fosforilado.

2. aminoácidos básicos (arginina, lisina [33] e histidina) en los que la formación de fosforamidatos da lugar al aminoácido fosforilado.

3. aminoácidos ácidos (aspártico y glutámico) en los que la formación de anhídridos mixtos da lugar al aminoácido fosforilado.

4. cisteína donde la formación del enlace fosfotioéster da lugar al aminoácido fosforilado [34, 35].
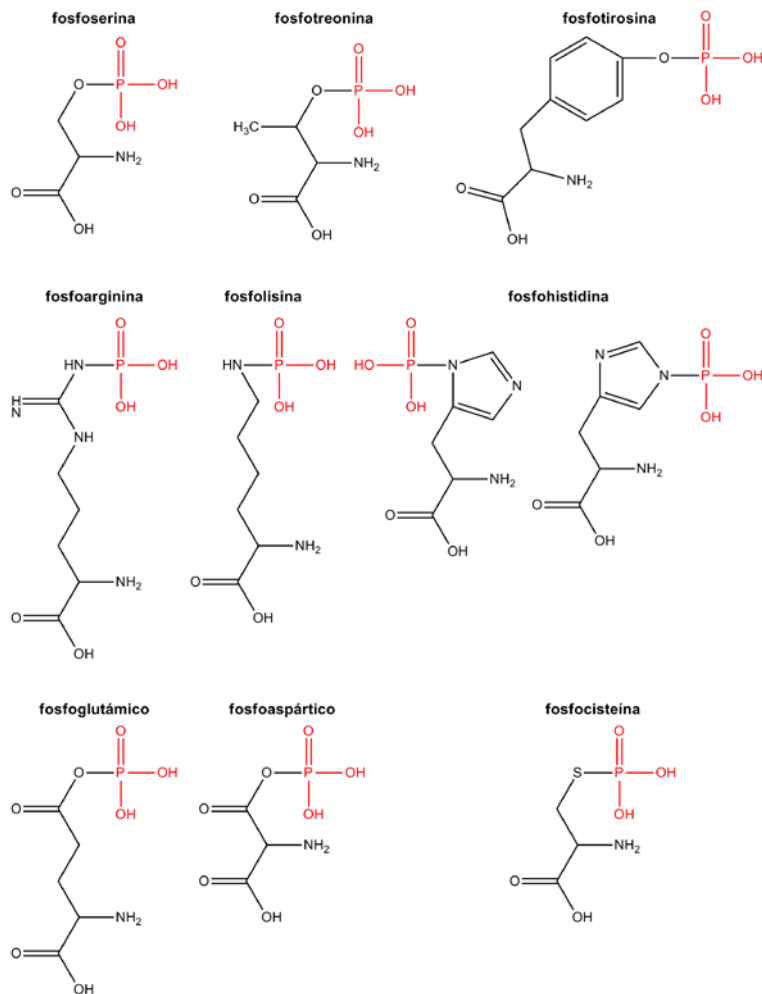
Figura 3. Estructura química de los residuos fosforilados presentes en sistemas biológicos.

2.1. *Sistemas de transducción de señales de dos componentes*

Estos sistemas regulatorios encargados del acoplamiento estímulo-respuesta han sido descriptos en *Bacteria*, *Archaea* e incluso *Eukaryota* aunque la abundancia relativa en cada uno de los dominios es sustancialmente diferente. Los sistemas de dos componentes participan en la mayoría de las vías de señalización en *Bacteria* pero son bastante infrecuentes en eucariotas en donde predominan los sistemas que involucran la fosforilación de Ser, Thr o Tyr [36].

Un sistema de dos componentes prototípico consta de una quinasa de proteínas en histidina (HK del inglés *histidine kinase*) que funciona como sensor de cambios en el ambiente extracelular y de un regulador de respuesta (RR del inglés *response regulator*). El mecanismo básico de acción de estos sistemas implica la autofosforilación de un residuo de histidina conservado en la HK seguido de la transferencia del grupo fosforilo a un residuo de aspártico en el dominio receptor del RR y la posterior activación del dominio efector del RR. Existen varias clases de dominios efectores aunque generalmente son dominios de unión al ADN, la presencia de estos dominios permite el

acoplamiento de la fosforilación del RR a cambios a nivel transcripcional. Además de la actividad quinasa, muchas HK poseen actividad fosfatasa lo que permite la desfosforilación de su regulador de respuesta en vías de transducción de señales que deben ser rápidamente apagadas (Fig. 4) [36-38].
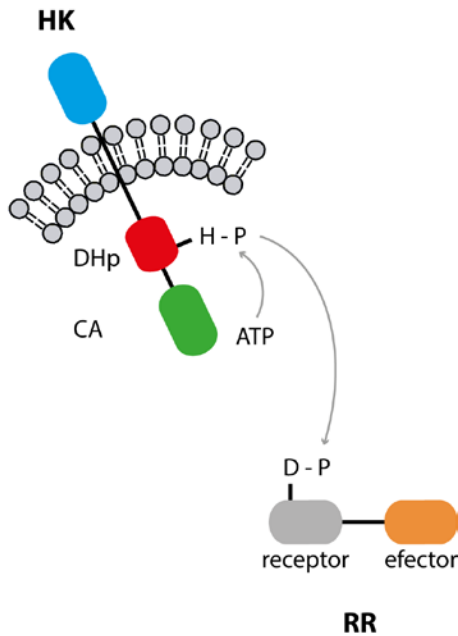
Figura 4. Representación esquemática de un sistema de dos componentes prototípico. En las vías de transducción de señales mediadas por HK la primera etapa consiste en la autofosforilación de un residuo de histidina conservado y la posterior transferencia del grupo fosforilo al RR.

El dominio sensor (indicado en celeste) es extremadamente variable entre distintas HK y según los reportes la homología estructural entre ellos es reducida. El dominio catalítico (CA, indicado en verde) une ATP y es responsable de la autofosforilación de un residuo de histidina conservado que se localiza en el dominio de dimerización (DHp, indicado en rojo). El dominio DHp es el encargado de la formación de un homodímero y funciona como dador de fosfato para el RR.

Los reguladores de respuesta generalmente constan de dos dominios: receptor (indicado en gris) que contiene al residuo de aspártico al que se transfiere el fosforilo proveniente de la histidina y efector (indicado en naranja) que es activado por fosforilación del dominio receptor y que típicamente está involucrado en la unión al ADN. Adaptado de [38].

El genoma de *M. tuberculosis* H37Rv codifica para 11 sistemas de dos componentes, algunos de los cuales regulan directamente factores de virulencia o están relacionados con la persistencia de la bacteria en el hospedero [37]. Sin embargo, solamente dos de ellos: *mtrA-mtrB* y *prrA-prrB* son esenciales para la viabilidad celular [39, 40].

Los genomas micobacterianos contienen un número relativamente bajo de sistemas de dos componentes si se los compara con genomas bacterianos de tamaño equivalente. Sin embargo, la reducción en el número de estos sistemas parece estar compensada por un mecanismo alternativo de transducción de señales que involucra la fosforilación de serinas y treoninas (Fig. 5) [30].
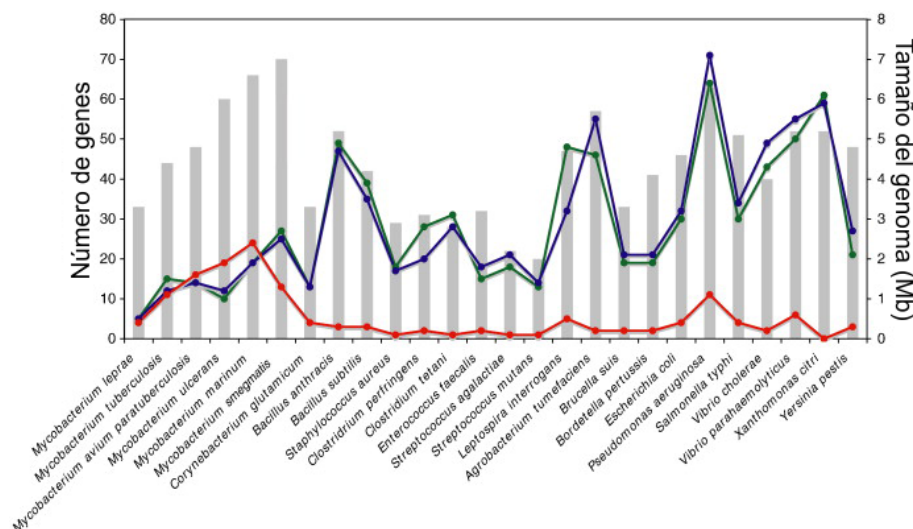
Figura 5. Elementos de señalización en bacterias. Se muestra el tamaño del genoma en Mb (barras grises) y el número de genes que codifican para sistemas de dos componentes (HK indicadas con puntos azules y RR con puntos verdes) y quinasas de proteínas en serinas y treoninas (puntos rojos). Tomado de [30].

2.2. *Quinasas de proteína en serina y treonina de tipo eucariota*

A pesar de que el primer residuo reportado como fosforilado en proteínas de origen procariota fue una serina [41] no se apreció la importancia de la O-fosforilación hasta dos décadas más tarde. Por el contrario, en procariotas los sistemas de dos componentes se convirtieron en sinónimo de vías de señalización mediadas por fosforilación [36]. Años más tarde se redescubrió la fosforilación en Ser, Thr y Tyr, así como las quinasas responsables de los mismos a las que se denominó de "tipo eucariota" por su aparente similitud con los sistemas descriptos en estos organismos. Hallazgos recientes han llevado a la revisión de la contribución relativa de estas dos clases de fosforilación en procariotas. Por una parte, datos metagenómicos muestran que el número de quinasas de proteína en serina y treonina (STPKs del inglés *serine/threonine protein kinases*) es mayor que el de los sistemas de dos componentes. Y por otro lado, desde el punto de vista evolutivo el término "tipo eucariota" se volvió anacrónico ya que datos genómicos demuestran que las STPKs aparecieron en un organismo ancestral anterior a la divergencia de procariotas y eucariotas [42]. De hecho, se postula que ocurrió transferencia horizontal de genes en ambas direcciones [43].

El genoma de *M. tuberculosis* incluye 11 genes que codifican para STPKs, a saber *pknA* a *pknL* [44] y el criterio utilizado para su anotación en el genoma como STPKs fue la presencia de 12 motivos conservados en el dominio catalítico [45]. El análisis de dominios indica que solo dos de estas enzimas (PknG y PknK) son proteínas citosólicas. El resto de las STPKs micobacterianas presentan un dominio quinasa N-terminal localizado dentro de la célula y que se conecta al dominio sensor C-terminal a través de una o varias hélices transmembrana [30].

2.3. *Fosforilación en tirosinas*

Esta clase de fosforilación fue la última en ser descubierta en procariotas y prácticamente todos los eventos de fosforilación de este tipo se atribuyen a la misma familia de proteínas: las quinasas de tirosina bacterianas o quinasas BY. Hasta hace pocos años se afirmaba que en *M. tuberculosis* no existía fosforilación de tirosinas y estaba bien establecido que las dos fosfatasas de proteínas en tirosinas: PtpA y PtpB (PTP del inglés *protein tyrosine phosphatase*) codificadas por el genoma se secretaban y afectaban vías de transducción de señales en el hospedero [42]. Sin embargo, en 2014 se identificaron residuos de tirosina fosforilados (pTyr) en un fosfoproteoma de *M. tuberculosis* a pesar de que este organismo no posee quinasas BY [46].

A partir de estos datos surge un nuevo nivel de señalización que puede tener implicancias por sí mismo y por intersección con las vías reguladas por STPKs dado que varias de las pTyr identificadas se encuentran en los segmentos de activación de algunas STPKs y son resultado de actividad

autocatalítica. La ausencia de quinasas BY sugiere que en *M. tuberculosis* la fosforilación en tirosinas pueda ocurrir por mecanismos no canónicos tales como STPKs de especificidad dual [42, 46].

2.4. *Desfosforilación*

La desfosforilación catalizada por fosfatasas balancea la acción de las quinasas. En la mayoría de las bacterias el número de quinasas se correlaciona con el de fosfatasas, sin embargo, *M. tuberculosis* es bastante atípica en este aspecto ya que una sola fosfatasa de proteínas en serinas y treoninas: PstP (del inglés *PP2C-family Ser/Thr phosphatase*) contrarresta el efecto de once STPKs [42]. Esto significa que PstP tiene la capacidad de revertir todos los eventos de fosforilación mediados por STPKs en *M. tuberculosis* y según las últimas estimaciones los sitios de fosforilación ascienden a 500 en esta bacteria [47]. Una pregunta relevante es ¿cómo se regula la desfosforilación en la célula? La regulación podría estar mediada por una región desestructurada del dominio extracelular que no tiene homología con ninguna secuencia conocida. Aunque es difícil predecir una función para esta parte de la proteína probablemente participe en la localización de la fosfatasa.

Con el advenimiento de la fosforilación en Tyr las enzimas PtpA (Rv2234) y PtpB (Rv0153c), previamente descriptas como factores de virulencia secretados, adquirieron un posible nuevo rol debido a la aparición de potenciales sustratos en los estudios fosfoproteómicos [48]. El análisis estructural de estas fosfatasas demostró que PtpA podría actuar como una fosfatasa de proteínas en tirosinas mientras que PtpB solamente sería capaz de desfosforilar fosfolípidos [42].

2.5. *Fosfoproteómica en Mycobacterium tuberculosis*

El conjunto de las proteínas fosforiladas en una célula en un momento dado constituye el **fosfoproteoma**. Para su análisis, se requiere la identificación de las fosfoproteínas así como de sus sitios de modificación, el seguimiento temporal de sus cambios y la cuantificación de sus niveles de expresión [49]. Este tipo de abordajes aportan una visión integral de la extensión y dinámica del proceso de fosforilación y fueron posibles gracias a avances tecnológicos tales como: la secuenciación completa de genomas, el desarrollo de herramientas bioinformáticas y de biología de sistemas y por sobre todo la espectrometría de masa (EM) aplicada a proteínas [50].

En *M. tuberculosis* los análisis fosfoproteómicos revelaron más de 500 sitios de modificación en serinas, treoninas y tirosinas (Tabla 1), evidenciando la amplia variedad de procesos en los que participan las STPKs y la complejidad de las cascadas de señalización mediadas por ellas. **Aún hoy se desconocen la mayoría de las quinasas responsables de estos de eventos de fosforilación *in vivo* así como el rol funcional de los mismos** [46, 47, 51-54].

Tabla 1. Recopilación de los estudios fosfoproteómicos en micobacterias. Adaptado de [55].

Especie	#fosfoproteínas identificadas	#sitios	%S	%T	%Y	Referencia
M. tuberculosis H37Rv	301	516	40	60	-	[47]
M. tuberculosis H37Rv	232	-	32	64	4	[46]
M. tuberculosis H37Rv ∆PknE	68	-	-	-	-	[54]
M. bovis BCG	203	289	35	61.6	3.1	[52]
M. smegmatis	76	106	39.47	57.02	3.51	[52]
M. tuberculosis (aislamiento clínico)	214	414	38	59	3	[51]
M. bovis BCG	398	659	39.5	48.7	11.8	[53]

Al día de hoy se han publicado seis estudios fosfoproteómicos en micobacterias y la comparación entre los mismos no es sencilla. Por una parte, es prácticamente imposible determinar si los sitios de fosforilación identificados, de forma única, en cada condición se pueden asignar a diferencias biológicas o a la naturaleza estocástica del proceso de medida. Por otra parte, para poder atribuir significado biológico a un evento discreto de fosforilación se debe validar dicha modificación. A pesar de las dificultades mencionadas y de la variedad de cepas y metodologías de análisis utilizadas, si un evento de fosforilación aparece en más de un estudio fosfoproteómico es poco probable que sea una identificación al azar. Tomando lo anterior en consideración, se pudo compilar una primera lista de sitios de fosforilación que podrá ser usada para la selección racional de sustratos a validar [55].

Se ha demostrado además que los patrones de fosforilación cambian drásticamente en respuesta a estímulos ambientales y que la fosforilación coordinada de múltiples STPKs permite dar una respuesta integrada a las señales externas. Sin embargo, **nuestro conocimiento respecto de la arquitectura de estas vías de señalización así como de las redes de interacciones de proteínas que las sustentan es escaso** [56].

3. Quinasas de proteínas en serinas y treoninas en *Mycobacterium tuberculosis*

La secuenciación del genoma de *M. tuberculosis* reveló la importancia en este microorganismo de los sistemas de fosforilación en serinas y treoninas. De hecho, el genoma de *M. tuberculosis* codifica para once STPKs (*pknA* a *pknL*) y una única fosfoserina/treonina fosfatasa (*pstP*) [44].

3.1. *Características de secuencia de las quinasas de proteínas en serinas y treoninas micobacterianas*

El análisis de secuencia del dominio catalítico de las STPKs micobacterianas, y su comparación con otras quinasas de proteína en serina y treonina procariotas, mostró que las primeras forman conglomerados de proteínas homólogas bien delimitados. Las nueve STPKs transmembrana

pertenecen a tres clados distintos pero cercanos: PknA/PknB/PknL; PknD/PknE/PknH y PknF/PknI/PknJ mientras que las enzimas citosólicas (PknG y PknK) se separan claramente de las demás. Esta observación permitió sugerir que los genes que codifican para las STPKs transmembrana derivarían de un ancestro común entretanto que PknG y PknK habrían sido adquiridas separadamente [57]. A pesar de estas diferencias, las STPKs de *M. tuberculosis* se parecen más entre ellas que a sus homólogas humanas, con las que tienen una homología de secuencia menor al 30% (Fig. 6). **La baja identidad de secuencia sumada a la experiencia acumulada en el diseño de inhibidores específicos de las STPKs eucariotas posicionan a esta familia de enzimas micobacterianas como blancos interesantes para el desarrollo de nuevos antibióticos** [30].
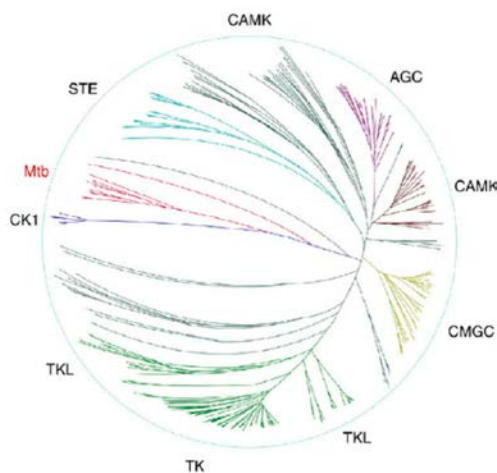
Figura 6. Dendograma comparativo de múltiples secuencias de dominios quinasa. Árbol de distancia donde se indican todas las STPKs micobacterianas (marcadas en rojo) y algunas enzimas humanas representativas de los distintos grupos definidos por Manning *et al.* [58]: AGC (conteniendo a las familias PKA, PKC y PKG), CAMK (proteína quinasa dependiente de calcio/calmodulina), CK1 (quinasa de caseína 1), CMGC (conteniendo a las familias CDK, MAPK, GSK3 y CLK), STE (homólogos del gen de levadura *Ste7*, *Ste11* y *Ste20*), TK (quinasa de proteína en tirosina) y TKL (proteínas del tipo quinasa de proteína en tirosina). Adaptado de [30].

A diferencia del dominio catalítico, los dominios extracelulares de las quinasas transmembrana no poseen similitud de secuencia (Fig. 7), indicando que probablemente respondan a distintas señales. En el caso de las quinasas citosólicas aparecen también dominios regulatorios por fuera del dominio quinasa [57].
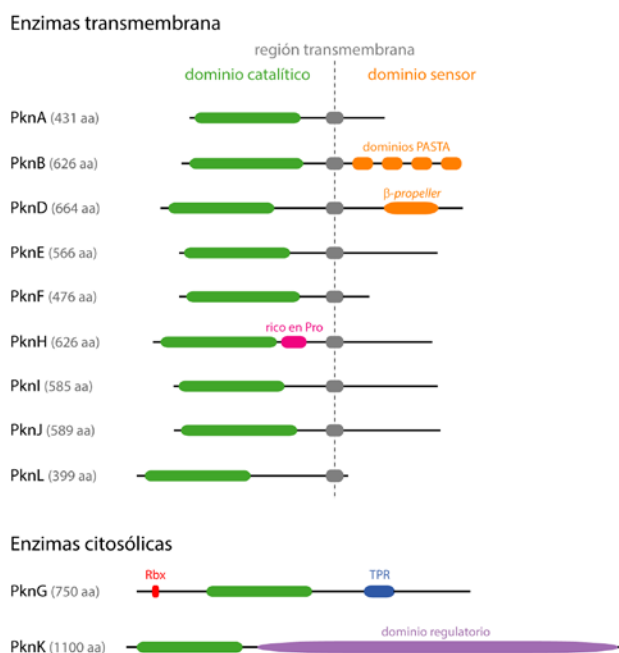
Figura 7. Estructura general de las STPKs de *M. tuberculosis*. Representación esquemática de las distintas regiones que componen a las STPKs de membrana y citosólicas. Dominios PASTA: del inglés ***P**enicilin **b**inding **p**rotein **A**nd **S**erine/**T**hreonine kinase **A**ssociated*, Rbx: dominio rubredoxina y TRP: motivo repetido tetratricopéptido. Adaptado de [59].

3.2. *Análisis estructural del dominio quinasa*

La elucidación estructural de los dominios catalíticos de tres STPKs (PknB [60, 61], PknE [62] y PknG [63]) permitió comenzar a comprender los mecanismos de activación y regulación de este tipo de quinasas. Estas proteínas exhiben una estructura bilobular donde la porción N-terminal está compuesta principalmente por láminas β y presenta una sola hélice α denominada hélice C; por el contrario, el lóbulo C-terminal contiene solamente hélices α. Esta organización espacial da lugar a una hendidura donde se localiza el sitio activo (Fig. 8). A pesar de la baja conservación de secuencia entre las proteínas micobacterianas y eucariotas, las estructuras cristalográficas demostraron una importante conservación del plegamiento global y de la maquinaria catalítica entre ambas clases de quinasas [30, 57]. A continuación se utilizará la estructura cristalográfica de PknB de *M. tuberculosis* formando un complejo con un análogo de ATP para ilustrar estas características.

La estructura de PknB que se muestra en la Figura 8 incluye todos los elementos que son comunes a las quinasas de proteínas eucariotas. La hélice C es un motivo estructural crítico ya que cuando se aproxima al sitio de unión del nucleótido da lugar a una enzima activa o "cerrada" [64]. En el caso de la estructura de PknB, la posición de la hélice C, próxima al sitio de unión del ATP, permite afirmar que se está frente a la conformación activa de la quinasa. La presencia de un bucle rico en glicinas en la región aminoterminal, que interactúa con el ATP, y de los residuos aspártico, fenilalanina y glicina (DFG), que contribuyen a posicionar a los iones Mg^{2+}, son también características conservadas en la estructura de esta quinasa [57].

El bucle de activación, secuencia de longitud y composición variable comprendida entre dos grupos de residuos altamente conservados (DFG y APE) y que participa en el control de la actividad enzimática [65], es visible en las estructuras cristalográficas de quinasas activas. Sin embargo, cuando se analiza la estructura de PknB en su conformación activa se determina que el bucle de activación no es completamente visible ya que está desordenado [60, 61] sugiriendo un modo de unión del sustrato del tipo encaje inducido [30]. Asimismo, se demostró que el bucle de activación está fosforilado sistemáticamente en dos residuos de treonina y que su estado de fosforilación regula la actividad quinasa de la misma forma que lo hacen esas secuencias en sus contrapartes eucariotas, es decir, se requiere la fosforilación de la misma para la completa activación de PknB [66]. Aunque la fosforilación del bucle de activación es necesaria, la primera etapa en la activación de PknB y otras quinasas (PknD [67] y PknE [62]) es la formación de un dímero por interacción de los dominios quinasa desfosforilados [68]. **En el caso de PknG el bucle de activación está ordenado y estabilizado a pesar de la ausencia de fosforilación, lo que sugiere un mecanismo de activación diferente [57].**
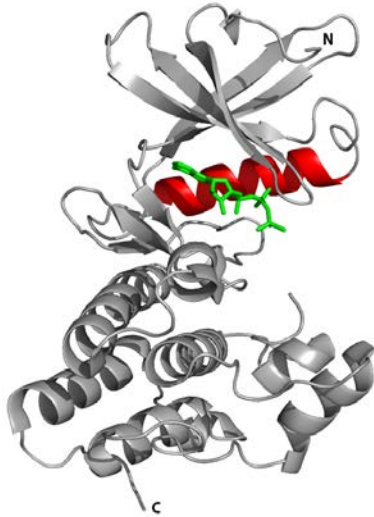
Figura 8. Representación tridimensional de la estructura de rayos X del dominio quinasa de PknB de *M. tuberculosis*. Se indican elementos estructurales importantes: el nucleótido (ATP-γ-S) ocupando el sitio activo se muestra en verde y la hélice C en rojo. La figura fue generada utilizando el programa PyMOL Molecular Graphics System (versión 1.3 Schrödinger, LLC), acceso PDB: 1MRU.

3.3. *Procesos regulados por quinasas de proteínas en serinas y treoninas*

Uno de los mayores desafíos en el área de la fosforilación es la elucidación de las redes de señalización mediadas por quinasas. Recientemente se han desarrollado aproximaciones fosfoproteómicas que han sido claves para el análisis global de los eventos de fosforilación que ocurren en las células; aunque no permiten la identificación de las quinasas directamente responsables de la modificación. Múltiples abordajes se han utilizado para la identificación de los pares quinasa-sustrato: modificación del sitio activo asociado a la utilización de análogos de ATP que "etiquetan" al sustrato, bibliotecas de péptidos y proteínas, sistemas de expresión ectópica o inhibidores específicos [69, 70].

En el caso de *M. tuberculosis*, el trabajo de numerosos grupos de investigación ha permitido comenzar a identificar los sustratos de las STPKs (Tabla 2) y a partir de ellos definir los roles fisiológicos de las mismas. No obstante, la búsqueda de sustratos se ha basado fuertemente en el supuesto de que la fosforilación *in vitro* proporciona información sobre la identidad de los sustratos *in vivo*. En las células, los sistemas de fosforilación están estrictamente regulados pero ese control se pierde en los experimentos *in vitro*, por lo que la relevancia fisiológica de los sustratos identificados debe ser confirmada ya que está bien establecida la promiscuidad de las STPKs *in vitro* [71, 72]. Esto ha llevado a que muchas de las proteínas identificadas, aparentemente, sean sustrato de más de una quinasa *in vitro*.

Tabla 2. Recopilación de los sustratos reportados de STPKs de *M. tuberculosis*. Adaptado de [73].

Quinasa	Sustrato	Identificación	Proceso en el que participa	Referencia
PknA	EmbR	*In vitro*	Metabolismo de la pared celular	[74]
	FabD	*In vitro*	Metabolismo de la pared celular	[75]
	FabH	*In vivo/in vitro*	Metabolismo de la pared celular	[76]
	FtsZ	*In vivo*	División celular y crecimiento	[77]
	GlmU	*In vivo/in vitro*	División celular y crecimiento	[78]
	GroEL1	*In vitro*	Respuesta frente al estrés	[79]
	KasA	*In vitro*	Metabolismo de la pared celular	[75]
	KasB	*In vitro*	Metabolismo de la pared celular	[75]
	MurD	*In vitro*	División celular y crecimiento	[80]
	Rv1422	*In vivo*	---	[81]
	Wag31	*In vivo*	División celular y crecimiento	[81]
PknB	EmbR	*In vitro*	Metabolismo de la pared celular	[74]
	FabD	*In vitro*	Metabolismo de la pared celular	[75]
	GarA	*In vivo*	Metabolismo energético	[82]
	GlmU	*In vivo/in vitro*	División celular y crecimiento	[78]
	GroEL1	*In vitro*	Respuesta frente al estrés	[79]
	KasA	*In vitro*	Metabolismo de la pared celular	[75]
	KasB	*In vitro*	Metabolismo de la pared celular	[75]
	FhaA	*In vitro*	División celular y crecimiento	[83, 84]
	FhaB	*In vitro*	División celular y crecimiento	[85]
	MviN	*In vitro*	División celular y crecimiento	[86]
	OprA	*In vitro*	Respuesta frente al estrés	[87, 88]
	PapA5	*In vitro*	División celular y crecimiento	[85]
	PbpA	*In vivo*	División celular y crecimiento	[89]
	RshA	*In vivo*	Respuesta al estrés	[90]
	Rv1422	*In vivo*	---	[81]
	Rv1747	*In vitro*	División celular y crecimiento	[83]
	SigH	*In vivo*	Respuesta al estrés	[90]
	Wag31	*In vivo*	División celular y crecimiento	[81]
PknD	FabD	*In vitro*	Metabolismo de la pared celular	[75]
	FabH	*In vivo/in vitro*	Metabolismo de la pared celular	[76]
	GarA	*In vitro*	Metabolismo energético	[82]
	GroEL1	*In vitro*	Respuesta frente al estrés	[79]
	KasA	*In vitro*	Metabolismo de la pared celular	[75]
	KasB	*In vitro*	Metabolismo de la pared celular	[75]
	MmpL7	*In vivo*	Metabolismo de la pared celular	[91]
	OprA	*In vivo*	Respuesta frente al estrés	[87, 88]
	Rv1747	*In vitro*	División celular y crecimiento	[83]
PknE	FabD	*In vitro*	Metabolismo de la pared celular	[75]
	FabH	*In vivo/in vitro*	Metabolismo de la pared celular	[76]
	GarA	*In vitro*	Metabolismo energético	[82]
	GroEL1	*In vitro*	Respuesta frente al estrés	[79]
	KasA	*In vitro*	Metabolismo de la pared celular	[75]
	KasB	*In vitro*	Metabolismo de la pared celular	[75]
	OprA	*In vitro*	Respuesta frente al estrés	[87, 88]
	RsfA	*In vitro*	Regulación de la expresión génica	[87]
	RshA	*In vitro*	Respuesta al estrés	[87]
	Rv1747	*In vitro*	División celular y crecimiento	[92]
	Rv1904	*In vitro*	---	[87]
PknF	FabD	*In vitro*	Metabolismo de la pared celular	[75]
	FabH	*In vivo/in vitro*	Metabolismo de la pared celular	[76]
	GarA	*In vitro*	Metabolismo energético	[82]
	GroEL1	*In vitro*	Respuesta frente al estrés	[79]

	KasA	*In vitro*	Metabolismo de la pared celular	75
	KasB	*In vitro*	Metabolismo de la pared celular	75
	FhaA	*In vitro*	División celular y crecimiento	83
	Rv1747	*In vitro*	División celular y crecimiento	92
PknG	GarA	*In vivo/in vitro*	Metabolismo energético	93
PknH	DacB1	*In vitro*	Metabolismo de la pared celular	94
	DosR	*In vitro*	Respuesta al estrés	95
	EmbR	*In vivo*	Metabolismo de la pared celular	96
	FabD	*In vitro*	Metabolismo de la pared celular	75
	FabH	*In vivo/in vitro*	Metabolismo de la pared celular	76
	GroEL1	*In vitro*	Respuesta frente al estrés	79
	KasA	*In vitro*	Metabolismo de la pared celular	75
	KasB	*In vitro*	Metabolismo de la pared celular	75
	Rv0681	*In vitro*	Regulación de la expresión génica	94
PknI	FabD	*In vitro*	Metabolismo de la pared celular	75
PknJ	PykA	*In vitro*	Metabolismo energético	97
PknK	FabD	*In vitro*	Metabolismo de la pared celular	75
	VirS	*In vivo*	Metabolismo de la pared celular	98
PknL	FabD	*In vitro*	Metabolismo de la pared celular	75
	GroEL1	*In vitro*	Respuesta frente al estrés	79
	KasA	*In vitro*	Metabolismo de la pared celular	75
	KasB	*In vitro*	Metabolismo de la pared celular	75
	Rv2175	*In vitro*	Regulación de la expresión génica	99

A continuación se hará una presentación somera del conocimiento actual de los procesos potencialmente regulados por cada una de las STPKs.

3.3.1   *PknA y PknB*

Estas dos quinasas están codificadas por un operón altamente conservado en actinobacterias y que está involucrado en el mantenimiento de la forma de la célula así como en el proceso de división celular [100]. Durante la etapa de crecimiento exponencial los niveles de expresión de este operón aumentan respecto a los de la fase estacionaria, lo que sugiere que la acción regulatoria de estas quinasas es necesaria durante la replicación. En la misma línea que lo anterior, datos de *microarrays* indican que la expresión de *pknA* y *pknB* disminuye en respuesta a la inanición [101]. Asimismo la inducción experimental de cambios en los niveles de expresión de estas dos quinasas esenciales altera el crecimiento y la morfología celular [81].

La identificación y caracterización de algunos sustratos de estas quinasas validó los datos previos que postulaban un rol en la síntesis de la pared celular. Algunas de las fosfoproteínas identificadas fueron: PBPA, una proteína de unión a penicilina (PBP del inglés *penicilin-binding protein*), que participa en el entrecruzamiento del PG durante la elongación y la división celular y que es un posible sustrato de PknB [89]. Wag31, proteína homóloga a DivIVA, un sustrato de PknA necesario para la síntesis del PG en micobacterias y para el correcto posicionamiento del septo durante la división [81]. Y MviN, que fue propuesta como una de las enzimas que participan en la inversión del

Glc-Nac/Mur-Nac-pentapéptido desde la cara citoplasmática de la membrana plasmática hacia el espacio extracelular (Fig. 1), como sustrato *in vitro* de PknB. Por otra parte, se demostró que FhaA reconoce una treonina fosforilada en MviN y que la formación de este complejo inhibe las etapas finales de la síntesis del PG [86]. FhaA, al igual que FhaB, es una proteína que posee dominios FHA (del inglés *ForkHead Associated*) codificada por el mismo operón que PknA y PknB y que ha sido reportada como sustrato *in vitro* de PknB [83-85]. Los dominios FHA son pequeños módulos proteicos que median interacciones proteína-proteína a través del reconocimiento de fosfotreoninas (pThr) [102] y en el caso de MviN y FhaA la interacción está mediada por dicha región [86]. El rol de PknA en el control de la división celular se confirmó en estudios que demuestran la interacción directa de la quinasa con FtsZ, una proteína central en la formación del septo bacteriano. La fosforilación de FtsZ disminuye su capacidad de polimerización dependiente de GTP [77].

PknA y PknB han sido vinculadas, directa o indirectamente, a la regulación de otros procesos celulares más allá de la síntesis de la pared celular: síntesis de lípidos, división celular y regulación de la transcripción (Tabla 2) [57]. En este sentido, nuestro grupo determinó que PknB fosforilaba eficientemente a GarA (del inglés *Glycogen accumulation regulator*) en extractos celulares de *M. tuberculosis* [82]. Posteriormente se demostró que la proteína ortóloga a GarA en *Corynebacterium glutamicum* inhibía a una enzima clave en el ciclo de los ácidos tricarboxílicos: la 2-oxoglutarato deshidrogenasa [103]. La participación de estas quinasas en múltiples procesos sugiere que podrían coordinar globalmente el crecimiento celular [57].

3.3.2 *PknD*

La comparación de los patrones de fosforilación *in vitro* de una cepa salvaje de *M. tuberculosis* con una mutante en la que el dominio quinasa de PknD había sido removido permitió identificar a MmpL7 (Rv2942), un transportador unido a membrana, como sustrato de la quinasa [91]. Esta proteína es esencial para la virulencia y se especula que esto se debe a su función como transportador de PDIM y otros glicolípidos fenólicos. Aunque el rol de la fosforilación de este transportador no se ha elucidado, se sugiere que podría regular el transporte de componentes a través de la membrana [104].

M. tuberculosis posee la habilidad de adaptarse a entornos de osmolaridad variable, desde las finas gotas en las que ingresa al hospedero hasta los granulomas, donde se altera la turgencia de la célula perjudicando el plegamiento de las proteínas así como su actividad metabólica. Recientemente se propuso un mecanismo mediado por PknD que regula la adaptación de *M. tuberculosis* al estrés osmótico. El incremento de la osmolaridad extracelular provoca la activación de PknD y la fosforilación de un factor anti-anti σ: OprA. La fosforilación de OprA permite que el factor σ SigF se

una a la ARN polimerasa lo que promueve la transcripción de *oprA*, *espA* y otros genes asociados a cambios en la estructura de la pared celular y el metabolismo [88].

3.3.3 *PknE*

Para asegurar su supervivencia en el hospedero, *M. tuberculosis* inhibe la apoptosis de macrófagos infectados. Se ha postulado que PknE sería un regulador antiapoptótico y que tiene un rol en la respuesta al estrés inducido por óxido nítrico (•NO) [105]. Estudios proteómicos y fosfoproteómicos de un mutante de deleción de esta quinasa mostraron que la misma participa en la respuesta a •NO regulando múltiples procesos que permiten la adaptación de la bacteria al estrés. Se observaron cambios en: proteínas controladas por el factor σ SigB, proteínas que participan en la entrada a quiescencia, sistemas de dos componentes, proteínas que participan en la transcripción y la traducción, proteínas antioxidantes, proteínas implicadas en división celular y metabolismo energético [54].

3.3.4 *PknF*

Experimentos *in vitro* e *in vivo* demostraron que Rv1747, un transportador ABC que presenta dos dominios FHA y que es necesario para el crecimiento de la bacteria en macrófagos, pertenece al mismo operón que PknF y es sustrato de la quinasa. Para evaluar el rol de la fosforilación sobre la actividad de Rv1747 se utilizó un mutante de deleción y esta cepa se complementó con un plásmido que codificaba para Rv1747 pero con los sitios de fosforilación mutados a alanina. Cuando se infectaron macrófagos murinos con la cepa complementada se observó atenuación del crecimiento, lo que indica que la fosforilación de Rv1747 es imprescindible para su actividad. Sin embargo, cuando se evalúa la infección de macrófagos por una cepa de *M. tuberculosis* Δ*pknF* no se observan alteraciones del fenotipo de crecimiento, por lo que se deduce que Rv1747 puede ser fosforilada por otras STPKs [92, 106]. Otro reporte indica que PknF regula el transporte de glucosa a través de la membrana, este dato asociado a la fosforilación del transportador Rv1747 podría sugerir que la quinasa regula negativamente el transporte de glucosa mediante la fosforilación del transportador [107].

3.3.5 *PknH*

Experimentos *in vitro* demostraron que PknH fosforila a EmbR, regulador transcripcional del operón *embCAB* que codifica para tres arabinosiltransferasas claves para la síntesis de AG y LAM. La interacción entre ambas proteínas está mediada por un dominio FHA presente en la región C-terminal de EmbR [104, 108]. Además, el análisis comparativo de una cepa de *M. tuberculosis* deficiente en *pknH* con una cepa salvaje mostró una disminución de PDIM específicos. Por lo que además de

regular los glicolípidos de la envoltura celular PknH contribuye a la síntesis de componentes de la pared celular [109].

El análisis proteómico comparativo de la cepa deficiente en PknH frente a la cepa parental mostró que varias proteínas pertenecientes al regulón DosR estaban aumentadas en la cepa mutante. Posteriormente, se demostró la fosforilación *in vitro* de DosR por PknH y que esta modificación promueve la unión al ADN de la primera. Tomando estos datos en conjunto es que se propone a la fosforilación catalizada por PknH como un mecanismo de regulación de la dormancia en *M. tuberculosis*. Asimismo, estos resultados mostraron por primera vez la convergencia entre los dos sistemas principales de transducción de señales en *M. tuberculosis*: STPKs y sistemas de dos componentes [95].

3.3.6   *PknI y PknJ*

Los resultados de infección de macrófagos con una cepa de *M. tuberculosis* deficiente en PknI mostraron que la activación de la quinasa podría estar mediada por señales generadas en el hospedero tales como pH ácido y baja disponibilidad de oxígeno, y que una vez activa esta enzima disminuiría la velocidad de crecimiento de la bacteria, conteniendo la infección [110].

Se postula que PknJ regula la actividad de la piruvato quinasa (PykA) por fosforilación. PykA es una enzima fundamental en el metabolismo energético de la bacteria ya que es la encargada de la producción de piruvato y ATP a partir de fosfoenolpiruvato y ADP. Además de la vía glicolítica, PykA participa del metabolismo de ácidos grasos y aminoácidos ya que genera piruvato, un precursor importante de estas vías [97].

3.3.7   *PknK y PknL*

El operón *mymA* (del inglés *mycobacterial monooxygenase*) es fundamental para el mantenimiento de la integridad de la pared celular en *M. tuberculosis* ya que codifica para varias proteínas involucradas en la biosíntesis de los ácidos micólicos. La transcripción de este operón depende del regulador transcripcional VirS que es asimismo sustrato de PknK, se demostró además que una vez fosforilado VirS aumenta su afinidad por el promotor de *mymA*. Ensayos *in vitro* indicaron que varias de las proteínas codificadas por *mymA* son también sustrato de PknK [98].

Una vía de transducción de señales que podría activarse en respuesta a cambios en el ambiente es la mediada por PknL y su sustrato Rv2175c. Rv2175c es una pequeña proteína de unión al ADN y la fosforilación de la T_9 modula negativamente su capacidad para formar complejos proteína-ADN [111].

En esta sección no se hará referencia alguna a **PknG** ya que por tratarse del sujeto de estudio de esta tesis se le dedicará la sección siguiente.

En el contexto de las vías de señalización mediadas por STPKs, la presencia de múltiples enzimas sugiere la existencia de cascadas y redes, sin embargo, en *M. tuberculosis* hay poca evidencia que apoye la existencia de un sistema organizado [42]. Recientemente se evalúo la interacción de las STPKs de *M. tuberculosis* entre ellas y se demostró que algunas fosforilan a otras *in vitro*, lo que constituye uno de los requisitos para la existencia de una cascada de señalización. Los autores del trabajo proponen una red de STPKs con una estructura en tres niveles: reguladores maestros, transductores de señales y sustratos terminales lo que permitiría plantear la existencia de amplificación y propagación de las señales en *M. tuberculosis* [56].

De las once STPKs de *M. tuberculosis* ocho se expresan en modelos de infección en cobayos [112] y entre ellas se destaca PknG como una de las moléculas centrales para la sobrevida de la bacteria en el hospedero.

4. PknG

PknG es una de las dos STPK citosólicas presentes en *M. tuberculosis* y es la única quinasa sin regiones transmembrana que se mantuvo en el genoma de *Mycobacterium leprae*. Este patógeno, estrechamente emparentado con *M. tuberculosis*, exhibe una reducción significativa del número total de genes y se postula que conservó el conjunto mínimo de genes necesarios para la vida [113].

4.1. *Análisis estructural*

Desde el punto de vista estructural PknG es una enzima singular ya que posee una combinación de dominios que no ha sido reportada en otras quinasas. A saber: una región N-terminal desestructurada (residuos 1-98), un dominio rubredoxina (Rbx, residuos 99-137), un dominio catalítico quinasa (residuos 138-405) y un dominio C-terminal que presenta un motivo repetido tetratricopéptido (TPR, residuos 406-750) (Fig. 9).
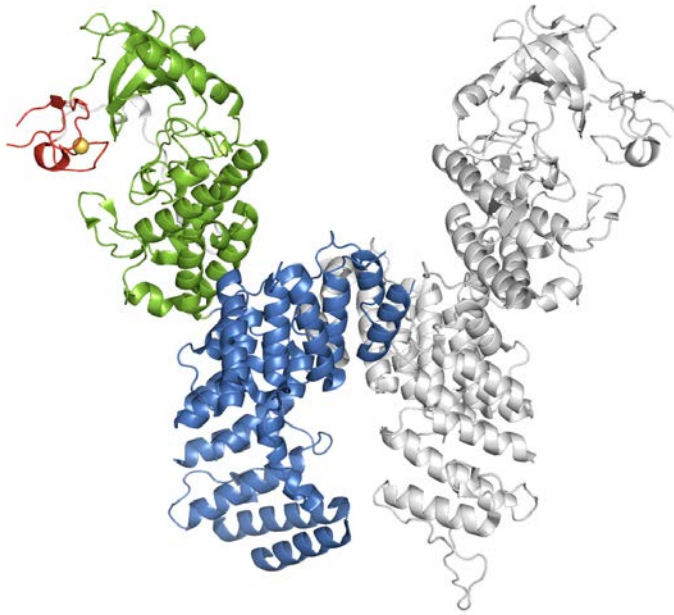
Figura 9. Representación tridimensional de la estructura de rayos X de PknG de _M. tuberculosis_. Esta estructura fue obtenida a partir de una construcción de PknG a la que le falta la región N-terminal desestructurada (PknG$_{\Delta73}$). Se indican los dominios rubredoxina (rojo), quinasa (verde) y TPR (azul). En gris se indica el segundo monómero y la esfera naranja representa el átomo de cadmio con el cual la proteína fue cristalizada. La figura fue generada utilizando el programa PyMOL Molecular Graphics System (versión 1.3 Schrödinger, LLC), acceso PDB: 2PZI.

4.1.1 _Región N-terminal (dominio rubredoxina)_

Este pequeño dominio globular que se extiende desde el aminoácido 99 al 137 contiene dos motivos CxxCG que unen hierro (donde x representa un aminoácido cualquiera) y que caracterizan a las **rubredoxinas** [114]. Las Rbx son pequeñas proteínas ferrosulfuradas que participan en la transferencia de electrones en microorganismos anaerobios o microaerofílicos y su sitio activo está constituido por un átomo de hierro, que alterna entre los estados de oxidación Fe^{+2} y Fe^{+3}, coordinado a cuatro residuos de cisteína [115].

En PknG el dominio Rbx interacciona con los dos lóbulos del dominio quinasa y se localiza sobre el sitio activo sin ocluir el acceso al mismo. La sustitución de todas las cisteínas del dominio Rbx (Cys$_{106}$, Cys$_{109}$, Cys$_{128}$ y Cys$_{131}$) por serinas provoca la disrupción completa de ese dominio y la consiguiente pérdida total de actividad quinasa [63]. Por otra parte, aunque los mecanismos de activación/inactivación de PknG no han sido elucidados, resultados de Tiwari _et al._ sugieren que la actividad de esta enzima podría estar modulada por el estado redox del microambiente celular a través del dominio Rbx [116] lo que justifica su estudio en mayor profundidad.

4.1.2 _Región central (dominio quinasa)_

El dominio quinasa se intercala entre los dominios Rbx y C-terminal adoptando la estructura clásica de dos lóbulos de las STPKs. Sin embargo, presenta diferencias respecto a la estructura de PknB mencionada en la _sección 3.2_. Primero, el bucle de activación que controla la actividad catalítica en otras quinasas es bastante atípico. Es más corto y no presenta las treoninas que están conservadas y fosforiladas en otras STPKs micobacterianas, además está estabilizado en una conformación

ordenada y extendida. Segundo, a pesar de que la fosforilación es uno de los mecanismos más frecuentes de posicionamiento del bucle de activación, en el caso de PknG esta sección no está fosforilada lo que sugiere un mecanismo de activación diferente. Tercero, se constató la ausencia de un residuo de arginina (R) que precede a un residuo de aspartato (D) altamente conservado en el bucle catalítico lo que clasifica a PknG como la única quinasa micobacteriana no RD. En esta clase de quinasas no se ha observado fosforilación del segmento de activación [65] y se ha postulado que en PknG la ausencia de fosforilación da lugar a una conformación única del bucle de activación [63].

4.1.3   *Región C-terminal (motivo repetido tetratricopéptido)*

La secuencia C-terminal se inicia en el aminoácido 406 y contiene un **motivo repetido tetratricopéptido** (TPR del inglés *tetratricopeptide repeat*) que interacciona con el extremo carboxiloterminal del dominio quinasa y contacta lateralmente con el dominio TPR del segundo monómero en la unidad asimétrica del cristal [63]. Lo observado en el cristal es consistente con la función reportada de los dominios TPR ya que estos repetidos de 34 aminoácidos, que usualmente se presentan en múltiples copias, median interacciones proteína-proteína [117]. En la *sección 3.2* se mencionó que PknB dimeriza a través de su dominio quinasa y que esto media la transfosforilación del bucle de activación [118]. En cambio, en PknG el dominio TPR es el que media la dimerización y como resultado los dominios quinasa quedan alejados entre sí. A pesar de que el rol del TPR no es claro, se determinó que la actividad quinasa disminuye a la mitad en una construcción a la que le falta esta región [116].

4.1.4   *Región desestructurada aminoterminal*

La secuencia previa al dominio Rbx constituye una región desestructurada que está ausente en la estructura cristalográfica. Nuestro grupo demostró que PknG tiene actividad autocatalítica tanto *in vitro* como *in vivo* y que se incorporan hasta cuatro fosfatos en serinas o treoninas de este segmento: T_{23}, S_{31} o T_{32} (el residuo modificado no pudo ser inequívocamente identificado) y T_{63}, T_{64} o S_{65} (los dos residuos modificados no pudieron ser inequívocamente identificados), no obstante este fenómeno no afecta la actividad quinasa de la enzima [93]. Para otras quinasas de *M. tuberculosis* está bien demostrado el papel que cumple la autofosforilación en la regulación de la actividad enzimática, por ejemplo en PknB la autofosforilación del bucle de activación es necesaria para obtener una enzima completamente activa [66, 119].

Al día de hoy no se comprende si esta combinación tan particular de dominios afecta la actividad catalítica de PknG o su interacción con otras moléculas.

4.2. *Relevancia funcional*

4.2.1   *Rol en la supervivencia en el hospedero*

La generación de una cepa de *M. tuberculosis* deficiente en *pknG* (Δ*pknG*) permitió evaluar el efecto de esta proteína sobre la viabilidad celular en cultivo y la virulencia *in vivo*. Se observó que en la fase de crecimiento exponencial el número de bacilos viables era diez veces menor al de la cepa salvaje (H37Rv). Para evaluar la virulencia de la cepa Δ*pknG* se utilizaron ratones con el síndrome de inmunodeficiencia combinada severa (SCID del inglés *severe combined immunodeficiency*) que se sabía eran susceptibles a la infección por *M. tuberculosis*. Se demostró que la mediana del tiempo de supervivencia de los ratones SCID infectados con Δ*pknG* era el doble que la de los infectados con la cepa salvaje. Se observó incluso atenuación del crecimiento bacteriano cuando se ensayó la cepa Δ*pknG* en ratones inmunocompetentes. **En resumen, la remoción de *pknG* origina una cepa atenuada de *M. tuberculosis* [120]**.

Un elemento central en la patogenicidad de *M. tuberculosis* es su capacidad para sobrevivir en el ambiente hostil de los macrófagos del hospedero, interfiriendo con el proceso de maduración del fagosoma. Mientras que una cepa salvaje de *M. tuberculosis* es capaz de sobrevivir en el macrófago infectado, micobacterias deficientes en PknG muestran un fenotipo atenuado y son rápidamente degradadas en fagosomas maduros, lo que ha permitido demostrar que **PknG participa, directa o indirectamente, en la inhibición de la fusión fagolisosomal [121]**.

El análisis de la sobrevida de una cepa de BCG deficiente en *pknG*, en un modelo de infección de macrófagos, mostró que los bacilos son rápidamente transferidos a los lisosomas y destruidos por la maquinaria bactericida del organelo. Lo mismo ocurre cuando la infección se realiza con una cepa de BCG que expresa una quinasa inactiva (PknG^{K181M}) [122] o una construcción que carece de los sitios de autofosforilación [123]. La modulación de la fusión fagolisosomal podría estar mediada por la fosforilación de sustratos en el hospedero pero para ello se requiere la secreción de PknG en el macrófago [121].

Asimismo ensayos de infección utilizando BCG viable e inactivada por calor mostraron que PknG es activamente secretada al citosol del macrófago [121]. Sin embargo, estos hallazgos han sido intensamente cuestionados ya que más allá de este reporte no se ha podido demostrar la secreción de PknG, ni tampoco la fosforilación de sustratos en el hospedero. Datos recientes sugieren que a pesar de la ausencia de péptido señal, PknG de *Mycobacterium marinum* se localizaría en la envoltura celular y que la translocación estaría mediada por el sistema de secreción SecA2. Se

propone incluso que sería la región N-terminal desestructurada la que mediaría el transporte a través de la membrana interna [124]. En definitiva, **el mecanismo por el cual PknG media la supervivencia de la bacteria en modelos celulares y animales es aún hoy desconocido.**

4.2.2   *Rol en el metabolismo bacteriano*

El gen *pknG* es parte de un operón, conservado en micobacterias y otros *Actinomycetales* estrechamente relacionados, que codifica para proteínas involucradas en la captación de glutamina [125]. Hasta la década del 2000, el rol directo de PknG en la regulación del metabolismo del glutamato en micobacterias era controversial [120, 126], sin embargo, nuestro grupo demostró que esta quinasa participa en la regulación del metabolismo intermediario a través de la fosforilación de una pequeña proteína: GarA [93].

GarA (del inglés *Glycogen accumulation regulator*) fue descripta en *Mycobacterium smegmatis* asociada al metabolismo del glucógeno [127]. Se la identificó además como sustrato de PknB en extractos celulares de *M. tuberculosis* [82] y como sustrato *in vitro* de la mayoría de las STPKs (Tabla 2). GarA es una pequeña proteína de 162 aminoácidos que en su región C-terminal presenta un dominio FHA y cuya extensión N-terminal contiene un motivo ETTS conservado [93]. Nuestro grupo de investigación demostró que **GarA es sustrato *in vitro* de PknG** y que la modificación ocurre en la T_{21} (perteneciente al motivo ETTS) [93], a diferencia de lo previamente demostrado para PknB donde la T_{22} es la fosforilada [82]. Asimismo se probó que estos eventos de fosforilación son mutuamente excluyentes lo que señalaría a GarA como intermediario en más de una vía de señalización [93]. Más aún, resultados de nuestro grupo mostraron que *in vivo* PknG y GarA interactúan y que la T_{21} es el residuo predominantemente fosforilado. Estos datos señalan a **GarA como un sustrato fisiológico de PknG**.

Se mencionó previamente que la autofosforilación del extremo N-terminal de PknG no es necesaria para la catálisis, sin embargo, cuando se determina la actividad frente a GarA de una construcción de la quinasa en la que estos sitios están ausentes (PknG$_{\Delta 73}$) se observa que es cinco veces menor a la de la forma nativa. Además, resultados de resonancia plasmónica de superficie (SPR) mostraron que los residuos fosforilados en el extremo N-terminal de la quinasa son necesarios para que ocurra la interacción entre PknG y GarA. Se demostró igualmente que la formación del complejo PknG-GarA solamente ocurre en ausencia de ATP, en presencia del nucleótido GarA es fosforilada por PknG y la interacción entre las proteínas no es posible. **Estos datos apuntan a una interacción de alta afinidad mediada por la unión específica del dominio FHA de GarA a residuos de fosfotreonina en PknG** [93].

El análisis estructural de GarA por resonancia magnética nuclear permitió proponer un modelo en el que el estado de fosforilación de esta proteína regula su interacción con otras. Se determinó que la forma desfosforilada presenta dos regiones disponibles para la interacción con proteínas: el péptido N-terminal flexible y el sitio de unión a fosfotreoninas localizado en el dominio FHA. La fosforilación de la extensión N-terminal de GarA por PknB (T_{22}) o PknG (T_{21}) promueve una fuerte interacción entre dicho segmento y el dominio FHA (Fig. 10). Este importante cambio conformacional bloquea el sitio de unión de FHA y por tanto inhibe la unión de las quinasas así como de otros eventuales blancos moleculares [128, 129].
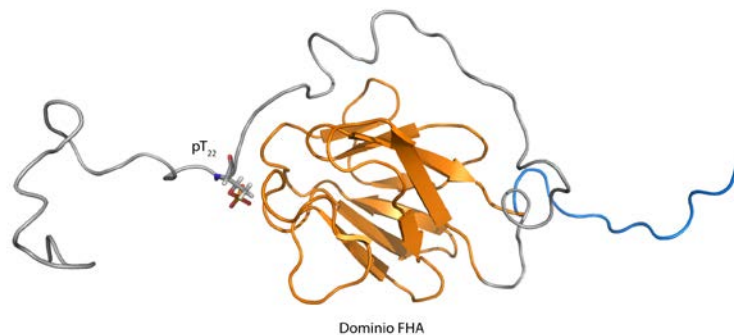
Figura 10. Representación tridimensional de la estructura en solución de GarA de *M. tuberculosis* fosforilada por PknB. La región N-terminal indicada en gris incluye a la T_{22} fosforilada por PknB. Entre los residuos 55 y 149 la proteína adopta el plegamiento clásico de 11 láminas β de los dominios FHA (indicado en naranja). La figura fue generada utilizando el programa PyMOL Molecular Graphics System (versión 1.3 Schrödinger, LLC), acceso PDB: 2KFU.

A continuación se presenta un modelo que considerando todos los datos experimentales explica la interacción entre PknG y GarA (Fig. 11).
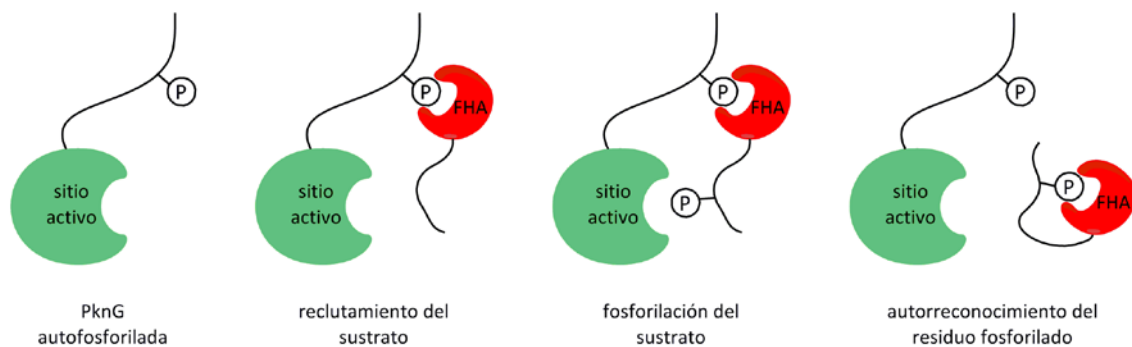
Figura 11. Modelo de interacción entre PknG y GarA. La autofosforilación de la secuencia N-terminal de PknG participa en el reclutamiento de GarA a través del dominio FHA. La fosforilación de la T_{21} en el sustrato y el reconocimiento por su propio dominio FHA constituirían un mecanismo posible de "llave" de encendido/apagado de la actividad de GarA como intermediario en vías de señalización.

Experimentos de sobrexpresión de GarA en *M. smegmatis* mostraron una reducción del tamaño de las colonias, conjuntamente se determinó que los defectos en el crecimiento se hacían aún más pronunciados cuando se utilizaban formas no fosforilables de GarA (GarA^{T21A} o GarA$^{T21A/T22A}$). Se observó igualmente que el fenotipo de la cepa que sobrexpresa GarA depende del medio de cultivo y que el suplemento con glutamato, succinato o α-cetoglutarato mitiga tal fenotipo. En resumen, **estos resultados sugieren un rol de PknG y GarA en la regulación del metabolismo bacteriano** [93].

El rol de PknG, mediado por la fosforilación de GarA, se elucidó a través de la identificación de proteínas interactoras. Nuestro grupo demostró que GarA se une a la α-cetoglutarato descarboxilasa (KGD), a la glutamato deshidrogenasa dependiente de NAD$^+$ (GDH) [93] y a la subunidad α del complejo glutamato sintasa (GOGAT) [129] mediante el dominio FHA y que esta interacción no es posible cuando GarA está fosforilada en la T_{21} o T_{22}. Además a partir de ensayos *in vitro* se determinó que la unión de GarA provoca la inhibición de las dos primeras enzimas [93] y la activación de la tercera [129]. El α-cetoglutarato, sustrato de KGD, se halla en la encrucijada del metabolismo del carbono y del nitrógeno ya que también es sustrato en la síntesis de glutamato. GarA regula el balance entre el ciclo de los ácidos tricarboxílicos y el metabolismo del glutamato inhibiendo a GDH y activando a GOGAT, implicadas en el catabolismo y anabolismo del glutamato respectivamente. **El resultado neto es que GarA desfosforilada inhibe el ciclo de los ácidos tricarboxílicos y promueve la síntesis de glutamato** (Fig. 12) [130].
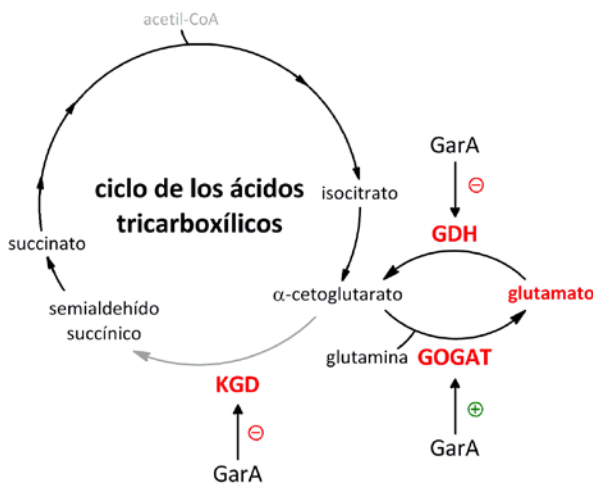
Figura 12. Regulación del ciclo de los ácidos tricarboxílicos por PknG. En ausencia de PknG (Δ*pknG*) GarA está predominantemente en su forma desfosforilada y es por tanto capaz de unirse e inhibir KGD y GDH. De esta manera se evita la conversión de glutamato a semialdehído succínico vía α-cetoglutarato. En simultáneo, GarA se une a GOGAT y aumenta su actividad lo que da lugar a la producción de glutamato. La suma de estos efectos origina la acumulación de glutamato dentro de la célula. Adaptado de [73, 129].

Considerando que: 1) PknG de *M. tuberculosis* es una enzima esencial para la infección; 2) que se ha identificado solamente un sustrato de esta quinasa y 3) en la actualidad se desconoce la identidad de las quinasas responsables de la mayoría de los eventos de fosforilación reportados en los estudios fosfoproteómicos, fue que nos propusimos un abordaje proteómico dirigido a la identificación de sustratos e interactores de PknG en la micobacteria.

4.3. *Inhibición de PknG*

La aparición de nuevas formas de tuberculosis resistentes al tratamiento con las drogas disponibles, así como la ausencia de compuestos capaces de lidiar con las micobacterias no proliferantes han desencadenado la búsqueda de nuevos fármacos antituberculosos.

Debido a su importancia en la patogenicidad de *M. tuberculosis*, PknG ha emergido como una diana relevante en esta búsqueda. Sin embargo, los esfuerzos realizados han redundado en escasas moléculas de moderada potencia capaces de inhibir PknG [29, 63, 131, 132]. Asimismo, la especificidad de estos compuestos se ha puesto en duda ya que todos se dirigen al sitio activo y como se mencionó previamente el plegamiento global y la maquinaria catalítica están muy conservados entre las STPKs eucariotas y procariotas [30, 57].

Una estrategia alternativa en la inhibición de PknG consiste en el diseño de inhibidores que se dirijan a los dominios no catalíticos, de forma tal de inhibir selectivamente la actividad quinasa de PknG explotando sus características estructurales únicas.

II. JUSTIFICACIÓN DEL TRABAJO DE TESIS

PknG de *M. tuberculosis* es una quinasa de proteínas en serinas y treoninas que ha sido reportada como uno de los factores fundamentales para la supervivencia del bacilo en el hospedero. Si bien este rol de PknG se describió hace más de una década, al día de hoy se desconocen los mecanismos moleculares subyacentes a su acción, los sustratos fosforilables tanto en la bacteria como en el hospedero, los mecanismos de activación/inhibición de esta quinasa y las señales que los desencadenan.

El objetivo general del presente trabajo es contribuir a elucidar el rol fisiológico de PknG y su regulación en las micobacterias. Para ello nos proponemos caracterizar el rol de los distintos dominios de PknG sobre su actividad enzimática, así como identificar nuevos mediadores y sustratos en las vías de señalización mediadas por esta quinasa.

La elucidación tanto de los mecanismos que regulan la actividad de PknG como de las vías de señalización en las que participa, contribuirán de forma significativa a la comprensión del **mecanismo de acción** de este importante **factor de virulencia** y a su emergencia como un interesante **blanco terapéutico**.

Objetivos específicos

1. Profundizar en el rol de los dominios no catalíticos sobre la actividad quinasa.

 1.1. Modulación de la actividad quinasa por modificación específica de residuos de cisteína del dominio rubredoxina.

 1.2. Modulación de la actividad quinasa por los dominios adyacentes al dominio quinasa.

2. Identificar blancos moleculares que actúen como mediadores en las vías de señalización de PknG.

 2.1. Análisis del interactoma de PknG en la micobacteria.

 2.2. Validación de los interactores y/o sustratos biológicamente relevantes.

III. RESULTADOS

Los resultados, relativos a cada uno de los objetivos planteados, que se obtuvieron durante el desarrollo de esta tesis se presentan en dos artículos ya publicados y un manuscrito que será prontamente enviado para su publicación.

En particular en relación al objetivo específico 1.1. Modulación de la actividad quinasa por modificación específica de residuos de cisteína del dominio rubredoxina, los resultados se muestran en el artículo "Inhibition of *Mycobacterium tuberculosis* PknG by non-catalytic rubredoxin domain specific modification: reaction of an electrophilic nitro-fatty acid with the Fe–S center".

Los resultados referentes al objetivo específico 1.2 Modulación de la actividad quinasa por los dominios adyacentes al dominio quinasa se incluyen en el artículo "Molecular Basis of the Activity and the Regulation of the Eukaryotic-like S/T Protein Kinase PknG from *Mycobacterium tuberculosis*".

Finalmente en el manuscrito "Unravelling *Mycobacterium tuberculosis* PknG interactome" se presentan los resultados referidos a los objetivos específicos 2.1 Análisis del interactoma de PknG en la micobacteria y 2.2. Validación de los interactores y/o sustratos biológicamente relevantes.

En los siguientes capítulos se discute brevemente los principales hallazgos de cada trabajo, así como los resultados obtenidos que no fueron incluidos en los mismos.

Capítulo 1

Original Contribution

Inhibition of *Mycobacterium tuberculosis* PknG by non-catalytic rubredoxin domain specific modification: reaction of an electrophilic nitro-fatty acid with the Fe–S center

CrossMark

Magdalena Gil [a,b], Martín Graña [c], Francisco J. Schopfer [d], Tristan Wagner [e,1], Ana Denicola [f], Bruce A. Freeman [d], Pedro M. Alzari [e], Carlos Batthyány [a,g,*,2], Rosario Durán [a,b,*,2]

[a] Unidad de Bioquímica y Proteómica Analíticas, Institut Pasteur de Montevideo, Uruguay
[b] Unidad de Bioquímica y Proteómica Analíticas, Instituto de Investigaciones Biológicas Clemente Estable, Ministerio de Educación y Cultura, Uruguay
[c] Unidad de Bioinformática, Institut Pasteur de Montevideo, Uruguay
[d] Department of Pharmacology and Chemical Biology, University of Pittsburgh School of Medicine, Pittsburgh, PA 15213, USA
[e] Unité de Microbiologie Structurale & CNRS URA 2185, Institut Pasteur, 25 rue du Dr. Roux, 75724 Paris Cedex 15, France
[f] Laboratorio de Fisicoquímica Biológica, Facultad de Ciencias, Universidad de la República, Uruguay
[g] Departamento de Bioquímica, Facultad de Medicina, Universidad de la República, Uruguay

ABSTRACT

PknG from *Mycobacterium tuberculosis* is a Ser/Thr protein kinase that regulates key metabolic processes within the bacterial cell as well as signaling pathways from the infected host cell. This multidomain protein has a conserved canonical kinase domain with N- and C-terminal flanking regions of unclear functional roles. The N-terminus harbors a rubredoxin-like domain (Rbx), a bacterial protein module characterized by an iron ion coordinated by four cysteine residues. Disruption of the Rbx-metal binding site by simultaneous mutations of all the key cysteine residues significantly impairs PknG activity. This encouraged us to evaluate the effect of a nitro-fatty acid (9- and 10-nitro-octadeca-9-*cis*-enoic acid; OA-NO$_2$) on PknG activity. Fatty acid nitroalkenes are electrophilic species produced during inflammation and metabolism that react with nucleophilic residues of target proteins (i.e., Cys and His), modulating protein functioCn and subcellular distribution in a reversible manner. Here, we show that OA-NO$_2$ inhibits kinase activity by covalently adducting PknG remote from the catalytic domain. Mass spectrometry-based analysis established that cysteines located at Rbx are the specific targets of the nitroalkene. Cys-nitroalkylation is a Michael addition reaction typically reverted by thiols. However, the reversible OA-NO$_2$-mediated nitroalkylation of the kinase results in an irreversible inhibition of PknG. Cys adduction by OA-NO$_2$ induced iron release from the Rbx domain, revealing a new strategy for the specific inhibition of PknG. These results affirm the relevance of the Rbx domain as a target for PknG inhibition and support that electrophilic lipid reactions of Rbx-Cys may represent a new drug strategy for specific PknG inhibition.

© 2013 Elsevier Inc. All rights reserved.

Introduction

Mycobacterium tuberculosis, the causative agent of tuberculosis, is a major public health problem that causes more than one million

deaths every year (http://www.who.int/tb/publications/global_report/gtbr12_executivesummary.pdf). Several factors comprise the efficacy of the available pharmacological treatments, including the emergence of drug-resistant strains, the lack of new drugs, and the bacilli's ability to persist inside the host macrophages by inhibiting phagosome maturation. One of the most promising strategies for drug discovery in tuberculosis is to interfere with bacterial and host signal transduction mechanisms [1,2].

Genomic studies of *M. tuberculosis* reveal the presence of 11 "eukaryotic-like" Ser/Thr protein kinases [3]. Among them, PknG, gains significance by regulating critical processes in mycobacterial survival [1,4,5]. Functional roles for PknG include regulation of metabolic processes and interference of signaling pathways from the infected host cell [4,5]. The inactivation of the *pknG* gene decreases cell viability and virulence in animal models, and suggests a central role in controlling intracellular glutamate/

glutamine levels [6]. It was further demonstrated that PknG participates in the regulation of glutamate metabolism via the phosphorylation of an endogenous substrate (GarA), with similar functions reported for PknG in the related actinomycete *Coryne-bacterium glutamicum* [5,7]. PknG also is a virulence factor that mediates *M. tuberculosis* survival within host cells by preventing macrophage phagosome-lysosome fusion [4]. In addition, inhibition of PknG activity yields bacteria more susceptible to intracellular degradation [4]. Due to the critical cellular processes that it controls, PknG inhibition has emerged as an attractive strategy for drug discovery. The main challenge is to achieve selectivity for PknG inhibition, since the catalytic mechanisms and the active Ser/Thr protein kinase fold are conserved from prokaryotes to eukaryotes.

M. tuberculosis PknG is a multidomain protein. The conserved canonical catalytic kinase domain is flanked by N- and C-terminal domains having undefined functional roles. The C-terminal domain of PknG contains a tetratricopeptide repeat motif (TPR), a domain that participates in protein–protein interactions in both eukaryotic and prokaryotic cells [8]. The TPR domain influences intermolecular interactions, as defined by the crystal structure of PknG. Notably, whether and how dimerization is linked to enzyme activity is currently unknown [8]. The N-terminal sequence preceding the kinase domain contains both autophosphorylation sites and a rubredoxin-like domain (Rbx) [5]. This latter domain is typified by an iron ion coordinated to four conserved cysteine residues that participates in electron transfer reactions [9,10]. The N-terminal sequence of PknG contains two CXXCG motifs involved in metal binding in Rbx domains. The crystal structure of a PknG construct confirmed the presence of a Rbx-like arrangement interacting with the kinase domain without occluding active site accessibility [8]. The role of the Rbx domain in PknG is still uncertain. Metal binding site disruption by simultaneous mutations of multiple key cysteine residues has a remarkable effect on PknG activity [8,11], pointing to a relevant functional or structural role of Rbx domain. This finding encouraged us to evaluate the effect of an electrophilic fatty acid nitroalkene on PknG activity.

Electrophilic unsaturated fatty acids are generated by diverse enzymatic and redox-related mechanisms, yielding signaling mediators that induce anti-inflammatory and chemotherapeutic responses [12]. In particular, nitration confers unsaturated fatty acids with an electrophilic reactivity that mediates the reversible nitroalkylation of proteins at nucleophilic Cys and His residues. This thiol-reversible posttranslational modification modulates protein function and distribution [13]. The reactivity of these molecules is directed by the electrophilic carbon β to the electron-withdrawing NO_2 group, allowing reversible Michael addition [12–14]. Compared with other electrophilic lipids, nitro-fatty acids (NO_2-FA) reversibly react with thiols at a greater rate constant [12,13,15].

Herein unique structural features of PknG are exploited to inhibit kinase activity by modification of its non-catalytic Rbx domain. These results reveal a novel mechanism for inducing kinase inhibition by iron release from the Rbx domain on covalent cysteine modification by nitrated fatty acids.

Materials and methods

Materials and chemicals

9-Octadecenoic acid (oleic acid; OA) was purchased from Nu-Check Prep (Elysian, MN). 9- and 10-nitro-octadeca-9-*cis*-enoic acid (OA-NO_2) were prepared as previously described [16]. GSH, DTT, bathophenanthrolinedisulfonic acid disodium salt (BPS), iodoacetamide (IAM), and 8-anilino-1-naphthalenesulfonic acid

ammonium salt (ANS) were from Sigma (St. Louis, MO). C18-Omix Pipette tips for sample preparation were from Varian (Lake Forest, CA). Sequencing grade modified trypsin was from Promega (Madison, WI).

Protein production and purification

Full-length PknG was overexpressed in *Escherichia coli* BL21 (DE3) cells grown for 24 h at 30 °C without IPTG and supplemented with 100 μM $FeCl_3$. PknG was purified as described before [5].

Site-directed mutagenesis of PknG was performed by PCR on pET-28a-Δ74PknG using a QuikChange II site-directed mutagenesis kit (Agilent Technologies). The sequence of the construct (PknG 74-406, Δ74/TPR) was verified by DNA sequencing. PknG Δ74/TPR was overexpressed in *E. coli* BL21(DE3) cells grown at 30 °C until $OD_{600}=0.6$, and then ON at 14 °C after addition of 1 mM IPTG. PknGΔ74/TPR purification was performed as previously described [5].

GarA (Rv1827) expression was performed in *E. coli* BL21(AI). Cells were grown in LB medium supplemented with 0.1% glucose and 10 μg/μL tetracycline at 37 °C until $OD_{600}=0.6$, and then for 18 h at 22 °C after addition of 1 mM IPTG and 0.02% arabinose. The cells were harvested by centrifugation and resuspended in buffer A (5 mM NaH_2PO_4, 50 mM Na_2HPO_4, 500 mM NaCl, 5% glycerol, 25 mM imidazole, pH 8.0) supplemented with Complete protease inhibitor cocktail (Roche). GarA was first purified by metal-affinity chromatography on a HisTrap Ni^{2+}-IMAC column (GE Healthcare) equilibrated in buffer A, using a linear imidazole gradient (20–400 mM). The GarA-containing fractions were dialyzed against buffer B (25 mM Tris-HCl, 150 mM NaCl, 5% glycerol, 1 mM DTT, pH 7.6), and the His_6 tag was removed by incubation for 24 h at 18 °C in the presence of His_6-tagged TEV endoprotease at a 1:30 ratio followed by separation on Ni-NTA agarose column (Qiagen). The untagged GarA was then further purified by size-exclusion chromatography on a Superdex 75 26/60 column (GE Healthcare) equilibrated in buffer B without DTT.

Protein kinase assay

Protein phosphorylation reactions were performed in 50 mM Hepes buffer, pH 7.0, containing 2 mM $MnCl_2$ and 100 μM ATP. Activity of PknG was assayed using recombinant GarA as substrate. The molar ratio of kinase:substrate ranged from 1:10 to 1:20. Reaction mixtures were incubated 30 min at 37 °C and substrate phosphorylation was evaluated by MALDI-TOF MS. The autophosphorylation activity of PknG was assessed by incubation of the enzyme in the presence of 2 mM $MnCl_2$ and 100 μM ATP for 40 min at 37 °C. The samples were then digested with trypsin and phosphopeptides were detected by MS.

The effect of OA-NO_2-PknG preincubation times on kinase inhibition was assayed. PknG and OA-NO_2 were incubated and at different time points aliquots of treated and control enzymes were removed for kinase activity determination. Activity assay was performed using Kinase Glo® Plus Luminescent Kinase Assay (Promega) according to the manufacturer guidelines. Briefly, activity was tested using GarA as substrate (kinase:substrate ratio was 1:25) and remaining ATP was quantified by luminescence after 30 min incubation time at 37 °C. For each time point the inhibition relative to control enzyme is plotted.

MS analysis

Proteolytic digestion was carried out by incubating the proteins with trypsin in 70 mM ammonium bicarbonate, pH 8.0, for 12 h at 37 °C (enzyme–substrate ratio 1:10). Mass spectra of peptides mixtures were acquired in a 4800 MALDI TOF/TOF instrument

(Applied Biosystems) in positive ion reflector or linear mode using a matrix solution of α-cyano-4-hydroxycinnamic acid in 0.2% trifluoroacetic acid in acetonitrile-H_2O (50%, v/v) and were externally calibrated using a mixture of standard peptides (Applied Biosystems). The molecular mass of the native and phosphorylated GarA was determined using a sinapinic acid matrix (10 mg/mL in acetonitrile-H_2O 50%, 0.2% trifluoroacetic acid). Alternatively, a linear ion trap mass spectrometer (LTQ Velos, Thermo) coupled on line with a nano-liquid chromatography system (easy-nLC, Proxeon-Thermo) was used for peptide mixtures analysis. Peptides were separated on a reversed-phase column (EASY-column[TM] 100 mm, ID 75 μm, 3 μm, C18-A2 from Proxeon) and eluted with a linear gradient of acetonitrile 0.1% formic acid (0–60% in 60 min) at a flow rate of 400 nL/min. Electrospray voltage was 1.40 kV and capillary temperature was 200 °C. Peptides were detected in the positive ion mode using a mass range of 300–2000 in the data-dependent triple play MS2 mode (full scan followed by zoom scan and MS/MS of the top 5 peaks in each segment).

Exposure of PknG to OA-NO₂ and other Cys-alkylating reagents

Native PknG (ranging from 5 to 10 μM) in 70 mM ammonium bicarbonate, pH 8.0, was incubated for 10 min at 25 °C with OA-NO_2 (0–100 μM) or IAM (0–500 μM) and kinase activity was immediately measured. PknG treatment with OA-NO_2 was also performed in 50 mM sodium bicarbonate, pH 7.2, or 50 mM Hepes buffer, pH 7.3. As a control, PknG was exposed to OA-NO_2 vehicle (methanol) under the same conditions. In addition for OA-NO_2 experiments, kinase activity in the presence of equivalent concentration of OA was determined. In some experiments, after nitroalkene treatment PknG was incubated with DTT (42 mM) or GSH (24 mM) for 10 min at 25 °C and enzymatic activity was redetermined. The activity of control PknG in the presence and absence of the thiol containing reagents was also assayed. Previous to enzymatic digestions, excess of DTT was removed by immobilization of PknG on reverse-phase Poros 10 R2 beads (Applied Biosystems).

PknG tryptic peptides were isolated by reverse-phase HPLC (Vydac® C18; 2.1 × 100 mm) and fractions including cysteine-containing peptides were selected after mass analysis by MALDI-TOF. Selected fractions were dried, resuspended in 70 mM ammonium bicarbonate, pH 8.0, and treated with OA-NO_2 (1:4) for 10 min at 25 °C. Peptide modification was analyzed by MALDI MS and ESI MS.

The experiments using rabbit muscle GAPDH as model protein were carried out as previously reported [13].

Iron determination

Control and OA-NO_2-treated PknG were loaded onto OMIX C4 pipette tips (Agilent Technologies) and flow through was collected. Non-bound fraction was incubated for 15 min at 25 °C with 40 mM DTT and iron was determined spectrophotometrically using BPS as previously described [17].

ANS fluorescence

Concentration of ANS was determined using the molar extinction coefficient $\varepsilon = 5000$ M^{-1} cm^{-1} at 350 nm [18]. PknG exposed to OA-NO_2 under the same experimental conditions stated above was diluted at 1 μM, dialyzed to remove excess OA-NO_2, and incubated with 16 μM ANS. As controls, PknG natively folded and thermally denatured was incubated with ANS. Fluorescence was collected on a Cary Eclipse fluorescence spectrophotometer (Varian, Inc.) with the excitation wavelength set on 350 nm and emission between 370 and 650 nm.

Circular dichroism (far-UV) spectroscopy

Far-UV circular dichroism spectra were performed on OA-NO_2-treated and OA-treated PknG. CD spectra were recorded between 190 and 260 nm on an Aviv 215 spectropolarimeter (Aviv Biomedical), using a cylindrical cell with a 0.02 cm path length and an averaging time of 1 s per step, with protein samples at 0.5 mg/mL in 25 mM Tris-HCl, 100 mM NaCl, glycerol 5%, pH 8.0. Five consecutive scans from each sample were merged to produce an averaged spectrum and corrected using buffer baselines measured under the same conditions. Data were normalized to the molar peptide bond concentration and path length and expressed as mean residue ellipticity ($[\theta]$ degree·cm^2·$dmol^{-1}$).

Bioinformatics procedures

Protein sequence homologs to PknG were searched at NCBI's NR database using CS-Blast [19], Psi-Blast [20], and HHsenser [21]. All searches were run performing both gapped and ungapped alignments, in order to selectively detect proteins carrying both Rbx and Ser/Thr protein kinase domains. Significant hits (Blast E values $< 1e^{-10}$) found with all methods with sequence coverage $> 75\%$ were kept. Multiple sequence alignments (MSAs) were computed with Mafft [22], T-Coffee [23], and Prank [24]. Such MSAs were manually analyzed in order to detect sequences with N-terminal Rbx motifs (all had the kinase domain). Distance-based phylograms for a subset of 652 sequences (lengths between 500 and 800 amino acids, pairwise identities $< 76\%$) were computed with BioNJ [25]. For proteins with the Rbx domain, maximum-likelihood phylogenetic trees were built by way of PhyML [26].

Results

PknG inhibition by OA-NO₂

To evaluate whether the electrophilic OA-NO_2 has an effect on PknG kinase activity, the enzyme was treated with different concentrations of OA-NO_2 (below its critical micelle concentration) and activity was measured using a recombinant protein substrate, GarA. Native and phosphorylated GarA were detected by MALDI-TOF MS as ions of m/z 17145 and m/z 17222, respectively [27]. The mass shift corresponds to the incorporation of one phosphate group (80 Da, Fig. 1A). Kinase activity of PknG exposed to micro-molar concentrations of nitro-oleic acid (molar excess ranging from 1:1 to 1:3) was significantly inhibited. The ability to phosphorylate GarA was decreased after 10 min incubation of the kinase with OA-NO_2 (50 μM, molar ratio OA-NO_2:PknG 5:1) (Fig. 1A). The same extent of enzyme inhibition was observed when PknG was exposed to OA-NO_2 under different experimental conditions (ammonium bicarbonate, pH 8.0, sodium bicarbonate, pH 7.2, and Hepes, pH 7.3). As controls, PknG was treated with equivalent concentrations of oleic acid or methanol (vehicle) and no effect on kinase activity was observed (data not shown).

The effect of OA-NO_2 on PknG autophosphorylation was also evaluated to confirm whether the inhibition observed reflects a general loss of the kinase activity as opposed to a substrate-specific effect (Fig. 1B). PknG can autophosporylate its N-terminal sequence at specific Thr residues [5]. The spectrum obtained in the linear mode after tryptic digestion of native PknG showed unphosphorylated (m/z 5395.8), monophosphorylated (m/z 5475.4), and diphosphorylated (m/z 5555.1) ions (sequence 10–60) as previously described (Fig. 1B, upper panel) [5]. When native PknG was incubated with ATP under phosphorylation conditions and then digested, the most intense ion observed corresponded to the diphosphorylated specie while unphosphorylated and
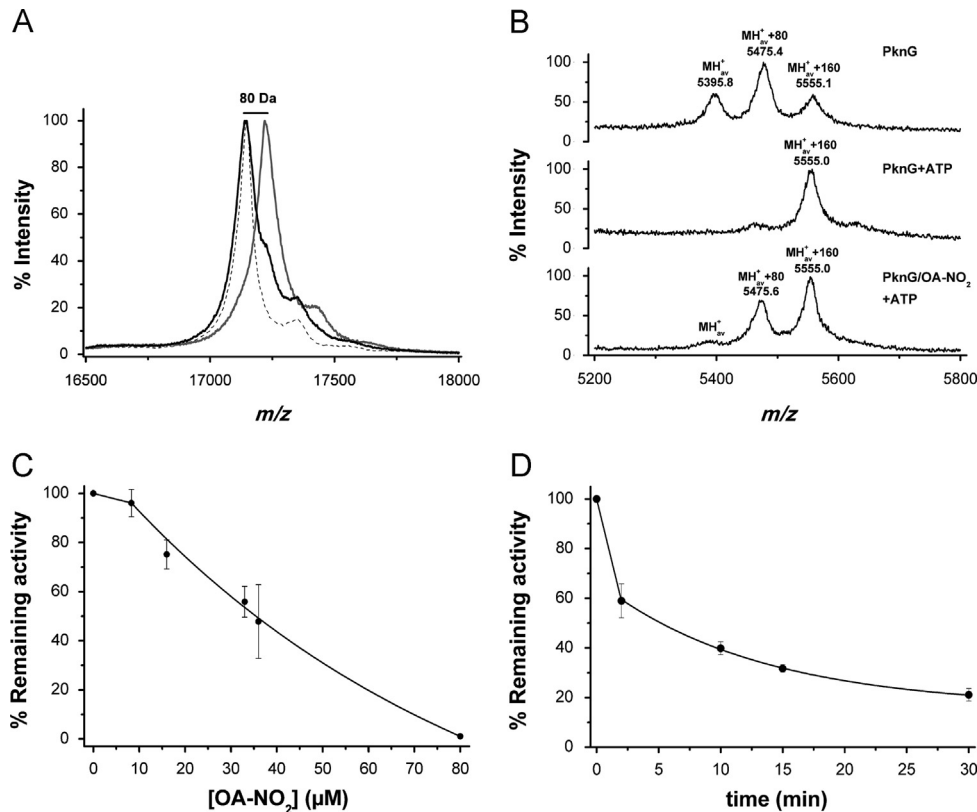
Fig. 1. OA-NO$_2$ inhibits PknG kinase activity. (A) GarA phosphorylation by PknG. Linear MALDI-TOF mass spectrum of GarA (m/z 17145, dashed line), GarA phosphorylated by PknG using a molar ratio PknG:GarA of 1:20 (m/z 17222, gray line), and GarA phosphorylated by PknG pretreated with 50 μM OA-NO$_2$ for 10 min (m/z 17143, black line). (B) PknG autophosphorylation. Upper panel: linear MALDI-TOF spectrum of native PknG after tryptic digestion showing basal autophosphorylation pattern. Unphosphorylated peptide with seq. 10–60 (m/z 5395.8), monophosphorylated (m/z 5475.4), and diphosphorylated species (m/z 5555.1) are detected as expected. Middle panel: linear MALDI-TOF spectrum of PknG after incubation with ATP and Mn^{2+} and further typtic digestion (autophoshorylation positive control). Diphosphorylated peptide (m/z 5555.0) is the most intense ion observed while unphosphorylated and monophosphorylated ions were almost undetectable. Lower panel: linear MALDI-TOF spectrum of tryptic digestion of PknG treated with 30 μM OA-NO$_2$ for 10 min before autophosporylation reaction. The detection of all three forms of the phosphorylatable peptide indicates that OA-NO$_2$ impairs conversion into the fully phosphorylated species. The spectra are representative of three independent experiments. (C) Dose-dependent inhibition of PknG by OA-NO$_2$. PknG (8 μM) was incubated with OA-NO$_2$ ranging from 0 to 80 μM in 70 mM ammonium bicarbonate, pH 8.0, at 25 °C. As a control, PknG was incubated with vehicle under the same experimental conditions. After 10 min of incubation enzyme activities were determined. Samples were analyzed in triplicates. (D) Time-dependent inhibition of PknG by OA-NO$_2$. PknG (10 μM) was incubated with 50 μM OA-NO$_2$ (treated PknG) or vehicle (control PknG) in 70 mM ammonium bicarbonate, pH 8.0, at 25 °C. At the indicated time points aliquots of control and treated PknG were removed, and enzyme inhibition was determined using Kinase Glo® by comparison with control activity. Samples were analyzed in triplicates.

monophosphorylated ions were almost undetectable (Fig. 1B, middle panel). On the other hand, OA-NO$_2$ (30 μM) impairs the conversion of the un- and monophosphorylated forms of the kinase into the fully phosphorylated species, indicating that autocatalytic activity is also inhibited by OA-NO$_2$ (Fig. 1B, lower panel).

The effect of OA-NO$_2$ on kinase activity was tested using inhibitor concentrations ranging from 0 to 80 μM. There was a dose-dependent inactivation of PknG on preincubation of the enzyme with different OA-NO$_2$ concentrations for 10 min (Fig. 1C). Under these conditions, 50% of enzyme inhibition was reached with 35 μM OA-NO$_2$.

To further characterize the effect of OA-NO$_2$ on PknG, the time course of enzyme inhibition was studied. PknG and OA-NO$_2$ (or vehicle control) were mixed for time periods ranging from 0 to 30 min, and kinase activity was then measured using a commercial luminescent kinase assay (Kinase Glo®). Under these conditions the control activity varies less than 10% during the 30 min incubation time. The percentage of remaining kinase activity was calculated with respect to control activity for the same time point. The exposure of PknG to OA-NO$_2$ decreased kinase activity in a time-dependent manner (Fig. 1D). The dependence of PknG inhibition on enzyme-OA-NO$_2$ reaction times suggests that the effect of OA-NO$_2$ is mediated by a covalent modification of PknG.

Cysteines in the Rbx domain are the molecular targets of OA-NO$_2$

In order to identify the PknG residues that account for the inhibition of the kinase, samples obtained by treatment with OA-NO$_2$ under different experimental conditions were digested with trypsin and peptide mixtures were analyzed by MS (Fig. 2). MALDI-TOF mass spectra of OA-NO$_2$-treated samples (OA-NO$_2$ concentration up to 35 μM) showed loss of only two peptides, each one containing a different CXXCG motif of the Rbx domain: m/z 1292.56 and 3530.62 (considering the intramolecular disulfide bond) (Figs. 2A and B). No other peptide showed a significant change in signal intensity compared with control spectra (data not shown). These observations were confirmed under all experimental conditions used (ammonium bicarbonate, pH 8.0, sodium bicarbonate, pH 7.2 and Hepes, pH 7.3).

The expected mass shift for the incorporation of a molecule of OA-NO$_2$ is 327 Da [13]. However, under this experimental conditions no peptide was detected with this mass increment. Analysis of MALDI-TOF/TOF mass spectra of OA-NO$_2$-treated samples showed the systematic appearance of a cluster of peaks with 198, 200, and 202 Da greater masses than the CXXCG-containing peptides (Fig. 2B, upper and middle panel, inset). MS/MS spectra of the observed ions (m/z 1492.54 and 3730.59) confirmed that these ions corresponded to modified peptides containing the sequences
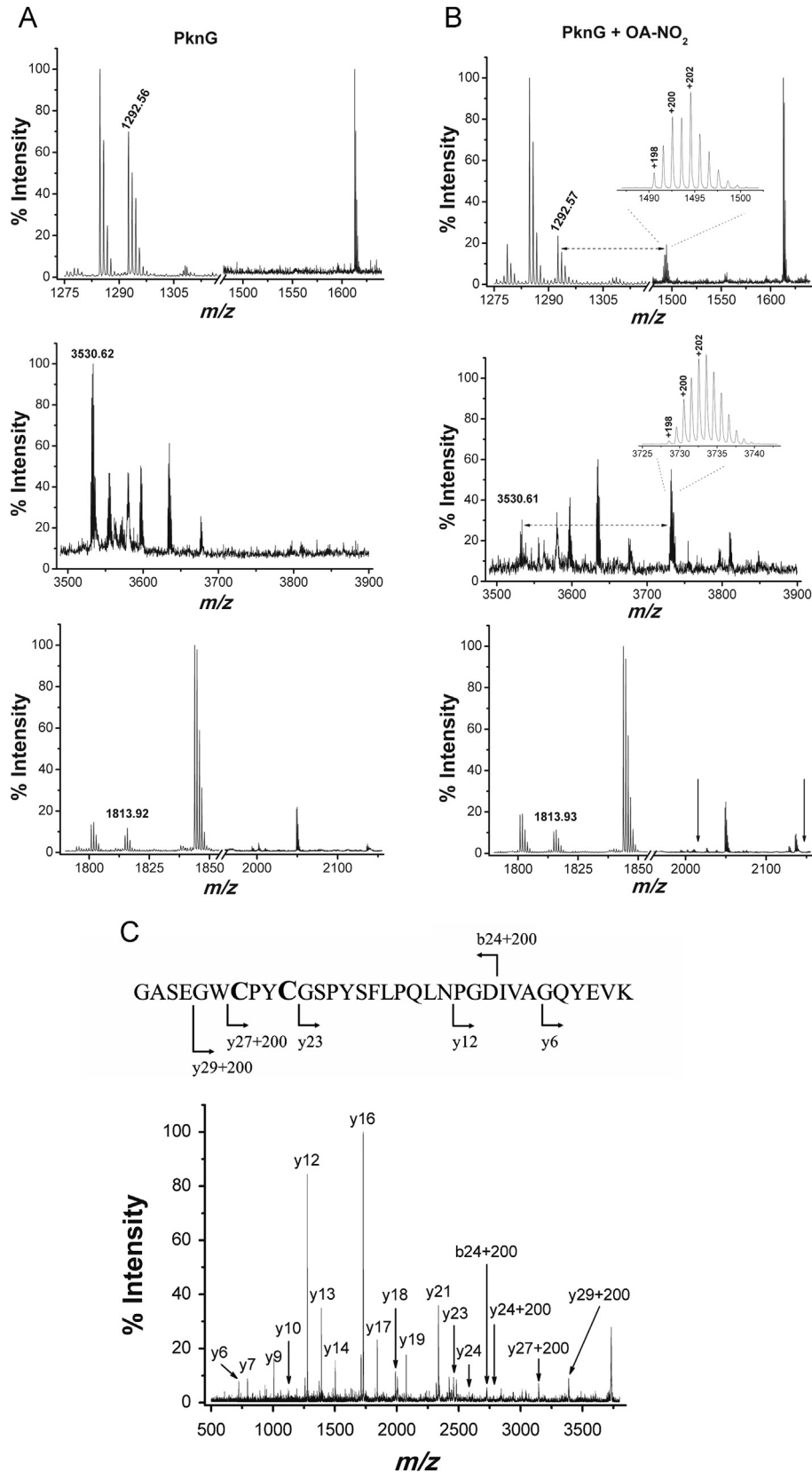
Fig. 2. Modification sites of PknG by OA-NO$_2$. (A) MALDI-TOF mass spectrum of Cys-containing peptides generated by tryptic digestion of untreated PknG. Upper panel: peak m/z 1292.56 corresponds to sequence FC$_{106}$WNC$_{109}$GRPVGR with an intramolecular disulphide bridge. Middle panel: peak m/z 3530.62 corresponds to sequence GASEGWC$_{128}$PYC$_{131}$G-SPYSFLPQLNPGDIVAGQYEVK with an intra molecular disulphide bridge. Lower panel: peak m/z 1813.92 (GC$_{156}$IAHGGLGWIYLALDR) is indicated. (B) MALDI-TOF mass spectrum of Cys-containing peptides generated by tryptic digestion of OA-NO$_2$-treated PknG. Upper panel: peak m/z 1292.57 is partially consumed and a new peak appears showing a mass increment of 198/200/202 Da. Inset: zoom in showing the 198/200/202 pattern and isotopic distribution. Middle panel: peak m/z 3530.61 is also partially consumed and again a new signal appears showing a mass increment of 198/200/202 Da. Inset: zoom in showing the 198/200/202 mass increment pattern. Lower panel: signal intensity of m/z 1813.93 is unchanged on treatment with OA-NO$_2$, and in agreement with this observation there were no new signals in this m/z range (the arrows indicate the expected m/z values: +327 for OA-NO$_2$ modification and +198/200/202 for the previously experimentally observed mass shift). (C) MALDI-TOF MS/MS spectrum of ion at m/z 3730.60 obtained from tryptic digestion of OA-NO$_2$-treated PknG. Sequence 122–154 showing main observed fragments (natives or modified). The spectra are representative of five independent experiments.

105–115 and 122–154, respectively. Fig. 2C shows the MS/MS spectra of the precursor ion corresponding to the modified sequence 122–154 (GASEGWC$_{128}$PYC$_{131}$GSPYSFLPQLNPGDIVAG-QYEVK) with a mass increment of 200 Da. The presence of native y-ion series up to y23 together with the detection of the modified y27 (+200 Da) indicates that any of the residues in between (i.e., $_{128}$CPYC$_{131}$) might be modified by OA-NO$_2$ (Fig. 2C). Based on the presence of small signals corresponding to unmodified y24, y24+200 Da, unmodified y26, and y26+200 Da we postulate that ion $m/z = 3730.59$ is actually a mixture of singly modified peptides where C$_{128}$ or C$_{131}$ have been individually modified by OA-NO$_2$ (Table S1). Similarly, MS/MS analysis of the other modified Rbx peptide (m/z 1492.54) showed that only daughter ions that contain the CWNC residues appeared with a modified mass (data not shown). Although there was not enough sequence information to identify the modified residue(s) within this motif, based on the previously reported reactivity of the nitrated fatty acid toward nucleophilic residues these results support that the Cys residues of those peptides are the main target of OA-NO$_2$ [13].

To further characterize this modification and the observed atypical mass shift, all three PknG native tryptic peptides that contained Cys residues (m/z 1292.57, m/z 1813.92, and m/z 3530.60) were isolated by RP-HPLC and then treated with OA-NO$_2$. The peptide with the sequence 155–171 (m/z 1813.92) is the only tryptic peptide of PknG that contains a single Cys. The modification pattern previously observed was found for all those three Cys-containing peptides treated with OA-NO$_2$ (Fig. 3). The mass shift for the peptide containing a single Cys was 196/198/200 Da. This observation is in agreement with the observed mass shift of 198/200/202 Da for the CXXCG containing peptide, with respect to the native peptide with an intramolecular disulfide bond. These results indicate that the presence of a single Cys residue is enough to generate the observed mass shift in MALDI-TOF MS analyses. In contrast, the analysis of those same peptides by electrospray ionization-MS showed the expected mass increment of 327 Da, but not the 198/200/202 Da pattern (Figs. 4A and B). MS/MS spectra of the modified peptides unequivocally identified Cys as the residue modified by OA-NO$_2$ (Figs. 4B and C). In aggregate, these results support the hypothesis of MALDI ionization-induced decomposition of Cys-nitro-fatty acid adducts. However, MALDI MS/MS spectra did not provide enough information to interpret the structure of the modification responsible for the 198/200/202 Da mass shift.

The PknG sequence contains a unique non-rubredoxin Cys (C$_{156}$). Noticeably, no consumption of the peptide containing this free Cys was observed on treatment of PknG with OA-NO$_2$ (Figs. 2A and B, lower panel). In agreement with this observation, no appearance of a new signal with the previously observed mass shift (198/200/202 Da) or the theoretical expected mass shift (+327 Da) was observed. To gain further insight into the reactivity of PknG-Cys residues, iodoacetamide was used to model cysteine alkylation. No modification of Rbx cysteines was observed for IAM concentrations up to 150 μM (Fig. S1). Conversely, the unbound Cys$_{156}$ was alkylated by IAM under similar conditions, revealing a different reactivity of OA-NO$_2$ vs other alkylating reagents for Cys residues in PknG. When treating PknG, nitro-fatty acid-modified cysteines tightly bound to the metal ion and inhibited the enzyme, while treatment with IAM only resulted in modification of free Cys without affecting kinase activity (Fig. S1). Thus Rbx-Cys, typically reported as non-reactive with alkylating reagents, is the principal target of reaction with an electrophilic nitroalkene.

At higher concentrations of OA-NO$_2$ (50–80 μM; molar ratio PknG:OA-NO$_2$ from 1:5 to 1:10) the modification of several His residues occurred at very low yields. In agreement with this observation, there was no significant consumption of His-containing peptides (Fig. S2 and Table 1). In contrast to cysteine

modification, alkylated histidine showed the expected mass shift of 327 Da when analyzed by MALDI MS and ESI MS. MS/MS analysis of these peptides confirmed that the incorporation of the nitrated fatty acid occurred at a His residue (Table 1 and Fig. S2). To confirm that inhibition is a consequence of Cys and not His modification, studies were repeated using an active PknG construct where N-terminal and TPR domains were removed (PknGΔ74/TPR). This truncated construct preserves only His185 out of the seven previously modified His residues. After the exposure of PknGΔ74/TPR to OA-NO$_2$ the enzyme is fully inhibited, the Cys-containing peptides from the Rbx domain appeared alkylated and another set of His residues was detected at low yields (Table S2 and Fig. S3).

Using both enzyme constructs, the Cys-Rbx domain encompasses the only OA-NO$_2$-modified residues using concentrations up to 35 μM that can render detectable loss of PknG activity. Altogether, these data indicate that Cys residues at the Rbx domain are the targets of OA-NO$_2$ and account for enzyme inhibition.

Irreversible PknG inhibition by the reversible OA-NO$_2$-mediated nitroalkylation of the kinase

We have previously reported that nitrated fatty acids inhibit GAPDH activity by modification of nucleophilic amino acid residues in a thiol-reversible manner [13]. Herein, the reversibility of PknG nitroalkylation was analyzed. Treatment of nitroalkylated PknG with DTT (42 mM) or GSH (24 mM) did not restore kinase activity (Fig. 5A). Under the same conditions GAPDH activity, used as a control, was restored (Fig. 5A). PknG control activity was measured in the presence of the thiol-containing reagents. Neither DTT nor GSH have an effect *per se* on kinase activity under these experimental conditions.

To evaluate the reversibility of the modification of His and Cys residues in PknG, OA-NO$_2$-treated samples were exposed to DTT and GSH before protein digestion and MS analysis. The peptide containing two Rbx-Cys residues (sequence 122–154) showed no change in mass after nitroalkylation and DTT treatment, supporting reversibility of this reaction of OA-NO$_2$. In addition, native cysteine-containing peptides were fully recovered after treatment with DTT (Fig. 5B). Moreover, there were no detectable Cys- or His-nitrated fatty acid adducts after treatment with DTT or GSH (data not shown). Although nitroalkylation was reversible as expected, the effect of OA-NO$_2$ on PknG activity was irreversible (Fig. 5A). These results indicate that OA-NO$_2$ is inactivating PknG by an unusual mechanism involving an irreversible change in PknG structure or function that persists after Cys nitroalkylation is reversed.

OA-NO$_2$ induces iron release from the Rbx domain

The rubredoxin domain contains an iron ion coordinated by the sulfur atoms of four conserved cysteine residues, forming an almost regular tetrahedron where Fe^{3+} is typically very stable [28,29]. To evaluate the effect of Cys nitroalkylation on the metal center of the Rbx domain, the amount of iron released on OA-NO$_2$ treatment was measured using a specific ligand that allows for spectrophotometric determinations. Non-protein-bound iron present in control and OA-NO$_2$-treated PknG samples was recovered, reduced with DTT, and quantified spectrophotometrically. The results showed that PknG nitroalkylation led to iron release from the Rbx domain (Fig. 6). These data showed that 70% of the iron present in the Rbx domain was released after protein exposure to 50 μM OA-NO$_2$. As a positive control, total iron in native protein was determined after Rbx domain disruption by protein digestion (data not shown).
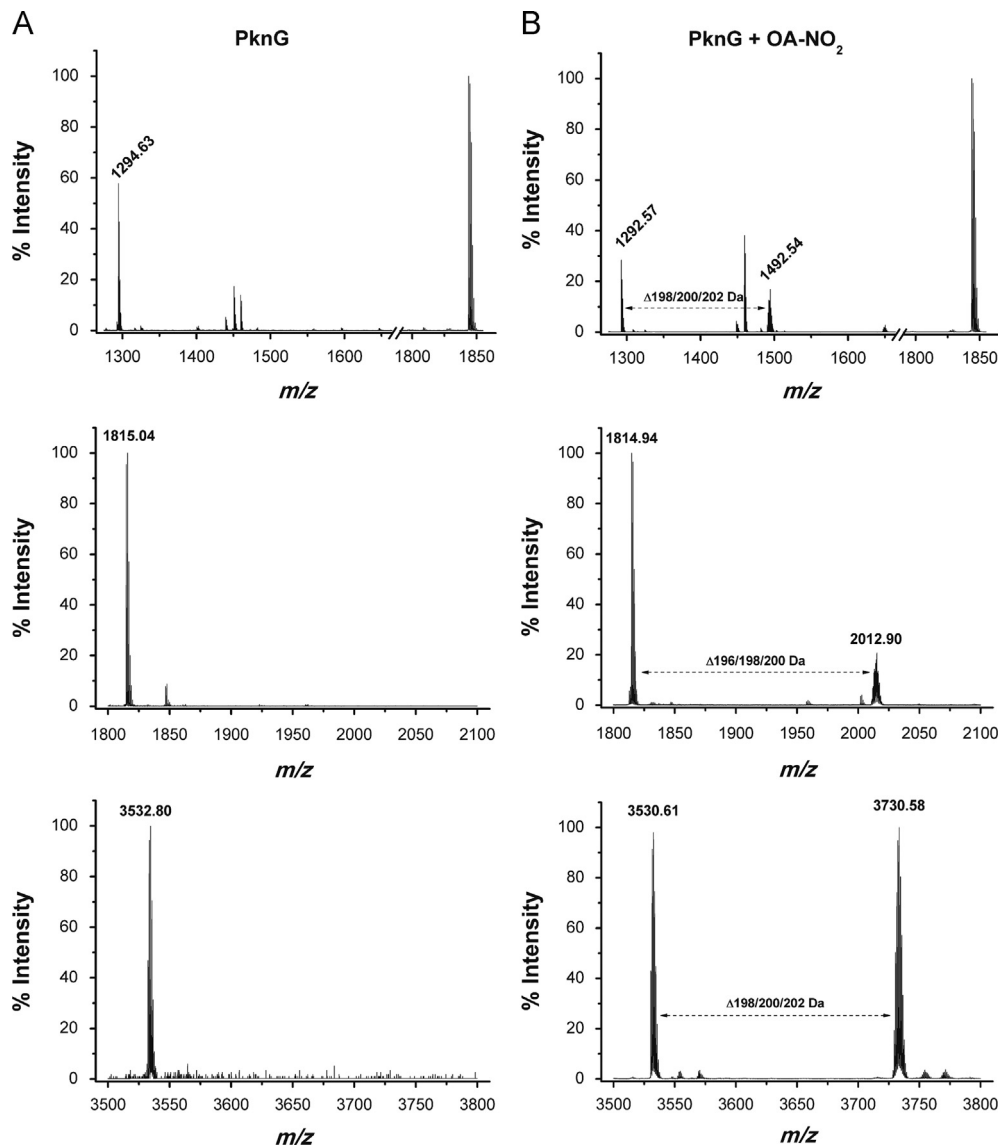
Fig. 3. Treatment of isolated Cys-containing peptides from PknG with OA-NO$_2$. (A) MALDI-TOF mass spectrum of peptides generated by tryptic digestion of native PknG. Peptides were purified by reverse-phase chromatography prior to MS analysis. Upper panel: peak m/z 1294.63 (FC$_{106}$WNC$_{109}$GRPVGR); middle panel: peak m/z 1815.04 (GC$_{156}$IAHGGLGWIYLALDR); and lower panel: peak m/z 3532.80 (GASEGWC$_{128}$PYC$_{131}$GSPYSFLPQLNPGDIVAGQYEVK). The reduced peptides containing the CXXCG sequence are detected after the chromatographic step but are spontaneously converted into the intermolecular disulphide bridged oxidized form. (B) MALDI-TOF mass spectrum of peptides isolated in (A), treated with OA-NO$_2$. Upper panel: a new peak appears showing a mass shift of 198/200/202 Da relative to m/z 1292.63; middle panel: a new peak appears showing a mass shift of 196/198/200 Da relative to m/z 1814.94; and lower panel: a new peak appears showing a mass shift of 198/200/202 Da relative to m/z 3530.61.

PknG inhibition by OA-NO$_2$ is not the consequence of a global change in PknG structure

In order to address if the effect of OA-NO$_2$ could be mediated by an unspecific global distortion of protein tridimensional structure as a consequence of the introduction of a quite large hydrophobic molecule, global changes in protein structure were analyzed by different approaches. Enzyme was first treated with OA-NO$_2$ and then ANS, a fluorescent probe that binds to hydrophobic patches on proteins with a concomitant change in emission spectrum (Fig. 7A). The same spectra were obtained with native and OA-NO$_2$-modified PknG, supporting that the hydrophobic surfaces of the protein were not affected by nitroalkylation (Fig. 7A). This suggests that the enzyme inhibition was not due to large protein structural changes. In order to discard minor changes in secondary structure circular dichroism experiments were performed on PknG treated with OA-NO$_2$ or OA (Fig. 7B). The comparison of the

spectra showed no difference between both samples. Overall, these results suggest that PknG inhibition by OA-NO$_2$ is not due to a general structural modification as a result of Cys nitroalkylation and iron release from the Rbx domain.

Rbx and kinase domain co-occurrence is restricted to few Actinomycetales

The specificity of OA-NO$_2$ reaction with Cys residues in the Rbx domain of PknG raised the possibility of a selective inhibition of certain kinases containing this domain. A bioinformatics analysis informed about the co-occurrence of Rbx and kinase domains. Multiple sequence alignments of PknG orthologs showed a minority of PknG-like kinases harboring the CXXCG motif linked to Rbx domains (Fig. 8). In a subset of 652 PknG-like kinases, the Rbx domain was present in less than 10% (shown in red in Fig. 8). The same trend was observed in more thorough data sets; e.g., from
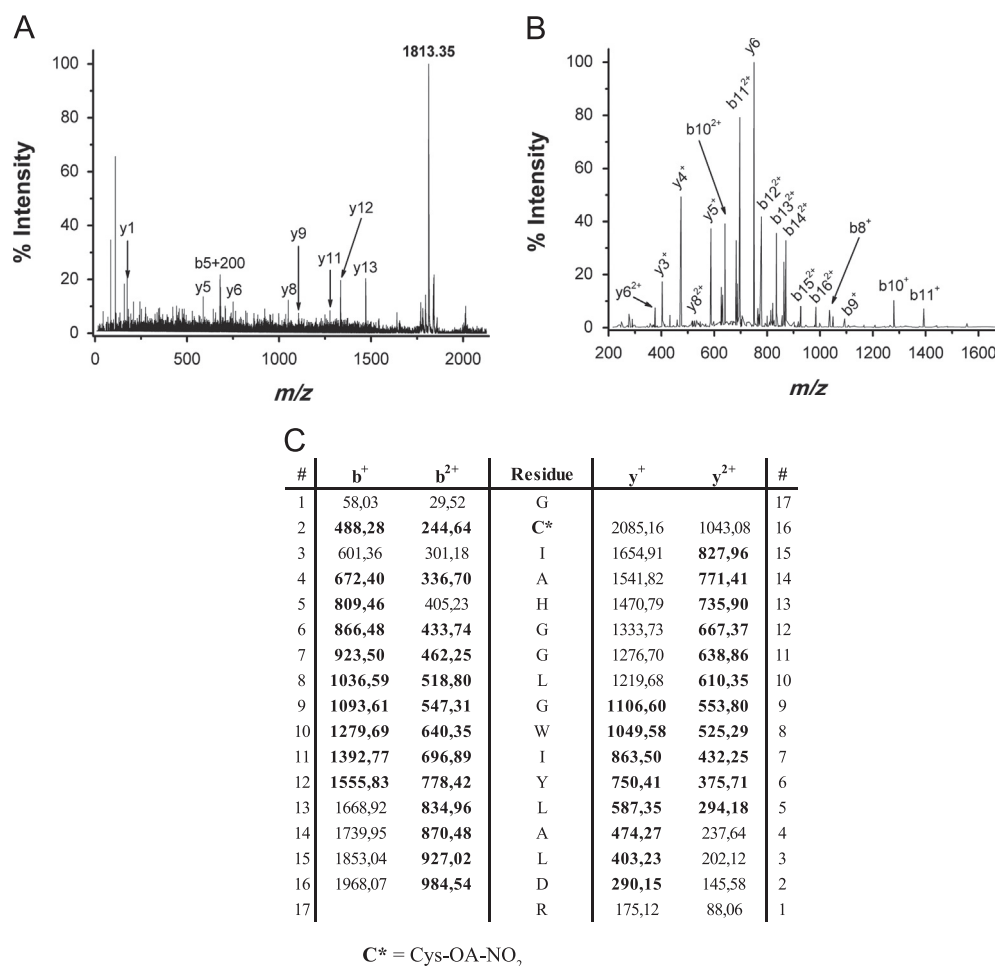
Fig. 4. MS/MS spectra of OA-NO$_2$ modified peptide 155–171. (A) MALDI ionization. MS/MS spectrum of ion at m/z 2014.31 (native peptide +200 Da) obtained using MALDI-TOF MS. (B) ESI ionization. MS/MS spectrum of ion at m/z 2141.85 (native peptide +327 Da) obtained using LC-MS. (C) List of theoretical m/z values of fragment ions from peptide of m/z 2141.85 (GC$_{156}$IAHGGLGWIYLALDR modified by OA-NO$_2$). y- and b-ions detected by LC-MS are highlighted in bold.

Table 1
Identification of OA-NO$_2$-modified residues in PknG.

Peptide from-to	Modified residue	Assigned sequence[b] (observed m/z)	OA-NO$_2$:PknG ratio[a]	
			Low	High
		Peptides containing modified cysteine residues ($\Delta m = 198/200/202$ Da)		
105–111	C$_{106}$ or C$_{109}$	FCWNCGR (1081.31/1083.30/1085.33)	x	x
105–115	C$_{106}$ or C$_{109}$	FCWNCGRPVGR (1490.51/1492.54/1496.55)	x	x
122–154	C$_{128}$ or C$_{131}$	GASEGWCPYCGSPYSFLPQLNPGDIVAGQYEVK (3728.55/3730.59/3732.60)	x	x
		Peptides containing modified histidine residues ($\Delta m = 327$ Da)		
182–199	H$_{185}$	GLVHSGDAEAQAMAMAER (2171.09)		x
420–443	H$_{430}$ or H$_{440}$	STFGVDLLVAHTDVYLDGQVHAEK (2941.56)		x
488–510	H$_{488}$	HGALDADGVDFSESVELPLMEVR (2813.41)		x
558–573	H$_{558}$	HFTEVLDTFPGELAPK (2128.15)		x
635–640	H$_{635}$	HFTTAR (1059.61)		x
698–714	H$_{702}$ or H$_{711}$	ASTNHILGFPFTSHGLR (2182.21)		x
733–743	H$_{733}$	HRYTLVDMANK (1674.91)		x

[a] Low OA-NO$_2$:PknG ratio means a molar ratio in between 1:1 and 3:1 ([OA-NO$_2$] between 8 and 30 μM). High OA-NO$_2$:PknG ratio means a molar ratio 10:1 ([OA-NO$_2$]=80 μM).
[b] The sequence and alkylation sites were confirmed by MS/MS analysis.

1294 proteins, just 88 had the N-terminal motif. The co-occurrence of Rbx and kinase domains in different species was also analyzed. Fig. 8 displays a maximum-likelihood phylogenetic tree for a set of 52 sequences with sequence identities below 90% and above 30%.

At sequence identity levels above 25%, all but two PknG homologs carrying both domains were confined to few suborders of the *Actinomycetales*. The two exceptions were *Acetivibrio cellulolyticus* (Firmicutes) and *Ktedonobacter racemifer* (Chloroflexi).

A

B
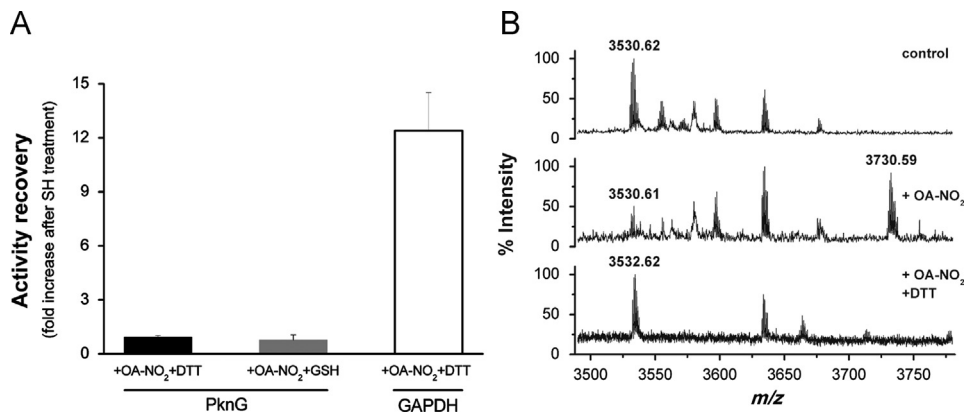
Fig. 5. Irreversible PknG inhibition by reversible OA-NO$_2$-mediated nitroalkylation. (A) PknG was inhibited by incubation with OA-NO$_2$ (50 μM) in 70 mM ammonium bicarbonate, pH 8.0, at 25 °C for 10 min, as previously described. Kinase activity of PknG inhibited by OA-NO$_2$ was measured (PknG+OA-NO$_2$). Subsequently, nitroalkylated samples were treated with DTT (42 mM) or GSH (24 mM) for 15 min at 25 °C and kinase activity was remeasured (PknG+OA-NO$_2$+RSH). Activity recovery was expressed as a ratio between [activity of PknG+OA-NO$_2$+RSH]/[activity PknG+OA-NO$_2$] where 1 means no activity recovery after thiol-containing agent treatment. GAPDH was used as a positive control: after DTT treatment 90% of the initial activity was recovered representing a 12-fold increase in the enzyme activity after the RSH treatment. Treatment of control PknG samples with 42 mM DTT or 24 mM GSH showed that these concentrations of the thiol reagents had no effect on kinase activity *per se*. Three independent experiments were performed. (B) MALDI-TOF mass spectrum of peptides generated by tryptic digestion of PknG. Upper panel: untreated PknG; middle panel: PknG exposed to OA-NO$_2$; and lower panel: idem middle panel+42 mM DTT. In each panel the native or nitroalkylated form of the peptide with the sequence 122–154 that includes the Rbx-Cys 128 and 131 is shown. All the spectra are representative of three independent experiments.
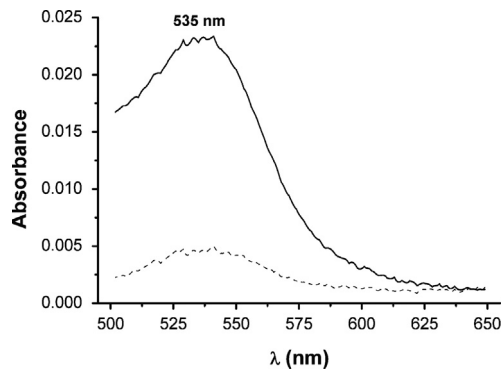
Fig. 6. OA-NO$_2$ induces iron release from Rbx domain. Non-protein-bound iron present in control and OA-NO$_2$-treated PknG samples was recovered, reduced with DTT, and quantified as Fe^{2+} using bathophenantrolinedisulfonic acid. The complex absorbs at 535 nm ($\varepsilon_{535\ nm}$ [BPS·Fe^{2+}]=22140 M^{-1} cm^{-1} [17]). Iron present on native PknG samples (dashed line) or from OA-NO$_2$-treated PknG (solid line) is shown.

Conservation of the catalytic domain, as well as variability of the N-terminus, has been described for a number of PknGs [30,31]. However, the presence or absence of the Rbx domain and sequence conservation levels had not been analyzed. Notably, (a) the joint occurrence of kinase and Rbx domains is restricted to bacteria from the *Actynomycetales* order, including pathogenic and nonpathogenic mycobacteria and (b) the sequences indicated by a dot in Fig. 8 correspond to different species of *Corynebacterium* lacking the CXXCG motifs of the Rbx domain.

Discussion

The analysis of the *M. tuberculosis* genome sequence predicted the presence of 11 eukaryotic-like Ser/Thr protein kinases denoted as *pknA* to *pknL* [3]. The importance of these enzymes in mycobacterial physiology and virulence has only recently been appreciated [1,32]. In particular, one of these enzymes, PknG, regulates critical processes in *M. tuberculosis*. Previous data support that PknG regulates glutamate metabolism in *M. tuberculosis* through the phosphorylation of GarA, an intermediate regulator of three metabolic enzymes [5,33]. PknG is also an important virulence

factor that inhibits phagosome maturation in infected macrophages through an unknown mechanism [4,11]. In agreement with this, inhibition of PknG activity yields bacteria more susceptible to macrophage killing.

In this scenario, PknG inhibitors represent a promising drug development strategy, with a few PknG inhibitors of moderate activity previously described [2,8,34]. These inhibitors are directed toward the kinase catalytic site. For example, AX20017 displays some degree of specificity toward PknG, based on structural features of its active site [8]. It is already appreciated that active "eukaryotic-like" kinase folding is highly conserved even among different kingdoms [35], making inhibitor specificity problematic. An alternative strategy for selective kinase inhibition is to target other protein domains besides active site. Herein, we report that the inhibition of PknG kinase activity by OA-NO$_2$ occurs by reversible alkylation of specific Cys residues of the Rbx domain, outside the catalytic domain.

The combination of Rbx and kinase domains has not been described for any other proteins than PknG-like kinases. PknG orthologs are present in all mycobacterial genomes sequenced to date, as well as in other related actinomycetes [31]. While the kinase domain is very well preserved, some other domains are not. In particular, the N-terminal Rbx domain appears in few PknG-like kinases from *Actinomycetales* (Fig. 8). Moreover, PknG ortholog from *Corynebacterium spp*, with the equivalent function toward glutamate metabolism, lacks the Rbx domain. This highlights the relevance of nitrated fatty acids as potential selective inhibitors acting on a subset of Rbx-containing PknG-like enzymes among prokaryotic and eukaryotic kinases.

The reported structure of a truncated form of PknG shows that the rubredoxin-like domain is closely associated with the N-terminal lobe of the catalytic domain of PknG, facing the kinase active site (Scheme 1). Thus, the modification of the Rbx domain could directly affect the active kinase conformation. Herein, the modification of Rbx-Cys by a nitroalkene was reversible and leads to a permanent effect on enzyme activity because of concomitant loss of an iron atom. Nitroalkylation of PknG did not induce a global change in structure (Fig. 7). In support of this, the conversion of bacterial rubredoxins to the apo form is followed by only minor structural changes [28], suggesting that minor local structural changes and/or the loss of iron could be mediating a regulatory role of kinase activity. To confirm the effect of iron loss
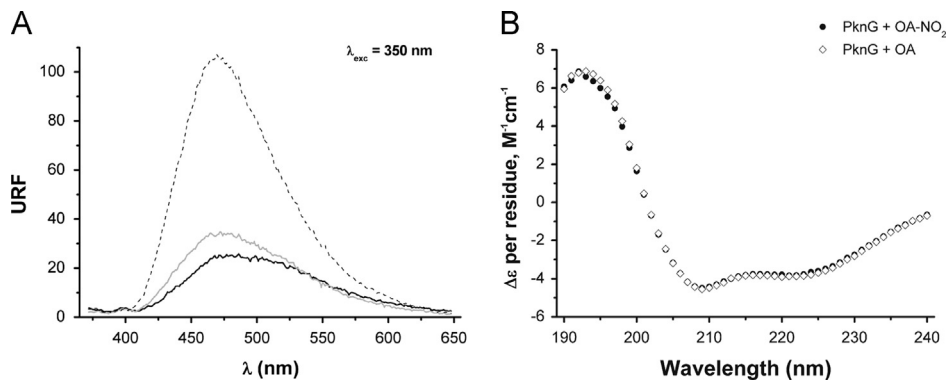
Fig. 7. OA-NO₂ modification does not generate a global change in PknG structure. (A) Fluorescence of PknG-ANS complexes was collected using excitation wavelength set on 350 nm and emission between 370 and 650 nm. PknG natively folded is represented as solid black line; PknG exposed to OA-NO₂ is shown as solid gray line; and PknG thermally denatured is indicated as dashed black line. (B) Far UV-CD spectra of OA-NO₂-treated PknG (black circles) and OA-treated PknG (white diamonds).
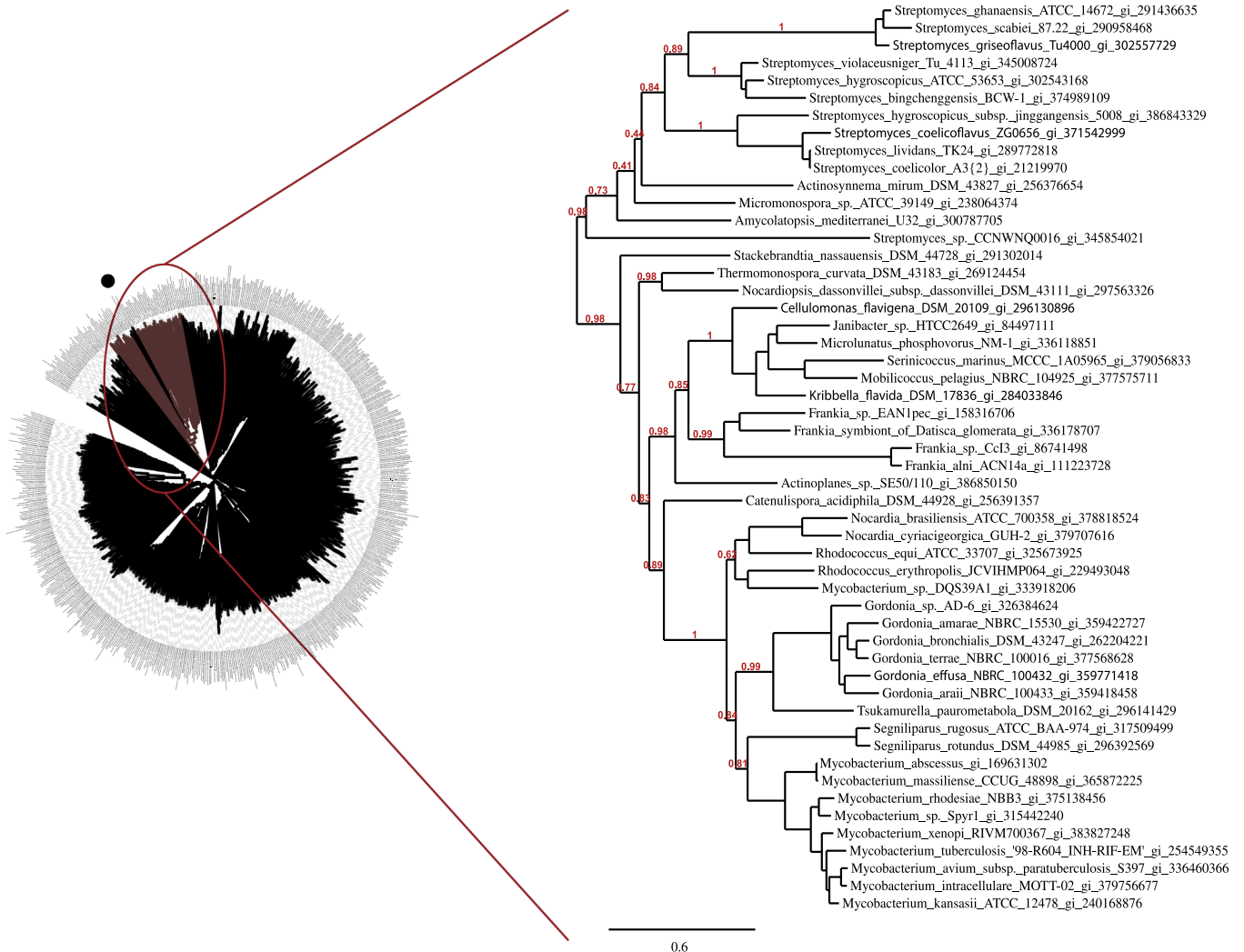
Fig. 8. Rbx and kinase domain co-occurrence is restricted to few *Actinomycetales*. Left: Distance-based tree of 652 sequence homologs with pairwise identities < 76%. The sequences were aligned with Mafft [22] and the tree was built with BioNJ [39]. Zoom in the clade containing sequences harboring both kinase and Rbx domains. They were multiply aligned using T-Coffee [23] for a maximum-likelihood reconstruction with PhyML [26]. The two non actinobacterial sequences (*A. cellulolyticus* and *K. racemifer*) were omitted. Right: Maximum-likelihood phylogenetic reconstruction for a set of 52 dissimilar homologous proteins. Species names and accession numbers are shown. Branch support values are shown in red, omitting a number of internal nodes for clarity. Scale bar indicates average substitutions per site.

on kinase activity, we attempted to reincorporate the iron ion into OA-NO₂-treated PknG. Nevertheless, we were not able to reassemble the Rbx domain by OA-NO₂ removal by thiol reagents in the presence of iron salts. In addition, previous attempts to

generate apo-Rbx-PknG under nondenaturing conditions failed, and thus did not permit more definitive conclusions. Elucidating the effect of iron release on more detailed Rbx structure and function will require further investigation. Significantly, PknG is
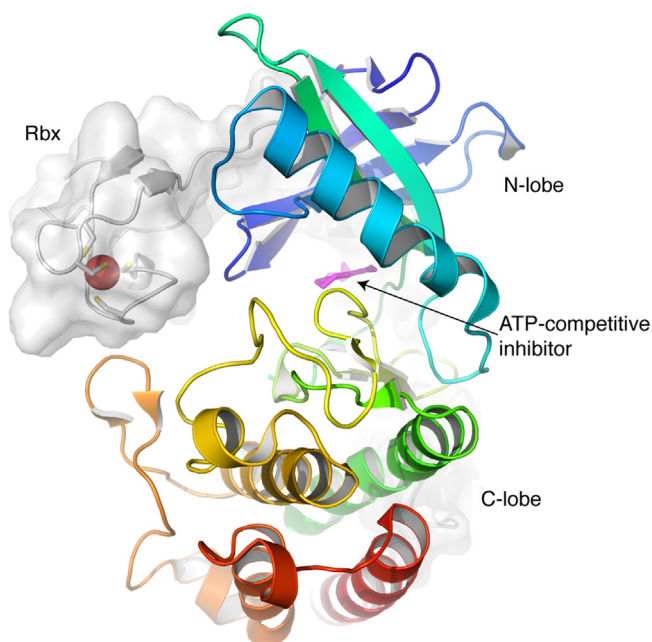
Scheme 1. Overall fold of PknG kinase and rubredoxin domains. The catalytic domain of *M. tuberculosis* PknG (PDB code 2PZI) shown in ribbon representation (colored from blue to red) interacts with the Rbx domain (ribbon+surface representation). The four Cys residues of the Rbx domain and the bound metal (red) are shown. The figure was drawn with Pymol.
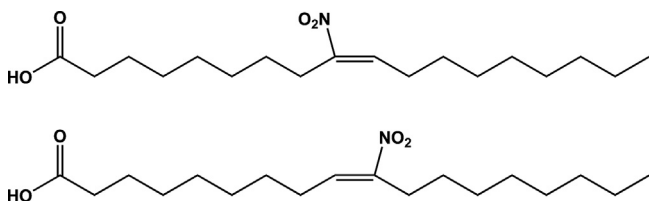
Scheme 2. Nitrated oleic acid (OA-NO$_2$). Two regioisomers of OA-NO$_2$ were synthesized by nitrosenylation of oleic acid yielding 9- and 10-nitro-9-*cis*-octadecenoic acids. Taken from [16].

one of the two soluble kinases present in *M. tuberculosis* lacking an extracellular sensor domain. The signals that activate/inactivate PknG and the mechanism for activity regulation are still unknown.

Fatty acid nitroalkenes are electrophilic species produced during inflammation and metabolism that react with nucleophilic amino acid residues of target proteins (i.e., Cys and His residues), reversibly modulating protein function and subcellular distribution. Nitroalkene reactivity is dictated by the electrophilic character of the β-carbon proximal to the alkenyl NO$_2$ group (Scheme 2) [16]. Nitroalkene fatty acids react with nucleophilic residues via a Michael addition to generate covalent adducts that can be reversed by thiol-containing molecules. This typically restores the native protein function [13]. Rbx-Cys are viewed as poor nucleophiles that do not react with low molecular weight alkylating reagents. However, mass spectrometry-based analysis confirmed that cysteines located at Rbx domain are the targets for nitroalkene reaction with PknG. While the kinase sequence contains reactive Cys and His residues, this is a novel reactivity of nitroalkenes with Rbx-Cys, wherein irreversible inhibition is achieved by reversible nitroalkylation of a redox-sensitive non-catalytic domain. The molecular basis of the selectivity of this reaction requires additional investigation.

Xanthine oxidoreductase (XOR) inhibition by OA-NO$_2$ was the first irreversible inhibition reported for fatty acid nitroalkenes. As in the case of PknG, XOR inhibition was not reversed by thiol

reagents [15]. At that time, we postulated that the inhibition of XOR was mediated via: (1) an irreversible covalent reaction between OA-NO$_2$ and XOR or (2) a reaction of the nitroalkene with the dithiolene of the pterin moiety and the concomitant loss of the molybdenum atom, with no direct evidence supporting either contention [15]. Inhibition of other enzymes by nitro-fatty acids was also previously reported [13,36]. OA-NO$_2$ inhibition of both GAPDH and PknG is achieved within the same range of micromolar concentrations. It is important to note that OA-NO$_2$ is considered a potent inhibitor of GAPDH, a key enzyme of intermediate metabolism that, due to a catalytic Cys residue, has also been postulated to be a redox sensor. OA-NO$_2$ is almost an order of magnitude more reactive toward GAPDH than hydrogen peroxide and peroxynitrite [13].

Covalent inhibitors display time-dependent inhibition and their potency must be characterized by the inactivation rate for different inhibitor concentrations. In the case of OA-NO$_2$, the determination of an inhibition constant for PknG is difficult for several reasons. OA-NO$_2$ concentration cannot be readily increased over its critical micelle concentration and increasing inhibitor concentration also increases the number of nonspecific covalent modifications, complicating kinetic analysis. There is a remarkable effect on kinase activity, with micromolar concentrations of OA-NO$_2$ at PknG:OA-NO$_2$ ratios of 1:3 and 1:5. Even though this does not allow us to evaluate inhibition constants, the data point to a potent effect of OA-NO$_2$ on kinase activity.

Analysis of Cys-alkylated peptides showed an unexpected mass shift by MALDI-TOF/TOF analysis. We hypothesized that laser ionization induces photodecomposition of cysteine-OA-NO$_2$ adducts, generating the observed pattern with a mass shift of 198/200/202 Da. In support of this, decomposition of nitro-compounds during MALDI analysis has been previously reported [37]. In contrast to cysteine modification, alkylated histidine showed the expected mass shift of 327 Da, indicating that the mass increment of 198/200/202 Da is a fingerprint of Cys modification when detected by MALDI MS. This can also explain why Cys-OA-NO$_2$ adducts were not detected in previous MALDI analyses [13]. MALDI MS/MS spectra of the modified peptides with this atypical mass increments did not show fragments of the 198/200/202 Da modification, thus not allowing a structural characterization.

This study exploited the unique structural characteristics of the multidomain protein kinase PknG for specific inhibition of enzymatic activity, revealing a new mechanism for PknG inhibition involving the Rbx domain rather than the catalytic domain. Electrophilic fatty acids thus represent a class of mediators that can induce the specific inhibition of a small subset of kinases containing a Rbx domain. Many *in vivo* studies have been performed with nitro oleic acid and are recently reviewed in Schopfer et al. [38]. This warrants further study of these PknG inhibitors in more complex model systems.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.freeradbiomed.2013.06.021.

References

[1] Wehenkel, A.; Bellinzoni, M.; Grana, M.; Duran, R.; Villarino, A.; Fernandez, P.; Andre-Leroux, G.; England, P.; Takiff, H.; Cervenansky, C.; Cole, S. T.; Alzari, P. M. Mycobacterial Ser/Thr protein kinases and phosphatases: physiological roles and therapeutic potential. *Biochim. Biophys. Acta* **1784:**193–202; 2008.

[2] Szekely, R.; Waczek, F.; Szabadkai, I.; Nemeth, G.; Hegymegi-Barakonyi, B.; Eros, D.; Szokol, B.; Pato, J.; Hafenbradl, D.; Satchell, J.; Saint-Joanis, B.; Cole, S. T.; Orfi, L.; Klebl, B. M.; Keri, G. A novel drug discovery concept for tuberculosis: inhibition of bacterial and host cell signalling. *Immunol. Lett.* **116:**225–231; 2008.

[3] Cole, S. T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S. V.; Eiglmeier, K.; Gas, S.; Barry 3rd C. E.; Tekaia, F.; Badcock, K.; Basham, D.; Brown, D.; Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Hornsby, T.; Jagels, K.; Krogh, A.; McLean, J.; Moule, S.; Murphy, L.; Oliver, K.; Osborne, J.; Quail, M. A.; Rajandream, M. A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, J.; Squares, R.; Squares, S.; Sulston, J. E.; Taylor, K.; Whitehead, S.; Barrell, B. G. Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature* **393:**537–544; 1998.

[4] Walburger, A.; Koul, A.; Ferrari, G.; Nguyen, L.; Prescianotto-Baschong, C.; Huygen, K.; Klebl, B.; Thompson, C.; Bacher, G.; Pieters, J. Protein kinase G from pathogenic mycobacteria promotes survival within macrophages. *Science* **304:**1800–1804; 2004.

[5] O'Hare, H. M.; Duran, R.; Cervenansky, C.; Bellinzoni, M.; Wehenkel, A. M.; Pritsch, O.; Obal, G.; Baumgartner, J.; Vialaret, J.; Johnsson, K.; Alzari, P. M. Regulation of glutamate metabolism by protein kinases in mycobacteria. *Mol. Biol.* **70:**1408–1423; 2008.

[6] Cowley, S.; Ko, M.; Pick, N.; Chow, R.; Downing, K. J.; Gordhan, B. G.; Betts, J. C.; Mizrahi, V.; Smith, D. A.; Stokes, R. W.; Av-Gay, Y. The Mycobacterium tuberculosis protein serine/threonine kinase PknG is linked to cellular glutamate/glutamine levels and is important for growth in vivo. *Mol. Biol.* **52:**1691–1702; 2004.

[7] Niebisch, A.; Kabus, A.; Schultz, C.; Weil, B.; Bott, M. Corynebacterial protein kinase G controls 2-oxoglutarate dehydrogenase activity via the phosphorylation status of the OdhI protein. *J. Biol. Chem.* **281:**12300–12307; 2006.

[8] Scherr, N.; Honnappa, S.; Kunz, G.; Mueller, P.; Jayachandran, R.; Winkler, F.; Pieters, J.; Steinmetz, M. O. Structural basis for the specific inhibition of protein kinase G, a virulence factor of Mycobacterium tuberculosis. *Proc. Natl. Acad. Sci. USA* **104:**12151–12156; 2007.

[9] Sieker, L. C.; Stenkamp, R. E.; LeGall, J. Rubredoxin in crystalline state. *Methods Enzymol.* **243:**203–216; 1994.

[10] van Beilen, J. B.; Neuenschwander, M.; Smits, T. H.; Roth, C.; Balada, S. B.; Witholt, B. Rubredoxins involved in alkane oxidation. *J. Bacteriol.* **184:**1722–1732; 2002.

[11] Tiwari, D.; Singh, R. K.; Goswami, K.; Verma, S. K.; Prakash, B.; Nandicoori, V. K. Key residues in Mycobacterium tuberculosis protein kinase G play a role in regulating kinase activity and survival in the host. *J. Biol. Chem.* **284:**27467–27479; 2009.

[12] Schopfer, F. J.; Batthyany, C.; Baker, P. R.; Bonacci, G.; Cole, M. P.; Rudolph, V.; Groeger, A. L.; Rudolph, T. K.; Nadtochiy, S.; Brookes, P. S.; Freeman, B. A. Detection and quantification of protein adduction by electrophilic fatty acids: mitochondrial generation of fatty acid nitroalkene derivatives. *Free Radic. Biol. Med.* **46:**1250–1259; 2009.

[13] Batthyany, C.; Schopfer, F. J.; Baker, P. R.; Duran, R.; Baker, L. M.; Huang, Y.; Cervenansky, C.; Branchaud, B. P.; Freeman, B. A. Reversible post-translational modification of proteins by nitrated fatty acids in vivo. *J. Biol. Chem.* **281:**20450–20463; 2006.

[14] Schopfer, F. J.; Cole, M. P.; Groeger, A. L.; Chen, C. S.; Khoo, N. K.; Woodcock, S. R.; Golin-Bisello, F.; Motanya, U. N.; Li, Y.; Zhang, J.; Garcia-Barrio, M. T.; Rudolph, T. K.; Rudolph, V.; Bonacci, G.; Baker, P. R.; Xu, H. E.; Batthyany, C. I.; Chen, Y. E.; Hallis, T. M.; Freeman, B. A. Covalent peroxisome proliferator-activated receptor gamma adduction by nitro-fatty acids: selective ligand activity and anti-diabetic signaling actions. *J. Biol. Chem.* **285:**12321–12333; 2010.

[15] Kelley, E. E.; Batthyany, C. I.; Hundley, N. J.; Woodcock, S. R.; Bonacci, G.; Del Rio, J. M.; Schopfer, F. J.; Lancaster Jr J. R.; Freeman, B. A.; Tarpey, M. M. Nitro-oleic acid, a novel and irreversible inhibitor of xanthine oxidoreductase. *J. Biol. Chem.* **283:**36176–36184; 2008.

[16] Baker, P. R.; Lin, Y.; Schopfer, F. J.; Woodcock, S. R.; Groeger, A. L.; Batthyany, C.; Sweeney, S.; Long, M. H.; Iles, K. E.; Baker, L. M.; Branchaud, B. P.; Chen, Y. E.; Freeman, B. A. Fatty acid transduction of nitric oxide signaling: multiple nitrated unsaturated fatty acid derivatives exist in human blood and urine and serve as endogenous peroxisome proliferator-activated receptor ligands. *J. Biol. Chem.* **280:**42464–42475; 2005.

[17] Blair, D.; Diehl, H. Bathophenanthrolinedisulphonic acid and bathocuproine-disulphonic acid, water soluble reagents for iron and cooper. *Talanta* **7:**163–174; 1961.

[18] Weber, G.; Young, L. B. Fragmentation of bovine serum albumin by pepsin. I. The origin of the acid expansion of the albumin molecule. *J. Biol. Chem.* **239:**1415–1423; 1964.

[19] Biegert, A.; Soding, J. Sequence context-specific profiles for homology searching. *Proc. Natl. Acad. Sci. USA* **106:**3770–3775; 2009.

[20] Altschul, S. F.; Madden, T. L.; Schaffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25:**3389–3402; 1997.

[21] Soding, J.; Remmert, M.; Biegert, A.; Lupas, A. N. HHsenser: exhaustive transitive profile search using HMM-HMM comparison. *Nucleic Acids Res.* **34:** W374–378; 2006.

[22] Katoh, K.; Toh, H. Recent developments in the MAFFT multiple sequence alignment program. *Briefings Bioinformatics* **9:**286–298; 2008.

[23] Notredame, C.; Higgins, D. G.; Heringa, J. T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* **302:**205–217; 2000.

[24] Loytynoja, A.; Goldman, N. An algorithm for progressive multiple alignment of sequences with insertions. *Proc. Natl. Acad. Sci. USA* **102:**10557–10562; 2005.

[25] Guindon, S.; Gascuel, O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *System. Biol* **52:**696–704; 2003.

[26] Guindon, S.; Delsuc, F.; Dufayard, J. F.; Gascuel, O. Estimating maximum likelihood phylogenies with PhyML. *Methods Mol. Biol.* **537:**113–137; 2009.

[27] Villarino, A.; Duran, R.; Wehenkel, A.; Fernandez, P.; England, P.; Brodin, P.; Cole, S. T.; Zimny-Arndt, U.; Jungblut, P. R.; Cervenansky, C.; Alzari, P. M. Proteomic identification of M. tuberculosis protein kinase substrates: PknB recruits GarA, a FHA domain-containing protein, through activation loop-mediated interactions. *J. Mol. Biol.* **350:**953–963; 2005.

[28] Bonomi, F.; Iametti, S.; Kurtz Jr D. M.; Ragg, E. M.; Richie, K. A. Direct metal ion substitution at the [M(SCys)4]2P site of rubredoxin. *J. Biol. Inorg. Chem.* **3:** 1998. 595–505.

[29] Messerschmidt, A., Huber, R., Poulos, T., Wieghardt, K., editors. *Handbook of metalloproteins.* Hoboken, NJ: Wiley; 2001.

[30] Fiuza, M.; Canova, M. J.; Zanella-Cleon, I.; Becchi, M.; Cozzone, A. J.; Mateos, L. M.; Kremer, L.; Gil, J. A.; Molle, V. From the characterization of the four serine/threonine protein kinases (PknA/B/G/L) of Corynebacterium glutamicum toward the role of PknA and PknB in cell division. *J. Biol. Chem.* **283:**18099–18112; 2008.

[31] Narayan, A.; Sachdeva, P.; Sharma, K.; Saini, A. K.; Tyagi, A. K.; Singh, Y. Serine threonine protein kinases of mycobacterial genus: phylogeny to function. *Physiol. Genomics* **29:**66–75; 2007.

[32] Alber, T. Signaling mechanisms of the Mycobacterium tuberculosis receptor Ser/Thr protein kinases. *Curr. Opin. Struct. Biol.* **19:**650–657; 2009.

[33] Nott, T.J.; Kelly, G.; Stach, L.; Li, J.; Westcott, S.; Patel, D.; Hunt, D.M.; Howell, S.; Buxton, R.S.; O'Hare, H.M.; Smerdon, S.J. An intramolecular switch regulates phosphoindependent FHA domain interactions in Mycobacterium tuberculosisSci. Signal.**2**ra12; 2009.

[34] Anand, N.; Singh, P.; Sharma, A.; Tiwari, S.; Singh, V.; Singh, D.K.; Srivastava, K.K.; Singh, B.N.; Tripathi, R.P. Synthesis and evaluation of small libraries of triazolylmethoxy chalcones, flavanones and 2-aminopyrimidines as inhibitors of mycobacterial FAS-II and PknG. Bioorg. Med. Chem.; 2012.

[35] Ortiz-Lombardia, M.; Pompeo, F.; Boitel, B.; Alzari, P. M. Crystal structure of the catalytic domain of the PknB serine/threonine kinase from Mycobacterium tuberculosis. *J. Biol. Chem.* **278:**13094–13100; 2003.

[36] Bonacci, G.; Schopfer, F. J.; Batthyany, C. I.; Rudolph, T. K.; Rudolph, V.; Khoo, N. K.; Kelley, E. E.; Freeman, B. A. Electrophilic fatty acids regulate matrix metalloproteinase activity and expression. *J. Biol. Chem.* **286:**16074–16081; 2011.

[37] Sarver, A.; Scheffler, N. K.; Shetlar, M. D.; Gibson, B. W. Analysis of peptides and proteins containing nitrotyrosine by matrix-assisted laser desorption/ionization mass spectrometry. *J. Am. Soc. Mass Spectrom* **12:**439–448; 2001.

[38] Schopfer, F. J.; Cipollina, C.; Freeman, B. A. Formation and signaling actions of electrophilic lipids. *Chem. Rev.* **111:**5997–6021; 2012.

[39] Desper, R.; Gascuel, O. Getting a tree fast: Neighbor Joining, FastME, and distance-based methods. *Curr. Protoc.Bioinformatics* **Chapter 6(Unit 6):**3; 2006.

Figure S1. **Cysteine reactivity against other alkylating agents.** PknG was exposed to IAM as a model reagent for cysteine alkylation. **A)** Mass spectrum of tryptic digest of native PknG. *Upper panel*: peak *m/z* 1292.58 (FC$_{106}$WNC$_{109}$GRPVGR) is shown, *middle panel*: peak *m/z* 3530.59 (GASEGWC$_{128}$PYC$_{131}$GSPYSFLPQLNPGDIVAGQYEVK) is indicated and *lower panel*: peak *m/z* 1814.94 (GC$_{156}$IAHGGLGWIYLALDR) is indicated. **B)** Mass spectrum of peptides obtained from tryptic digestion of PknG (10 µM) exposed to IAM (150 µM). *Upper panel*: no new signal appeared in the spectra, *middle panel*: no modification appeared in the spectra and *left panel*: after the treatment with IAM a new peak showing a mass increase of 57 Da appeared. **C)** Fragmentation pattern of *m/z* 1871.95. The increment of 57 Da in ion fragments corresponds to the incorporation of IAM. **D)** Mass spectrum of: native GarA (*upper panel*), GarA phosphorylated by native PknG (*middle panel*) and GarA phosphorylated by PknG previously treated with IAM (*lower panel*).

C)

D)
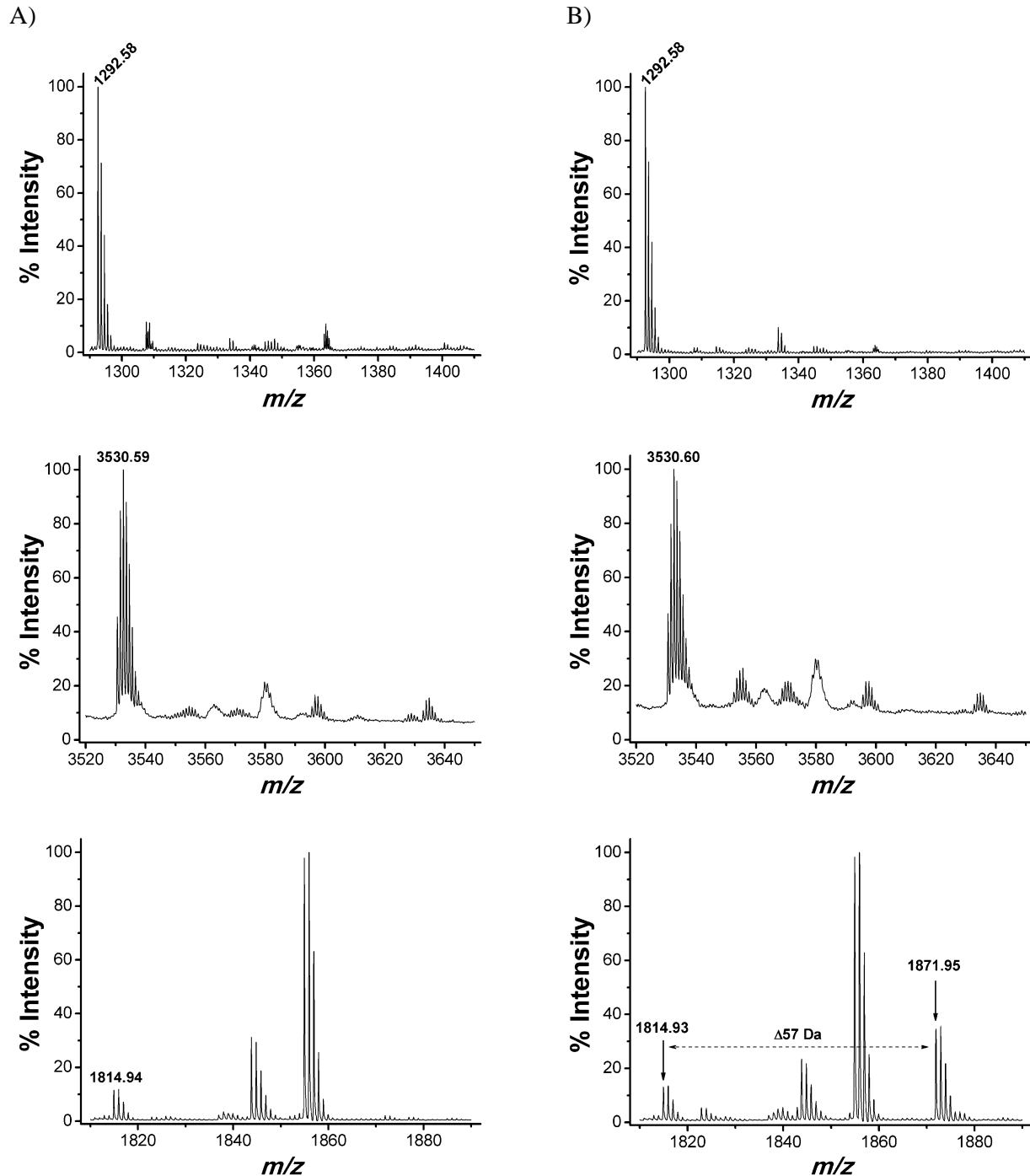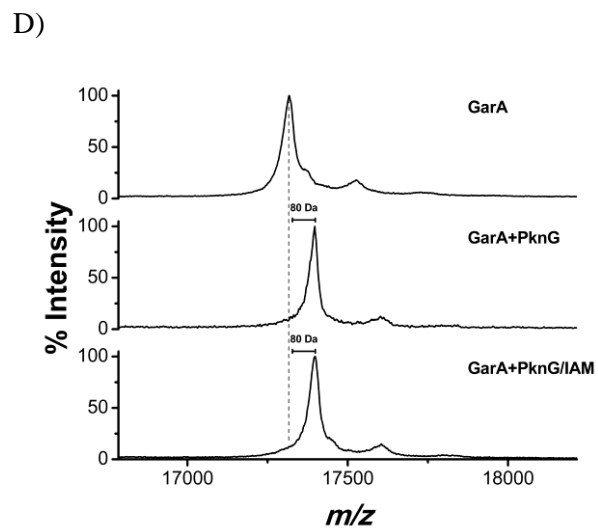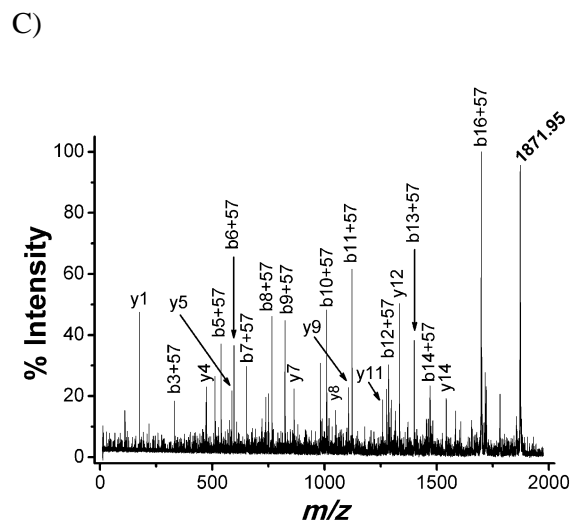
Fig. S1

Figure S2. **Identification of modified histidine residues on nitroalkylated PknG.** PknG was exposed to OA-NO$_2$ in a molar ratio PknG:OA-NO$_2$ of 1:5. **A)** *Left panel*: mass spectrum of peptides generated by tryptic digestion of native PknG, peak *m/z* 732.39 (**H**$_{635}$FTTAR) is indicated. *Right panel*: mass spectrum of peptides generated by tryptic digestion of PknG treated with OA-NO$_2$. **B)** Fragmentation pattern of modified peptide **H**$_{635}$FTTAR obtained from tryptic digest of OA-NO$_2$-treated PknG. H$_{635}$ was identified as the modified residue **C)** List of MS/MS fragment ions *m/z* of OA-NO$_2$ modified peptide 635-640. Detected ions are highlighted in bold. **D)** *Left panel*: mass spectrum of peptides generated by tryptic digestion of native PknG, peak *m/z* 1800.90 (**H**$_{558}$FTEVLDTFPGELAPK), *m/z* 1843.86 (GLV**H**$_{185}$SGDAEAQAMAMAER) and *m/z* 1854.95 (ASTN**H**$_{702}$ILGFPFTS**H**$_{711}$GLR) are indicated. *Right panel*: mass spectrum of peptides generated by tryptic digestion of PknG treated with OA-NO$_2$. **E)** *Left panel*: mass spectrum of peptides generated by tryptic digestion of native PknG, peak *m/z* 2486.19 (**H**$_{488}$GALDADGVDFSESVELPLMEVR) and *m/z* 2614.30 (STFGVDLLVA**H**$_{430}$TDVYLDGQV**H**$_{440}$AEK) are indicated. *Right panel*: mass spectrum of peptides generated by tryptic digestion of PknG treated with OA-NO$_2$.

A)
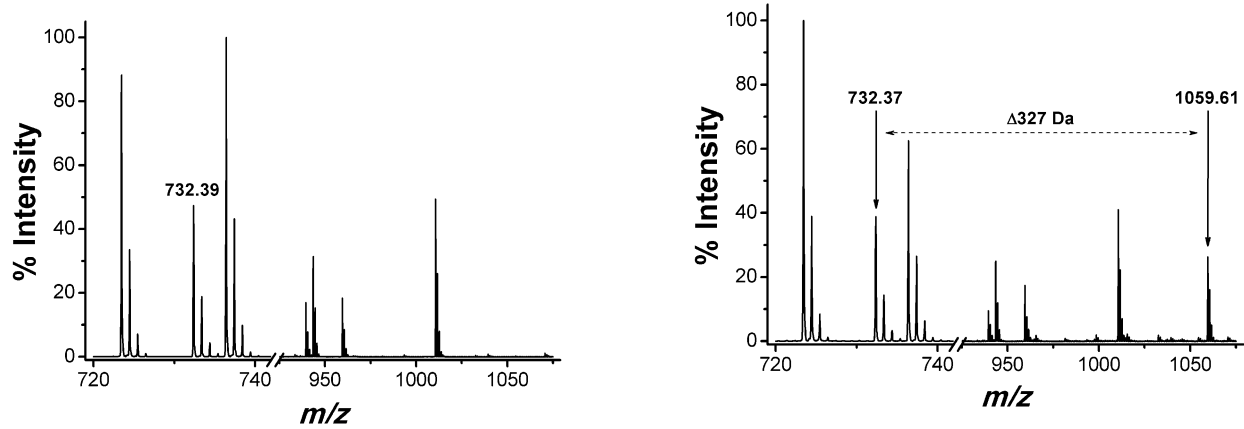
B)
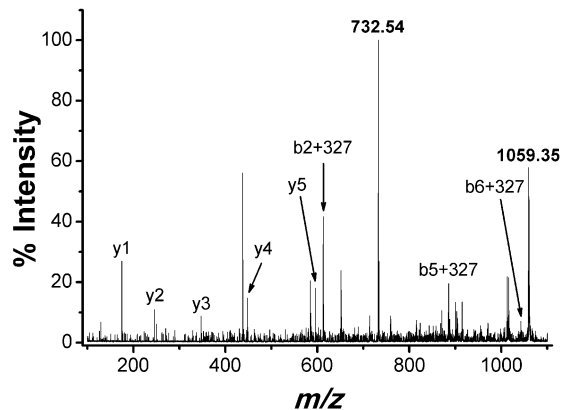
C)

#	b	Residue	y	#
1		**H***		6
2	**612,40**	F	**595,68**	5
3	**713,50**	T	**448,50**	4
4	814,61	T	**347,40**	3
5	**885,69**	A	**246,29**	2
6	**1041,87**	R	**175,21**	1

H* = His-OA-NO$_2$

D)
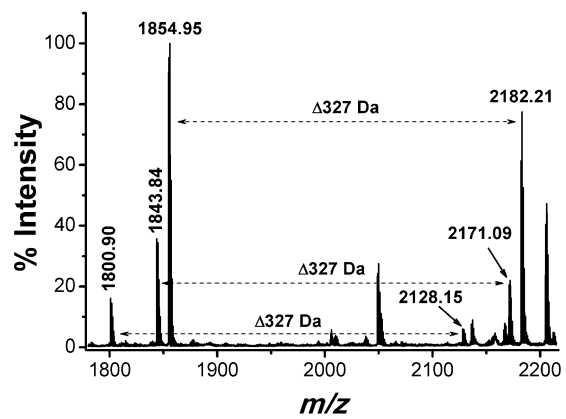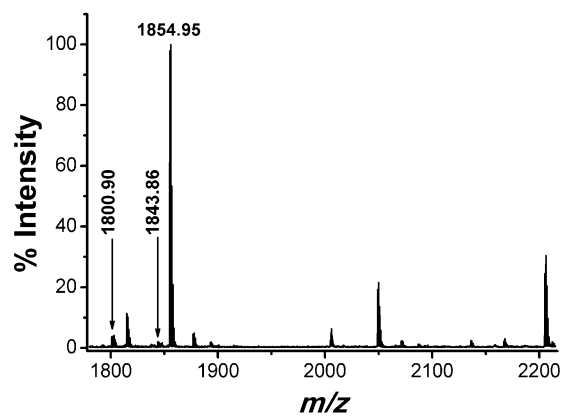
E)
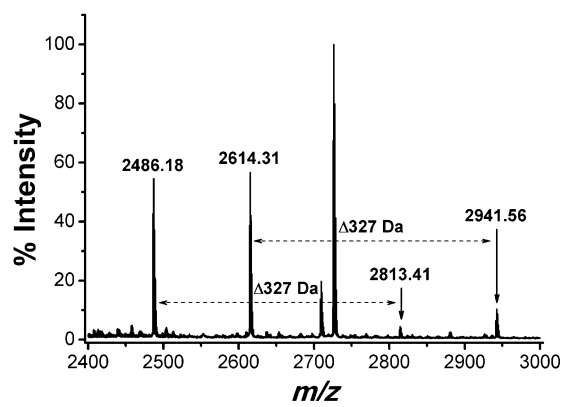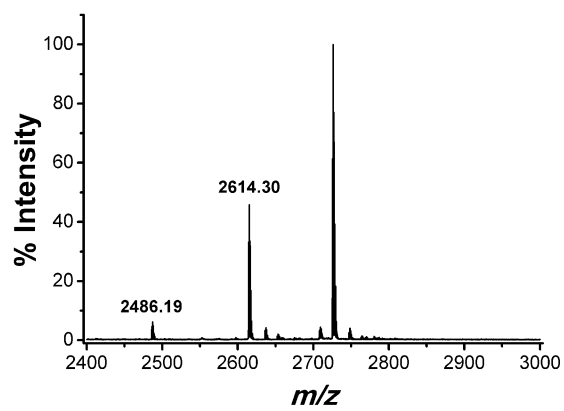
Fig. S2

Figure S3. **Modification sites of PknGΔ74/TPR by OA-NO₂.**
Left panel: mass spectrum of tryptic digest of native PknGΔ74/TPR, cysteine-containing peptide (GASEGW**C**$_{128}$PY**C**$_{131}$GSPYSFLPQLNPGDIVAGQYEVK) is indicated. *Right panel*: mass spectrum of tryptic digest of OA-NO₂-treated PknGΔ74/TPR. A new signal appears showing a mass increment of 198/200/202 Da with respect to peak *m/z* 3530.63.
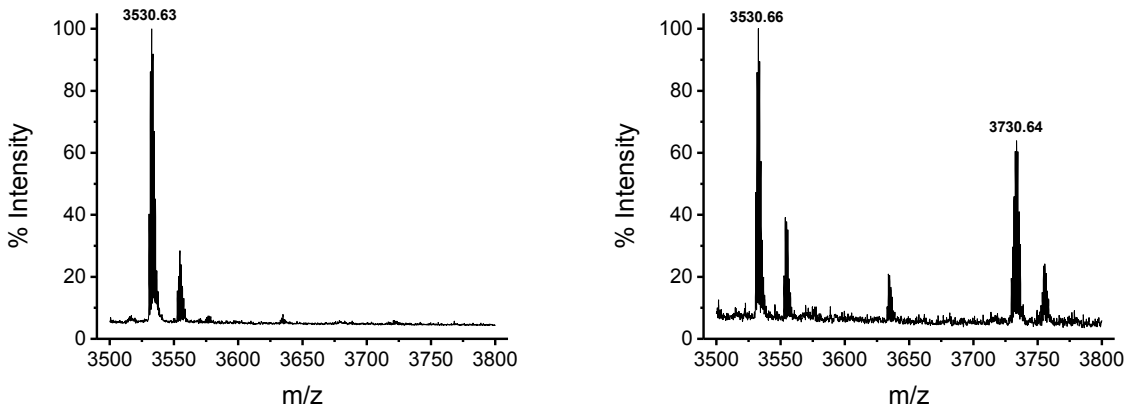
Fig. S3

Table S1. **MS/MS data of OA-NO$_2$ modified peptide 122 - 154.** List of theoretical *m/z* values of fragment ions from peptide GASEGW**C**$_{128}$PY**C**$_{131}$GSPYSFLPQLNPGDIVAGQYEVK and its corresponding ion +200Da. y- and b-ions detected by MALDI MS are highlighted in bold. The observed mass shift after OA-NO$_2$ treatment was 198/200 Da.

#	b	b + 200	Residue	y	y + 200	#
1	58,03	258,03	G			33
2	129,07	329,07	A	3475,60	3675,60	32
3	216,10	416,10	S	3404,56	**3604,56**	31
4	345,14	545,14	E	3317,53	**3517,53**	30
5	402,16	602,16	G	3188,49	**3388,49**	29
6	588,24	788,24	W	3131,47	3331,47	28
7	690,24	890,24	**C**	2945,39	**3145,39**	27
8	787,30	987,30	P	**2842,38**	**3042,38**	26
9	950,36	1150,36	Y	2745,32	2945,32	25
10	1052,36	1252,36	**C**	**2582,26**	**2782,26**	24
11	1109,38	1309,38	G	**2479,25**	2679,25	23
12	1196,41	1396,41	S	**2422,23**	2622,23	22
13	1293,47	1493,47	P	**2335,20**	2535,20	21
14	1456,53	1656,53	Y	2238,15	2438,15	20
15	1543,56	1743,56	S	**2075,08**	2275,08	19
16	1690,63	1890,63	F	**1988,05**	2188,05	18
17	1803,72	2003,72	L	**1840,98**	2040,98	17
18	1900,77	2100,77	P	**1727,90**	1927,90	16
19	2028,83	2228,83	Q	1630,84	1830,84	15
20	2141,91	2341,91	L	**1502,79**	1702,79	14
21	2255,95	**2455,95**	N	**1389,70**	1589,70	13
22	2353,01	2553,01	P	**1275,66**	1475,66	12
23	2410,03	2610,03	G	1178,61	1378,61	11
24	2525,05	**2725,05**	D	**1121,58**	1321,58	10
25	2638,14	**2838,14**	I	**1006,56**	1206,56	9
26	2737,21	**2937,21**	V	893,47	1093,47	8
27	2808,24	**3008,24**	A	**794,41**	994,41	7
28	2865,27	3065,27	G	**723,37**	923,37	6
29	2993,32	3193,32	Q	**666,35**	866,35	5
30	3156,39	3356,39	Y	538,29	738,29	4
31	3285,43	**3485,43**	E	375,22	575,22	3
32	3384,50	3584,50	V	246,18	446,18	2
33	3512,59	3712,59	K	147,11	347,11	1

Table S1

55

Table S2. **Identification of modified residues within PknGΔ74/TPR sequence.**

Peptide from-to	Modified residue	Assigned sequence[a]
		Peptides containing modified cysteine residues (Δm=198/200/202 Da)
105 – 111	C_{106} or C_{109}	FCWNCGR
105 - 115	C_{106} or C_{109}	FCWNCGRPVGR
122 - 144	C_{128} or C_{131}	GASEGWCPYCGSPYSFLPQLNPGDIVAGQYEVK
		Peptides containing modified histidine residues (Δm=327 Da)
182 - 199	H_{185}	GLVHSGDAEAQAMAMAER
200 - 222	H_{207} or H_{219}	QFLAEVVHPSIVQIFNFVEHTDR

[a]The sequences and nitroalkylation sites were confirmed by MS/MS analysis.

Modulación de la actividad quinasa por modificación específica de residuos de cisteína del dominio rubredoxina

Además del dominio catalítico quinasa, PknG de *M. tuberculosis* posee dominios N- y C- terminales con roles poco definidos. El extremo amino-terminal presenta un motivo rubredoxina, caracterizado por la presencia de un átomo de hierro coordinado a cuatro residuos de cisteínas y que en muchas bacterias anaerobias participa en la transferencia de electrones. A nuestro mejor saber y entender, esta combinación de dominios no se encuentra en ninguna otra quinasa descripta hasta el momento y por tanto es una característica singular de las PknG de micobacterias de función desconocida.

Trabajos previos habían descripto la conservación del dominio catalítico así como la variabilidad de la región N-terminal de PknG [133, 134] pero no se había evaluado la asociación entre los dominios quinasa y Rbx. El alineamiento múltiple de secuencias ortólogas de PknG, realizado en el presente trabajo, mostró que la coocurrencia de los dominios quinasa y rubredoxina se restringía a unos pocos miembros del orden *Actinomycetales* (Fig. 8 en artículo 1). Además, estos datos permitieron descartar el aparente vínculo entre la presencia del dominio Rbx en PknG y la patogenicidad ya que tanto micobacterias patógenas como no patógenas presentan ambos dominios. Una importante conclusión que surge de este análisis es que el dominio rubredoxina está ausente en PknG de *C. glutamicum*, organismo estrechamente emparentado a *M. tuberculosis*. En este organismo, PknG regula el metabolismo del glutamato a través de la fosforilación de una pequeña proteína ortóloga a GarA: OdhI, de forma similar a lo que ocurre en *M. tuberculosis* [103]. Este resultado plantea la interrogante de si PknG está cumpliendo otras funciones o actuando en respuesta a otros estímulos en las micobacterias.

Sabiendo que la disrupción del dominio Rbx por sustitución de los residuos de Cys disminuye significativamente la actividad quinasa de PknG [63] y que la asociación de dominios quinasa y Rbx está confinada a un grupo reducido de organismos, nos propusimos evaluar el efecto de la modificación química de este dominio sobre la actividad quinasa de PknG. Para ello estudiamos la reactividad de derivados nitrados de ácidos grasos insaturados (moléculas electrofílicas capaces de reaccionar específicamente con residuos de His y Cys en las proteínas) con PknG y su efecto sobre la catálisis.

Nuestros resultados indican que el ácido nitro-oleico (ácido 9- y 10-nitro-octadeca-9-*cis*-enoico; OA-NO$_2$) es un inhibidor que reacciona covalentemente con la enzima (Fig. 1 en artículo 1). Utilizando una aproximación por espectrometría de masa demostramos que las Cys localizadas en el dominio rubredoxina son el blanco preferencial del ácido nitro-oleico, y que esta modificación es además la responsable de la pérdida de actividad enzimática (Fig. 2A y 2B en artículo 1). Si bien

demostramos que la modificación puede ocurrir en los dos motivos CxxCG que forman parte de este dominio, los datos de MS/MS no fueron concluyentes al momento de establecer la identidad del residuo modificado (Fig. 2C en artículo 1).

La secuencia de PknG contiene una sola Cys que no forma parte del dominio Rbx y en nuestro trabajo mostramos que el péptido conteniendo este residuo no se modifica como resultado del tratamiento con OA-NO$_2$ (Fig. 2A y 2B en artículo 1). Para ahondar en la comprensión de este fenómeno, evaluamos a la iodoacetamida (IAM) como agente alquilante modelo. Observamos que las Cys del dominio Rbx no se modifican por IAM mientras que la Cys libre si lo hace y que la actividad quinasa de PknG no se ve afectada por el tratamiento (Fig. S1 en artículo 1). En consecuencia, las Cys de los dominios Rbx, típicamente reportadas como no reactivas frente a agentes alquilantes, aparecen como los blancos preferenciales de la reacción con nitroalquenos electrofílicos.

Los derivados nitrados de ácidos grasos insaturados son capaces de adicionar compuestos nucleofílicos mediante adición de Michael, reacción típicamente revertida por reactivos con grupos sulfhidrilo. Si bien demostramos que en el caso de PknG, la modificación química sigue el comportamiento usual de los aductos de Michael (Fig. 5B en artículo 1), no fue posible recuperar la actividad de PknG luego del tratamiento con ditiotreitol (Fig. 5A en artículo 1) como había sido reportado para otras enzimas con tioles críticos [135].

Con el fin de comprender cómo una modificación química reversible generaba una inhibición irreversible, analizamos el efecto de la misma sobre la estructura de la proteína y el dominio rubredoxina. Mientras que la modificación covalente de las cisteínas no altera globalmente la estructura secundaria o terciaria de PknG (Fig. 7 en artículo 1) observamos que provoca la salida del hierro del dominio Rbx (Fig. 6 en artículo 1).

Para elucidar el posible efecto de la pérdida del ion sobre la estructura del sitio catalítico de PknG, nos propusimos llevar a cabo ensayos de cristalogénesis de la enzima inhibida por OA-NO$_2$. Brevemente, las diferentes versiones de PknG (esquema a continuación) fueron purificadas, diluidas hasta una concentración final de 25 μM, tratadas con ácido nitro-oleico 100 μM y posteriormente preparadas en el amortiguador adecuado para los ensayos de cristalogénesis (Tris 50 mM, NaCl 500 mM, glicerol 5% pH 8.0) en presencia o no de otros componentes como ser ADP e incluso la proteína sustrato de PknG, GarA. Las proteínas fueron luego concentradas por ultrafiltración y la evaluación de distintas condiciones de cristalización se efectúo de forma automatizada usando kits comerciales

que exploran hasta 576 condiciones de cristalización por vez. Sin embargo, en la mayoría de las condiciones evaluadas no se obtuvieron formas cristalinas.
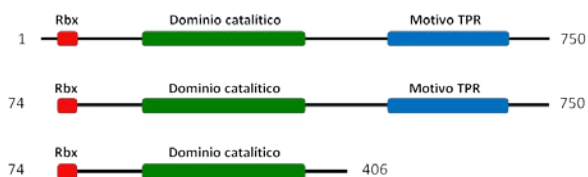
Se observó el crecimiento de cristales en forma de placas para la condición de PknG$_{\Delta73/\Delta TPR}$ 30 mg/mL pretratada con OA-NO$_2$ 100 μM, ADP 3.5 mM en Hepes 100 mM, PEG 4000 30% w/v, CaCl$_2$ 200 mM pH 7.5 (Fig. 13). La difracción de rayos X de estos cristales se realizó en el Swiss Light Source del Paul Scherrer Institut en Villigen (Suiza) y se observó que los cristales analizados difractan a muy baja resolución (10-15 Å). Estos resultados en sí mismo son desalentadores, no obstante, si se compara la resolución obtenida en la misma condición pero en ausencia de tratamiento con ácido nitro-oleico se observa que la difracción en ese caso es 1.5 Å. Estos datos estarían indicando que el OA-NO$_2$ de alguna forma altera la estructura de los cristales.

A pesar de los esfuerzos realizados en la optimización de estos cristales no fue posible mejorar la resolución. Una explicación posible a la baja resolución obtenida es que la modificación de los residuos de cisteína del dominio Rbx sea heterogénea. Una estrategia alternativa es el agregado de ligandos una vez que el cristal está formado (del inglés *soaking*). En este caso sumergimos cristales de PknG$_{\Delta73/\Delta TPR}$ obtenidos en la siguiente condición de cristalización: Hepes 100 mM, PEG 4000 35% w/v, CaCl$_2$ 200 mM pH 7.25 en presencia de ADP 2 mM en una mezcla conteniendo solución de cristalización, glicerol 10% y OA-NO$_2$ 200 μM durante 10, 60, 120 y 210 minutos. Estos cristales fueron llevados a European Synchrotron Radiation Facility en Grenoble (Francia) y se obtuvieron datos de muy buena calidad. Sin embargo, la resolución de esta estructura confirmó que el ácido nitro-oleico no había modificado ningún residuo de la proteína y que por tanto no se habían producido cambios estructurales. En resumen, a partir de los datos cristalográficos no fue posible extraer conclusiones respecto al estado del dominio rubredoxina una vez ocurrida la modificación de las cisteínas.
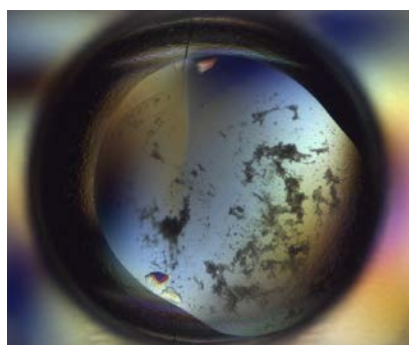
Figura 13. Imagen de los cristales de PknG$_{\Delta73/\Delta TPR}$ tratada con OA-NO$_2$. Aparición de cristales en forma de placa de PknG$_{\Delta73/\Delta TPR}$ 30 mg/mL pretratada con OA-NO$_2$ en presencia de ADP 3.5 mM en Hepes 100 mM pH 7.5, PEG 4000 30% w/v, CaCl$_2$ 200 mM luego de tres semanas a 20 °C.

En conjunto nuestros resultados permiten describir un nuevo mecanismo para la inhibición de PknG que involucra **la liberación del ion presente en el dominio rubredoxina por modificación específica de los residuos de Cys** que lo coordinan. Este hallazgo es especialmente relevante ya que reportes previos demuestran que la inhibición de PknG genera bacterias más suceptibles a ser eliminadas por los macrófagos [121]. Si bien las quinasas de proteínas han sido consideradas blancos farmacológicos muy interesantes, los inhibidores utilizados son de muy baja especifidad ya que generalmente están dirigidos al dominio catalítico, estructura altamente conservada entre quinasas procariotas y eucariotas. La principal interrogante que surge a partir de nuestros datos es si estos compuestos podrán ser efectivos en modelos biológicamente relevantes.

Por tanto, en la última etapa de este trabajo comenzamos a evaluar el efecto del OA-NO$_2$ sobre la viabilidad de las micobacterias en colaboración con la Dra. Mariana Piuri del Departamento de Química Biológica (Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Argentina). Para abordar este objetivo se utilizaron fluoromicobacteriófagos (virus que infectan de manera específica bacterias del género *Mycobacterium* [136] y que han sido modificados para expresar genes reporteros fluorescentes [137-139]) ya que las micobacterias una vez infectadas comienzan a expresar los genes virales. Este tipo de abordajes constituyen una manera simple de revelar la viabilidad celular. Luego de la infección, las células viables pueden ser fácilmente detectadas por microscopía de fluorescencia o citometría de flujo. Una ventaja adicional de esta técnica es que las muestras pueden ser fijadas con paraformaldehído, sin pérdida de intensidad de fluorescencia por un período de al menos 14 días, lo que facilita su manipulación [137].

La cepa no virulenta *M. tuberculosis* MC^{2}6230 [140] (1×10^8 células) se trató con OA-NO$_2$ 20 μM en dimetilsulfóxido (DMSO) 2% durante 24 horas a 37 °C. Transcurrido ese período, las células se infectan con el fásmido termosensible pHA159::*hsp60-mCherrybomb* con una multiplicidad de infección de 100 y se registra la emisión de fluorescencia (λ_{exc} = 584 nm, λ_{em} = 620 nm). En la Figura 14 se muestra un gráfico de intensidad de fluorescencia en función del tiempo para las cepas control y tratadas con ácido nitro-oleico. Como claramente se ve en la figura, el tratamiento con el inhibidor provoca una disminución de la viabilidad celular de las micobacterias respecto a las expuestas solamente al vehículo (DMSO) (Fig. 14). Estos resultados, aunque preliminares, son muy auspiciosos y plantean la pregunta si la inhibición de PknG podría ser la responsable de la disminución de la viabilidad celular.
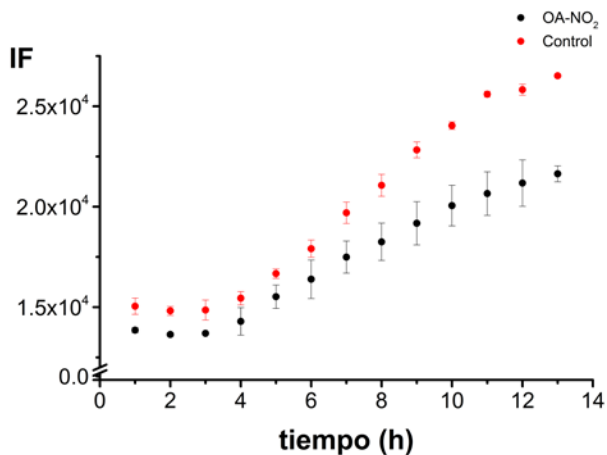
Figura 14. Curva de crecimiento de *M. tuberculosis* MC^{2}6230 tratada con OA-NO$_2$. 1×10^8 células de *M. tuberculosis* MC^{2}6230 fueron tratadas con OA-NO$_2$ 20 µM en DMSO 2% durante 24 horas a 37 °C. Posteriormente, las células tratadas se infectan con pHA159::*hsp60-mCherrybomb* con una multiplicidad de infección de 100 y se registra la emisión de fluorescencia a 620 nm.

Por tanto, nuestros resultados nos permiten postular no solo que los ácidos grasos nitrados representan una nueva familia de inhibidores de PknG sino que constituyen promisorios agentes terapéuticos que merecen seguir siendo evaluados en modelos celulares y animales. En este sentido, una ventaja adicional es que estos compuestos han demostrado ser inocuos en ensayos clínicos (Dr. Francisco J. Schopfer, comunicación personal). Finalmente cabe destacar que en vista de la posible importancia de estos compuestos para el tratamiento de la tuberculosis, este trabajo ha dado lugar por una parte a una Patente de Invención en Estados Unidos (Composition and Method for Inhibition of PknG from *Mycobacterium tuberculosis*", co-inventors: C. Batthyány & R. Durán, US20150051283 A1) y por otra al desarrollo de una segunda generación de nitroalquenos en base a un diseño racional y que está siendo evaluada en nuestro laboratorio.

Capítulo 2

Structure

Molecular Basis of the Activity and the Regulation of the Eukaryotic-like S/T Protein Kinase PknG from *Mycobacterium tuberculosis*

Graphical Abstract
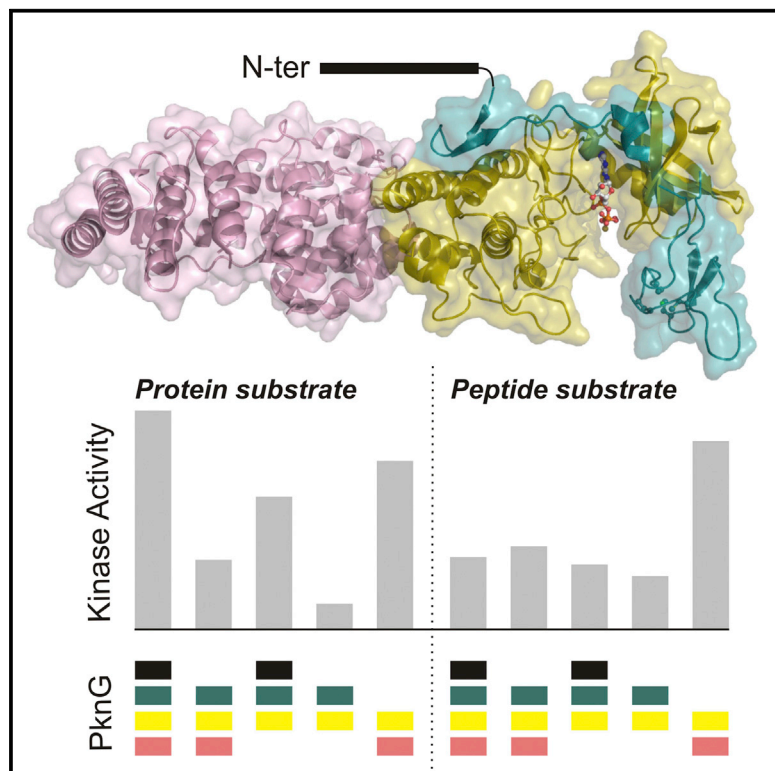
Authors

María-Natalia Lisa, Magdalena Gil, ..., Ricardo M. Biondi, Pedro M. Alzari

Correspondence

pedro.alzari@pasteur.fr

In Brief

In this work, Lisa and colleagues investigate the molecular determinants of enzyme activity, substrate specificity, and regulation of the protein kinase PknG from *Mycobacterium tuberculosis*, providing alternatives for the design of selective inhibitors with potential therapeutic applications.

Highlights

- The active site of PknG preferentially interacts with extended peptide substrates

- PknG domains flanking the kinase core define substrate selectivity

- A rubredoxin-like metal-binding domain downregulates PknG kinase activity

- A K_{181}-E_{198} salt bridge is not required for PknG kinase activity

Accession Numbers

4Y0X
4Y12

CrossMark

CellPress

Molecular Basis of the Activity and the Regulation of the Eukaryotic-like S/T Protein Kinase PknG from *Mycobacterium tuberculosis*

María-Natalia Lisa,[1,2,3] Magdalena Gil,[4] Gwénaëlle André-Leroux,[1,2,3,5] Nathalie Barilone,[1,2,3,7] Rosario Durán,[4] Ricardo M. Biondi,[6] and Pedro M. Alzari[1,2,3,*]

[1]Institut Pasteur, Unité de Microbiologie Structurale, 25 rue du Docteur Roux, 75724 Paris Cedex 15, France
[2]CNRS UMR 3528, Biologie structurale des processus cellulaires et maladies infectieuses, Institut Pasteur, 75724 Paris Cedex 15, France
[3]Université Paris Diderot, Sorbonne Paris Cité, Microbiologie structurale, 75724 Paris Cedex 15, France
[4]Unidad de Bioquímica y Proteómica Analíticas, Instituto de Investigaciones Biológicas Clemente Estable e Institut Pasteur de Montevideo, Mataojo 2020, Montevideo 11400, Uruguay
[5]INRA, UR1404 Unité Mathématiques et Informatique Appliquées du Génome à, l'Environnement, 78352 Jouy-en-Josas, France
[6]Research Group PhosphoSites, Medizinische Klinik 1, Universitätsklinikum Frankfurt, Theodor-Stern-Kai 7, 60590 Frankfurt, Germany
[7]Present address: Institut Pasteur, Unité de Neuro-Immunologie viral, 25 rue du Docteur Roux, 75724 Paris Cedex 15, France
*Correspondence: pedro.alzari@pasteur.fr
http://dx.doi.org/10.1016/j.str.2015.04.001

SUMMARY

Tuberculosis remains one of the world's deadliest human diseases, with a high prevalence of antibiotic-resistant *Mycobacterium tuberculosis* (*Mtb*) strains. A molecular understanding of processes underlying regulation and adaptation of bacterial physiology may provide novel avenues for the development of antibiotics with unconventional modes of action. Here, we focus on the multidomain S/T protein kinase PknG, a soluble enzyme that controls central metabolism in Actinobacteria and has been linked to *Mtb* infectivity. Our biochemical and structural studies reveal how different motifs and domains flanking the catalytic core regulate substrate selectivity without significantly affecting the intrinsic kinase activity, whereas a rubredoxin-like domain is shown to downregulate catalysis through specific intramolecular interactions that modulate access to a profound substrate-binding site. Our findings provide the basis for the selective and specific inhibition of PknG, and open new questions about regulation of related bacterial and eukaryotic protein kinases.

INTRODUCTION

Reversible protein phosphorylation has evolved as a ubiquitous molecular mechanism of protein regulation. In eukaryotes, protein phosphorylation commonly takes place on serine, threonine, and tyrosine residues and is catalyzed by a large group of enzymes, the eukaryotic protein kinases (ePKs) (Huse and Kuriyan, 2002; Kornev and Taylor, 2010). Evolutionarily related S/T protein kinases (STPKs) also exist in many sequenced bacterial genomes, including important human pathogens such as *Mycobacterium tuberculosis* (*Mtb*), *Staphylococcus aureus*, *Pseudomonas aeruginosa*, and *Bacillus anthracis*, where they fulfill

important roles in bacterial physiology and pathogenesis (Wehenkel et al., 2008). While kinase-mediated signaling mechanisms in bacteria share important common features with their eukaryotic counterparts, specific differences are emerging as novel functions and components of bacterial phosphosystems are being identified (Sherman and Grundner, 2014). The genome of *Mtb*, the causative agent of tuberculosis, codes for 11 STPKs (Cole et al., 1998), some of which have been extensively investigated and shown to be either essential for and/or involved in the regulation of metabolic processes, gene transcription, cell division, and host-pathogen interactions (Wehenkel et al., 2008; Av-Gay and Everett, 2000; Sassetti et al., 2003). Among these, the protein kinase PknG is exclusively found in Actinobacteria and displays a unique modular organization, which seems to be absent in known ePKs.

PknG is of particular interest because of its roles in *Mtb* pathogenicity, where it has been proposed to inhibit the phagosome-lysosome fusion within the infected macrophage (Walburger et al., 2004), and in bacterial metabolism, where it regulates the fate of α-ketoglutarate, a key metabolic intermediate at the crossroads of the tricarboxylic acid cycle (C metabolism) and glutamate synthesis (N metabolism) (Cowley et al., 2004; Ventura et al., 2013). While the signaling pathway mediating infectivity has not been unequivocally determined, PknG is known to phosphorylate in vivo the *Mtb* regulator GarA (O'Hare et al., 2008), an FHA (Forkhead Associated) protein (England et al., 2009; Barthe et al., 2009) that controls the accumulation of glutamate by direct binding to three metabolic enzymes that use α-ketoglutarate as substrate (O'Hare et al., 2008; Nott et al., 2009). Thus, while the knockdown of GarA induces a glutamate/glutamine/asparagine auxotrophy (these amino acids are required as N donors) (Ventura et al., 2013), *Mtb* Δ*PknG* shows increased levels of intracellular glutamate (Cowley et al., 2004), and the related *Corynebacterium glutamicum* Δ*PknG* cannot grow in a minimal medium containing glutamine as the sole C source (Niebisch et al., 2006).

Despite these dual roles in bacterial physiology and pathogenesis, very little is known about the molecular basis of PknG activity and regulation. PknG is a multidomain protein, with a unique domain organization. Flanking the kinase catalytic core,
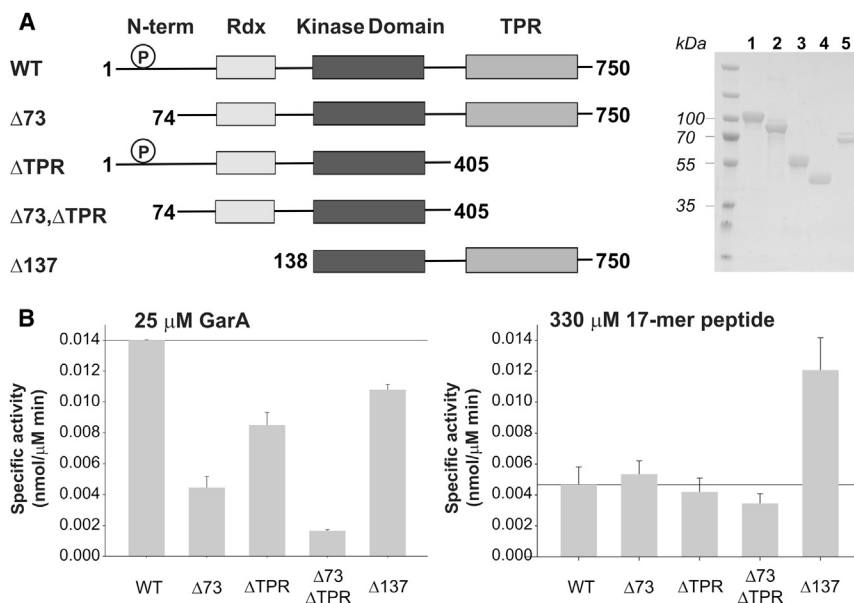
Figure 1. PknG Deletion Mutants and their Relative Kinase Activities

(A) The left panel shows a schema representing the different PknG constructions (wild-type and deletion mutants) used in this study. The cluster of autophosphorylation sites in the N-terminal region of PknG is indicated by a (P) symbol. The right panel shows a Coomassie-stained SDS-PAGE of the purified PknG constructions. Lanes 1 to 4 correspond to H$_6$-tagged wild-type PknG, PknG$_{\Delta73}$, PknG$_{\Delta TPR}$, and PknG$_{\Delta73,\Delta TPR}$, respectively; lane 5 corresponds to TEV-treated PknG$_{\Delta137}$.

(B) Relative kinase activity of PknG variants against the physiological substrate GarA (left panel) and the GarA-derived 17-mer peptide SDEVT-VET$_{21}$TSVFRADFL (right panel). Measurements were performed at least twice; error bars represent the scattering of measured values. See also Figures S1, S3, and S5.

the protein contains a possibly unstructured N-terminal extension, a rubredoxin (Rdx)-like domain immediately adjacent to the catalytic core, and a C-terminal dimerization domain composed of tetratricopeptide repeats (TPR) (Scherr et al., 2007) (Figure 1A). In contrast to most other *Mtb* kinases (Duran et al., 2005; Prisic et al., 2010), the activation loop (and the entire catalytic domain) of PknG is not phosphorylated on serine or threonine residues. Instead, four autophosphorylation sites (T$_{23}$, T$_{32}$, T$_{63}$, and T$_{64}$) were identified within the N-terminal extension of the protein (O'Hare et al., 2008; Scherr et al., 2009; Prisic et al., 2010), where they provide pT-dependent anchoring sites for high-affinity interactions with the FHA domain of GarA (O'Hare et al., 2008). The Rdx domain contains the two characteristic CxxCG motifs in which the cysteine residues act as ligands of a divalent metal ion (Scherr et al., 2007), and is tightly associated with the small lobe of the kinase domain, giving rise to a deep substrate-binding cleft. This metal-binding ability and the absence of pT residues in the activation loop led some authors to suggest a possible redox-dependent regulation mechanism of PknG (Scherr et al., 2007; Tiwari et al., 2009; Gil et al., 2013; Chao et al., 2010), but no strong evidence is available in support of this hypothesis, and the underlying mechanisms involved are yet to be elucidated.

Here, we investigate the molecular mechanisms of PknG regulation by its N-terminal region, the C-terminal TPR domain, and the helix αC, which act jointly in determining the selective phosphorylation of the physiological substrate GarA, together with the molecular details for the inhibition of PknG activity by the Rdx domain. Furthermore, we have found that PknG interacts with an extended peptide sequence at the protein substrate-binding site, but that, contrary to the knowledge about ePKs (Huse and Kuriyan, 2002; Kornev and Taylor, 2010), the highly conserved salt bridge between a lysine side chain that coordinates the α and β phosphates of ATP and a glutamate residue from helix αC is not essential for the kinase activity. The protein kinase PknG from *Mtb* represents a detailed example regarding

molecular determinants of activity, interaction with the substrates, and mechanism of regulation within the family of STPKs, providing alternatives for the design of selective inhibitors with potential therapeutic applications.

RESULTS

PknG Domains Flanking the Catalytic Core Contribute to the Selective Phosphorylation of GarA

Full-length PknG from *Mtb* displays a unique multidomain organization, containing a possibly unstructured N-terminal extension (residues 1–98), a rubredoxin (Rbx)-like domain (residues 99–137), the kinase catalytic core (residues 138–405), and a C-terminal TPR domain (residues 406–750) (Figure 1A). To understand the roles of these flanking regions in modulating PknG kinase activity, we first measured the specific activity of PknG against two different substrates: (1) the small protein GarA, a physiological substrate of PknG containing an FHA domain with high affinity for phosphorylated threonine residues (O'Hare et al., 2008; England et al., 2009) (Figure 1B, left panel), and (2) a 17-mer peptide (SDEVTVET$_{21}$TSVFRADFL) derived from GarA (residues 14–30), centered around the phosphorylatable T$_{21}$ (O'Hare et al., 2008) and lacking the FHA domain, to serve as a reporter of the intrinsic activity of the PknG catalytic domain (Figure 1B, right panel). The kinase activity of full-length PknG varied linearly with the concentration of the 17-mer peptide up to 1 mM (Figure S1A), indicating a high K_M (>1 mM), and the slope (k_{cat}/K_M) providing a measure of the catalytic efficiency for this substrate ((1.9 ± 0.1) × 10^{-5} nmol/μM^2.min). By comparison, the phosphorylation of GarA by wild-type PknG was approximately 3-fold higher than for the 17-mer substrate even though a ca. 15-fold lower concentration of GarA was used (Figure 1B), indicating a 45-fold higher activity toward GarA due to the selective docking interaction granted by the FHA domain. Furthermore, when taking into account the K_M value of 2.1 μM previously reported for GarA phosphorylation

Table 1. Data Collection and Refinement Statistics

	PknG$_{\Delta73,\Delta TPR}$ + ADP (PDB code: 4Y0X)	PknG$_{\Delta73,\Delta TPR}$ + ATPγS (PDB code: 4Y12)
Data Collection		
Space group	C121	C121
Cell dimensions		
a, b, c (Å)	75.92, 37.25, 108.52	75.45, 37.18, 108.03
α, β, γ (°)	90.00, 97.74, 90.00	90.00, 97.93, 90.00
Resolution (Å)	37.61–1.74 (1.77–1.74)[a]	37.37–1.90 (2.00–1.90)
R_{sym} or R_{merge}	0.067 (0.794)	0.073 (0.774)
I/σI	18.4 (2.1)	12.8 (2.6)
Completeness (%)	99.9 (99.9)	98 (96.4)
Redundancy	5.0 (3.7)	7.4 (7.2)
Refinement		
Resolution (Å)	37.61–1.74 (1.77–1.74)	37.37–1.90 (2.00–1.90)
No. of reflections	31,242	23,246
R_{work}/R_{free}	19.72/20.87	19.56/21.16
No. of atoms		
Protein	2,429	2,379
Ligand/Mg(II)/Zn(II)	27(ADP)/2/1	31(ATPγS)/2/1
Water	229	171
B factors		
Protein	38.17	44.61
Ligand/ion	28.80	40.50
Water	44.88	47.78
Root-mean-square deviations		
Bond lengths (Å)	0.01	0.01
Bond angles (°)	1.02	1.05

[a]One protein crystal was used for structure determination in each case. Values in parentheses are for highest-resolution shell.

by PknG (Tiwari et al., 2009), the catalytic efficiency would be ~0.007 nmol/μM^2.min, ca. 400 times higher than that estimated for the 17-mer peptide substrate.

We then tested the selectivity of GarA phosphorylation by a PknG deletion mutant lacking the N-terminal autophosphorylated threonines (PknG$_{\Delta73}$) (O'Hare et al., 2008). This mutant presented a ca. 3-fold lower kinase activity against GarA compared with the wild-type enzyme, while phosphorylation of the 17-mer peptide was unaffected (Figure 1B). These results thus indicated that segment 1–73, including all PknG pT sites (O'Hare et al., 2008; Scherr et al., 2009), had no effect on the intrinsic activity of the kinase domain but was an important determinant of the kinase affinity and specificity toward its physiological substrate GarA. In line with this observation, PknG ability to phosphorylate the 17-mer peptide remained unaffected after addition of the isolated FHA domain of GarA (Figure S1B), indicating that the activity of PknG is not modulated by interactions of the N-terminal pT residue(s) with the FHA domain. It should be noted, however, that the moderate decrease in GarA phosphorylation due to PknG mutation Δ73 did not fully explain the much larger difference observed between the phosphorylation of GarA and the 17-mer peptide by full-length PknG,

suggesting that additional motifs may also contribute to GarA selectivity.

We therefore investigated the effect of the C-terminal TPR domain on PknG kinase activity. Interestingly, both phosphorylation of the 17-mer peptide and PknG autophosphorylation remained unchanged upon removal of the C-terminal region 406–750, encompassing a linker and the TPR domain (mutant PknG$_{\Delta TPR}$) (Figure 1B; Figure S1C). In contrast, the kinase activity of PknG$_{\Delta TPR}$ against GarA was lower than that of the wild-type enzyme. Together, these results indicated that the TPR domain did not affect the intrinsic kinase activity of PknG but that it contributed to the specific phosphorylation of the physiological substrate GarA.

The available crystal structure of PknG (PDB code 2PZI, comprising residues 74–750) suggested a role of the TPR domains in protein dimerization (Scherr et al., 2007). However, our analytical ultracentrifugation experiments (Figure S1D) indicated that full-length PknG behaves mainly as a monomer in solution (70%) even at high protein concentration (1 mg/ml). It is therefore unclear how the TPR domain could exert the observed modulation of PknG activity. To further investigate this issue, we studied a PknG deletion mutant lacking both the first 73 N-terminal residues and the whole C-terminal TPR domain (PknG$_{\Delta73,\Delta TPR}$). Of note, the kinase activity of PknG$_{\Delta73,\Delta TPR}$ toward the 17-mer peptide was similar to those of the wild-type enzyme, PknG$_{\Delta73}$, and PknG$_{\Delta TPR}$ (Figure 1B), whereas the phosphorylation of GarA was still further reduced in comparison with PknG$_{\Delta73}$ and PknG$_{\Delta TPR}$. This indicated a cumulative effect of both PknG regions on substrate selectivity, and suggested that the role of the TPR domain was independent of the pT-dependent interactions between the kinase and the FHA domain of GarA.

The Rdx Domain Downregulates the Kinase Activity of PknG

While all known bacterial PknG orthologs conserve the N-terminal region and the TPR domain, the Rdx domain is present only in PknGs within few suborders of the Actinomycetales, including the genera *Streptomyces* and *Mycobacterium* (Gil et al., 2013). Interestingly, while the deletion of the N-terminal extension and/or the TPR domain did not influence the intrinsic kinase activity of PknG against the 17-mer peptide, deletion of the Rdx domain (mutant PknG$_{\Delta137}$) promoted a significantly higher activity against the peptide substrate (Figure 1B), clearly evidencing a release of an inhibition of PknG activity mediated by the Rdx domain. Consistent with this, PknG$_{\Delta137}$ showed a ca. 2.5-fold higher activity than PknG$_{\Delta73}$ against GarA (Figure 1B).

To further investigate the molecular basis of this regulatory process, we determined the high-resolution crystal structures of PknG$_{\Delta73,\Delta TPR}$ in complex with either ADP or ATPγS, respectively, to 1.74- and 1.9-Å resolutions (Table 1). The final atomic models comprise PknG residues 83–404, including a fragment of the N-terminal segment, the Rdx domain, and the kinase catalytic core. In both structures, *mFo − DFc* sigma-A-weighted electron density maps clearly revealed the bound nucleotide and two Mg(II) cations at the active site, and one cation identified as a Zn(II) ion (Figure S2) coordinated by residues C$_{106}$, C$_{109}$, C$_{128}$, and C$_{131}$ in the Rdx domain. The close association between the Rdx domain and the N lobe of the catalytic core
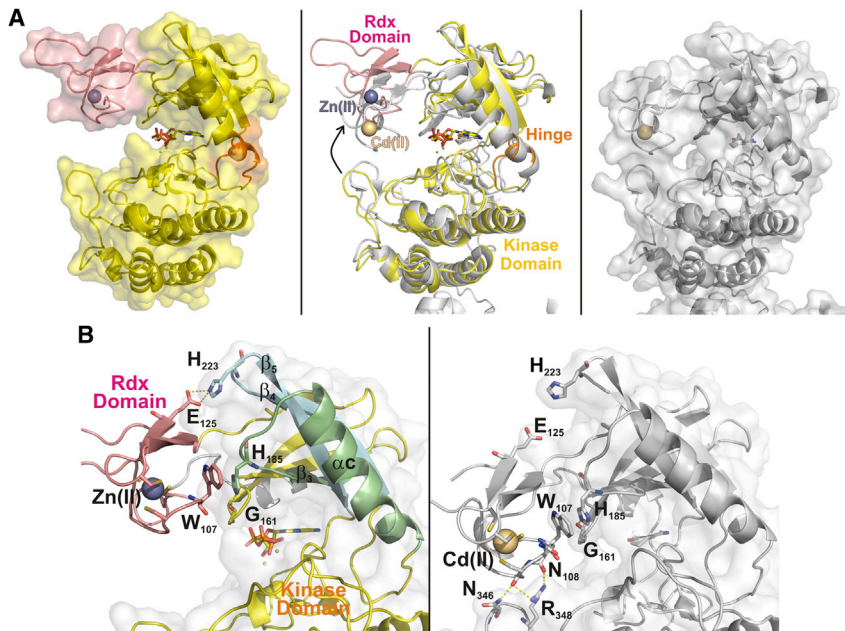
Figure 2. The Structure of PknG Suggests a Regulatory Role for the Rdx Domain by Opening/Closing the Active Site Entrance

(A) Left: overall structure of PknG$_{\Delta73\Delta TPR}$ (this work) with the Rdx domain (in pink) bound to the N-terminal lobe of the catalytic domain (in yellow). At the bottom of the open substrate-binding cleft, the ATPγS molecule is shown in stick representation. Center: comparison of the PknG$_{\Delta73\Delta TPR}$ structure (pink and yellow ribbons) with the structure of PknG bound to inhibitor AX20017 (PDB code 2PZI [Scherr et al., 2007], shown in gray), illustrating the movement of the Rdx domain. Right: molecular surface representation of 2PZI (in gray) with a closed substrate-binding cleft. In each case, the metal ion within the Rdx domain is shown as a sphere.

(B) Interactions between the Rdx domain and the kinase core of PknG. Left: the structure of PknG$_{\Delta73\Delta TPR}$ is shown in pink (Rdx domain), gray (N-terminal tail), green (β$_3$ strand, αC helix, and the loop connecting these motifs), cyan (β$_4$ and β$_5$ strands and the loop between them), and yellow (the remainder of the kinase N lobe). Right: structure 2PZI is shown in gray. In sticks, N, O, P, and S atoms are colored in blue, red, orange and dark yellow, respectively. Dashed lines represent atomic interactions. See also Figures S2 and S3.

(Figure 2A) generates a much deeper substrate-binding cleft compared with other Hank-type protein kinases, suggesting that the active site of PknG could selectively bind extended peptide substrates. When comparing the available PknG crystal structures (Figure 2A), important differences in the position of the Rdx domain become apparent. In particular, the Rdx domain can adopt different positions in front of the active site of PknG, interacting or not with the kinase C lobe, without affecting the overall kinase fold and topology. Differences between the open and closed states in the different crystal structures are not due to conformational changes in the Rdx domain itself, as a root-mean-square deviation (RMSD) of only 0.31 Å is found among 34 alpha carbons in segment 99–137. Instead, a hinge motion takes place involving residues 200–205, 210–213, and 236–237 in the kinase domain, leaving a more open active site in the structures of PknG$_{\Delta73,\Delta TPR}$, more permissive for substrate binding (Figure 2A).

In available PknG crystal structures, the interface area between the Rdx domain and the kinase catalytic core spans 300–350 Å^2, and there is a conserved hydrogen bond between the carbonyl oxygen of residue F$_{137}$ and the side chain of W$_{164}$ (sequence adjacent to the G-rich loop, in the β$_2$ strand). Furthermore, in the crystal structures of PknG$_{\Delta73,\Delta TPR}$, the Rdx domain contacts the kinase catalytic core through the ionic interaction between the conserved residues E$_{125}$ and H$_{223}$ (Figure 2B, left panel). Residue H$_{223}$ is located in the loop connecting the β$_4$ and β$_5$ strands in the kinase N lobe. This loop is particularly long in PknG orthologs containing an Rdx domain (Figure S3), suggesting that the entire region may be functionally relevant. Interestingly, in model 2PZI (Scherr et al., 2007), where the Rdx domain approaches the kinase C lobe and occludes the active site (Figure 2A, right panel), residues E$_{125}$ and H$_{223}$ are distant from each other and do not interact (Figure 2B, right panel), indicating that these residues may contribute to regulate the position

of the Rdx domain. Another contact with the kinase N lobe involves the stacking of the conserved W$_{107}$ (sequence adjacent to the metal ligand C$_{106}$) and H$_{185}$ (in the loop connecting the β$_3$ strand and the αC helix) (Figure 2B). These motifs harbor the catalytically relevant residues K$_{181}$ (K$_{72}$ in the ePK protein kinase A [PKA]) and E$_{198}$ (E$_{91}$ in PKA), respectively. Moreover, the N$_{\epsilon2}$ atom of H$_{185}$ hydrogen bonds the carbonyl oxygen of residue G$_{161}$ in the kinase G-rich loop (Figure 2B and next section), involved in positioning the ATP phosphates within the kinase active site. Thus, the PknG Rdx domain not only restricts the substrate entrance into the active site but also links catalytically relevant structural motifs within the kinase core. These observations therefore provide a structural framework to explain the regulation of the intrinsic kinase activity of PknG by the Rdx domain.

The ATP-binding Site of PknG

Most PknG residues within the ATP-binding site regions (as defined by Traxler and Furet, 1999) are functionally conserved and adopt conformations very similar to those observed for PKA and other ePKs. Thus, the adenine moiety is buried in a hydrophobic pocket and makes two direct hydrogen bond interactions with the enzyme: the amine N at position 6 binds the backbone carbonyl oxygen of E$_{233}$, and the N$_1$ atom binds the amide of V$_{235}$ (Figure 3, left panel). Furthermore, the N$_7$ atom establishes a water-mediated interaction with the amide of D$_{293}$. Additional indirect contacts take place between the nucleoside and the protein: two water molecules link the adenine N$_3$ atom, the ribose O$_2$* position, and the backbone carbonyl oxygens of V$_{235}$ and G$_{237}$, and another water molecule contacts both the ribose O$_2$* and O$_3$* atoms and the amide of S$_{239}$. In addition, van der Waals contacts are made with the side chains of I$_{157}$, A$_{158}$, I$_{165}$, V$_{179}$, V$_{211}$, M$_{232}$, Y$_{234}$, V$_{235}$, M$_{283}$, and I$_{292}$. Most PknG residues that bind the inhibitor AX20017 in 2PZI (Scherr

Figure 3. The ATP-binding Site of PknG

The protein is depicted as yellow ribbons. The ATPγS molecule and the protein residues interacting with it are shown in sticks with atoms colored by type (C, yellow; N, blue; 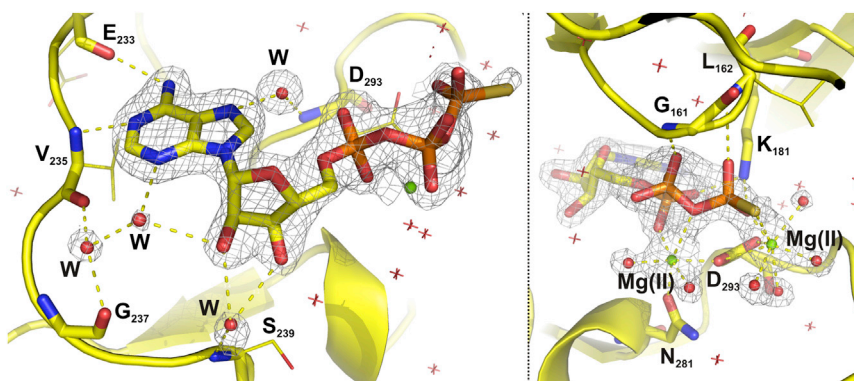O, red; P, orange; S, dark yellow). Water molecules are presented as red spheres or stars, and magnesium atoms are shown as green spheres. The *2mFo-DFc* electron density is contoured to 1.5 σ and represented as a blue mesh. Dashed lines represent atomic interactions.

et al., 2007) are also involved in nucleotide binding, except for I_{87}, A_{92}, and G_{236} that interact with the cyclopropyl ring of the molecule.

The G-rich motif (A_{158}HGGLGW) within the N lobe of the PknG kinase domain acts as a flap that covers the nucleotide (Figure 3, right panel). The main-chain amides of G_{161} and L_{162} hydrogen bonds the ATPγS β- and γ-phosphates, respectively, and the ATPγS α- and γ-phosphates both establish hydrogen bonds with the side chain of K_{181}. In addition, two magnesium ions are coordinated to the phosphates of the nucleotide. One Mg(II) displays a well-defined octahedral coordination sphere, including the α- and β-phosphates, the side chains of the conserved N_{281} and D_{293}, and two water molecules. The other Mg(II) is hexacoordinated and binds to the ATPγS γ-phosphate, the side chain of D_{293}, and four water molecules (an additional water molecule occupies the place of the γ-phosphate in the ADP-containing PknG structure). This coordination is reminiscent of that observed for human kinase Aurora A (PDB code 1MQ4; Nowakowski et al., 2002).

PknG Binds Extended Peptide Substrates

The absence of an arginine residue preceding the invariant catalytic aspartate in the catalytic loop (D_{276} in PknG) classifies PknG as a non-RD kinase (Johnson et al., 1996). Accordingly, and as predicted for ePKs (Huse and Kuriyan, 2002; Kornev and Taylor, 2010), the activation loop is stabilized in an open and extended conformation permissive for substrate binding in the absence of phosphorylation. Notably in the present crystal structures, the N-terminal extended segment comprising residues 83–89 (RAPDIDP) of the single PknG molecule in the asymmetric unit reaches into the active site of a crystallographic symmetry mate (Figure 4A, left panel). This interaction resembles the mode of binding of the peptide PKI within the active site of the mouse kinase PKA (PDB code 1ATP; Zheng et al., 1993) (Figure 4A, right panel), suggesting that in this crystal lattice PknG residues 83–89 act as a substrate mimic or pseudosubstrate with residue D_{86} occupying the place of the GarA nucleophilic threonine T_{21}. According to this model (Figure 4B), the side chain of E_{280} in the catalytic domain of PknG would interact with the main-chain amide of GarA residue V_{19}, the amide and the carbonyl oxygen of G_{308} would contact the carbonyl oxygen and the amide of GarA T_{22}, respectively, and the side chain of T_{309} would form a hydrogen bond with the carbonyl oxygen of GarA E_{20}. This pattern is reminiscent of the binding mode of pep-

tide substrates to ePKs, where residues equivalent to T_{309}, located within the p+1 loop, are known to be important for the interaction at the active site (Huse and Kuriyan, 2002; Kornev and Taylor, 2010). In fact, residue T_{309} has been previously shown to be critical for PknG kinase activity, and has been predicted to be involved in substrate binding (Tiwari et al., 2009).

The preceding findings suggested that, if the interaction between $PknG_{\Delta73,\Delta TPR}$ molecules takes place in solution as observed in the crystal structures, a phosphorylation event should occur given a threonine residue in position 86 of the PknG protein sequence. To test this hypothesis, we mutated the $PknG_{\Delta73,\Delta TPR}$ segment 74–90 (LGGGLVEIPRAPD$_{86}$IDPL) into GarA residues 9–25 (EKDQTSDEVTVE**T**TSVF) or EKDQTSDEVTVE**A**TSVF, and then investigated the autophosphorylation activity of these PknG variants (named respectively $PknG_{\Delta73,\Delta TPR,E}$**TTS** and $PknG_{\Delta73,\Delta TPR,E}$**ATS**). As determined by using radiolabeled [γ-^{32}P] ATP and mass spectrometry (Figure 4C; Figure S4A), $PknG_{\Delta73,\Delta TPR,E}$**TTS** incorporated one phosphate equivalent while $PknG_{\Delta73,\Delta TPR}$ and $PknG_{\Delta73,\Delta TPR,E}$**ATS** failed to show significant autophosphorylation activity, confirming our hypothesis. Taken together, our results provide a structural framework to explain the phosphorylation of extended peptide substrates by PknG.

The Role of Helix αC in PknG

The catalytic core of PknG adopts the typical two-lobed structure found in STPKs, with most functionally important and conserved residues in the active site exhibiting conformations compatible with a standard ePK active state (Scherr et al., 2007; Huse and Kuriyan, 2002; Kornev and Taylor, 2010). Nevertheless, an overlay with the structure of the prototypic active state of the cAMP-dependent PKA (PDB code 1ATP; Zheng et al., 1993) shows a clearly distinct position of PknG helix αC (Figure 5, left panel). This helix acts as mediator of multiple regulatory mechanisms in numerous ePKs, outstandingly via a conserved glutamate residue (E_{91} in PKA) that forms an invariable salt bridge with the ATP-binding lysine (K_{72} in PKA). In many cases the activation of ePKs is mediated by the formation of this salt bridge, which is usually considered a hallmark of a kinase active state (Huse and Kuriyan, 2002; Kornev and Taylor, 2010). Curiously, in all available PknG crystal structures the equivalent residues, E_{198} and K_{181}, are more than 15 Å apart. To investigate the role of PknG residue E_{198}, we generated and analyzed the mutant $PknG_{E198A}$. Surprisingly, this point mutant displayed the same activity as the wild-type enzyme when using
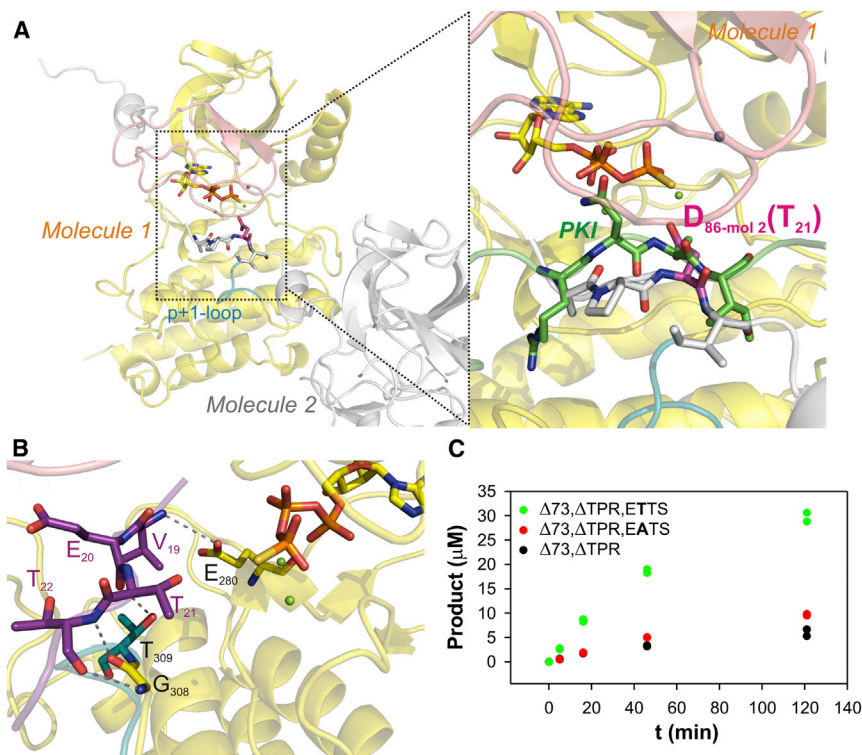
Figure 4. A Peptide Substrate Mimic within the Active Site of PknG

(A) A portion of the N-terminal segment (residues shown in sticks) of a PknG molecule in the crystal gets into the active site of a symmetry-related molecule. The right panel shows a zoom into PknG active site. The conformation of the PknG peptide substrate mimic resembles that of peptide PKI (in sticks with green C atoms) within the active site of mouse PKA (PDB code 1ATP). Residue D_{86} occupies the position of the phosphorylatable T_{21} in GarA.

(B) Model of peptide $TVET_{21}TSV$, corresponding to GarA residues 18–24, within the substrate-binding site of PknG. The protein is shown as yellow ribbons, except for the p+1 loop which is depicted in cyan, and the peptide as a violet ribbon. Residues $VET_{21}T$ in the peptide substrate, PknG residues E_{280}, G_{308}, and T_{309}, and the nucleotide are shown in sticks and atoms are colored by type: C, violet in the peptide substrate, cyan in PknG T_{309}, and yellow in PknG G_{308}, PknG E_{280}, and the nucleotide; N, blue; O, red; P, orange; S, dark yellow. Magnesium atoms are shown as green spheres. Dashed lines represent atomic interactions. The total potential energy of complex $PknG_{\Delta73,\Delta TPR}$/TVETTSV was −10,255 kcal/mol and the interaction energy was −54 kcal/mol, similar to the values found for complex $PknG_{\Delta73,\Delta TPR}$/RAPDIDP (−10,374 and −68 kcal/mol, respectively).

(C) Autophosphorylation activity of PknG mutants. Measurements were performed using 30 μM enzyme in all cases. See also Figure S4.

GarA or the 17-mer peptide as substrates (Figure 5, right panel). This result conclusively established that neither residue E_{198} nor, consequently, the K_{181}-E_{198} salt bridge is required for PknG activity. Therefore, these results indicated that the available PknG crystal structures are descriptive of a productive kinase conformation in the absence of the K_{181}-E_{198} salt bridge. This would imply that the active state of PknG would be more permissive than typical ePKs regarding the actual positioning of the helix αC.

DISCUSSION

The extended knowledge about ePKs structures and regulation mechanisms has enlightened the research on the evolutionarily related bacterial STPKs (Wehenkel et al., 2008). However, recent results uncover a broader view of bacterial phosphosignaling systems with particular unprecedented features (Sherman and Grundner, 2014). In this work we describe novel molecular mechanisms of regulation found for a bacterial STPK, the multi-domain kinase PknG from Mtb. We show that the N-terminal extended region and the TPR domain are both involved in the specific phosphorylation of the physiological substrate GarA. Our observations indicate a remarkable distinction between PknG and ePKs in that much of PknG machinery seems to be dedicated to assure substrate selectivity and not to tune the specific kinase activity. We suggest that this notion might represent a more extended regulatory strategy among bacterial STPKs.

The mechanisms leading to substrate selectivity described here for PknG resemble those mediated by docking interactions in ePKs, for example among MAP and AGC kinases (Biondi and

Nebreda, 2003). However, the most characterized feature in ePKs is by far the regulation of the intrinsic kinase activity and the concomitant phosphorylation of substrates that possess a rather simple recognition motif surrounding the phosphorylation site (Pinna and Ruzzene, 1996). Interestingly, we found that the peptide-binding site in PknG is very similar to that of ePKs. Among the interactions predicted between the peptide substrate and the active site of PknG (Figure 4B), one involves the side chain of residue T_{309}, located in the p+1 loop, which has been shown to be important for kinase activity in ePKs (Huse and Kuriyan, 2002; Kornev and Taylor, 2010) and PknG itself (Tiwari et al., 2009). In addition, all known PknG phosphorylation sites contain a non-polar amino acid two positions before the phosphorylatable threonine (Figure S4B). This residue (V_{19} in GarA) is predicted to be stabilized by van der Waals interactions within a small pocket mainly comprising hydrophobic groups. This may explain why PknG phosphorylates specifically GarA T_{21} and not T_{22}, since the latter would require a charged side chain, GarA E_{20}, binding to the hydrophobic pocket. Also, our results do not indicate strong interactions between the peptide substrate and PknG, in agreement with the high K_M exhibited by the 17-mer substrate (Figure S1A) and with the FHA domain of GarA being the major determinant of substrate selectivity (Figure 1B).

PknG has been proposed to be active in the cytoplasm of mammalian cells (Walburger et al., 2004). It is thus tempting to speculate that the PknG machinery leading to the selective phosphorylation of the Mtb substrate GarA may also be relevant for the phosphorylation of mammalian substrates. As revealed by structure 2PZI (Scherr et al., 2007), the N-terminal segment
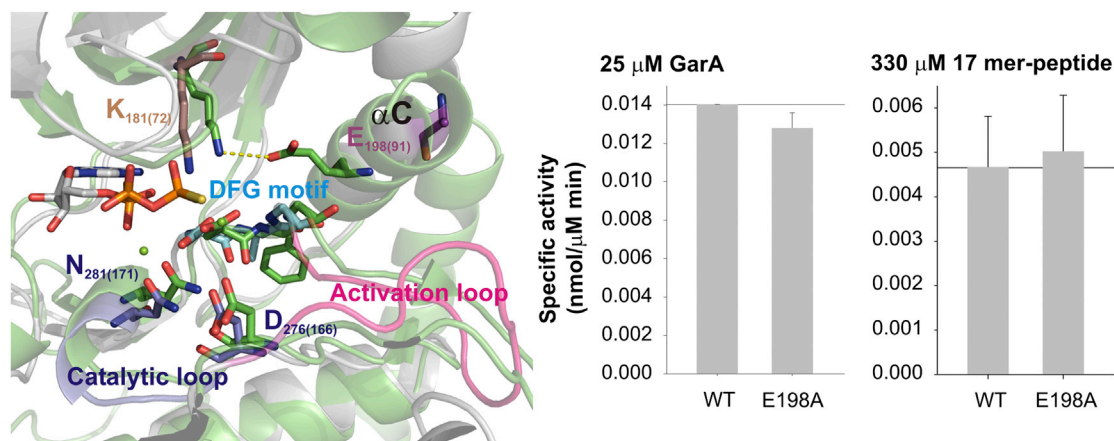
Figure 5. Functionally Important and Conserved Residues within the Active Site of PknG

Left: comparison of PknG$_{\Delta 73,\Delta TPR}$ active site with that of PKA (PDB code 1ATP) in its active "on" state. PknG$_{\Delta 73,\Delta TPR}$ is shown as gray ribbons, except for the highlighted motifs, and sticks with C atoms colored as the label of the corresponding motif or residue. PKA is presented as green ribbons and sticks with green C atoms. The ATPγS molecule bound to the PknG$_{\Delta 73,\Delta TPR}$ active site is shown in sticks with gray C atoms, and magnesium ions are depicted as green spheres. The other atoms in sticks are colored by type: N, blue; O, red; P, orange; S, dark yellow. The dashed line represents the salt bridge between PKA residues K$_{72}$ and E$_{91}$. PKA residue numbers are shown in parentheses. Right: relative kinase activity of wild-type PknG and PknG$_{E198A}$ against GarA and the GarA-derived 17-mer peptide SDEVTVET$_{21}$TSVFRADFL. Measurements were performed at least twice; error bars represent the scattering of measured values. See also Figure S3.

of PknG contains a P-rich region (largely conserved in the PknG family, Figure S3) reminiscent of that present in the C-terminal tail of kinase PKA (Figure S5A). This region has been proposed to be functionally relevant for the AGC kinases (Kannan et al., 2007), and might also be important for PknG function. Based on these data, we speculate that STPKs substrates could be better identified as interacting partners rather than by a consensus sequence surrounding the phosphorylation site.

All available PknG crystal structures show residue E$_{198}$ pointing out of the active site due to a tilted and outward conformation of the helix αC in the kinase N lobe (Figure 5, left panel). According to the criteria normally taken into account to define an STPK conformation as productive or not, the observed PknG conformations would then be associated to kinase "off" states given the absence of a salt bridge between E$_{198}$ and the catalytic K$_{181}$. However, PknG$_{E198A}$ retained the wild-type PknG activity (Figure 5, right panel), clearly showing that neither residue E$_{198}$ nor, consequently, the K$_{181}$-E$_{198}$ salt bridge is essential for kinase activity. In the context of the overall mechanism of PknG regulation, largely dedicated to control substrate specificity, the K$_{181}$-E$_{198}$ salt bridge may have evolved as a mechanism enabling a higher degree of modulation of the activity toward a physiological substrate rather than as a structural requirement for kinase activity. It is tempting to speculate that such an additional level of modulation might be mediated by tyrosine phosphorylation, recently reported for several *Mtb* proteins (including PknG) (Kusebauch et al., 2014).

PknG is a non-RD kinase (Johnson et al., 1996) and, consistently, the kinase activity is not regulated by S/T phosphorylation in the activation loop. Instead, all PknG pT sites are located within the N-terminal extension of the protein, which is crucial for recruitment of the physiological substrate GarA (O'Hare et al., 2008). Here we show that PknG activity is not modulated by S/T phosphorylation, as its N-terminal extension does not affect the intrinsic activity of the kinase domain (Figure 1B).

Also, the evidence presented indicates that the main role of the FHA domain of Gar is to facilitate the docking of the substrate through FHA-pT interactions, but not to exert an allosteric effect on the PknG catalytic domain (Figure S1B).

Our kinetic data indicate that the TPR domain of PknG has no effect on the intrinsic activity of the kinase core (Figure 1B; Figure S1C). However, mutant PknG$_{\Delta TPR}$ displayed a reduced ability to phosphorylate GarA (Figure 1B), suggesting that the TPR domain contributes to stabilizing the complex PknG-GarA. We propose that this could occur via the stabilization of a β hairpin in the N-terminal segment of PknG (Figure S5B), which is crucial for GarA recruitment. Consistently, all residues involved in the series of TPR-mediated interactions observed in structure 2PZI (Scherr et al., 2007) are strictly conserved among mycobacterial PknGs (Figure S3), suggesting that they are functionally relevant. In previous work, the TPR domain has been found to mediate the interaction between two PknG monomers in the crystalline state (Scherr et al., 2007). However, data obtained by analytical ultracentrifugation indicated that this interaction was weak in solution, even at high protein concentrations (Figure S1D). Thus, even though TPR domains are known to be involved in protein-protein interactions in different systems (D'Andrea and Regan, 2003), it is unclear if and how this contact may be important for PknG function, and we have failed to identify elements in support of a PknG regulatory mechanism mediated by TPR-mediated kinase dimerization.

The Rdx domain is the only region identified to directly regulate the intrinsic activity of PknG (Figure 1B). Interestingly this domain is present in PknG orthologs within the genera *Streptomyces* and *Mycobacterium*, including several pathogenic species (Gil et al., 2013), but is missing in *Corynebacterium* PknGs (Figure S3). In all available PknG crystal structures the Rdx domain interacts indirectly with the G-rich loop of the catalytic domain through residue H$_{185}$, which is located in the loop connecting the β_3 strand and the αC helix (Figure 2B). This contact would also restrict the motion of

the αC helix, which was systematically found in a conformation that maintains the E_{198} out of the active site. However, the Rdx domain can adopt different positions relative to the PknG active site, mediated by the interplay of distinct interdomain interactions. Thus, in the present crystal structures residues E_{125} (in the Rdx domain) and H_{223} (within the kinase N lobe) interact with each other, while no contact is observed between the Rdx domain and the kinase C lobe (Figure 2B, left panel). This open conformation allows the active site to accept a nucleotide and a peptide substrate mimic (Figure 4A). On the other hand, replacement of the nucleotide by the small molecule inhibitor AX20017 in structure 2PZI (Scherr et al., 2007) brings the Rdx domain into contact with the kinase C lobe, whereas residues E_{125} and H_{223} are now distant from each other (Figure 2B, right panel). In this closed conformation, the ATP phosphates would clash with the G-rich loop, and no place is left for a peptide substrate. In line with these observations, our data clearly show that the Rdx domain downregulates the kinase activity of PknG (Figure 1B). This domain has been proposed to modulate PknG activity by sensing the redox status of the surroundings (Scherr et al., 2007; Tiwari et al., 2009). However, some controversy remains regarding the sign of this regulation. While the deletion of the Rdx domain derepresses PknG activity (as shown herein), kinase activity is reduced (Tiwari et al., 2009) or completely abolished (Scherr et al., 2007) after perturbation (and probably disruption) of the metal center by substitution of the protein ligands. It is worth noting that the structural integrity of those PknG mutants was not opportunely assessed. In addition, a MBP-PknG$_{151-750}$ deletion mutant was analyzed and only a marginal activity was detected (Tiwari et al., 2009). The differences found between this variant and PknG$_{\Delta 137}$ can be attributed to the constructions themselves and/or to the alternative set of experimental conditions used in kinase assays. In any case, our results clearly show that the Rdx domain of PknG is not essential for the kinase activity, in contrast to previous conclusions (Scherr et al., 2007).

In conclusion, based on the presented data we propose a model for the modulation of PknG kinase activity on its physiological substrate GarA. According to this model, the N-terminal extension and the TPR domain of PknG work jointly in optimizing the affinity for the *Mtb* substrate GarA, whereas the Rdx domain (which may respond to the redox state of the environment) regulates the intrinsic PknG kinase activity by restricting the accessibility of the active site.

Very few representative members of the bacterial STPKs group have been crystallized and investigated in detail so far. We envisage that the regulation of the kinase activity based on substrate selectivity might be more widely distributed among bacterial STPKs. Such mechanisms may, in turn, enable the development of antibiotics with novel modes of action targeting different STPKs in *Mtb* and other pathogenic bacteria.

EXPERIMENTAL PROCEDURES

Cloning and Mutagenesis
Plasmids pET28a-PknG, pET28a-PknG$_{\Delta 73}$, and pET28a-PknG$_{\Delta 73,\Delta TPR}$ were already available (O'Hare et al., 2008; Gil et al., 2013). The construction of plasmids pET28a-PknG$_{\Delta TPR}$, pET28a-PknG$_{\Delta 137}$, and pET28a-PknG$_{E198A}$ is detailed in the Supplemental Experimental Procedures. Plasmids pET28a-PknG$_{\Delta 73,\Delta TPR,E}$**TT**S and pET28a-PknG$_{\Delta 73,\Delta TPR,E}$**A**TS were constructed by GenScript using pET28a-PknG$_{\Delta 73,\Delta TPR}$ as template.

Protein Production and Purification
All PknG variants were overproduced in *Escherichia coli* BL21(DE3) cells and then purified following the same protocol. Cells were harvested by centrifugation and sonicated. After clarification by centrifugation, the supernatant was loaded onto a HisTrap HP column (GE Healthcare) and the His-tagged protein was purified applying a linear imidazole gradient (20–500 mM). The PknG-containing fractions were pooled and the protein was further purified by size-exclusion chromatography on a Superdex 200 column (GE Healthcare). PknG$_{\Delta 137}$ was also prepared adding an additional purification step. Prior to the size-exclusion chromatography, the H$_6$ tag in PknG$_{\Delta 137}$ was removed by incubation with His$_6$-tagged TEV protease (van den Berg et al., 2006), followed by separation on an Ni-NTA agarose column (Qiagen). Following the purification step by size-exclusion chromatography, fractions corresponding to the PknG peak were pooled and concentrated, flash-frozen in liquid nitrogen, and stored at $-80°$C. See also Supplemental Experimental Procedures.

The substrate GarA and its FHA domain alone (GarA$_{\Delta 43}$) were prepared as previously described (England et al., 2009).

Protein Kinase Activity Assay
Kinase activity assays were performed in 96-well plates. Each activity measurement was performed in a final volume of 20 μl, containing 50 mM Tris-HCl (pH 7.4), 0.1% v/v 2-mercaptoethanol, 10 mM MnCl$_2$, 100 μM [γ-^{32}P] ATP (5–50 cpm/pmol), and 330 μM 17-mer peptide or 25 μM Gar as substrate. The enzyme concentration in the assays was 0.3–1.5 μM and 0.15–0.45 μM when using the 17-mer peptide or GarA, respectively, as substrates. The kinase reactions were started by the addition of 4 μl [γ-^{32}P]ATP-Mn^{2+} and were performed at room temperature. The reactions were stopped by the addition of phosphoric acid, and 4 μl of each reaction was spotted on P81 phosphocellulose papers (Whatman) using the epMotion 5070 (Eppendorf) workstation. The papers were washed in 0.01% phosphoric acid, dried, then measured and analyzed using PhosphorImager (FLA-9000 Starion, Fujifilm). Each reaction was performed in duplicate (<5% variation). In all cases, specific activity values were derived from reactions performed using three different enzyme concentrations within the indicated ranges (<10% variation), verifying a linear dependence of activity with PknG concentration. Each assay was performed at least twice. The proportion of 17-mer peptide or GarA consumed in the reactions was lower than 10% and 30%, respectively. GarA consumption was verified to be linear in time up to 50% of its initial concentration. Under the experimental conditions used to test phosphorylation of the 17-mer peptide or GarA, PknG autophosphorylation represented less than 5% of the total signal. The measured signal was at least five times higher than the background.

H$_6$-tagged PknG$_{\Delta 137}$ and untagged PknG$_{\Delta 137}$ showed the same kinetic behavior.

Crystallization and Data Collection
Crystallization screenings were carried out using the sitting-drop vapor diffusion method and a Mosquito nanoliter-dispensing crystallization robot (TTP Labtech). Following optimization, crystals of PknG$_{\Delta 73,\Delta TPR}$ + ADP and PknG$_{\Delta 73,\Delta TPR}$ + ATPγS grew after 10–15 days from a solution of 20–30 mg/ml protein supplemented with 2–4 mM ADP or ATPγS, respectively, by mixing 1 μl of protein solution and 1 μl of mother liquor (100 mM HEPES, 35% w/v PEG 4000, 200 mM CaCl$_2$, pH 7.5), in a hanging-drop setup with 1 ml of mother liquor in the reservoir, at 18°C. Single crystals reaching a size of ~300 μm were cryoprotected in mother liquor containing 25% glycerol and flash-frozen in liquid nitrogen. X-ray diffraction data were collected at the synchrotron beamlines Proxima 1 (Synchrotron Soleil, Saint-Aubin, France) and ID23-1 (ESRF, Grenoble, France) at 100 K. Wavelengths used were 0.973 Å and 1.28189 Å for PknG$_{\Delta 73,\Delta TPR}$ + ADP and PknG$_{\Delta 73,\Delta TPR}$ + ATPγS crystals, respectively. The diffraction data were processed using XDS (Kabsch, 2010) and scaled with Aimless or Scala from the CCP4 program suite (Winn et al., 2011).

Structure Determination and Refinement
The crystal structure of PknG$_{\Delta 73,\Delta TPR}$ + ADP was solved by molecular replacement using the program Phaser (McCoy et al., 2007) and the atomic coordinates of PknG residues 74–405 from PDB 2PZI as search probe. The rubredoxin domain and some fractions of the kinase N lobe were removed during the firsts rounds of refinement, and were subsequently rebuilt through iterative cycles of manual model building with Coot (Emsley et al., 2010) and refinement

with BUSTER (Bricogne et al., 2011). The structure of PknG$_{\Delta73,\Delta TPR}$ + ATPγS was solved using the refined coordinates of protein atoms in PknG$_{\Delta73,\Delta TPR}$ + ADP. Ligand molecules were manually placed in $mFo - DFc$ sigma-A-weighted electron density maps employing Coot, and the resulting models were refined as described above. The final models were validated through the Molprobity server (http://molprobity.biochem.duke.edu) (Chen et al., 2010). The two crystal structures are very similar to each other, with an RMSD of 0.13 Å for 321 alpha carbons. In each case, the final model contained 98% of residues within favored regions of Ramachandran plot, with no outliers. Figures were generated and rendered with PyMOL 1.5.0.2 (Schrödinger LLC).

ACCESSION NUMBERS

The PDB accession numbers for the PknG$_{\Delta73,\Delta TPR}$ crystal structures reported in this paper are 4Y0X and 4Y12.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures and five figures and can be found with this article online at http://dx.doi.org/10.1016/j.str.2015.04.001.

AUTHOR CONTRIBUTIONS

M.N.L. designed experiments, prepared proteins, carried out crystallographic studies and structural analysis, performed kinase activity assays, analyzed data, and wrote the paper. M.G. prepared constructs pET28a-PknG$_{\Delta TPR}$ and pET28a-PknG$_{\Delta137}$ and carried out mass spectrometry analyses. G.A.L. performed molecular modeling. N.B. prepared plasmid pET28a-PknG$_{E198A}$ and performed analytical ultracentrifugation experiments. R.D. designed and performed mass spectrometry studies. R.M.B. designed kinase activity assays, analyzed data, and wrote the paper. P.M.A. designed research, analyzed data, and wrote the paper.

ACKNOWLEDGMENTS

REFERENCES

Av-Gay, Y., and Everett, M. (2000). The eukaryotic-like Ser/Thr protein kinases of *Mycobacterium tuberculosis*. Trends Microbiol. *8*, 238–244.

Barthe, P., Roumestand, C., Canova, M.J., Kremer, L., Hurard, C., Molle, V., and Cohen-Gonsaud, M. (2009). Dynamic and structural characterization of a bacterial FHA protein reveals a new autoinhibition mechanism. Structure *17*, 568–578.

Biondi, R.M., and Nebreda, A.R. (2003). Signalling specificity of Ser/Thr protein kinases through docking-site-mediated interactions. Biochem. J. *372*, 1–13.

Bricogne, G., Blanc, E., Brandl, M., Flensburg, C., Keller, P., Paciorek, W., Roversi, P., Sharff, A., Smart, O.S., Vonrhein, C., and Womack, T.O. (2011). BUSTER Version 2.11.4. (Global Phasing Ltd).

Chao, J., Wong, D., Zheng, X., Poirier, V., Bach, H., Hmama, Z., and Av-Gay, Y. (2010). Protein kinase and phosphatase signaling in *Mycobacterium*

tuberculosis physiology and pathogenesis. Biochim. Biophys. Acta *1804*, 620–627.

Chen, V.B., Arendall, W.B., III, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. *66*, 12–21.

Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S.V., Eiglmeier, K., Gas, S., Barry, C.E., 3rd., et al. (1998). Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature *393*, 537–544.

Cowley, S., Ko, M., Pick, N., Chow, R., Downing, K.J., Gordhan, B.G., Betts, J.C., Mizrahi, V., Smith, D.A., Stokes, R.W., and Av-Gay, Y. (2004). The *Mycobacterium tuberculosis* protein serine/threonine kinase PknG is linked to cellular glutamate/glutamine levels and is important for growth in vivo. Mol. Microbiol. *52*, 1691–1702.

D'Andrea, L.D., and Regan, L. (2003). TPR proteins: the versatile helix. Trends Biochem. Sci. *28*, 655–662.

Duran, R., Villarino, A., Bellinzoni, M., Wehenkel, A., Fernandez, P., Boitel, B., Cole, S.T., Alzari, P.M., and Cervenansky, C. (2005). Conserved autophosphorylation pattern in activation loops and juxtamembrane regions of *Mycobacterium tuberculosis* Ser/Thr protein kinases. Biochem. Biophys. Res. Commun. *333*, 858–867.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. *66*, 486–501.

England, P., Wehenkel, A., Martins, S., Hoos, S., Andre-Leroux, G., Villarino, A., and Alzari, P.M. (2009). The FHA-containing protein GarA acts as a phosphorylation-dependent molecular switch in mycobacterial signaling. FEBS Lett. *583*, 301–307.

Gil, M., Grana, M., Schopfer, F.J., Wagner, T., Denicola, A., Freeman, B.A., Alzari, P.M., Batthyany, C., and Duran, R. (2013). Inhibition of *Mycobacterium tuberculosis* PknG by non-catalytic rubredoxin domain specific modification: reaction of an electrophilic nitro-fatty acid with the Fe-S center. Free Radic. Biol. Med. *65*, 150–161.

Huse, M., and Kuriyan, J. (2002). The conformational plasticity of protein kinases. Cell *109*, 275–282.

Johnson, L.N., Noble, M.E., and Owen, D.J. (1996). Active and inactive protein kinases: structural basis for regulation. Cell *85*, 149–158.

Kabsch, W. (2010). XDS. Acta Crystallogr. D Biol. Crystallogr. *66*, 125–132.

Kannan, N., Haste, N., Taylor, S.S., and Neuwald, A.F. (2007). The hallmark of AGC kinase functional divergence is its C-terminal tail, a cis-acting regulatory module. Proc. Natl. Acad. Sci. USA *104*, 1272–1277.

Kornev, A.P., and Taylor, S.S. (2010). Defining the conserved internal architecture of a protein kinase. Biochim. Biophys. Acta *1804*, 440–444.

Kusebauch, U., Ortega, C., Ollodart, A., Rogers, R.S., Sherman, D.R., Moritz, R.L., and Grundner, C. (2014). *Mycobacterium tuberculosis* supports protein tyrosine phosphorylation. Proc. Natl. Acad. Sci. USA *111*, 9265–9270.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J. Appl. Crystallogr. *40*, 658–674.

Niebisch, A., Kabus, A., Schultz, C., Weil, B., and Bott, M. (2006). Corynebacterial protein kinase G controls 2-oxoglutarate dehydrogenase activity via the phosphorylation status of the OdhI protein. J. Biol. Chem. *281*, 12300–12307.

Nott, T.J., Kelly, G., Stach, L., Li, J., Westcott, S., Patel, D., Hunt, D.M., Howell, S., Buxton, R.S., O'Hare, H.M., and Smerdon, S.J. (2009). An intramolecular switch regulates phosphoindependent FHA domain interactions in *Mycobacterium tuberculosis*. Sci. Signal. *2*, ra12.

Nowakowski, J., Cronin, C.N., McRee, D.E., Knuth, M.W., Nelson, C.G., Pavletich, N.P., Rogers, J., Sang, B.C., Scheibe, D.N., Swanson, R.V., et al. (2002). Structures of the cancer-related Aurora-A, FAK, and EphA2 protein kinases from nanovolume crystallography. Structure *10*, 1659–1667.

O'Hare, H.M., Durán, R., Cerveñansky, C., Bellinzoni, M., Wehenkel, A.M., Pritsch, O., Obal, G., Baumgartner, J., Vialaret, J., Johnsson, K., and Alzari,

P.M. (2008). Regulation of glutamate metabolism by protein kinases in mycobacteria. Mol. Microbiol. *70*, 1408–1423.

Pinna, L.A., and Ruzzene, M. (1996). How do protein kinases recognize their substrates? Biochim. Biophys. Acta *1314*, 191–225.

Prisic, S., Dankwa, S., Schwartz, D., Chou, M.F., Locasale, J.W., Kang, C.M., Bemis, G., Church, G.M., Steen, H., and Husson, R.N. (2010). Extensive phosphorylation with overlapping specificity by *Mycobacterium tuberculosis* serine/threonine protein kinases. Proc. Natl. Acad. Sci. USA *107*, 7521–7526.

Sassetti, C.M., Boyd, D.H., and Rubin, E.J. (2003). Genes required for mycobacterial growth defined by high density mutagenesis. Mol. Microbiol. *48*, 77–84.

Scherr, N., Honnappa, S., Kunz, G., Mueller, P., Jayachandran, R., Winkler, F., Pieters, J., and Steinmetz, M.O. (2007). Structural basis for the specific inhibition of protein kinase G, a virulence factor of *Mycobacterium tuberculosis*. Proc. Natl. Acad. Sci. USA *104*, 12151–12156.

Scherr, N., Muller, P., Perisa, D., Combaluzier, B., Jeno, P., and Pieters, J. (2009). Survival of pathogenic mycobacteria in macrophages is mediated through autophosphorylation of protein kinase G. J. Bacteriol. *191*, 4546–4554.

Sherman, D.R., and Grundner, C. (2014). Agents of change—concepts in *Mycobacterium tuberculosis* Ser/Thr/Tyr phosphosignalling. Mol. Microbiol. *94*, 231–241.

Tiwari, D., Singh, R.K., Goswami, K., Verma, S.K., Prakash, B., and Nandicoori, V.K. (2009). Key residues in *Mycobacterium tuberculosis* protein kinase G play a role in regulating kinase activity and survival in the host. J. Biol. Chem. *284*, 27467–27479.

Traxler, P., and Furet, P. (1999). Strategies toward the design of novel and selective protein tyrosine kinase inhibitors. Pharmacol. Ther. *82*, 195–206.

van den Berg, S., Lofdahl, P.A., Hard, T., and Berglund, H. (2006). Improved solubility of TEV protease by directed evolution. J. Biotechnol. *121*, 291–298.

Ventura, M., Rieck, B., Boldrin, F., Degiacomi, G., Bellinzoni, M., Barilone, N., Alzaidi, F., Alzari, P.M., Manganelli, R., and O'Hare, H.M. (2013). GarA is an essential regulator of metabolism in *Mycobacterium tuberculosis*. Mol. Microbiol. *90*, 356–366.

Walburger, A., Koul, A., Ferrari, G., Nguyen, L., Prescianotto-Baschong, C., Huygen, K., Klebl, B., Thompson, C., Bacher, G., and Pieters, J. (2004). Protein kinase G from pathogenic mycobacteria promotes survival within macrophages. Science *304*, 1800–1804.

Wehenkel, A., Bellinzoni, M., Graña, M., Duran, R., Villarino, A., Fernandez, P., Andre-Leroux, G., England, P., Takiff, H., Cerveñansky, C., et al. (2008). Mycobacterial Ser/Thr protein kinases and phosphatases: physiological roles and therapeutic potential. Biochim. Biophys. Acta *1784*, 193–202.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A., et al. (2011). Overview of the CCP4 suite and current developments. Acta Crystallogr. D Biol. Crystallogr. *67*, 235–242.

Zheng, J., Trafny, E.A., Knighton, D.R., Xuong, N.H., Taylor, S.S., Ten Eyck, L.F., and Sowadski, J.M. (1993). 2.2 A refined crystal structure of the catalytic subunit of cAMP-dependent protein kinase complexed with MnATP and a peptide inhibitor. Acta Crystallogr. D Biol. Crystallogr. *49*, 362–365.

Supplemental Information

Molecular Basis of the Activity and the Regulation

of the Eukaryotic-like S/T Protein Kinase

PknG from *Mycobacterium tuberculosis*

María-Natalia Lisa, Magdalena Gil, Gwénaëlle André-Leroux, Nathalie Barilone, Rosario Durán, Ricardo M. Biondi, and Pedro M. Alzari

SUPPLEMENTAL DATA

Figure S1, related to Figure 1. *PknG domains flanking the catalytic core contribute to the selective phosphorylation of GarA.* (A) A high K_M and a low k_{cat}/K_M for the phosphorylation of the GarA derived 17-mer peptide substrate (SDEVTVET$_{21}$TSVFRADFL) by wild type PknG. The kinase activity of full-length PknG (0.8 μM) varied linearly with the concentration of the 17-mer peptide up to 1 mM ($K_M > 1$ mM) and the slope of the curve (k_{cat}/K_M) provides a measure of the catalytic efficiency for this substrate ((1.9 ± 0.1) 10^{-5} nmol/μM^2.min). (B) The FHA domain of GarA does not affect the phosphorylation of the 17-mer peptide substrate by wild type PknG. Substrate concentration was 1 mM. (C) The TPR domain of PknG has no effect on autophosphorylation. Protein concentration was 30 μM for the wild type enzyme and the deletion mutant. (D) Full-length PknG behaves mainly as a monomer (70%) at 1 mg/ml. Analytical ultracentrifugation, result of analysis by sedimentation velocity.
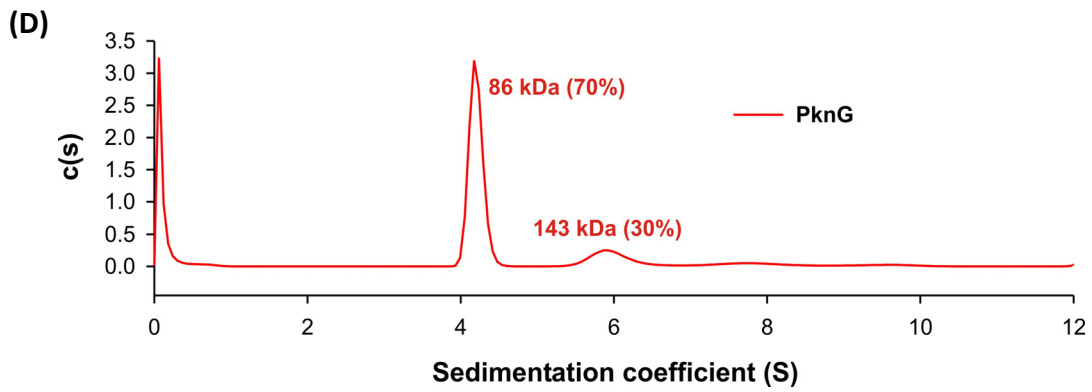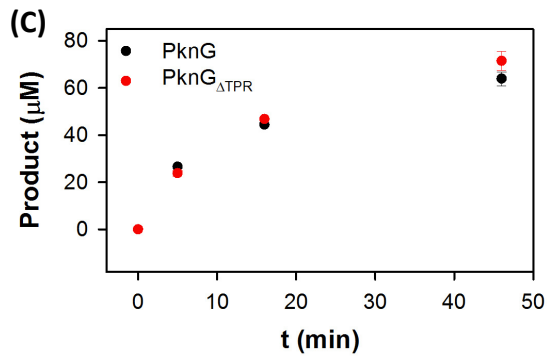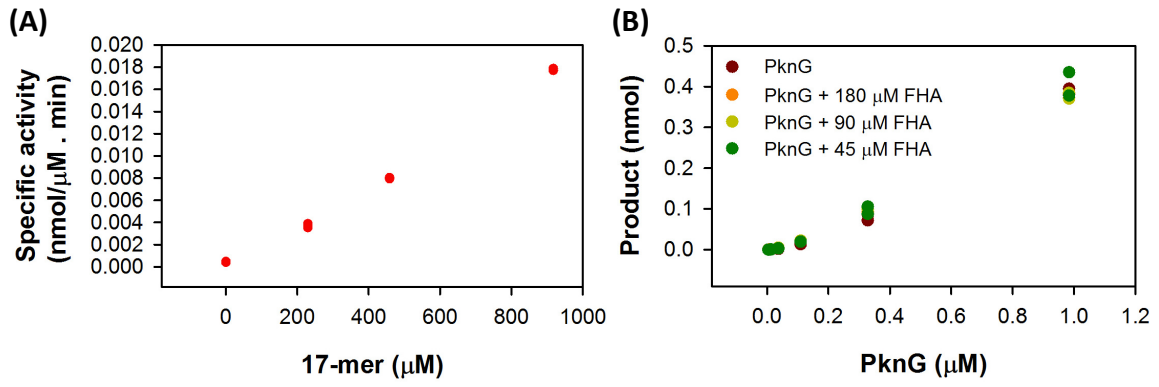
Figure S2, related to Figure 2. *The Rdx domain of PknG.* (A) A fluorescence emission spectrum obtained from a PknG$_{\Delta73,\Delta TPR}$ crystal pointed out the presence of significant amounts of zinc. Indeed, the analysis of X-ray anomalous diffraction data just above and bellow the Zn K edge confirmed the presence of a Zn(II) metal ligand in the Rdx domain of PknG. It is shown in pink ribbons and sticks colored by atom type (C, pink; N, blue; O, red; S, dark yellow). The Zn(II) metal ligand is shown as a grey sphere. The *2mFo–DFc* electron density is contoured to 2 σ and represented as a blue mesh. A double difference anomalous map calculated with diffraction data above (9672 eV) and below (9640 eV) the Zn K-edge is shown in cyan and contoured at 4 σ. Dashed lines represent atomic interactions. PDB code 4Y12. (B) Rdx domains normally bind iron as metal cofactor (Sieker et al., 1994). In effect, an UV-vis electronic spectrum of PknG$_{\Delta73,\Delta TPR}$ in solution revealed an absorption band centered at 490 nm, consistent with a Cys-Fe(III) ligand to metal charge transfer band (Bonomi et al., 1998). On the other hand, the low molar extinction coefficient observed (~ 500 M^{-1} cm^{-1}) indicated that only a small fraction of the protein sample (5-10%) contained iron as ligand. All PknG variants containing the Rdx domain presented similar results. It has been shown that in some rubredoxin containing proteins iron and zinc exchange as an artifact of protein overproduction in *E. coli* cells, without substantially affecting the coordination geometry in the metal binding site (Bonomi et al., 1998).

(A)

(B)
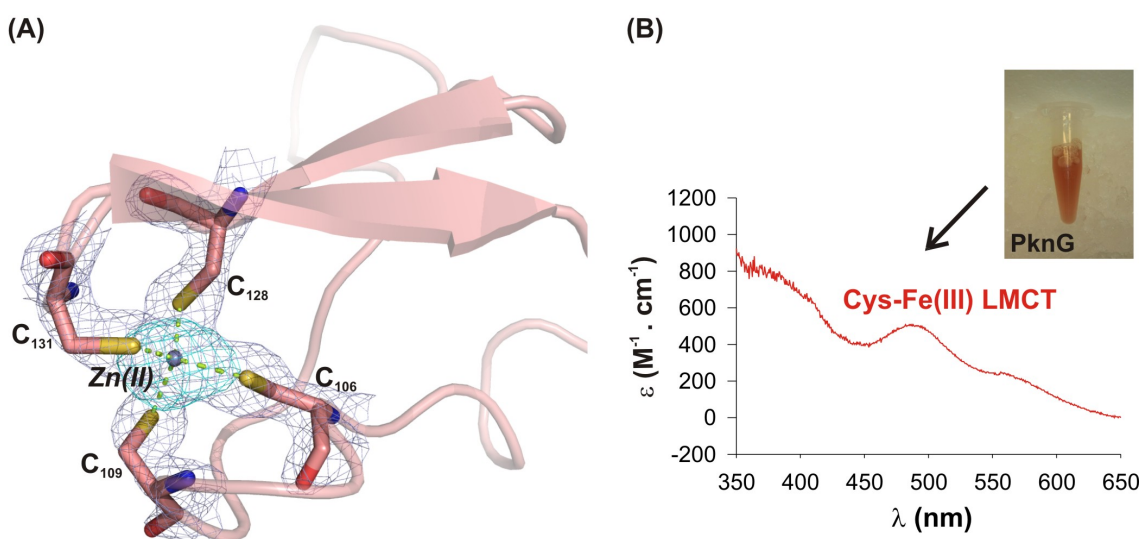
Figure S3, related to Figures 1, 2 and 5 and section Discussion. *Multiple sequence alignment of PknG homologs.* Residues of PknG from *M. tuberculosis* involved in interactions (as described in the main text) between the Rdx domain and the kinase core are signaled in red. Similarly, residues involved in interactions between the TPR domain, the kinase core, and the N-terminal segment are signaled in green. Some motifs and secondary structure elements are indicated.
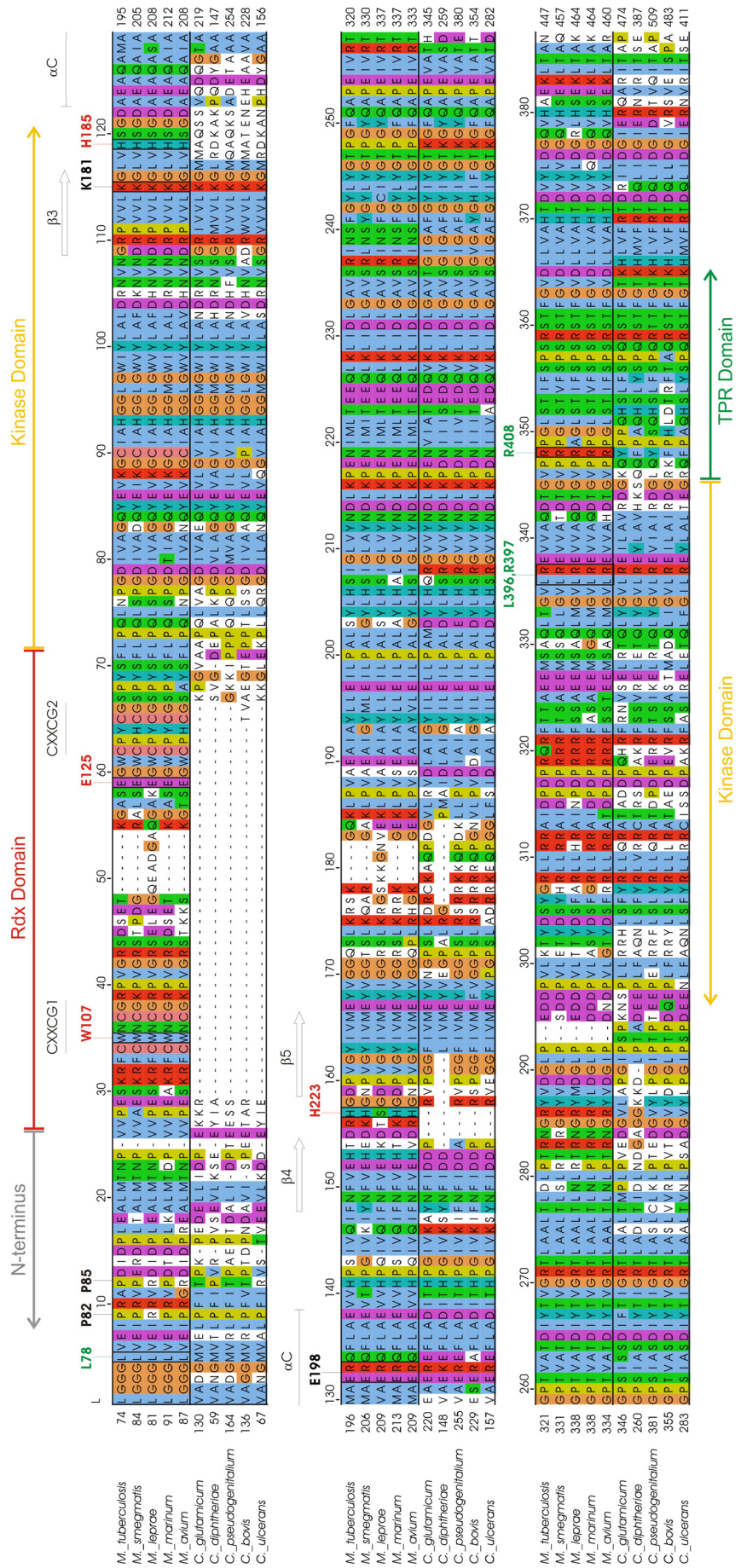
Figure S4, related to Figure 4. *The peptide binding site of PknG.* (A) Identification of PknG$_{\Delta73,\Delta TPR,E\mathbf{TT}S}$ phosphorylation site. MS/MS spectrum of doubly charged ion at *m/z* 596.70 (MH$^+$ 1192.4) from an endoproteinase GluC digestion of PknG$_{\Delta73,\Delta TPR,E\mathbf{TT}S}$ previously incubated with ATP and MnCl$_2$. The N-terminal (*b*, red-labeled) and C-terminal (*y*, blue-labeled) fragment ions that allowed the sequence 52-61 assignment and that includes a phosphorylated threonine residue (T56) are shown. Ion fragments showing neutral losses (-H$_2$O and -H$_3$PO$_4$) are also indicated. *Inset*, amino acid sequence of peptide VTVEpTTSVFE, indicating major *b* and *y* ions detected by full scan MS/MS. (B) GarA residue V$_{19}$ is predicted to be stabilized by van der Waals interactions within a small hydrophobic pocket in PknG active site. Left: model of peptide TVE**T**TSV, corresponding to GarA residues 18-24, within the active site of PknG. The peptide substrate is shown as a ribbon and sticks (residues VE**T**T) with violet C carbons. The surface of the protein is depicted with C carbons colored in yellow in the kinase core and in pink in the Rdx domain. The nucleotide within the active site is shown in sticks with yellow carbons and magnesium ions are depicted as small green spheres. N, O, P and S atoms are colored in blue, red, orange and dark yellow, respectively. Right: peptides known to be phosphorylated by PknG. They contain a non polar aminoacid two positions before the phosphorylatable threonine (in red).

(A)
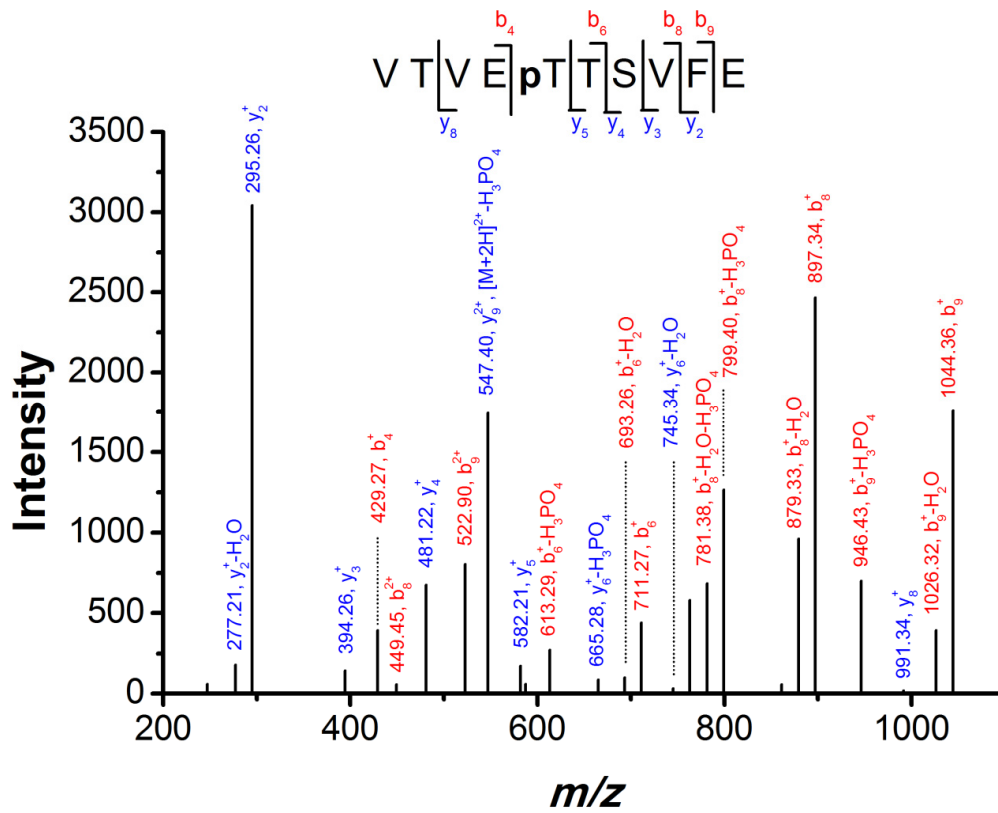
(B)
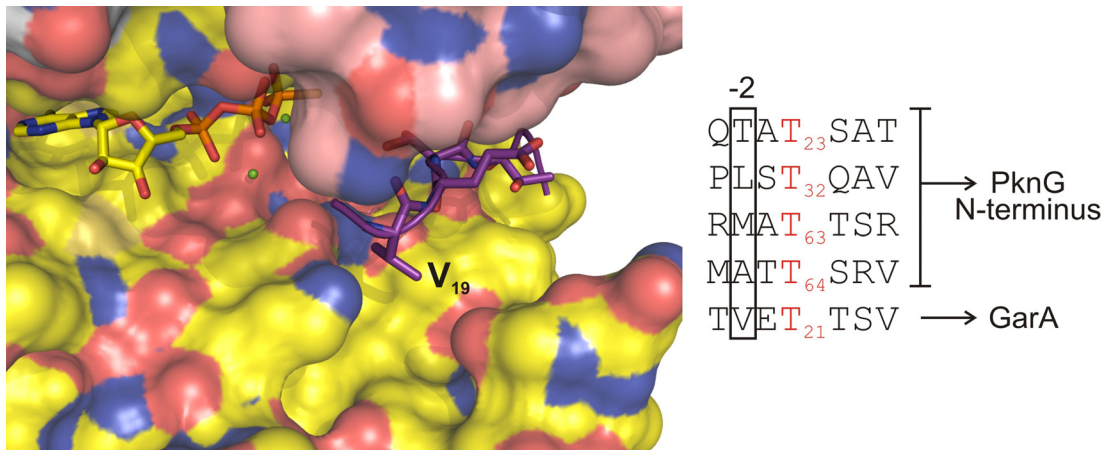
Figure S5, related to Figure 1 and section Discussion. *The N-terminal segment and the TPR domain of PknG.* (A) PknG (left panel, yellow surface) contains a P-rich region in its N-terminal tail which closely resembles the P-rich region in the C-terminal extension on protein kinasa A (right panel, pink surface). The proline residues involved in each case are shown in sticks. (B) Structure 2PZI shows a series of interactions between the C-terminal portion of the PknG N-terminal segment (gray ribbons and C atoms), the C-terminus of the catalytic core (yellow ribbons and C atoms) and the linker connecting the kinase C-lobe and the TPR domain (green ribbons and C atoms), whose conformation is determined by interaction with the concave surface of the TPR domain. These interactions contribute to stabilize a hairpin encompassing residues 75-78 (GGGL) by leading L_{78} to establish van der Waals contacts with L_{396} (kinase C-lobe) and the aliphatic portions of R_{397} (kinase C-lobe) and R_{408} (linker connecting the kinase C-lobe and the TPR domain). N and O atoms are colored in blue and red, respectively.

(A) 2PZI   1ATP

P85

P82

P316

P313

(B) Rdx Domain

Kinase Domain

N-terminal segment

TPR Domain

G75

G76

G77

L396

R397

L78

R408
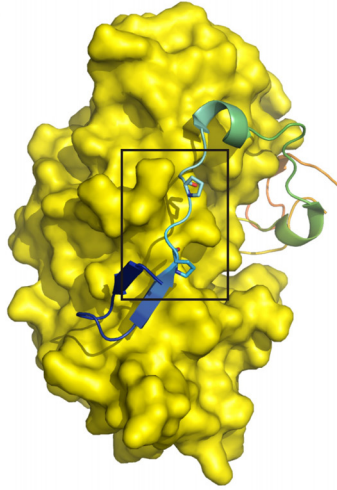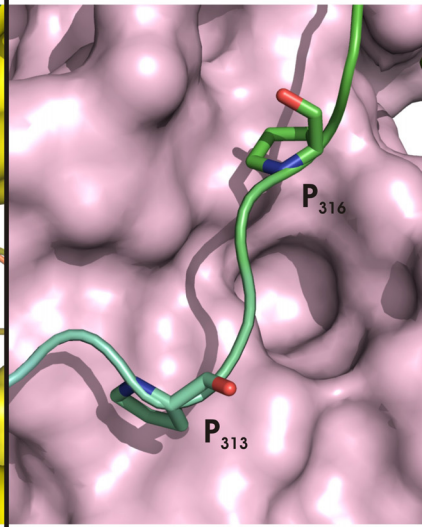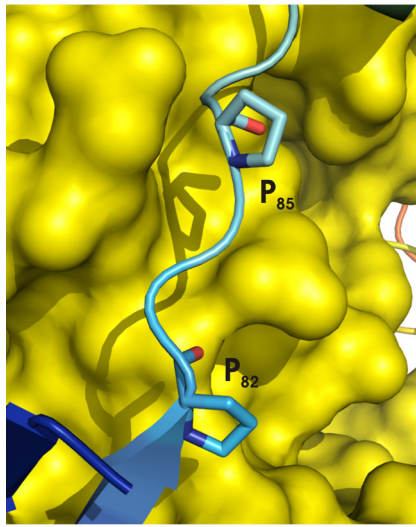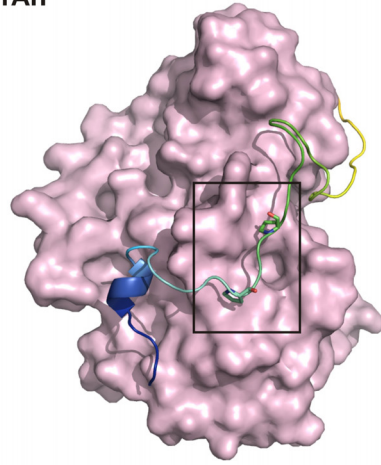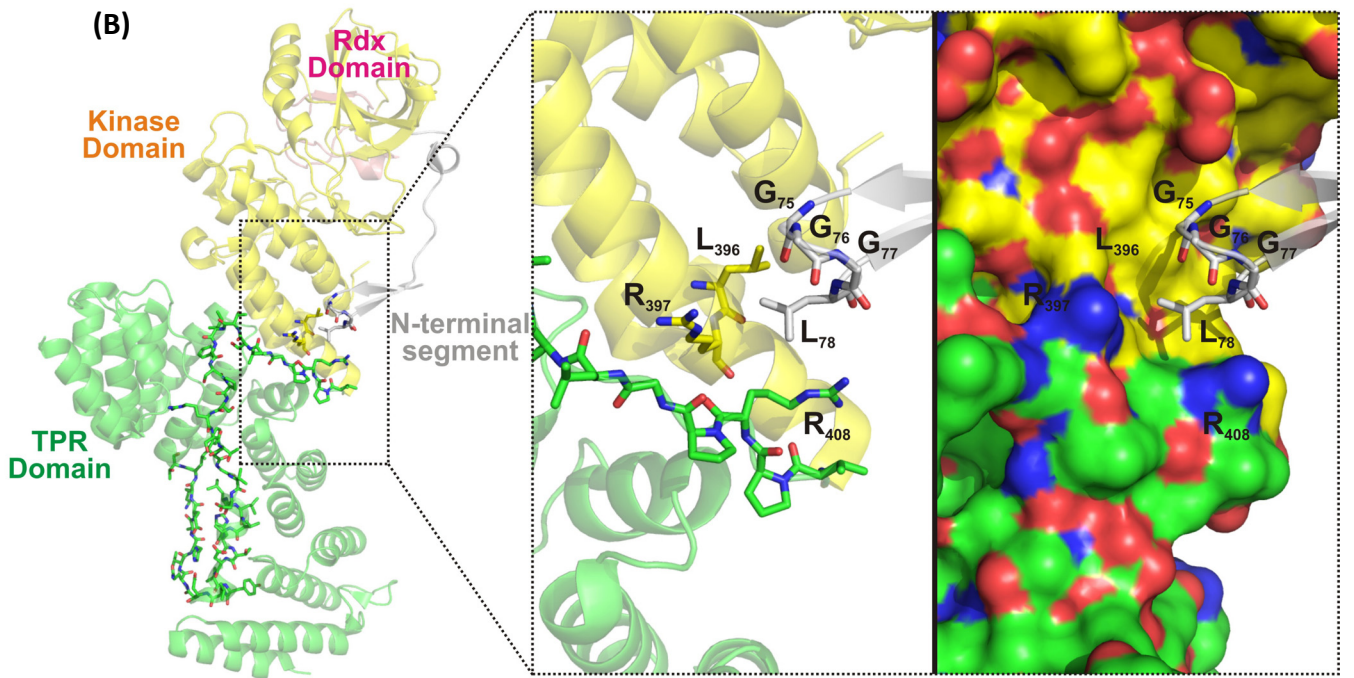
SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Chemicals

The synthetic 17-mer peptide SDEVTVET$_{21}$TSVFRADFL, corresponding to residues 14-30 of the protein GarA, was produced with a purity >98% by Thermo Fisher Scientific.

Cloning and mutagenesis

Plasmid pET28a-PknG$_{\Delta TPR}$ was generated by site directed mutagenesis of pET28a-PknG, using a pair of complementary primers to introduce a stop codon at position 406 of PknG sequence. Plasmid pET28a-PknG$_{\Delta 137}$, for the expression of the mutant with a hexahistidine tag, was constructed by PCR amplification of *pknG* region 138-750 from vector pET28a-PknG, digestion and ligation into the EcoRI site of plasmid pET28a (Novagen). Plasmid pET28a-PknG$_{E198A}$ was generated by site directed mutagenesis of pET28a-PknG, employing a pair of complementary primers to encode the desired mutation. The oligonucleotides employed are:

ΔTPR-F: GTCGCCCAGGACACCGGGTAACCGCGGCCAGGGCTATC

ΔTPR-R: GATAGCCCTGGCCGCGGTTACCCGGTGTCCTGGGCGAC

Δ137-F: ATTATCATATGGAGAATCTTTATTTTCAAGGTCTGCCGCAGCTAAATCCCG

Δ137-R: ATTAGGAATTCTTAGAACGTGCTGGT

E$_{198}$A-F: CAGGCAATGGCGATGGCCGCTCGCCAGTTCCTGGCCGAGGTGG

E$_{198}$A-R: CCACCTCGGCCAGGAACTGGCGAGCGGCCATCGCCATTGCCTG

Protein production and purification

Wild type PknG and PknG$_{E198A}$ were expressed for 18 h at 30ºC without IPTG. PknG$_{\Delta 73}$ was expressed after 18 h induction at 14ºC with 100 μM IPTG. PknG$_{\Delta 137}$, PknG$_{\Delta TPR}$, PknG$_{\Delta 73,\Delta TPR}$, PknG$_{\Delta 73,\Delta TPR,E TTS}$ and PknG$_{\Delta 73,\Delta TPR,E ATS}$ were produced similarly but employing 1mM IPTG. All the proteins were then purified following the same protocol. *E. coli* cells were harvested by centrifugation, re-suspended in lysis buffer (25 mM Hepes, 500 mM NaCl, 20% glycerol, 20 mM imidazole, pH 8) supplemented with Complete protease inhibitor cocktail (Roche) and sonicated. After clarification by centrifugation, the supernatant was loaded onto a 5 ml HisTrap HP column (GE

Healthcare) equilibrated with lysis buffer and the His-tagged protein was then purified applying a linear imidazole gradient (20–500 mM). The PknG-containing fractions, as verified by SDS-PAGE, were pooled and the protein was further purified by size-exclusion chromatography on a Superdex 200 column (GE Healthcare) equilibrated in either 50 mM Tris-HCl, 250 mM NaCl, 5% glycerol, pH 8 (wild type PknG and PknG$_{E198A}$) or 50 mM Tris-HCl, 500 mM NaCl, 5% glycerol, pH 8 (truncated PknG variants). PknG$_{\Delta137}$ was also prepared adding an additional purification step. Prior to the size-exclusion chromatography, the H$_6$ tag in PknG$_{\Delta137}$ was removed by incubating for 18 hours at 18ºC in the presence of His$_6$-tagged TEV protease at a 1/30 ratio (w/w) in buffer 50 mM Tris-HCl, 500 mM NaCl, 5% glycerol, 1 mM DTT, pH 8, followed by separation on a Ni-NTA agarose column (Qiagen). This step was omitted for the other PknG variants in order to avoid the contact of the protein with DTT (required for TEV activity), which might perturb the metal binding site in the Rdx domain.

Following the purification step by size-exclusion chromatography, fractions corresponding to the PknG peak, as confirmed by SDS-PAGE, were pooled and concentrated up to 30 mg/ml, flash-frozen in liquid nitrogen and stored at -80ºC.

Proteins were quantified by using the molar absorption coefficient predicted from the aminoacid sequence by the ProtParam tool (http://web.expasy.org/protparam/).

Electronic absorption spectroscopy

Electronic absorption spectra were recorded by employing a Beckman Coulter Du 800 spectrophotometer operating at room temperature.

Analytical ultracentrifugation

Sedimentation velocity experiments were carried out at 20 °C in an XL-I analytical ultracentrifuge (Beckman Coulter). Samples were spun using an An60Ti rotor and 12-mm double sector epoxy centerpieces. The partial specific volume of PknG (0.732 ml·g^{-1}) was estimated from their amino acid sequences using the software Sednterp. The same software was used to estimate the buffer viscosity (η = 1.032 centipoises) and density (ρ = 1.009 g·ml^{-1}). PknG (400 µl at 1 mg/ml) was spun at 42,000 rpm, and

absorbance profiles were recorded every five minutes. Sedimentation coefficient distributions, c(s), were determined using the software Sedfit 14.1 (Schuck, 2000).

Mass spectrometry analysis

The autophosphorylation activity of PknG$_{\Delta73,\Delta TPR}$, PknG$_{\Delta73,\Delta TPR,E TTS}$ and PknG$_{\Delta73,\Delta TPR,E ATS}$ was assessed by mass spectrometry after incubation of the different PknG constructs in the presence of 2 mM MnCl$_2$ and 500 µM ATP for 30 min at 37ºC. Protein samples were digested with different proteolytic enzymes (sequencing grade trypsin and endoproteinase GluC from Promega and Roche, respectively) by overnight incubation at 37ºC. Peptides were separated in a reversed-phase column (PepMap RSLC, C18, 75 µm x 500 mm, Thermo) using a linear gradient of B (from 5 to 55%), in 70 min, at 250 nl/min (solvent A: 0.1 % formic acid in H$_2$0; solvent B: 0.1 % formic acid in acetonitrile).

On line mass spectrometry analysis of peptides was performed using a linear ion trap mass spectrometer (LTQ Velos, Thermo) in data dependent acquisition mode (full scan followed by MS/MS of the top 10 peaks in each segment, using a dynamic exclusion list). Raw MS/MS spectra were interpreted with the Proteome Discoverer software package (v.1.3.0.339, Thermo) using Sequest as search engine. The following parameters were used for searching: *Escherichia coli* (strain K12) reference proteome database (Uniprot_2014-12, 4.305 sequences) with the sequences of PknG constructs incorporated; peptide tolerance: 1.5 Da; MS/MS tolerance: 0.8 Da; methionine oxidation and Ser/Thr phosphorylation as allowed variable modifications. Phosphopeptide validation and phosporylation site identification were performed using the PhosphoRS algorithm.

Molecular modeling

In the crystal structures of PknG$_{\Delta73,\Delta TPR}$, the segment encompassing residues 83-89 (RAPDIDP) of the single PknG molecule within the asymmetric unit gets into the active site of a crystallographic symmetry mate. From this structure, the complex composed by one molecule of PknG plus the fragment RAPDIDP from the symmetry mate was retrieved in order to evaluate their binding. Also, the fragment RAPDIDP provided the starting coordinates for its *in silico* mutation into TVETTSV, corresponding to GarA residues 18-24 and containing the residue T$_{21}$ phosphorylatable by PknG. Binding was

evaluated for the two peptides after energy optimization of complexes PknG$_{\Delta 73,\Delta TPR}$/RAPDIDP and PknG$_{\Delta 73,\Delta TPR}$/TVETTSV by using CHARMM implemented in Discovery Studio 2.5© (Brooks et al., 2009) with the following protocol: 10,000 iterations using Conjugate Gradient algorithm, constraints on the protein backbone, ATPγS, water molecules and cations, no constraint on the peptide and 0.01 RMS gradient as stop criterion. The total potential energy and the binding energy between the protein and the peptide were calculated for each complex.

SUPPLEMENTAL REFERENCE LIST

Bonomi,F., Iametti,S., Kurtz,J., Richie,K.A., and Ragg,E. (1998). Direct metal ion substitution at the [M(SCys)4]2− site of rubredoxin. JBIC *3*, 595-605.

Brooks,B.R. et al (2009). CHARMM: the biomolecular simulation program. J. Comput. Chem. *30*, 1545-1614.

Schuck,P. (2000). Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. Biophys. J. *78*, 1606-1619.

Sieker,L.C., Stenkamp,R.E., and LeGall,J. (1994). Rubredoxin in crystalline state. Methods Enzymol. *243*, 203-216.

Modulación de la actividad quinasa por los dominios adyacentes al dominio quinasa

La actividad de las quinasas de proteína puede estar regulada mediante diversos mecanismos entre los que se incluyen: subunidades y dominios reguladores, mecanismos alostéricos, segundos mensajeros, dominios y proteínas inhibitorias, localización subcelular mediada por dominios de localización y proteínas de anclaje, acilación de ácidos grasos, isoprenilación y fosforilación [141, 142]. Asimismo, la fosforilación de un segmento particular del dominio catalítico denominado bucle de activación, es una forma muy importante de regulación de la actividad enzimática de un número significativo de STPK de tipo eucariota [143].

Aunque en muchas quinasas la regulación de la actividad por dominios proteicos específicos está bien demostrada, el rol de los dominios N- y C- terminales de PknG, y su posible participación en la regulación de la actividad enzimática, no ha sido establecido. En este sentido surge la necesidad de ahondar en el papel de los dominios no catalíticos presentes en PknG de *M. tuberculosis*. Para ello se generaron formas truncas de la quinasa: PknG$_{\Delta 137}$ (PknG$_{\Delta Rbx}$), PknG$_{\Delta TPR}$ y PknG$_{\Delta 73,\Delta TPR}$ (Fig. 1A en artículo 2) y se evalúo el efecto de la remoción de los dominios sobre la actividad específica de PknG frente a dos sustratos: 1. GarA, que constituye el sustrato fisiológico de la quinasa y contiene un dominio FHA que reconoce fosfotreoninas con alta afinidad; 2. un péptido sintético derivado de GarA que corresponde a la secuencia 14-30 (SDEVTVET$_{21}$TSVFRADFL) y que permite evaluar la actividad intrínseca de la quinasa ya que no posee el dominio FHA involucrado en el reclutamiento de GarA (Fig. 1B en artículo 2).

La remoción del dominio TPR no tuvo consecuencias sobre la fosforilación del péptido sintético ni tampoco sobre la autofosforilación de PknG (Fig. S1C en artículo 2), no obstante provoca la disminución de la actividad frente a GarA. Tomados en conjunto, estos resultados indican que el dominio TPR no afecta la actividad intrínseca de PknG pero contribuye a la fosforilación específica de GarA. La estructura tridimensional de PknG (Fig. 9) sugiere que el dominio TPR estaría implicado en la dimerización de la quinasa empero nuestros resultados de ultracentrifugación analítica muestran que PknG se comporta mayoritariamente como un monómero en solución (Fig. S1D en artículo 2). Por lo tanto, no está claro cómo el dominio TPR modula la actividad catalítica de PknG.

Para profundizar en el rol del dominio TPR, la construcción PknG$_{\Delta 73,\Delta TPR}$ fue analizada. Los datos muestran que la actividad quinasa frente al péptido sintético es similar a la de la proteína nativa, mientras que frente a GarA la actividad es menor comparada con la enzima completa (Fig. 1B en artículo 2). A partir de estos datos puede sugerirse que la especificidad de sustrato proviene del efecto acumulado de las dos regiones de PknG (N-terminal y dominio TPR) y que el papel del dominio

TPR es independiente de las interacciones mediadas por el dominio FHA de GarA y las fosfotreoninas del extremo N-terminal de PknG.

Un hallazgo interesante de este trabajo fue que mientras la remoción de la extensión N-terminal y/o del dominio TPR no modifica la actividad intrínseca de PknG frente al péptido sintético, la deleción del dominio Rbx provoca un aumento significativo de la actividad frente a este sustrato. Este resultado indica que la actividad quinasa de PknG estaría inhibida y la eliminación del dominio Rbx provocaría la supresión de esa inhibición. El mismo tipo de comportamiento se observó cuando se utilizó GarA como sustrato; en este caso la actividad aumentó 2.5 veces respecto a la proteína nativa (Fig. 1B en artículo 2).

Para analizar las bases moleculares de la regulación de la actividad quinasa por el dominio Rbx, se determinó la estructura de $PknG_{\Delta73,\Delta TPR}$ en complejo con ATPγS. El modelo generado abarca desde el residuo 83 al 404 y en ese tramo se incluye un fragmento del segmento N-terminal, el dominio Rbx y el dominio quinasa (Fig. 2A en artículo 2). La estrecha asociación entre el dominio Rbx y el lóbulo aminoterminal del dominio catalítico da origen a un surco de unión al sustrato más profundo que el de otras quinasas tipo Hanks [45]. Lo anterior sugiere que el sitio activo de PknG podría unir selectivamente sustratos peptídicos extendidos. La comparación de la estructura obtenida en este trabajo con la estructura previamente reportada de PknG reveló importantes diferencias en la posición del dominio Rbx. En particular, este dominio puede adoptar distintas posiciones respecto al sitio activo de PknG, pudiendo interactuar o no con el lóbulo carboxiloterminal del dominio quinasa. El plegamiento y la topología del dominio catalítico no se ven afectados por la posición del dominio Rbx y las diferencias observadas entre las estructuras no se deben a cambios conformacionales del dominio Rbx. Por otra parte, el análisis detallado de los contactos en la estructura cristalina resuelta en este trabajo mostró que el dominio Rbx vincula motivos estructurales catalíticamente relevantes en el seno del dominio quinasa.

PknG de *M. tuberculosis* se clasifica como una quinasa no RD debido a la ausencia de un residuo de arginina precediendo al residuo de aspartato invariable en el sitio catalítico (D_{276}) [65]. Esto da lugar a un bucle de activación estabilizado en una conformación abierta y extendida que permite la unión de sustratos en ausencia de fosforilación. En el cristal de $PknG_{\Delta73,\Delta TPR}$ se observó que el segmento N-terminal de una molécula de PknG, que comprende del residuo 83 al 89 ($RAPD_{86}IDP$), ocupaba el sitio activo de la otra molécula presente en la unidad asimétrica (Fig. 4A en artículo 2). Esta interacción sugiere que en el entramado del cristal los residuos 83-89 se comportan como un sustrato mimético o pseudosustrato donde el D_{86} ocupa el lugar de la T_{21} de GarA.

Si la interacción entre las moléculas de PknG$_{\Delta73,\Delta TPR}$ tuviese lugar en solución de la misma manera en la que ocurre en el cristal, la introducción de un residuo de treonina en la posición 86 debería dar lugar a un evento de fosforilación. Para evaluar esta hipótesis se mutó el segmento 74-90 de PknG$_{\Delta73,\Delta TPR}$ (LGGGLVEIPRAPD$_{86}$IDPL) por los residuos 9-25 de la secuencia de GarA (EKDQTSDEVTVE**T**TSVF) o la misma secuencia donde el residuo fosforilable fue sustituido por A (EKDQTSDEVTVE**A**TSVF). Se evalúo la autofosforilación de estas variantes de PknG y se observó que solamente PknG$_{\Delta73,\Delta TPR,ETTS}$ presentaba actividad autocatalítica (Fig. 4C y S4A en artículo 2). Estos datos otorgan el marco estructural que explica la fosforilación de sustratos peptídicos extendidos por PknG.

En base a los datos aquí presentados se propone un modelo para explicar la modulación de la actividad quinasa de PknG respecto de su sustrato fisiológico: GarA. La extensión N-terminal y el dominio TPR actúan conjuntamente para optimizar la afinidad por el sustrato mientras que el dominio Rbx regularía la actividad intrínseca de PknG restringiendo la accesibilidad al sitio activo. De hecho, la estructura resuelta en este trabajo reveló una estrecha asociación entre el dominio Rbx y el lóbulo N-terminal del dominio quinasa que origina un profundo surco de unión al sustrato, mientras que en la estructura de PknG reportada previamente [63] se determinó que el dominio Rbx estaba próximo al lóbulo C-terminal del dominio quinasa restringiendo la entrada al sitio activo. La comparación de estas estructuras sugiere un movimiento tipo bisagra que modula el acceso al sitio activo [144].

Este trabajo se realizó en estrecha colaboración con la Dra. Natalia Lisa de la Unité de Microbiologie Structurale del Institut Pasteur (París, Francia).

Capítulo 3

Unravelling *Mycobacterium tuberculosis* PknG interactome

Magdalena Gil[1], Analía Lima[1], Estefanía Urdániz[2], Bernardina Rivera[1], Federico Carrión[3], Annemarie Wehenkel[4], María-Natalia Lisa[4,7], Otto Pritsch[3], Ana Denicola[5], Carlos Batthyány[1,6], Pedro M. Alzari[4], Mariana Piuri[2], Rosario Durán[1]

[1] Unidad de Bioquímica y Proteómica Analíticas, Institut Pasteur de Montevideo e Instituto de Investigaciones Biológicas Clemente Estable, Uruguay

[2] Departamento de Química Biológica, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Argentina

[3] Unidad de Biofísica de Proteínas, Institut Pasteur de Montevideo, Uruguay

[4] Unité de Microbiologie Structurale & CNRS URA 2185, Institut Pasteur, 25 rue du Dr. Roux, 75724 Paris Cedex 15, France

[5] Laboratorio de Fisicoquímica Biológica, Facultad de Ciencias, Universidad de la República, Uruguay

[6] Departamento de Bioquímica, Facultad de Medicina, Universidad de la República, Uruguay.

[7] Present address: Laboratory of Molecular and Structural Microbiology, Institut Pasteur de Montevideo, Uruguay

Abstract

PknG from *Mycobacterium tuberculosis* is a eukaryotic-like Ser/Thr protein kinase that regulates key bacterial metabolic processes as well as the pathogen's ability to handle macrophage's bactericidal effectors by a still unknown mechanism. To start unraveling the physiological role of this kinase we developed a new affinity purification-mass spectrometry (AP-MS) strategy based on the unique domain combination found in PknG to provide a precise and global view of PknG protein interactome in the mycobacteria. This multidomain protein has a conserved canonical kinase domain with N- and C-terminal flanking regions that constitutes docking sites controlling kinase activity and protein-protein interactions. Initially we identify 63 proteins representing putative direct or indirect PknG interactors. Using sequential elution steps we further separate PknG interactome to specifically recover substrates and phosphorylation-dependent interactors. This approach allowed us to identify nine putative new substrates of PknG as well as four proteins intreacting with PknG N-terminal autophosphorylated docking site. One of these proteins, FhaA, was further validated as a PknG substrate *in vitro* and the its *in vivo* phosphorylation sites were identified. We observed that FhaA overexpression in *Mycobacterium smegmatis* impairs mycobacterial growth. Interestingly the effect on growth rate could be partially reverted by substitution of each of the PknG phosphorylated residues in FhaA, pointing to the fact that these phosphorylation events have a physiological role in the FhaA control of cell growth. Overall our results indicate that PknG is playing a key roles in the regulation of intermediate metabolism, nitrogen assimilation and cell wall synthesis. These represent vital mycobacterial processes not only for bacterial physiology but also for survival in the nutrient deficient host environment.

Keywords: PknG; Ser/Thr kinase; FhaA; AP-MS; sequential elution; *Mycobacterium tuberculosis*.

Introduction

Mycobacterium tuberculosis, the etiological agent of tuberculosis (TB), is a major public health problem that caused 1.5 million deaths in 2014 [1]. This highly versatile bacteria can exist in different physiological states ranging from dormant bacilli in asymptomatic latent TB infection to actively replicating bacilli in active TB disease [2]. To ensure its survival in distinct environments, the tubercule bacilli senses the milieu and adapt itself to changes [3]. One of the major mechanisms that mediates the adaptation of this pathogen to its environment is the transduction of signals via serine/threonine protein kinases (STPKs) [4].

Genomic analysis of *M. tuberculosis* revealed the presence of 11 STPKs [5]. Nine of them are receptor-like proteins with an intracellular kinase domain connected through a single transmembrane helix to an extracellular sensor domain, whereas the two other members of the family (PknG and PknK) are soluble proteins [6]. These protein kinases have been related to several processes in mycobacteria including gene transcription, cell division and host-pathogen interactions [7]. Among them, PknG is of special interest as it plays dual roles in mycobacterial metabolism and pathogenesis.

From the structural point of view, PknG is unique among *M. tuberculosis* STPKs. In addition to the conserved catalytic kinase domain, it has a N-terminal rubredoxin-like domain (Rbx) and a C-terminal domain containing a tetratricopeptide repeat motif (TPR). Rubredoxin domains are bacterial protein modules characterized by the presence of an iron ion coordinated by four conserved cysteines [8]. In the case of PknG our previous results indicated that Rbx is involved in the modulation of catalytic activity [9, 10]. TPR domains mediate protein-protein interactions in both, eukaryotic and prokaryotic organisms. The reported crystal structure of a truncated form of PknG showed that the TRP domain is indeed involved in protein dimerization [11]. Preceding the Rbx domain, PknG has an unstructured N-terminal extension that contains up to four autophosphorylation sites. For many kinases, including mycobacterial kinases, autophosphorylation is a key process that controls kinase activity [12, 13]. However, in the case of PknG these phosphorylated sites participates mainly in substrate recruitment [14].

The physiological roles of PknG are still poorly understood. On one hand, disruption of *pknG* gene decreases *M. tuberculosis* viability and causes delayed mortality of infected mice suggesting that PknG has a crucial role in pathogenesis [15]. Furthermore, previous reports showed that a *pknG* knockout strain of *M. tuberculosis* is unable to block phagolysosme maturation, as is the case of WT strain. It has been suggested that secretion of PknG into the cytosol of infected macrophages mediates intracellular survival of the pathogen by inhibition of phagosome-lysosome fusion [16]. On the other hand, our group has demonstrated that PknG

also participates in the regulation of the bacterial metabolism. We have shown that PknG regulates the activity of several enzymes that controls glutamate metabolism, via the phosphorylation of an endogenous substrate (GarA) [14].

GarA, is a small forkhead-associated domain (FHA) protein that has been reported as substrate of several kinases, including PknB and PknG [13, 14]. FHA domains are protein modules that specifically recognize phosphorylated Thr residues and are known to play important roles in phosphorylation dependent signal transduction [17]. GarA is phosphorylated at the N-terminal sequence preceding the FHA domain by PknG and PknB. While PknG phosphorylates specifically T_{21}, PknB phosphorylates the adjacent residue, T_{22} [13]. Moreover, phosphorylation in these two threonine residues is mutually exclusive *in vitro*, pointing to GarA as a putative phosphorylated intermediate in different mycobacterial signaling pathways [14]. The phosphorylation of a threonine residue within the conserved N-terminal motif of GarA triggers the self-recognition of the phosphorylated residue by its own FHA domain. This interaction acts as an intramolecular switch that controls GarA activation/inhibition of enzymes that regulates glutamate metabolism. In particular, it has been shown that GarA acts in a synergic manner on enzymes that control glutamate intracellular levels, inhibiting α-ketoglutarate decarboxylase (KGD) and NAD^+-specific glutamate dehydrogenase (GDH) while activating α-subunit of the glutamate synthase complex (GOGAT) [14, 18].

Mass spectrometry-based phosphoproteomic approaches allowed the identification of more than 500 Ser- and Thr-phosphorylated residues in *M. tuberculosis* and showed that phosphorylation pattern in mycobacteria changed dramatically in response to environmental stimuli [19, 20]. However, the identification of cognate STPKs responsible of each phosphorylation event, the characterization of molecular interactions underlying each signaling cascade and ultimately the characterization of the functional role of each phosphoprotein still lags behind [4]. In this sense, GarA is the only well-characterized substrate of PknG up to now, pointing to the fact that most substrates and interactors of PknG as well as the identity of the processes regulated by this kinase are yet to be discovered.

In this work we developed a new affinity purification-mass spectrometry (AP-MS) strategy based on the unique domain combination found in PknG from *M. tuberculosis* to provide a precise and global view of PknG protein interactome in the mycobacteria. This approach allowed us to identify putative new substrates and new processes regulated by PknG. Furthermore one of these proteins: FhaA, was validated as a PknG substrate, and *in vitro* and in *vivo* phosphorylation sites were determined. Overall our results indicate that PknG plays key roles in the regulation of intermediate metabolism, nitrogen assimilation and cell wall synthesis.

Experimental procedures

Cloning and mutagenesis of Rv0020c

Plasmid pLAM12-fhaA, for *Strep*-tag® II-FhaA overexpression in mycobacteria, was constructed by PCR amplification of Rv0020c from genomic DNA of *M. tuberculosis* H37Rv and cloned under the control of the *M. smegmatis* acetamidase promoter on an extrachromosomally-replicating parent vector (pLAM12) cut with NdeI and EcoRI [21]. To introduce the chosen point mutations into the FhaA sequence site directed mutagenesis was used. The PCR reaction for all the mutations was done with the Phusion High-fidelity DNA polymerase (Thermo Scientific) in GC rich buffer plus 3% DMSO using pLAM12-fhaA as a template. All the oligonucleotides used in this study are listed in Table S1.

Recombinant protein production and purification

Full-length PknG was overexpressed in *Escherichia coli* BL21(DE3) cells grown for 24 h at 30 °C without IPTG and supplemented with 100 μM $FeCl_3$. The truncated $PknG_{\Delta 73}$ was expressed in the same strain after 20 h induction at 14 °C with 1 mM IPTG. Both proteins were purified as described before [14].

Mycobacterium smegmatis lysate preparation

Mycobacterium smegmatis MC^2155 were cultured in 200 mL of Sauton's medium supplemented with 0.05% Tween® 80 until early-logarithmic phase (OD_{600} between 0.6 and 0.8). Mycobacterial cells were harvested by centrifugation ($2000 \times g$ for 10 minutes at 4 °C) and washed with PBS. The supernatant was discarded and the pellet was resuspended in 3 - 4 mL of 25 mM Hepes, 150 mM NaCl, 1% glycerol, 1 mM EDTA, pH 7.4 (protein interaction buffer) plus Complete EDTA-free Protease Inhibitor Cocktail (Roche). An equal amount of acid-washed glass beads (≤ 106 μm, Sigma) were added to the cell pellet and resuspended in protein interaction buffer. Lysis was achieved by vortexing at top speed for 10 minutes. Cell debris and beads were removed by centrifugation at $1000 \times g$ for 5 minutes at 4 °C, supernatant were collected and stored at -80 °C until use. The protein extracts from three independent biological replicates were prepared.

Covalent protein immobilization

Full-length PknG or the truncated form was immobilized on NHS-Activated Sepharose 4 Fast Flow (GE Healthcare) according to manufacturer's guidelines. Briefly, 140 nmoles of PknG or $PknG_{\Delta 73}$ diluted in coupling buffer (0.2 M $NaHCO_3$, 0.5 M NaCl, pH 8.3) were incubated with activated agarose beads overnight at 4 °C with gentle end-over-end agitation (50 nmoles of NHS active groups per nmol of ligand). After the coupling step, non-reacted groups were blocked by alternating buffers with high and low pH; the buffers used were 0.5 M ethanolamine,

0.5 M NaCl, pH 8.3 (high pH) and 0.1 M sodium acetate, 0.5 M NaCl, pH 4.0 (low pH). The resin with the immobilized kinase was finally equilibrated in phosphorylation buffer (50 mM Hepes, pH 7.0). In parallel, all the steps previously described were performed using the same amount of matrix incubated in coupling buffer without protein, which was used as a control for unspecific binding.

Autophosphorylation assay

The matrix with or without immobilized kinase was incubated with 1 mM $MnCl_2$ and 0.5 mM ATP in phosphorylation buffer for 30 minutes at 37 °C under 1000 rpm agitation. After the treatment, the resin was washed twice with protein interaction buffer. PknG autophosphorylation was confirmed by MALDI-TOF MS [9].

Affinity purification

Immobilized PknG, PknG$_{\Delta73}$ or control resin was incubated with 800 µg of *M. smegmatis* protein extract, diluted to 1.7 mg/mL in protein interaction buffer and supplemented with 2 mM EDTA, overnight at 4 °C with gentle end-over-end agitation. The matrix was collected by centrifugation at 3500 × g and packed into Pierce™ Micro-Spin Columns (Thermo Scientific) and then washed five times with 50 mM Hepes, pH 7.0. The retained proteins were eluted using one of the two approaches described below.

One-step elution: one column volume of 70% acetonitrile, 0.1% formic acid was added to the matrix, incubated for 10 minutes at 25 °C and collected by centrifugation. This step was repeated twice.

Sequential elution: *elution of substrates under phosphorylation conditions (E1)*: the matrix was incubated with one column volume of 2 mM $MnCl_2$ and 500 µM ATP in phosphorylation buffer for 30 minutes at 37 °C and the eluted proteins were recovered by centrifugation. *Elution of phospho-residues recognizing proteins under dephosphorylation conditions (E2)*: the remaining resin was further incubated with 0.18 U/µL of calf intestine alkaline phosphatase (Roche) for 40 minutes at 37 °C. Afterward the eluted proteins were recuperated by centrifugation. After each specific elution stage, the matrix was washed with 0.5 column volume of ultrapure water and this wash was mixed with the corresponding elution (**Scheme 1**).

Sample preparation for mass spectrometry analysis

The protein mixtures were reduced by 10 mM DTT at 56 °C for 60 minutes, slowly cooled to room temperature and alkylated with 55 mM iodoacetamide at 25 °C for 45 minutes. Protein samples were digested with 1 µg of Sequencing Grade Modified Trypsin (Promega) by overnight incubation at 37 °C. Prior to proteolytic digestion the samples containing organic solvents (i.e. one-step elution) were dried in a SpeedVac concentrator and resuspended in 50

mM Hepes, pH 7.0. Peptide mixtures were desalted and concentrated using C18 reverse phase micro-columns (Omix®Tips, Varian), eluted with 2 × 20 µL of 70% acetonitrile, 0.1% formic acid, concentrated to dryness in a SpeedVac concentrator, and then resuspended in 10 µL of buffer A (0.1% of formic acid (v/v) in water). Peptide quantitation was performed at 280 nm using a NanoDrop 1000 spectrophotometer (Thermo Scientific).

LC MS/MS data acquisition

Each sample was injected into a nano-HPLC system (EASY-nLC 1000, Thermo Scientific) fitted with a reverse-phase column (EASY-Spray column, 50 cm × 75 µm ID, PepMap RSLC C18, 2 µm, Thermo Scientific). Peptides were separated using one of the following protocols: a linear gradient of buffer B (acetonitrile, 0.1% formic acid (v/v)) from 0 % to 35 % in 40 minutes at a constant flow of 300 nl/min at 45 °C for samples derived from one-step elution or a linear gradient of buffer B from 5 % to 55 % in 75 minutes at a constant flow of 250 nl/min at 45 °C for samples obtained from sequential elution.

Peptide analysis was carried out in a LTQ Velos nano-ESI linear ion trap instrument (Thermo Scientific) set in a data-dependent acquisition mode with a dynamic exclusion list of 45 s. Peptide ions were analyzed with Xcalibur 2.1 in two steps: 1. acquisition of full MS scan in the positive ion mode with *m/z* between 400 and 1200 Da, 2. sequential isolation of the ten most intense ions with a target value of 1000 ions and isolation width of 2 *m/z*, and fragmentation using collision-induced dissociation with a normalized collision energy of 35. The activation Q was set on 0.25 and the activation time on 15 ms. Standard MS parameters were set as follows: 2.3 kV electrospray voltage and 260 °C capillary temperature.

Proteomic data analysis

PatternLab for Proteomics (version 3.2.0.3) was employed to generate a target-decoy database using sequences from *M. smegmatis* strain ATCC 700084/MC[2]155 downloaded from Uniprot consortium in November 2014. In addition, 127 common mass spectrometry contaminants were included [22]. The Comet search engine was operated using the following parameters: tryptic peptides; oxidation of methionine and phosphorylation of serine, threonine or tyrosine as variable modifications and carbamidomethylation as fixed modification; and 700 ppm of tolerance from the measured precursor *m/z*. XCorr and Z-Score were used as the primary and secondary search engine scores, respectively [23].

Peptide spectrum matches were filtered using the Search Engine Processor (SEPro) and acceptable FDR criteria was set on 1% at the protein level. PatternLab's Approximately Area Proportional Venn Diagram module was used to perform comparisons between conditions and to determine proteins uniquely identified in each situation [24]. Proteins found in at least two

biological replicates of one condition were considered as uniquely identified when absent in at least two replicates of the other condition.

Functional analysis

The list of *M. smegmatis* identified proteins was converted to *M. tuberculosis* orthologs and classified using TubercuList (*M. tuberculosis* H37Rv database, March 2013 release 27 [25]). If a protein ortholog was absent in TubercuList, a blastp sequence and similarity search was performed using *M. smegmatis* sequence as query and *M. tuberculosis* complex (taxid: 77643) as database [26]. For protein classification of PknG putative interactors PANTHER with default settings were used [27].

FhaA overexpression in M. smegmatis

Electrocompetent *M. smegmatis* MC^{2}155 were transformed with pLAM12 constructs. After recovery in 1 mL of Middlebrook 7H9 broth supplemented with ADS at 37 °C for 3 h, cells were spotted onto solid medium (7H10) containing kanamycine (50 µg/mL). Plates were incubated at 37 °C for 2 - 5 days and then a colony was picked up and grown to mid-logarithmic phase in Middlebrook 7H9 broth supplemented with ADS, 0.05% Tween® 80 and kanamycine. Acetamide was added at a final concentration of 0.2% during exponential growth in liquid medium to induce FhaA (or FhaA mutants) expression.

For bacterial growth curve experiments overnight starter culture was diluted to OD$_{600}$ at 0.05 in the presence or absence of 0.2% acetamide for pLAM12 constructs, and every three hours the turbidity was determined.

Mapping of FhaA phosphorylation sites

The mycobacterial cell lysate was prepared as described above. *Strep*-tag® II-FhaA was purified using Strep-Tactin® Sepharose® (IBA) according to manufacturer's guidelines and eluted in a competitive manner with D-desthiobiotin.

For *in vivo* identification of FhaA phosphorylation sites, peptides from trypsin digested FhaA were separated on nano HPLC and analyzed in a LTQ Velos nano-ESI linear ion trap instrument.

To evaluate FhaA as PknG substrate *in vitro*, cell lysates of *M. smegmatis* MC^{2}155 overexpressing *Strep*-tag® II-FhaA were loaded on Strep-Tactin® Sepharose® (IBA) and immobilized FhaA was dephosphorylated with 0.07 U/µL of calf intestine alkaline phosphatase (Roche) for 40 minutes at 37 °C under 1000 rpm agitation. The supernatant was discarded and *Strep*-tag® II-FhaA was eluted with D-desthiobiotin according to manufacturer's guidelines. The eluate was divided in two fractions and one of them was incubated with 1 mM MnCl$_2$ and 250

µM ATP in the presence of PknG 10 nM in phosphorylation buffer for 30 minutes at 37 °C under 1000 rpm agitation. The same experiment was performed in absence of PknG (control).

Protein mixtures were digested with 0.5 µg of Sequencing Grade Modified Trypsin (Promega) by overnight incubation at 37 °C. Peptides were separated on nano HPLC and analyzed with LTQ Velos, as described in previous sections. In addition, raw MS/MS spectra were interpreted with the Proteome Discoverer software package (v.1.3.0.339, Thermo) using Sequest as search engine. We considered positive phosphosite identification when more than one spectra for the phosphopeptide was obtained, pRS probability was >95% and manual inspection of the MS/MS spectra showed at least two confirmatory fragment ions.

SPR analysis

For SPR analysis, FhaA was diluted in 10 mM sodium acetate, pH 4.5 at a concentration of 5 µg/mL and immobilized on a CM5 sensorchip by standard amine coupling (activation with a 1:1 mixture of EDC/NHS, injection of FhaA at a flow rate of 5 µL/min and inactivation with ethanolamine) using a BIACORE 3000 (Biacore AB, Uppsala, Sweden) achieving a final density of 25 RU.

After immobilization phosphorylated PknG and PknG$_{\Delta73}$ were diluted in HBS-EP buffer (0.01 M HEPES pH 7.4, 0.15 M NaCl, 3 mM EDTA, 0.005% v/v Surfactant P20) to a final concentration of 30 nM and injected during 3 minutes at a flow rate of 80 µL/min over the immobilized and a reference surface. Regeneration was achieved by extensively washing with running buffer. All injections were done at 25 °C and double referenced by subtracting the reference cell signal and a buffer injection.

Results

Identification of PknG protein interaction partners by AP-MS in a one-step elution protocol

In order to unveil PknG interactome, we carried out affinity purification experiments to recover kinase direct or indirect interacting partners. Thus, PknG was first immobilized on NHS-Activated Sepharose and incubated with ATP under phosphorylation conditions. This step is relevant for complete N-terminal sequence autophosphorylation, as recombinant PknG is produced as a mixture of partially phosphorylated isoforms [14]. We verified PknG coupling to the resin and the kinase phosphorylation state by digesting a fraction of the resin with trypsin and analyzing the resultant peptides by MS. Before autophosphorylation unphosphorylated (*m/z* 5396.2) and monophosphorylated (*m/z* 5477.0) ions of sequence 10-60 were found, similar to previous reports [14] (**Figure S1A**). After incubation with ATP, the most intense ion observed corresponded to the diphosphorylated specie (*m/z* 5556.3). Besides incubation of immobilized PknG with its substrate, GarA [14], under phosphorylation conditions followed by MALDI TOF MS analysis allowed the detection of phosphorylated GarA (mass increase of 80 Da), confirming that the immobilized kinase was an active enzyme (**Figure S1B**).

Immobilized PknG and control resin were used to trap binding partners from soluble cell lysates of *M. smegmatis* MC[2]155, a non pathogenic bacteria extensively used as a model in *M. tuberculosis* research. Retained proteins were eluted using the one-step elution protocol described above and digested with trypsin for further nano-LC MS/MS analysis. Data analysis allowed the identification of 78, 82 and 65 proteins in three independent biological replicates (**Table S2**). After an extensive protocol optimization, 13 proteins were recovered from control experiments and represent proteins that interact nonspecifically with the mock resin. These proteins are abundant and belong to protein families frequently detected as contaminants in affinity purification experiments [28]. In our approach, a protein was considered a contaminant if detected in at least two mock experiments. All the contaminant proteins were manually removed from the lists of PknG interacting proteins and the ultimate inventory included 69, 71 and 55 proteins, respectively (**Figure 1A**). 63 proteins were common to at least two of three replicates of PknG interactome, and 42 were present in all three analyzed replicates thus representing a confident and reproducible list of PknG interactors (**Table S3a**).

Table S3a displays the complete list of 63 PknG putative interactors, identified in at least two of three replicates (shaded grey in **Figure 1A**) but absent from mock experiments, grouped according to their functional category. The largest groups were composed of proteins functionally related to the intermediary metabolism and respiration (24%) and lipid metabolism (16%) (**Figure 1B**).

With this experimental approach it is impossible to discriminate between PknG direct interactors and protein recovered as part of protein complexes with PknG interacting proteins. In fact, some of the proteins presents in PknG interactome have been previously reported to interact with each other and/or to be part of the same operon. Analysis of PknG interactome employing DOOR^2 (Database of prOkaryotic OpeRons, version 2.0, [29]), a database containing computationally predicted operons in *M. smegmatis* genome, clearly showed that some of the proteins recovered are coded by genes that are part of the same operon and in several cases the interaction at the protein level was also demonstrated [30] (**Table S3b**). In particular, MSMEG_2080 and MSEG_2081 belong to the same operon and encode the proteins subunits of an heterotetrameric acyl coenzyme A dehydrogenase [31]. Genes MSMEG_2351 and MSMEG_2352 are in the same operon and code for different subunits of an electron transfer flavoprotein [5]. Moreover, MSMEG_4327 and MSMEG_4328 are within a four-gene operon and a strong interaction between the proteins was reported by yeast two-hybrid approach [32]. MSMEG_4636 to MSMEG_4640 genes are located in the middle of a eight-gene cluster that encode 4 of the 5 subunits presents in the cytosolic ATP synthase F_1 moiety [33]. And finally MSMEG_6759 and MSMEG_6761 are included in a three-gene operon related with glycerol degradation. The presence of different proteins of a complex among PknG interactors can shed light on the processes regulated by this kinase though the direct binding partner cannot be ruled out.

Since the immobilized protein is a Ser/Thr kinase, the possibility that some of the interacting partners were also substrates of PknG was considered. Comparison of our protein list with *M. smegmatis* phosphoproteome data [34] revealed that only two putative PknG interactors were previously identified as phosphoproteins: FHA domain protein (MSMEG_0035) and 30S ribosomal protein S5 (MSMEG_1472) (**Table S3c**). Moreover, PknG interactome in *M smegmatis* was matched up with other mycobacterial phosphoproteomes in order to further identify phosphoproteins among interactors [19, 20, 35]. Five proteins identified in PknG interactome have orthologs that appeared phosphorylated in Ser, Thr or Tyr in other members of *Mycobacterium* genus: Rv0020c (FHA domain-containing protein, FhaA), Rv0350 (chaperone protein DnaK), Rv2187 (long-chain-fatty-acid-CoA ligase, FadD15), Rv2744c (35 kDa protein), and Rv3849 (uncharacterized protein) (**Table S3c**). All these phosphoproteins were found in two or more sets of phosphoproteomic data (different MS-based methodologies and in different strains of mycobacteria) [36].

The diversity of the proteins identified as putative interactors of PknG suggest that the signal transduction pathways mediated by this kinase affect a great variety of biological processes.

This could be explained inasmuch as bacterial serine/threonine kinases constitute signal integration hubs that coordinate different cellular processes [37].

Mapping interactors to PknG domains

Notably, GarA, the only well-characterized substrate of PknG [14], was not present in PknG interactome. One possible strategy to improve the detection of less abundant proteins in proteomic analysis is to reduce the complexity of samples. To achieve this aim we developed a novel sequential elution approach based on the singular domain organization of PknG (**Scheme 1**). We decided to fractionate PknG interactome using sequential elution steps in order to separately recover substrates and phosphorylation-dependent interactors from the rest of the partners. Additionally the interactome of PknG was compared with the interactome of a truncated form of the kinase that lacks the N-terminal autophosphorylation sites that mediate GarA recruitment (PknG$_{\Delta73}$). Thus, we allowed partners to interact with full-length or the truncated mutant PknG$_{\Delta73}$ and performed a first elution step using phosphorylation conditions. This approach would allow to discern between substrates requiring or not the docking sites of the N-extension of PknG, and to recover GarA only when using full-length PknG as a bait. In a second step we used a phosphatase treatment for the specific disruption of interactions mediated by phosphoresidues. Comparison of the interactors recovered using full-length and PknG$_{\Delta73}$ would let to specifically identify binding partners that recognize PknG autophosphorylation sites.

E1: putative substrates of PknG

When *M. smegmatis* MC^{2}155 proteins retained on whole length PknG-coupled resin were eluted under phosphorylation conditions (*E1*), eleven proteins representing putative PknG substrates were recovered in at least two of three replicates. Using immobilized PknG$_{\Delta73}$, seven of the eleven previously identified PknG putative substrates were recovered. PknG form *M. tuberculosis* (Rv0410c) was identified in all elution steps using both full-length and truncated PknG constructions as baits, clearly corresponding to PknG non-covalently bound to the resin that is washed away during the elution steps. Overall, we have identified 10 putative substrates of full-length PknG and only 6 of them are also putative substrates of PknG$_{\Delta73}$ (**Table S4**). Interestingly, five out of ten proteins identified as putative substrates of PknG were previously reported as phosphoproteins in phosphoproteome studies [19, 20, 34, 35] (**Table S4c**).

Thus, four proteins were identified in *E1* only when using whole length PknG as bait and represent putative substrates interacting through N-terminal region of the kinase (**Figure 2A** and **Tables 1** and **S4a**). As expected, GarA (MSMEG_3647) was recovered in all three biological

replicates within this group, in full agreement with our previous knowledge of PknG-GarA interaction. [38]. The presence of GarA in this elution step not only validates our experimental approach but also supports the list of putative substrates identified.

The additional three substrates (or interactors of substrates) of PknG that required the N-terminal phosphorylation of the kinase for its interaction are: integration host factor, electron transfer flavoprotein beta subunit and glutamine synthetase 1. In these cases a phosphorecognition domain is not present in the substrates; therefore the interaction that supports the specific recovery in full-length PknG interactome deserves further study. In particular, integration host factor is a nucleoid associated protein that participates in DNA compactation and relaxation by a DNA bending mechanism [39]. This protein has been identified as a phosphoprotein in *M. tuberculosis* but the physiological role of this phosphorylation is still unknown [19]. Electron transfer flavoprotein beta subunit has not been characterized and is the only putative substrate of full-length PknG that has not been ever reported as a phosphoprotein. Glutamine synthetase 1 catalyzes the condensation of ammonia and glutamate to form glutamine and is a fundamental component of bacterial nitrogen metabolism. Moreover, it is involved in the synthesis of poly(L-glutamic acid/glutamine), a cell wall element exclusively present in virulent mycobacteria [40]. This protein has been identified as a phosphoprotein and unlike the other putative substrates in this case the modified residue is proposed to be a Tyr [35].

Additionally six proteins were recovered using both PknG constructions, and represent putative substrates (or substrate interactors) whose interaction does not rely on PknG N-terminal sequence. Among them, four have been related to lipid metabolism or membrane permeability (MSMEG_5122, MSMEG_0709, MSMEG_0880 and MSMEG_1807) [41-45] and other two: MSMEG_5122 and MSMEG_6242 are iron-sulfur proteins that participates in electron transfer reactions [41, 42, 46]. In addition, three of them have been reported as phosphoproteins (**Table S4c**).

Analysis of *E1* by MS/MS did not allowed the identification of phosphopeptides among tryptic peptides of PknG putative substrates. Several factors preclude phosphopeptide identification in complex mixtures without a pre-enrichment step. Phosphopeptides have lower ionization efficiencies and also generate low quality MS/MS spectra dominated by the neutral loss of the phosphate group, resulting in lower confidence in spectral matching [47, 48]. Even when we cannot conclude if all of the proteins recovered in *E1* are direct substrates of PknG, some evidence support this hypothesis. On one hand, many of them have been reported as substrates of kinases in phosphoproteomics or *in vitro* phosphorylation assays [19, 20, 34, 35]. Moreover,

the recovered putative substrates have not been reported or predicted to interact with each other. In any case, up to now only two PknG substrates have been well-characterized: PknG itself and GarA. Our interactomic results expand this list to include nine new putative PknG substrates that deserved further *in vitro* and *in vivo* validation.

<u>*E2*: phosphorylation-dependent interactors</u>

PknG N-terminal phosphorylation sites constitute docking sites that may be responsible not only for the recruitment of substrates as GarA, but also for other interacting partners in PknG signaling pathways. In order to characterize partners that could be interacting in a phosphorylation dependent manner, we performed a second elution step in which *M. smegmatis* MC^2155 proteins retained on PknG or $PknG_{\Delta73}$-coupled matrix were eluted under dephosphorylation conditions (*E2*).

Four proteins were systematically recovered using whole length PknG but not the truncated form, representing proteins that may interact with phosphoresidues in the N-terminal sequence of the kinase (**Table 1**). Very interestingly, among them we recovered FHA domain protein (MSMEG_0035), a FHA-containing protein as GarA. The function of the protein ortholog in *M. tuberculosis* has been related to cell wall biosynthesis and/or its regulation [49]. As in *E1*, the integration host factor was also recovered as an interactor of full-length PknG together with two ribosomal proteins: 30S ribosomal protein S17 and 50S ribosomal protein L29. Recently, other ribosomal protein (50S ribosomal protein L13) was proposed as a PknG substrate and a functional correlation was suggested between the phosphorylation of L13 by PknG and mycobacterial biofilm growth [50].

Even though no additional proteins should appear in *E2* when using $PknG_{\Delta73}$ as bait, deletion of the N-terminal region of PknG could expose surfaces able to interact with proteins, explaining why a new group of interactors was indeed detected (**Table S5b**). However, these proteins are not further discussed as they probably constitute an artifact. In addition several proteins were recovered with alkaline phosphatase treatment when using both, full-length PknG and $PknG_{\Delta73}$ as baits. Most of them are ribosomal proteins (**Table S5c**), and even when we cannot rule out other phosphorylation dependent interactions among PknG and its partners (or among direct and indirect partners), the conditions employed for dephosphorylation can disturb protein interactions, in particular ribosomal protein complexes [51].

The elution scheme developed on this work let us to deepen into PknG domain interactions. For example FhaA was recovered in one step elution as PknG interactor, and using the sequential

strategy we could further dissect that this interaction involves N-terminal phosphorylated PknG residues.

FhaA is a substrate of PknG

One of the most striking results arising from both one-step and sequential elution protocols is the systematic recovery of the protein FhaA as PknG interactor. *M. tuberculosis* genome encodes several FHA-containing proteins and all of them have been identified as substrates of multiple STPKs *in vitro* [13, 52-55]. In particular, FhaA has been reported as an *in vitro* substrate of the essential STPK PknB and systematically appeared highly phosphorylated in mycobacterial phosphoproteomes [36, 49]. The structural analysis of FhaA showed an N-terminal domain of unknown function and a C-terminal FHA domain linked by an intermediate unfolded domain of approximately 300 residues. This structural arrangement is characteristic of scaffold proteins that tether signaling components through independent binding domains [49]. Based on this, we selected FhaA for further validation as substrate and/or interactor of PknG.

FhaA was clearly identified using both approaches, however, the sequential elution protocol allowed us to identify this protein as PknG interactor via phosphoresidues in the N-terminal sequence of the kinase. Even when FhaA was not eluted in the first step (*E1*), where GarA and other putative substrates of PknG were recovered, we wondered if FhaA could be a substrate of PknG that has a different behavior than GarA. In fact, FhaA can represent a substrate that is released from the active site once phosphorylated but keeps the interaction through the N-terminal phosphoresidue. Alternatively FhaA might represent a scaffold protein, and binding to the phosphorylated N-terminal region of PknG can participate in recruiting downstream mediators in PknG signaling cascade.

To investigate the interaction between PknG and FhaA we initially analyzed *in vitro* FhaA phosphorylation by PknG. For that purpose we have overexpressed *M. tuberculosis* FhaA protein carrying a *Strep*-tag® II on *M. smegmatis* strain. While no phosphopeptides were found in LC MS/MS analysis of phosphatase-treated FhaA, two phosphopeptides were confidently identified after incubation with PknG, and the phosphorylated residues were unequivocally localized. In particular peptide of *m/z* 1630.8 (corresponding to phosphorylated sequence 108-120) and peptide of *m/z* 1415.0 (corresponding to phosphorylated sequence 14-25) were systematically detected in our phosphorylation assays. These two sequences mapped to the N-terminal domain of FhaA. MS/MS spectra allowed the unequivocal identification of T_{18} and T_{116} as the phosphorylated residues (100% pRS probability in both cases) (**Figure S2**).

In order to determine if these sites were phosphorylated in living mycobacteria recombinant FhaA produced in *M. smegmatis* was digested and analyzed by mass spectrometry. Thus, three FhaA phosphopeptides were found, two of them matching those previously identified in *in vitro* phosphorylation assays (sequences 14-25 and 108-120) and an additional one corresponding to sequence 368-378, with residue T_{377} identified as the phosphorylation site (**Figure 3**). Previous reports demonstrated the phosphorylation of FhaA T_{116} *in vitro* by PknB*,* and *in vivo* in phosphoproteomic studies [36, 49]. The presence of the same phosphorylated residues in PknG phosphorylation assay as *in vivo* experiments supports the idea that PknG can be kinase responsible for FhaA phosphorylation. However, the identification of FhaA as a physiological substrate of PknG requires further experimentation. In any case, the presence of extra phosphorylation sites point to the fact that other kinases might also be regulating FhaA phosphorylation state.

FhaA interacts in vitro with PknG

The sequential elution results suggested that PknG-FhaA interaction mainly relies on the FHA domain of FhaA and the N-terminal phospho-Thr residues of PknG. To confirm the direct interaction between these proteins we have used plasmon resonance (SPR) biosensor technology. As shown in **Figure 4**, phosphorylated PknG was able to interact and form a stable complex with FhaA immobilized on a chip. Interestingly enough this interaction could not be observed when the truncated form of PknG was used. Thus plasmon resonance experiments allowed us to recapitulate the results of sequential elution, and clearly demonstrated that N-terminal autophosphorylation sites of PknG play a key role in FhaA recruitment.

FhaA phosphorylation affects the mycobacterial growth

We observed that *M smegmatis* transformed with plasmid pLAM12-fhaA formed colonies after *ca.* 4 days, twice the time required by cells containing the empty vector. To quantify this alteration of the growth rate we followed the mycobacterial growth in liquid medium. While the exponential and stationary phases of both cultures did not vary considerably, a longer lag phase was observed in the case of cells transformed with pLAM12-fhaA as compared to cells containing the empty vector (**Figure 5**). Therefore, FhaA overexpression appears to negatively affect the bacterial growth rate. Further, in order to evaluate the effect of FhaA phosphorylation on this trend, FhaA threonines susceptible of phosphorylation of PknG were individually mutated to alanine, generating plasmids pLAM12-fhaA $T_{18}A$ and pLAM12-fhaA $T_{116}A$. A differential growth phenotype was found among cells containing different *fhaA* alleles (wild-type or phosphoablative) (**Figure 5**), with a lag phase of an intermediate duration between those observed for cells containing pLAM12-fhaA or the empty vector, pointing to a role of FhaA phosphorylation events on growth rate.

The bacterial cell adaptation required to exploit new environmental conditions occurs during the lag phase of growth [56]. Thus, a change in the length of this phase is indicative of differences in adaptability. According to our data, the overproduction of FhaA (and its phosphoablative mutants) impairs the mycobacterial fitness to the environment. Despite the importance of these results, we might have guessed an important effect of FhaA overexpression because this protein was proposed to participate in the maintenance of proper cell morphology [57].

Discussion

The identification of protein interacting partners is a critical step in the understanding of protein function, because the biological attributes of a protein depend on its physical interaction with other molecules. Taking into account that protein complexes are the functional units responsible for driving biological processes in cells we employed an affinity purification-mass spectrometry approach to unravel signaling pathways regulated by *M. tuberculosis* PknG [58, 59]. AP-MS is a widely used tool for the study of protein interaction networks, because it couples the specificity of the pray:bait protein isolation procedure with the sensitivity of high-throughput mass spectrometry analysis [60]. Even when this approach has been successfully used to identify protein partners in several biological systems it has several well-known drawbacks [59, 61]. On one hand, cell disruption can lead to formation of protein-protein interactions not actually present in the living cell. Additionally, using high sensitive MS analysis, unspecific protein binding to affinity resins has systematically led to false positive identifications. This is a very well studied phenomenon that has been the subject of many reports [28, 62]. In order to overcome some of these technical limitations we combined the classical approach with a sequential elution protocol using different PknG constructions as baits. Overall this approach gave us a global picture of kinase protein interactors, and allowed us to specifically elute substrates and phosphorylation-dependent interactors, while reducing the identification of nonspecifically bound proteins.

One-step elution protocol let us to recover 63 partners that form part of protein complexes around PknG, representing putative functional multiprotein units. To understand the biological relevance of these candidates we examined their distribution in specific protein classes using PANTHER [27]. The interactors of PknG classified into 10 protein classes with almost two-thirds of the proteins belonging to oxidoreductase (27%), transferases (19%), and nucleic acid binding proteins (17%) (**Table S3d**). Recently, Deng *et al.* built up a *M. tuberculosis* proteome microarray and reported 59 new PknG interactors [63]. About half of the putative PknG interactors identified in that work belong to three protein classes: hydrolases (17.8%), transferases (17.8%), and oxidoreductases (13.3%). Thus PknG interactome analyzed using very different methodologies point to the same processes being regulated by this kinase. However, only three proteins appeared in both interactomes: HBHA-like protein (MSMEG_0919), ATP synthase gamma chain (MSMEG_4937) and cyclopropane fatty-acyl-phospholipid synthase (MSMEG_6284). This little overlap is not surprising as protein microarrays only identify direct interactors, whereas our approach was designed to recover protein complexes mediated by PknG and give a more general view of PknG regulated processes. In addition, Deng´s approach is a high throughput strategy where more than 3000 proteins were produced without final

functional or structural controls. However, the only well-established PknG interactor: GarA [14] is absent both in Deng´s work and our one-step elution protocol.

According to TubercuList classification (*M. tuberculosis* H37Rv database, march 2013 release 27 [25]) PknG interactome is composed mainly by proteins included in the following categories: intermediary metabolism and respiration, lipid metabolism, information pathways, conserved hypotheticals and unknown, and cell wall and cell processes.

In particular, we identified several members of the mycobacterial ATP synthase multisubunit complex as PknG interactors, and one of them was also previously identified by microarrays as direct PknG interactor (ATP synthase gamma chain). ATP synthase complex is composed of a membrane-integral F_0 sector (subunits ab_2c_{10-15}) and a hydrophilic F_1 section (subunits α3β3γδε) that catalyzes the synthesis of ATP using the proton-motive force across the membrane [33]. All the F_1 subunits, excluding ε, were identified in PknG interactome (**Table S3a**).

In addition, the interactome of PknG includes several enzymes involved in central carbon metabolism (**Table S3a** and **Scheme 2**) [64], which generates energy in the form of reducing equivalents, ATP, and biosynthetic precursors [65]. Thus, our results suggest that PknG has a role in the modulation of energy metabolism. Remarkably the only well characterized substrate of PknG acts synergically on three enzymes to regulate the crossroads between nitrogen and carbon metabolism [14].

In this line, PknG interactome includes enzymes in every step of fatty acid β-oxidation: the acyl-CoA dehydrogenase *fadE23* (MSMEG_2080) and the putative acyl-CoA dehydrogenase (MSMEG_2081), the enoyl-CoA hydratase/3-hydroxyacyl-CoA dehydrogenase *fadB* (MSMEG_5720) and the 3-ketoacyl-CoA thiolase *fadA2* (MSMEG_0373). Even an AMP-binding enzyme (long-chain fatty-acid-CoA ligase, MSMEG_4254) was identified; the product of the reaction catalyzed by this enzyme is a fatty acyl-CoA that enters directly into the β-oxidation-cycle (**Table S3a** and **Scheme 2**). Moreover, the ortholog in *Mycobacterium avium* of the electron transfer flavoprotein beta subunit (MSMEG_2351) has been implicated in fatty acid β-oxidation [66]. This protein was detected in PknG interactome and as a PknG putative substrate that requires the N-terminal extension of the kinase for its interaction (**Table S3a** and **S4a**). The possible role of PknG-mediated phosphorylation of this protein on the control of fatty acid β-oxidation requires further validation.

Among PknG direct or indirect interacting proteins we also identified proteins that participate in fatty acid synthase type II elongation pathway (FASII): two 3-oxoacyl-[ACP] reductases (*fabG4* and *fabG*, MSMEG_0372 and MSMEG_3150 respectively) that catalyze the first reaction of this pathway, and two 3-oxoacyl-[ACP] synthase 1 (*kasA* and *kasB,* MSMEG_4327 and MSMEG_4328) involved in the last reaction step. Another protein related to this biosynthetic route is the cyclopropane fatty-acyl-phospholipid synthase (MSMEG_6284) responsible for the introduction of a cyclopropane ring in an α-mycolic acid [67] (**Table S3a**). This protein was also identified as a direct interactor of PknG in previous works [63]. The synthesis of mycolic acid is a crucial process for cell envelope structure as well as *M tuberculosis* pathogenesis. Thus ours and previous results indicate that PknG might regulate the synthesis of this lipid moiety that represents one of the most relevant characteristics of this group of pathogenic bacteria. Besides, another enzyme recovered as PknG substrate: the acetyl-/propionyl-coenzyme A carboxylase alpha chain (MSMEG_1807) is also involved in mycolic acid synthesis, this protein catalyzes the first step in FASI elongation pathway (**Table S3a** and **S4b**) [68].

Other interesting group of proteins identified in PknG interactome were the ATP binding cassette (ABC) transporters, permeases that transport molecules across biological membranes [69]. We distinguished three proteins involved in the transport of phosphate (MSMEG_5782), sulfate (MSMEG_4533) and glutamate (MSMEG_2727). In the same functional category we found a previously reported PknG direct interactor: HBHA-like protein (MSMEG_0919) [63]. This is a cell-surface protein involved in bacterial aggregation and cell adhesion and an important player in the dissemination of *M. tuberculosis* from the site of primary infection [70, 71].

The sequential elution procedure allowed us to isolate PknG protein interactors according to its site of interaction and to identify putative substrates that may mediate the effects of PknG on mycobacterial physiology. As expected, using this experimental approach we were able to recover GarA in the elution corresponding to PknG substrates (*E1*). As mentioned above GarA phosphorylation state controls the activation/inhibition of enzymes that regulates ammonium assimilation and glutamate metabolism [14, 18]. GarA inhibits glutamate degradation by acting on GDH and KGD and at the same time activates GOGAT. Thus GarA act in a very coordinated manner on three of the four enzymes responsible of the balance between ammonium assimilation and glutamate oxidative deamination. Strikingly, the fourth enzyme that participates in this process, glutamine synthetase (GS), was identified in this study as a putative PknG substrate. This enzyme incorporates ammonia by converting glutamate to glutamine, which is the first step in the pathway of ammonia assimilation, that results in the incorporation of ammonia into an aminoacidic backbone [72, 73]. GS is a key enzyme whose synthesis and

activity are regulated at different levels [74]. In particular the reversible adenylylation of a conserved Tyr residue is a well established activity regulation mechanism [75]. Our results suggest that protein phosphorylation may represent a new mechanism for glutamine synthetase activity control, and the identification of this enzyme in phosphoproteomic analysis of mycobacteria further support this notion. In addition this result is in very good agreement with previous proposed roles of PknG and indicates that through the synergic action on GarA and glutamine synthetase, PknG is regulating the four enzymes involved in nitrogen assimilation in mycobacteria.

The FHA-containing protein (FhaA) was initially identified as a phospho-dependent PknG interactor, we also demonstrated that FhaA is an *in vitro* substrate of PknG. The recovery of this substrate specifically in *E2* indicated that the interaction with PknG and its further phosphorylation relies in a different mechanism than GarA. FhaA (Rv0020c) is a protein of 527 residues with a C-terminal FHA domain connected to a N-terminal domain of unknown function by an unfolded 300 residues domain [49]. Our data supports a model for PknG-FhaA interaction involving two different regions in each protein. The identification of phosphorylated sites in FhaA indicates that the active site of PknG interacts with the phosphorylatable N-terminal domain of FhaA. Additionally, the autophosphorylated N-terminal extension of the kinase interacts with the C-terminal FHA domain of FhaA since this protein is eluted in *E2* and the SPR data clearly demonstrated the absence of interaction when the kinase lacks the N-terminal. Therefore, our data support a model in which the self-recognition of the pThr residue by the FHA domain is not a favored process and thus phosphorylation does not trigger disruption of PknG-FhaA interaction (**Scheme 3**).

Although the function of FhaA remains poorly understood, it has been postulated that this protein could participate in cell shape and cell division processes. The *fhaA* gene is encoded by a highly conserved operon involved in the control of these processes that also harbours two STPKs (*pknA* and *pknB*) and the solely Ser/Thr phosphatase encoded by *M. tuberculosis* [76]. Recent studies further support the role of FhaA in mycobacterial physiology. It has been described the recruitment of FhaA by the phosphorylated threonine of the pseudokinase MviN and the involvement of this complex in mycobacterial cell wall biosynthesis [57]. Moreover, *in vitro* experiments showed that PknB specifically phosphorylates FhaA at T_{116} [49]. Altogether, these results suggest that PknB, a peptidoglycan responsive STPK, phosphorylates MviN initiating a complex assembly among MviN and FhaA. In fact the observation that in *M. smegmatis* FhaA localizes in septa and poles is compatible with its proposed role in cell wall synthesis [57]. Global phosphoproteomics studies in mycobacteria demonstrated that FhaA appeared phosphorylated *in vivo* both in threonines and tyrosines but the kinase(s)

responsible(s) of each phosphorylation event are still unknown [36]. In this work we have identified three phosphorylation sites in FhaA overproduced in *M. smegmatis* (T_{18}, T_{116} and T_{377}); one of them has been previously reported in phosphoproteomics studies (T_{116}) while the remaining sites represents new phosphorylation sites. It is difficult to exclude the possibility of unspecific phosphorylation events as a result of protein overexpression. However, we have identified the same phosphoT residue as a result of PknG phosphorylating activity, pointing to this residue as an specific point of PknG control.

In agreement with the proposed function for FhaA we have shown that protein overexpression generates a growth phenotype. Interestingly the effect on growth rate could be partially reverted by substitution of each of the PknG phosphorylated residues in FhaA, pointing to the fact that these phosphorylation events have a physiological role in the FhaA control of cell growth. Altogether our results provide evidence that PknG could be contributing to the control of cell division and cell wall synthesis in mycobacteria through the regulation of FhaA phosphorylation state. Based on the result presented here and the model for PknG–FhaA interaction, our working hypothesis is that FhaA is a scaffold protein that tethers PknG signaling complexes through phosphorylation-dependent recruitment of protein partners. To further validate FhaA as a physiological substrate of PknG and to identify the downstream interactors comparative proteomics studies using PknG knockout as well as *in vivo* crosslinking experiments are being performed.

Here we presented a reliable AP-MS protocol to identify substrate and specific interactors of PknG from *M. tuberculosis* that may be applicable not only to other mycobacterial kinases but also to unrelated protein kinases. Moreover, we shed light on new processes regulated by PknG and uncover new PknG substrate. In particular, our results suggest that PknG plays key roles in the regulation of intermediate metabolism, nitrogen assimilation and cell wall synthesis. These represent vital mycobacterial processes not only for bacterial physiology but also for survival in the nutrient deficient host environment. These results open the possibility that PknG mediates survival inside macrophages by controlling bacterial metabolic adaptation in the host. This warrants further studies on PknG signaling in mycobacteria.

Table 1. List of identified proteins in *E1* and *E2*.

Proteins identified as exclusive interactors of whole length PknG in *E1* (4)		
UniProt identifier	**Description**	**Gene name**
A0QYG2	Glycogen accumulation regulator GarA	MSMEG_3647
A0QWS8	Integration host factor	MSMEG_3050
A0QUV6	Electron transfer flavoprotein beta subunit	MSMEG_2351
A0R079	Glutamine synthetase	MSMEG_4290

Proteins identified as common interactors of PknG and PknG$_{\Delta73}$ in *E1* (6)		
UniProt identifier	**Description**	**Gene name**
A0R2I1	4Fe-4S ferredoxin, iron-sulfur binding protein	MSMEG_5122
A0R597	Inorganic pyrophosphatase	MSMEG_6114
A0R5M3	Alcohol dehydrogenase, iron-containing	MSMEG_6242
A0QQC8	Chaperone protein DnaK	MSMEG_0709
A0QQU5	60 kDa chaperonin 1	MSMEG_0880
A0QTE1	Acetyl-/propionyl-coenzyme A carboxylase alpha chain	MSMEG_1807

Proteins identified as exclusive interactors of whole length PknG in *E2* (4)		
UniProt identifier	**Description**	**Gene name**
A0QNG7	FHA domain protein	MSMEG_0035
A0QSE0	30S ribosomal protein S17	MSMEG_1445
A0QWS8	Integration host factor	MSMEG_3050
A0QSD9	50S ribosomal protein L29	MSMEG_1444

Figures
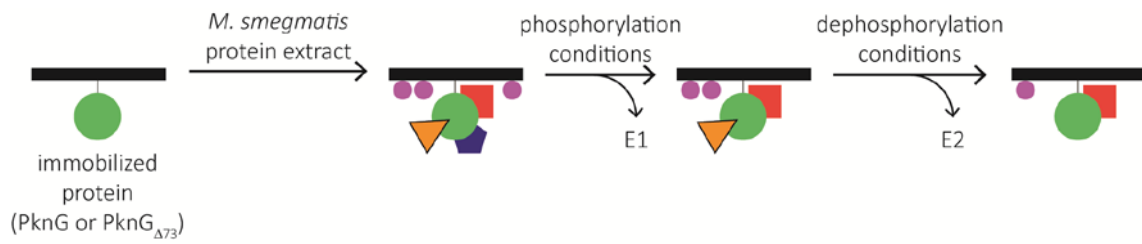
Scheme 1. Affinity purification approach for isolation of PknG protein interactors. Immobilized PknG, PknG$_{\Delta73}$ or control resin were incubated with *M. smegmatis* protein extract and domain specific PknG-interacting proteins were sequentially eluted. *E1* corresponds to substrates eluted under phosphorylation conditions (ATP/Mn^{2+})and *E2* comprises protein partners interacting through phospho-residues eluted under dephosphorylation conditions (calf intestine alkaline phosphatase).
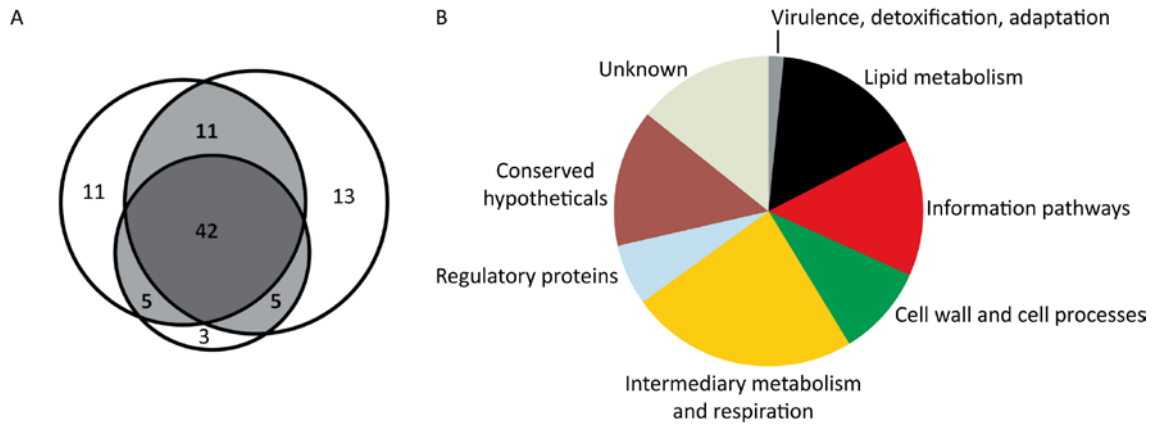
Figure 1. PknG interactome. A. Venn diagram comparing proteins identified in the three replicates of the one-step elution experiment. 69, 71 and 55 proteins were identified in each replicate. 63 proteins present in at least two of three replicates (shaded grey) represent PknG interactome and are listed in **Table S3a**.

B. Functional classification of PknG interactome. Proteins were classified using TubercuList (*M. tuberculosis* H37Rv database, march 2013 release 27 [25]) and all of them belong to one of the following categories: virulence, detoxification, adaptation (1 protein), lipid metabolism (10 proteins), information pathways (9 proteins), cell wall and cell processes (6 proteins), intermediary metabolism and respiration (15 proteins), regulatory proteins (4 proteins), conserved hypotheticals (9 proteins) and unknown (9 proteins).
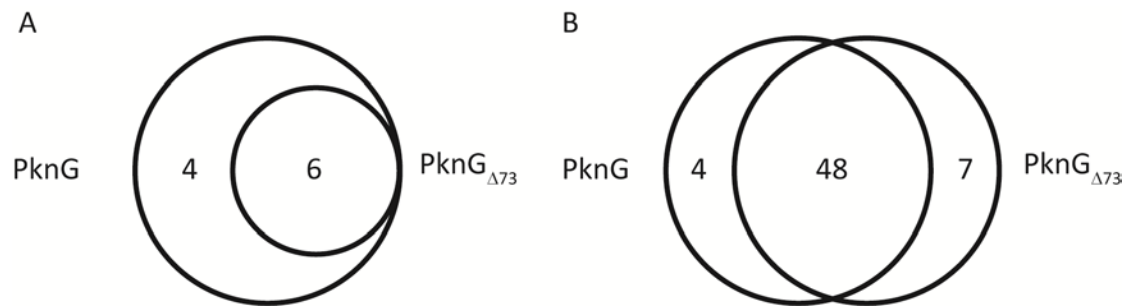
Figure 2. Venn diagrams summarizing PknG-interacting proteins recovered in sequential elution approach. A. The overlap of proteins identified in *E1* (under phosphorylation conditions) using PknG or PknG$_{\Delta73}$ as baits. Six proteins were identified in at least two replicates of both conditions and represent putative PknG substrates. Four additional proteins were only identified in *E1* when using PknG as a bait, and represent putative substrates interacting through N-terminal region of PknG. The list of the eleven putative substrates is summarized in **Table 1** and **Table S4**.

B. The overlap of proteins identified in *E2* (under dephosphorylation conditions) using PknG or PknG$_{\Delta73}$ as baits. Protein list was generated by substraction of 10 proteins nonspecifically interacting with the resin (mock). Four proteins were only recovered using full length PknG, and represent proteins that may interact with phosphoresidues in the N-terminal sequence of the kinase. These putative interactors are listed in **Table 1** and supplementary information regarding all the interactors recovered in *E2* are included in **Table S5**.
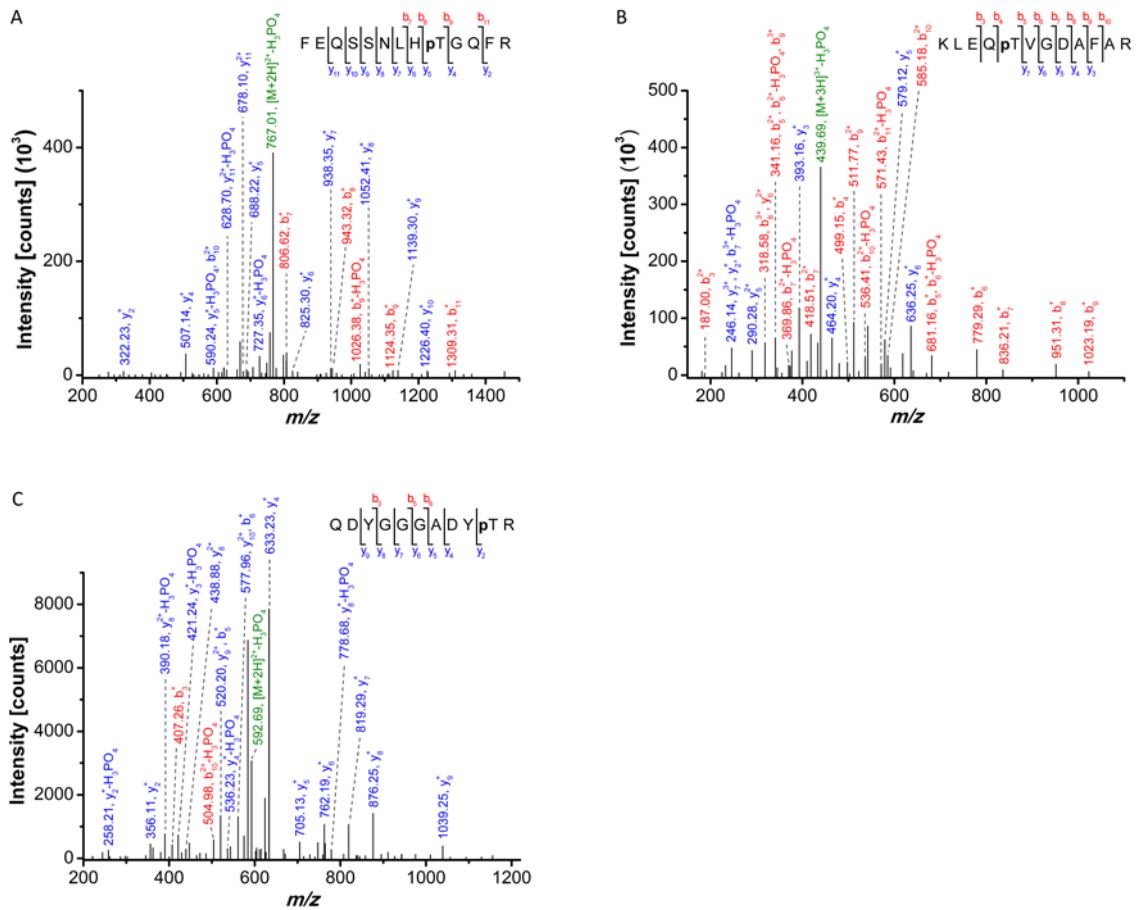
Figure 3. Identification of FhaA phosphorylation sites *in vivo*. *Strep*-tag® II-FhaA was purified from *M. smegmatis* and digested previous to mass spectrometry analysis. y- (blue) and b-(red) ions assigned to the sequence are indicated. Ions presenting phosphorylation characteristic neutral loss (green) are shown. **A.** MS/MS spectrum of peptide FEQSSNLH**p**T_{116}GQFR (MH$^+$ = 1630.90 Da). pRS score was 106 and allowed the identification of T_{116} as an *in vivo* phosphorylation site. **B.** MS/MS spectrum of peptide KLEQ**p**T_{18}VGDAFAR (MH$^+$ = 1414.32 Da). pRS score was 84 and allowed the identification of T_{18} as an *in vivo* phosphorylation site. **C.** MS/MS spectrum of peptide QDYGGGADY**p**T_{377}R (MH$^+$ = 1282.77 Da. pRS score was 69 and allowed the identification of T_{377} as an *in vivo* phosphorylation site.
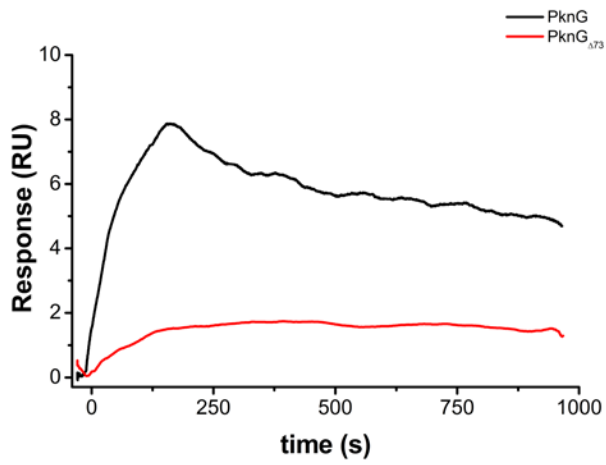
Figure 4. SPR analysis of the interaction between PknG and FhaA. FhaA was immobilized on a CM5 sensorchip and PknG or PknG$_{\Delta73}$ was injected for 180 s.
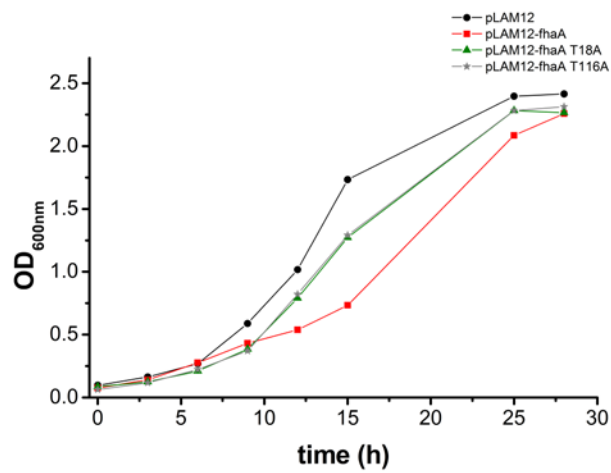
Figure 5. Overexpression of FhaA in *M. smegmatis* slows down cell growth. Overnight starter culture in Middlebrook 7H9 broth supplemented with ADS, 0.05% Tween® 80 and kanamycine was diluted to OD_{600} at 0.05 in the presence of 0.2% acetamide for pLAM12 constructs.
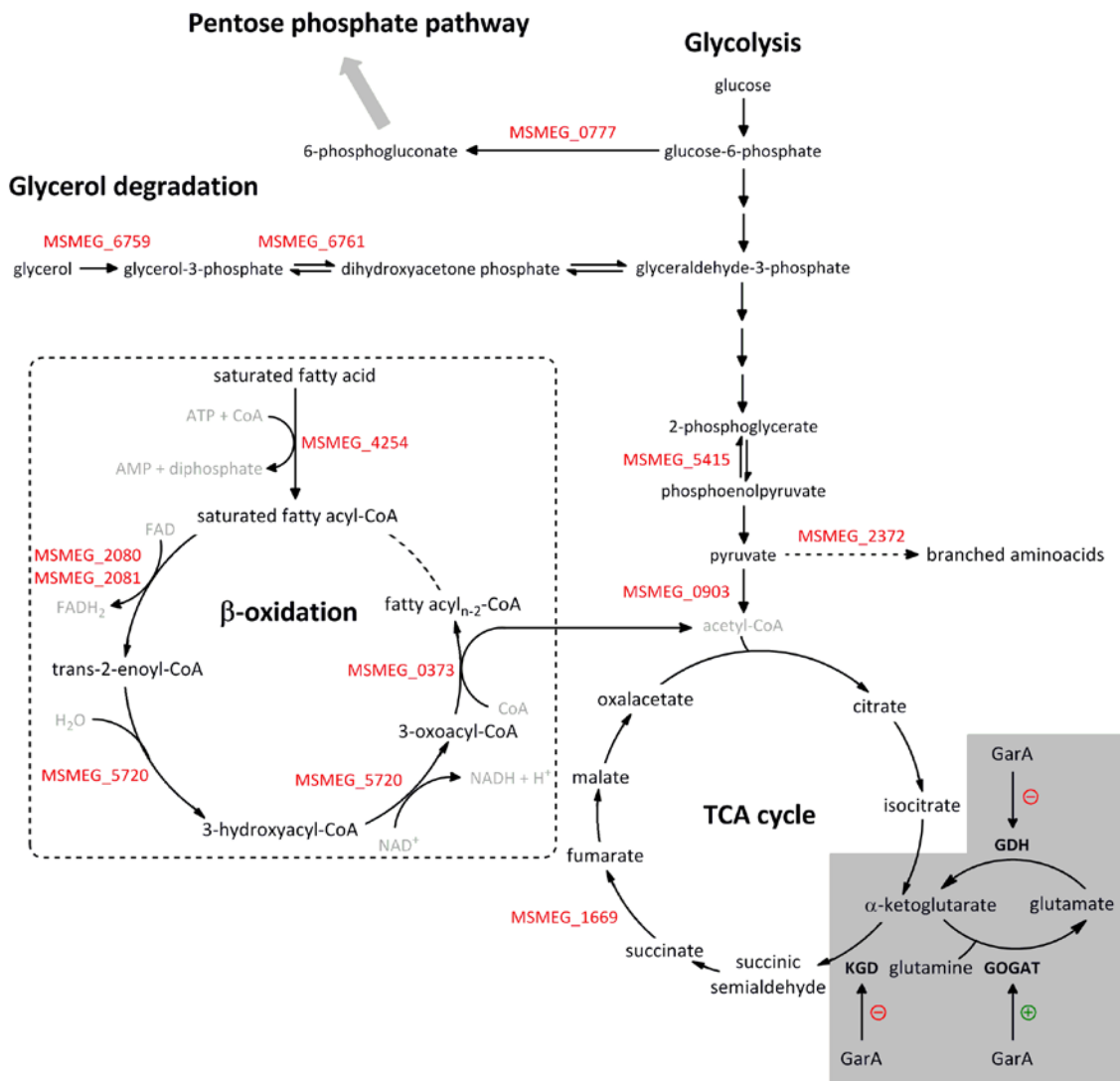
Scheme 2. Central carbon metabolism in *M. smegmatis*. Schematic representation of part of the central carbon metabolism in *M. smegmatis*, the enzymes identified as putative PknG-interactors are indicated in red. β-oxidation is not included in central carbon metabolism (dashed rectangle), however, is depicted here to show another source of acetyl-CoA. The picture was generated using BioCyc Database Collection [64] and [65, 77]. In *M. tuberculosis*, PknG indirectly regulates the glutamate metabolism via the phosphorylation of GarA. The unphosphorylated form of GarA acts in a synergic manner on enzymes that control glutamate intracellular levels, inhibiting α-ketoglutarate decarboxylase and NAD^+-specific glutamate dehydrogenase while activating α-subunit of the glutamate synthase complex (grey rectangle) [14, 18].

Scheme 3. Proposed model of interaction between PknG and FHA-containing substrates.
Upper panel: phosphorylation of the N-terminal extension of PknG is required for substrate recruitment through the C-terminal FHA domain of FhaA. Besides, the active site of PknG interacts with the phosphorylatable N-terminal domain of FhaA and catalyzes the phosphorylation of threonine residues. Unlike GarA, phosphorylation does not trigger disruption of PknG-FhaA interaction, pointing to the fact that self-recognition of pThr by its own FHA domain does not take place.

Lower panel: previously reported model of GarA interaction with PknG [14].
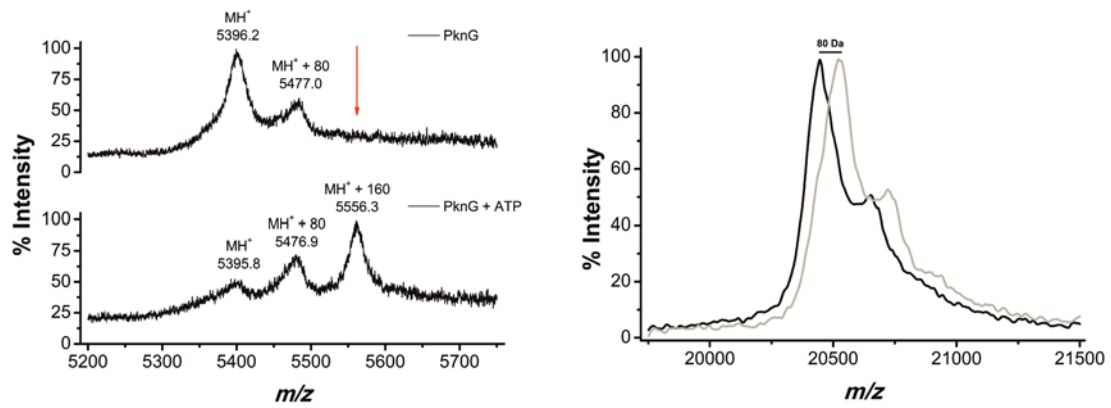
Figure S1. Kinase activity of immobilized PknG. A. PknG autophosphorylation. Linear MALDI-TOF spectra of tryptic digestion of immobilized PknG (upper panel) and immobilized PknG previously incubated with ATP under phosphorylation conditions (bottom panel). Non-phosphorylated sequence 10-60 (theoretical *m/z* 5395.7), and its mono and diphosphorylated forms are indicated (theoretical *m/z* 5475.7 and 5555.7, respectively). **B.** GarA phosphorylation by PknG. Linear MALDI-TOF mass spectrum of: recombinant GarA (*m/z* 20446, black line), GarA phosphorylated by PknG immobilized on NHS-Activated Sepharose 4 Fast Flow (*m/z* 20526, grey line). The mass shift corresponds to the incorporation of one phosphate group (80 Da).
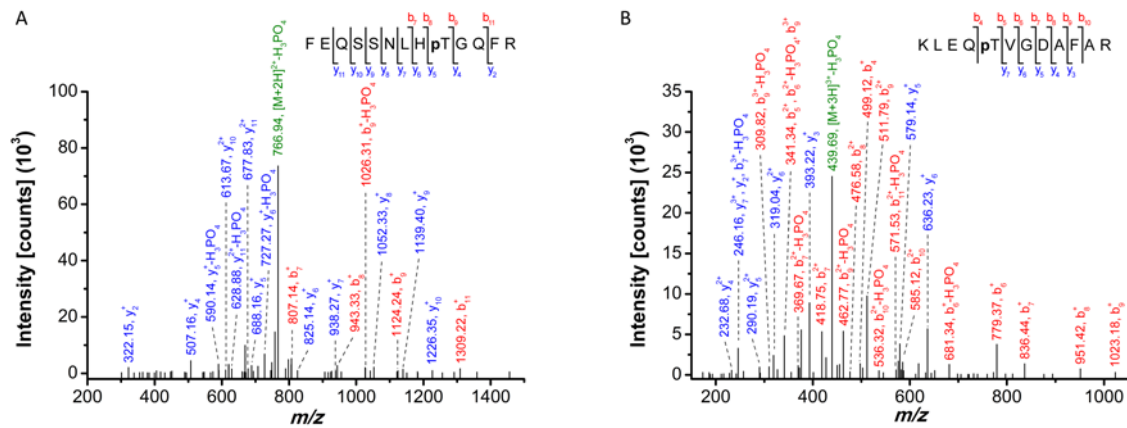
Figure S2. Identification of FhaA phosphorylation sites *in vitro*. *Strep*-tag® II-FhaA was purified from *M. smegmatis*, dephosphorylated with alkaline phosphatase, submitted to phosphorylation experiment using PknG as kinase and digested previous to mass spectrometry analysis. y- (blue) and b-(red) ions assigned to the sequence are indicated. Ions presenting phosphorylation characteristic neutral loss (green) are shown. **A.** MS/MS spectrum of peptide FEQSSNLH**p**T$_{116}$GQFR (MH$^+$ = 1630.79 Da). pRS score was 110 and allowed the identification of T$_{116}$ as phosphorylated residue. **B.** MS/MS spectrum of peptide KLEQ**p**T$_{18}$VGDAFAR (MH$^+$ = 1414.94 Da). pRS score was 79 and allowed the identification of T$_{18}$ as phosphorylated residue.

References

1   (2015) Global tuberculosis report.   ed.)^eds.), World Health Organization, Geneva, Switzerland

2   Russell, D. G., Cardona, P. J., Kim, M. J., Allain, S. and Altare, F. (2009) Foamy macrophages and the progression of the human tuberculosis granuloma. Nature immunology. **10**, 943-948

3   Agarwal, N., Lamichhane, G., Gupta, R., Nolan, S. and Bishai, W. R. (2009) Cyclic AMP intoxication of macrophages by a Mycobacterium tuberculosis adenylate cyclase. Nature. **460**, 98-102

4   Prisic, S. and Husson, R. N. (2014) Mycobacterium tuberculosis Serine/Threonine Protein Kinases. Microbiol Spectr. **2**

5   Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S., Barry, C. E., 3rd, Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, K., Osborne, J., Quail, M. A., Rajandream, M. A., Rogers, J., Rutter, S., Seeger, K., Skelton, J., Squares, R., Squares, S., Sulston, J. E., Taylor, K., Whitehead, S. and Barrell, B. G. (1998) Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. Nature. **393**, 537-544

6   Wehenkel, A., Bellinzoni, M., Grana, M., Duran, R., Villarino, A., Fernandez, P., Andre-Leroux, G., England, P., Takiff, H., Cervenansky, C., Cole, S. T. and Alzari, P. M. (2008) Mycobacterial Ser/Thr protein kinases and phosphatases: physiological roles and therapeutic potential. Biochim Biophys Acta. **1784**, 193-202

7   Av-Gay, Y. and Everett, M. (2000) The eukaryotic-like Ser/Thr protein kinases of Mycobacterium tuberculosis. Trends in microbiology. **8**, 238-244

8   Sieker, L. C., Stenkamp, R. E. and LeGall, J. (1994) Rubredoxin in crystalline state. Methods in enzymology. **243**, 203-216

9   Gil, M., Grana, M., Schopfer, F. J., Wagner, T., Denicola, A., Freeman, B. A., Alzari, P. M., Batthyany, C. and Duran, R. (2013) Inhibition of Mycobacterium tuberculosis PknG by non-catalytic rubredoxin domain specific modification: reaction of an electrophilic nitro-fatty acid with the Fe-S center. Free Radic Biol Med. **65**, 150-161

10   Lisa, M. N., Gil, M., Andre-Leroux, G., Barilone, N., Duran, R., Biondi, R. M. and Alzari, P. M. (2015) Molecular Basis of the Activity and the Regulation of the Eukaryotic-like S/T Protein Kinase PknG from Mycobacterium tuberculosis. Structure. **23**, 1039-1048

11   Scherr, N., Honnappa, S., Kunz, G., Mueller, P., Jayachandran, R., Winkler, F., Pieters, J. and Steinmetz, M. O. (2007) Structural basis for the specific inhibition of protein kinase G, a virulence factor of Mycobacterium tuberculosis. Proc Natl Acad Sci U S A. **104**, 12151-12156

12   Boitel, B., Ortiz-Lombardia, M., Duran, R., Pompeo, F., Cole, S. T., Cervenansky, C. and Alzari, P. M. (2003) PknB kinase activity is regulated by phosphorylation in two Thr residues and dephosphorylation by PstP, the cognate phospho-Ser/Thr phosphatase, in Mycobacterium tuberculosis. Mol Microbiol. **49**, 1493-1508

13   Villarino, A., Duran, R., Wehenkel, A., Fernandez, P., England, P., Brodin, P., Cole, S. T., Zimny-Arndt, U., Jungblut, P. R., Cervenansky, C. and Alzari, P. M. (2005) Proteomic identification of M. tuberculosis protein kinase substrates: PknB recruits GarA, a FHA domain-containing protein, through activation loop-mediated interactions. Journal of molecular biology. **350**, 953-963

14   O'Hare, H. M., Duran, R., Cervenansky, C., Bellinzoni, M., Wehenkel, A. M., Pritsch, O., Obal, G., Baumgartner, J., Vialaret, J., Johnsson, K. and Alzari, P. M. (2008) Regulation of glutamate metabolism by protein kinases in mycobacteria. Mol Microbiol. **70**, 1408-1423

15   Cowley, S., Ko, M., Pick, N., Chow, R., Downing, K. J., Gordhan, B. G., Betts, J. C., Mizrahi, V., Smith, D. A., Stokes, R. W. and Av-Gay, Y. (2004) The Mycobacterium tuberculosis

protein serine/threonine kinase PknG is linked to cellular glutamate/glutamine levels and is important for growth in vivo. Mol Microbiol. **52**, 1691-1702

16   Walburger, A., Koul, A., Ferrari, G., Nguyen, L., Prescianotto-Baschong, C., Huygen, K., Klebl, B., Thompson, C., Bacher, G. and Pieters, J. (2004) Protein kinase G from pathogenic mycobacteria promotes survival within macrophages. Science. **304**, 1800-1804

17   Durocher, D. and Jackson, S. P. (2002) The FHA domain. FEBS letters. **513**, 58-66

18   Nott, T. J., Kelly, G., Stach, L., Li, J., Westcott, S., Patel, D., Hunt, D. M., Howell, S., Buxton, R. S., O'Hare, H. M. and Smerdon, S. J. (2009) An intramolecular switch regulates phosphoindependent FHA domain interactions in Mycobacterium tuberculosis. Sci Signal. **2**, ra12

19   Prisic, S., Dankwa, S., Schwartz, D., Chou, M. F., Locasale, J. W., Kang, C. M., Bemis, G., Church, G. M., Steen, H. and Husson, R. N. (2010) Extensive phosphorylation with overlapping specificity by Mycobacterium tuberculosis serine/threonine protein kinases. Proc Natl Acad Sci U S A. **107**, 7521-7526

20   Fortuin, S., Tomazella, G. G., Nagaraj, N., Sampson, S. L., Gey van Pittius, N. C., Soares, N. C., Wiker, H. G., de Souza, G. A. and Warren, R. M. (2015) Phosphoproteomics analysis of a clinical Mycobacterium tuberculosis Beijing isolate: expanding the mycobacterial phosphoproteome catalog. Front Microbiol. **6**, 6

21   van Kessel, J. C., Marinelli, L. J. and Hatfull, G. F. (2008) Recombineering mycobacteria and their phages. Nat Rev Microbiol. **6**, 851-857

22   Carvalho, P. C., Fischer, J. S., Xu, T., Yates, J. R., 3rd and Barbosa, V. C. (2012) PatternLab: from mass spectra to label-free differential shotgun proteomics. Curr Protoc Bioinformatics. **Chapter 13**, Unit13 19

23   Carvalho, P. C., Lima, D. B., Leprevost, F. V., Santos, M. D., Fischer, J. S., Aquino, P. F., Moresco, J. J., Yates, J. R., 3rd and Barbosa, V. C. (2016) Integrated analysis of shotgun proteomic data with PatternLab for proteomics 4.0. Nature protocols. **11**, 102-117

24   Carvalho, P. C., Yates Iii, J. R. and Barbosa, V. C. (2010) Analyzing shotgun proteomic data with PatternLab for proteomics. Curr Protoc Bioinformatics. **Chapter 13**, Unit 13 13 11-15

25   Kapopoulou, A., Lew, J. M. and Cole, S. T. (2011) The MycoBrowser portal: a comprehensive and manually annotated resource for mycobacterial genomes. Tuberculosis (Edinb). **91**, 8-13

26   Gish, W. and States, D. J. (1993) Identification of protein coding regions by database similarity search. Nat Genet. **3**, 266-272

27   Mi, H., Muruganujan, A. and Thomas, P. D. (2013) PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. Nucleic acids research. **41**, D377-386

28   Mellacheruvu, D., Wright, Z., Couzens, A. L., Lambert, J. P., St-Denis, N. A., Li, T., Miteva, Y. V., Hauri, S., Sardiu, M. E., Low, T. Y., Halim, V. A., Bagshaw, R. D., Hubner, N. C., Al-Hakim, A., Bouchard, A., Faubert, D., Fermin, D., Dunham, W. H., Goudreault, M., Lin, Z. Y., Badillo, B. G., Pawson, T., Durocher, D., Coulombe, B., Aebersold, R., Superti-Furga, G., Colinge, J., Heck, A. J., Choi, H., Gstaiger, M., Mohammed, S., Cristea, I. M., Bennett, K. L., Washburn, M. P., Raught, B., Ewing, R. M., Gingras, A. C. and Nesvizhskii, A. I. (2013) The CRAPome: a contaminant repository for affinity purification-mass spectrometry data. Nat Methods. **10**, 730-736

29   Mao, F., Dam, P., Chou, J., Olman, V. and Xu, Y. (2009) DOOR: a database for prokaryotic operons. Nucleic acids research. **37**, D459-463

30   Wells, J. N., Bergendahl, L. T. and Marsh, J. A. (2016) Operon Gene Order Is Optimized for Ordered Protein Complex Assembly. Cell Rep

31   Wipperman, M. F., Yang, M., Thomas, S. T. and Sampson, N. S. (2013) Shrinking the FadE proteome of Mycobacterium tuberculosis: insights into cholesterol metabolism through identification of an alpha2beta2 heterotetrameric acyl coenzyme A dehydrogenase family. J Bacteriol. **195**, 4331-4341

32   Veyron-Churlet, R., Guerrini, O., Mourey, L., Daffe, M. and Zerbib, D. (2004) Protein-protein interactions within the Fatty Acid Synthase-II system of Mycobacterium tuberculosis are essential for mycobacterial viability. Mol Microbiol. **54**, 1161-1172

33   Lu, P., Lill, H. and Bald, D. (2014) ATP synthase in mycobacteria: special features and implications for a function as drug target. Biochim Biophys Acta. **1837**, 1208-1218

34   Nakedi, K. C., Nel, A. J., Garnett, S., Blackburn, J. M. and Soares, N. C. (2015) Comparative Ser/Thr/Tyr phosphoproteomics between two mycobacterial species: the fast growing Mycobacterium smegmatis and the slow growing Mycobacterium bovis BCG. Front Microbiol. **6**, 237

35   Kusebauch, U., Ortega, C., Ollodart, A., Rogers, R. S., Sherman, D. R., Moritz, R. L. and Grundner, C. (2014) Mycobacterium tuberculosis supports protein tyrosine phosphorylation. Proc Natl Acad Sci U S A. **111**, 9265-9270

36   Calder, B., Albeldas, C., Blackburn, J. M. and Soares, N. C. (2016) Mass Spectrometry Offers Insight into the Role of Ser/Thr/Tyr Phosphorylation in the Mycobacteria. Front Microbiol. **7**, 141

37   Cousin, C., Derouiche, A., Shi, L., Pagot, Y., Poncet, S. and Mijakovic, I. (2013) Protein-serine/threonine/tyrosine kinases in bacterial signaling and regulation. FEMS Microbiol Lett. **346**, 11-19

38   Barthe, P., Roumestand, C., Canova, M. J., Kremer, L., Hurard, C., Molle, V. and Cohen-Gonsaud, M. (2009) Dynamic and structural characterization of a bacterial FHA protein reveals a new autoinhibition mechanism. Structure. **17**, 568-578

39   Mishra, A., Vij, M., Kumar, D., Taneja, V., Mondal, A. K., Bothra, A., Rao, V., Ganguli, M. and Taneja, B. (2013) Integration host factor of Mycobacterium tuberculosis, mIHF, compacts DNA by a bending mechanism. PLoS One. **8**, e69985

40   Krajewski, W. W., Jones, T. A. and Mowbray, S. L. (2005) Structure of Mycobacterium tuberculosis glutamine synthetase in complex with a transition-state mimic provides functional insights. Proc Natl Acad Sci U S A. **102**, 10499-10504

41   Ouellet, H., Johnston, J. B. and Ortiz de Montellano, P. R. (2010) The Mycobacterium tuberculosis cytochrome P450 system. Archives of biochemistry and biophysics. **493**, 82-95

42   Ricagno, S., de Rosa, M., Aliverti, A., Zanetti, G. and Bolognesi, M. (2007) The crystal structure of FdxA, a 7Fe ferredoxin from Mycobacterium smegmatis. Biochemical and biophysical research communications. **360**, 97-102

43   Fay, A. and Glickman, M. S. (2014) An essential nonredundant role for mycobacterial DnaK in native protein folding. PLoS genetics. **10**, e1004516

44   Ojha, A., Anand, M., Bhatt, A., Kremer, L., Jacobs, W. R., Jr. and Hatfull, G. F. (2005) GroEL1: a dedicated chaperone involved in mycolic acid biosynthesis during biofilm formation in mycobacteria. Cell. **123**, 861-873

45   Xu, W. X., Zhang, L., Mai, J. T., Peng, R. C., Yang, E. Z., Peng, C. and Wang, H. H. (2014) The Wag31 protein interacts with AccA3 and coordinates cell wall lipid permeability and lipophilic drug resistance in Mycobacterium smegmatis. Biochemical and biophysical research communications. **448**, 255-260

46   Reid, M. F. and Fewson, C. A. (1994) Molecular characterization of microbial alcohol dehydrogenases. Critical reviews in microbiology. **20**, 13-56

47   Dephoure, N., Gould, K. L., Gygi, S. P. and Kellogg, D. R. (2013) Mapping and analysis of phosphorylation sites: a quick guide for cell biologists. Molecular biology of the cell. **24**, 535-542

48   Mann, M., Ong, S. E., Gronborg, M., Steen, H., Jensen, O. N. and Pandey, A. (2002) Analysis of protein phosphorylation using mass spectrometry: deciphering the phosphoproteome. Trends in biotechnology. **20**, 261-268

49   Roumestand, C., Leiba, J., Galophe, N., Margeat, E., Padilla, A., Bessin, Y., Barthe, P., Molle, V. and Cohen-Gonsaud, M. (2011) Structural insight into the Mycobacterium

tuberculosis Rv0020c protein and its interaction with the PknB kinase. Structure. **19**, 1525-1534

50   Wolff, K. A., de la Pena, A. H., Nguyen, H. T., Pham, T. H., Amzel, L. M., Gabelli, S. B. and Nguyen, L. (2015) A Redox Regulatory System Critical for Mycobacterial Survival in Macrophages and Biofilm Development. PLoS Pathog. **11**, e1004839

51   Blobel, G. (1971) Isolation of a 5S RNA-protein complex from mammalian ribosomes. Proc Natl Acad Sci U S A. **68**, 1881-1885

52   Grundner, C., Gay, L. M. and Alber, T. (2005) Mycobacterium tuberculosis serine/threonine kinases PknB, PknD, PknE, and PknF phosphorylate multiple FHA domains. Protein Sci. **14**, 1918-1921

53   Gupta, M., Sajid, A., Arora, G., Tandon, V. and Singh, Y. (2009) Forkhead-associated domain-containing protein Rv0019c and polyketide-associated protein PapA5, from substrates of serine/threonine protein kinase PknB to interacting proteins of Mycobacterium tuberculosis. J Biol Chem. **284**, 34723-34734

54   Molle, V., Soulat, D., Jault, J. M., Grangeasse, C., Cozzone, A. J. and Prost, J. F. (2004) Two FHA domains on an ABC transporter, Rv1747, mediate its phosphorylation by PknF, a Ser/Thr protein kinase from Mycobacterium tuberculosis. FEMS Microbiol Lett. **234**, 215-223

55   Sharma, K., Gupta, M., Krupa, A., Srinivasan, N. and Singh, Y. (2006) EmbR, a regulatory protein with ATPase activity, is a substrate of multiple serine/threonine kinases and phosphatase in Mycobacterium tuberculosis. The FEBS journal. **273**, 2711-2721

56   Madigan, M. T., Martinko, J. M. and Parker, J. (1998) Brock Biología de los microorganismos. Prentice Hall Iberia, Madrid

57   Gee, C. L., Papavinasasundaram, K. G., Blair, S. R., Baer, C. E., Falick, A. M., King, D. S., Griffin, J. E., Venghatakrishnan, H., Zukauskas, A., Wei, J. R., Dhiman, R. K., Crick, D. C., Rubin, E. J., Sassetti, C. M. and Alber, T. (2012) A phosphorylated pseudokinase complex controls cell wall synthesis in mycobacteria. Sci Signal. **5**, ra7

58   Srihari, S., Yong, C. H., Patil, A. and Wong, L. (2015) Methods for protein complex prediction and their contributions towards understanding the organisation, function and dynamics of complexes. FEBS letters. **589**, 2590-2602

59   Dunham, W. H., Mullin, M. and Gingras, A. C. (2012) Affinity-purification coupled to mass spectrometry: basic principles and strategies. Proteomics. **12**, 1576-1590

60   Piazzi, M., Blalock, W. L., Bavelloni, A., Faenza, I., D'Angelo, A., Maraldi, N. M. and Cocco, L. (2013) Phosphoinositide-specific phospholipase C beta 1b (PI-PLCbeta1b) interactome: affinity purification-mass spectrometry analysis of PI-PLCbeta1b with nuclear protein. Mol Cell Proteomics. **12**, 2220-2235

61   Oeffinger, M. (2012) Two steps forward--one step back: advances in affinity purification mass spectrometry of macromolecular complexes. Proteomics. **12**, 1591-1608

62   Trinkle-Mulcahy, L., Boulon, S., Lam, Y. W., Urcia, R., Boisvert, F. M., Vandermoere, F., Morrice, N. A., Swift, S., Rothbauer, U., Leonhardt, H. and Lamond, A. (2008) Identifying specific protein interaction partners using quantitative mass spectrometry and bead proteomes. The Journal of cell biology. **183**, 223-239

63   Deng, J., Bi, L., Zhou, L., Guo, S. J., Fleming, J., Jiang, H. W., Zhou, Y., Gu, J., Zhong, Q., Wang, Z. X., Liu, Z., Deng, R. P., Gao, J., Chen, T., Li, W., Wang, J. F., Wang, X., Li, H., Ge, F., Zhu, G., Zhang, H. N., Gu, J., Wu, F. L., Zhang, Z., Wang, D., Hang, H., Li, Y., Cheng, L., He, X., Tao, S. C. and Zhang, X. E. (2014) Mycobacterium tuberculosis proteome microarray for global studies of protein function and immunogenicity. Cell Rep. **9**, 2317-2329

64   Caspi, R., Altman, T., Billington, R., Dreher, K., Foerster, H., Fulcher, C. A., Holland, T. A., Keseler, I. M., Kothari, A., Kubo, A., Krummenacker, M., Latendresse, M., Mueller, L. A., Ong, Q., Paley, S., Subhraveti, P., Weaver, D. S., Weerasinghe, D., Zhang, P. and Karp, P. D. (2014) The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. Nucleic acids research. **42**, D459-471

65   Baughn, A. D. and Rhee, K. Y. (2014) Metabolomics of Central Carbon Metabolism in Mycobacterium tuberculosis. Microbiol Spectr. **2**

66   Brunori, L., Giannoni, F., Bini, L., Liberatori, S., Frota, C., Jenner, P., Thoresen, O. F., Orefici, G. and Fattorini, L. (2004) Induction of Mycobacterium avium proteins upon infection of human macrophages. Proteomics. **4**, 3078-3083

67   Takayama, K., Wang, C. and Besra, G. S. (2005) Pathway to synthesis and processing of mycolic acids in Mycobacterium tuberculosis. Clinical microbiology reviews. **18**, 81-101

68   Forrellad, M. A., Klepp, L. I., Gioffre, A., Sabio y Garcia, J., Morbidoni, H. R., de la Paz Santangelo, M., Cataldi, A. A. and Bigi, F. (2013) Virulence factors of the Mycobacterium tuberculosis complex. Virulence. **4**, 3-66

69   Braibant, M., Gilot, P. and Content, J. (2000) The ATP binding cassette (ABC) transport systems of Mycobacterium tuberculosis. FEMS microbiology reviews. **24**, 449-467

70   Esposito, C., Marasco, D., Delogu, G., Pedone, E. and Berisio, R. (2011) Heparin-binding hemagglutinin HBHA from Mycobacterium tuberculosis affects actin polymerisation. Biochemical and biophysical research communications. **410**, 339-344

71   Esposito, C., Cantisani, M., D'Auria, G., Falcigno, L., Pedone, E., Galdiero, S. and Berisio, R. (2012) Mapping key interactions in the dimerization process of HBHA from Mycobacterium tuberculosis, insights into bacterial agglutination. FEBS letters. **586**, 659-667

72   Tullius, M. V., Harth, G. and Horwitz, M. A. (2003) Glutamine synthetase GlnA1 is essential for growth of Mycobacterium tuberculosis in human THP-1 macrophages and guinea pigs. Infect Immun. **71**, 3927-3936

73   Gouzy, A., Poquet, Y. and Neyrolles, O. (2014) Nitrogen metabolism in Mycobacterium tuberculosis physiology and virulence. Nat Rev Microbiol. **12**, 729-737

74   Stadtman, E. R. (2001) The story of glutamine synthetase regulation. J Biol Chem. **276**, 44357-44364

75   Leigh, J. A. and Dodsworth, J. A. (2007) Nitrogen regulation in bacteria and archaea. Annual review of microbiology. **61**, 349-377

76   Fernandez, P., Saint-Joanis, B., Barilone, N., Jackson, M., Gicquel, B., Cole, S. T. and Alzari, P. M. (2006) The Ser/Thr protein kinase PknB is essential for sustaining mycobacterial growth. J Bacteriol. **188**, 7778-7784

77   Rhee, K. Y., de Carvalho, L. P., Bryk, R., Ehrt, S., Marrero, J., Park, S. W., Schnappinger, D., Venugopal, A. and Nathan, C. (2011) Central carbon metabolism in Mycobacterium tuberculosis: an unexpected frontier. Trends in microbiology. **19**, 307-314

Análisis del interactoma de PknG en la micobacteria y validación de los interactores y/o sustratos biológicamente relevantes

En los últimos años se han reportado fosfoproteomas de algunas especies del género *Mycobacterium* realizados en distintas situaciones metabólicas. Estos estudios dan cuenta de la enorme variedad de proteínas fosforiladas en residuos hidroxilados que presentan estos organismos [46, 47, 51, 52]. No obstante, en la actualidad se desconoce la identidad de las quinasas responsables de la mayoría de los eventos de fosforilación reportados en tales estudios. Es por ello que, sabiendo que PknG de *M. tuberculosis* es una enzima esencial para la infección y que solamente se ha identificado un sustrato de esta quinasa, nos propusimos un abordaje proteómico del tipo purificación por afinidad seguida de espectrometría de masa (AP-MS) dirigido a la identificación de nuevos sustratos e interactores de PknG en la micobacteria.

Para revelar nuevos interactores proteicos de PknG de *M. tuberculosis*, se inmovilizó la enzima recombinante en un soporte sólido y se incubó con un extracto citosólico de *M. smegmatis* MC^{2}155, utilizándose una resina blanco como control de unión inespecífica [145]. Las proteínas retenidas en la matriz fueron eluídas utilizando un protocolo estándar para la disrupción de interacciones hidrofóbicas. En la siguiente etapa, estas proteínas se analizaron por espectrometría de masa de tipo trampa iónica de alta sensibilidad acoplada a un nano-LC y usando una aproximación de tipo *shotgun* (nano Easy-HPLC/LTQ-Velos, Proxeon, Thermo). Finalmente, se realizó el análisis bioinformático de los espectros de masa en tándem resultantes del experimento de EM con PatternLab for proteomics [146].

A partir de tres réplicas biológicas se generó una lista de 63 interactores putativos de PknG (Tabla S3A en artículo 3, las tablas suplementarias se encuentran en el Anexo). Las proteínas identificadas fueron agrupadas de acuerdo a su categoría funcional y se determinó que los grupos mayoritarios estaban constituidos por proteínas relacionadas al metabolismo intermediario y la respiración, al metabolismo lipídico, a las vías de la información y a la pared celular y procesos celulares (Fig. 1B en artículo 3). La información que se desprende de este interactoma es sumamente valiosa, sin embargo, el abordaje experimental utilizado no permite discriminar entre interactores directos e indirectos de PknG. En efecto, algunas de las proteínas presentes en este interactoma ya habían sido reportadas formando parte de complejos o como integrantes del mismo operón (Tabla S3B en artículo 3).

GarA, el único sustrato bien caracterizado de PknG [93], no se encontró entre las proteínas identificadas como interactores putativos de la quinasa. Esta observación nos llevó a desarrollar una

estrategia secuencial de elución de sustratos y/o interactores basada en la particular combinación de dominios de PknG (esquema 1 en artículo 3). Esta aproximación experimental fue diseñada para permitir el aislamiento de sustratos o interactores dependientes de fosforilación del resto de las proteínas interactoras. Además del interactoma de PknG, evaluamos el de una forma trunca de la quinasa (PknG$_{\Delta 73}$) a la que le falta la región N-terminal involucrada en el reclutamiento de GarA.

Someramente, se incubó la resina con PknG o PknG$_{\Delta 73}$ inmovilizada con el extracto de *M. smegmatis* en condiciones favorables a la formación de complejos proteína-proteína. La primera elución se realizó en condiciones de fosforilación (*E1*), de forma tal de recuperar sustratos de la quinasa que al igual que GarA se liberaran una vez fosforilados. Resultados previos de nuestro grupo demostraron que los sitios de autofosforilación localizados en la extensión N-terminal de PknG reclutan a GarA a través de su dominio FHA y que esta interacción desaparece una vez que GarA es fosforilada [93]. La comparación de las proteínas recuperadas usando una u otra construcción de PknG nos permitió discernir entre sustratos que requieran o no del sitio de anclaje aminoterminal. En la segunda etapa utilizamos condiciones de desfosforilación (*E2*) para abolir las interacciones mediadas por fosforresiduos. Al igual que en *E1*, la comparación de los interactores recuperados utilizando las dos construcciones de PknG nos permitió identificar cuáles son las proteínas que reconocen los sitios de autofosforilación de PknG.

A partir de los datos de *E1* se identificaron diez posibles sustratos de PknG y solamente seis de ellos son asimismo sustratos de PknG$_{\Delta 73}$ (Tabla S4 en artículo 3). Más aún, cinco de estas diez proteínas habían sido ya reportadas como fosfoproteínas en estudios fosfoproteómicos (Tabla S4C en artículo 3). La comparación de la lista de sustratos identificados, utilizando PknG o PknG$_{\Delta 73}$ como cebo, permitió establecer que la interacción de cuatro de ellos ocurre a través del extremo N-terminal de la quinasa (Fig. 2A, Tabla 1 y Tabla S4A en artículo 3). De acuerdo a lo esperado, GarA pertenece a este grupo de sustratos. La presencia de GarA no solo valida nuestro abordaje experimental sino que respalda la lista de sustratos recuperados.

Los tres sustratos adicionales que requieren de la presencia del extremo N-terminal para interaccionar con PknG son: el factor de integración al hospedero, la subunidad beta de la *electron transfer flavoprotein* y la glutamina sintetasa 1. Estas proteínas no poseen dominios de reconocimiento de fosforresiduos por lo que aún desconocemos cuáles son los patrones moleculares que participan en la interacción con PknG.

Para el caso de *E2*, cuatro proteínas se recuperan sistemáticamente como interactores de PknG pero están ausentes cuando se utiliza la forma trunca como cebo. Estos resultados indican que la interacción de estas proteínas con PknG está mediada por los fosforresiduos de la secuencia N-terminal de la quinasa (Tabla 1 en artículo 3). Entre las proteínas recuperadas se encuentra una que, al igual que GarA, contiene un dominio FHA. Esta proteína se denomina FhaA en *M. tuberculosis* y se ha postulado que participa en el mantenimiento de la morfología celular [86].

Un punto a destacar es la presencia sistemática de FhaA como interactor putativo de PknG independientemente de qué método de elución se utilice. Esta proteína fue reportada como sustrato *in vitro* de PknB y aparece profusamente fosforilada en los fosfoproteomas micobacterianos [55, 84]. Teniendo en cuenta lo anterior nos planteamos entonces evaluar a FhaA como sustrato de PknG. Nuestra hipótesis de trabajo es que FhaA podría ser un sustrato de PknG que, a diferencia de GarA, una vez fosforilado permanecería unido a la extensión N-terminal de la quinasa. Lo anterior explicaría su recuperación en *E2* y no en *E1*.

Para evaluar la interacción entre estas proteínas purificamos la proteína FhaA de *M. tuberculosis* a partir de una cepa de *M. smegmatis* y analizamos si ocurría fosforilación *in vitro* catalizada por PknG. Digestiones trípticas de FhaA, previamente expuesta a PknG en condiciones de fosforilación, permitieron la identificación de manera estadísticamente significativa de dos fosfopéptidos de FhaA. A partir de datos de MS/MS se pudo identificar inequívocamente a la T_{18} y a la T_{116} como residuos fosforilados (Fig. S2 en artículo 3). Posteriormente se analizaron los sitios de fosforilación de FhaA *in vivo* y se hallaron tres secuencias modificadas; dos de ellas correspondían a las secuencias identificadas en el ensayo *in vitro* mientras que la tercera solamente fue detectada *in vivo*. De los residuos fosforilados, dos coincidían con los ya identificados en el caso de la fosforilación *in vitro* en tanto que la T_{377} es exclusiva del experimento *in vivo* (Fig. 3 en artículo 3). La presencia de los mismos residuos fosforilados tanto *in vitro* como *in vivo* señala a PknG como una posible quinasa responsable de la fosforilación de FhaA. Sin embargo, el sitio extra (T_{377}) apunta a que *in vivo* otras quinasas podrían estar también regulando el estado de fosforilación de FhaA.

En una etapa posterior se caracterizaron las cepas de *M. smegmatis* que habían sido transformadas con los vectores conteniendo el gen fhaA y el plásmido control. Se observó que una vez inducida la expresión de dicho gen la velocidad de crecimiento de la micobacteria se veía seriamente disminuida. Mientras que las fases de crecimiento exponencial y estacionaria prácticamente no se ven afectadas por la sobreexpresión de la proteína FhaA, la fase de adaptación se extiende de forma marcada. Más adelante evaluamos si la fosforilación de FhaA tenía algún efecto sobre el crecimiento bacteriano.

Para ello preparamos mutantes puntuales de las treoninas identificadas como blanco de PknG ($T_{18}A$ y $T_{116}A$). Se realizaron las curvas de crecimiento y se observó que la fase de adaptación presenta una duración intermedia respecto a la de la cepa transformada con el plásmido vacío y a la que sobreexpresa FhaA (Fig. 5 en artículo 3). Estos resultados apuntan a un rol de la fosforilación de FhaA en el crecimiento bacteriano.

Para empezar a comprender a qué se debía la disminución de la velocidad de crecimiento se realizó la observación de las bacterias por criomicroscopía electrónica de transmisión. Estos experimentos se efectuaron en colaboración con la Dra. Annemarie Wehenkel de la Unité de Microbiologie Structurale del Institut Pasteur (París, Francia) y el Dr. Sylvain Trépout del Institut Curie (Orsay, Francia). Se observó que la sobrexpresión de la proteína FhaA provoca defectos en la morfología celular que se indican con flechas en la Figura 15. El análisis manual de las imágenes nos permitió corroborar la presencia sistemática de bacterias engrosadas en los polos y/o bacilos "doblados". Sin embargo la evaluación de estos fenotipos a nivel poblacional para obtener datos estadísticamente significativos fue difícil debido a que al momento del análisis no existían algoritmos capaces de cuantificar este tipo de deformaciones. Para resolver esto establecimos una colaboración con el Dr. Federico Lecumberry y el MSc. Matías Tailanián del Laboratorio de Procesamiento de Señales (Facultad de Ingeniería - Institut Pasteur de Montevideo) con el objetivo de desarrollar un software dirigido a la segmentación semiautomática de esta clase de imágenes. La primera versión de este software está actualmente disponible y nos permitirá el análisis de la deformación de las células que sobreexpresan FhaA de manera semiautomática para evitar la subjetividad del operador, así como la determinación de parámetros celulares de interés.

M. smegmatis transformada con pLAM12 | *M. smegmatis* transformada con pLAM12-fhaA
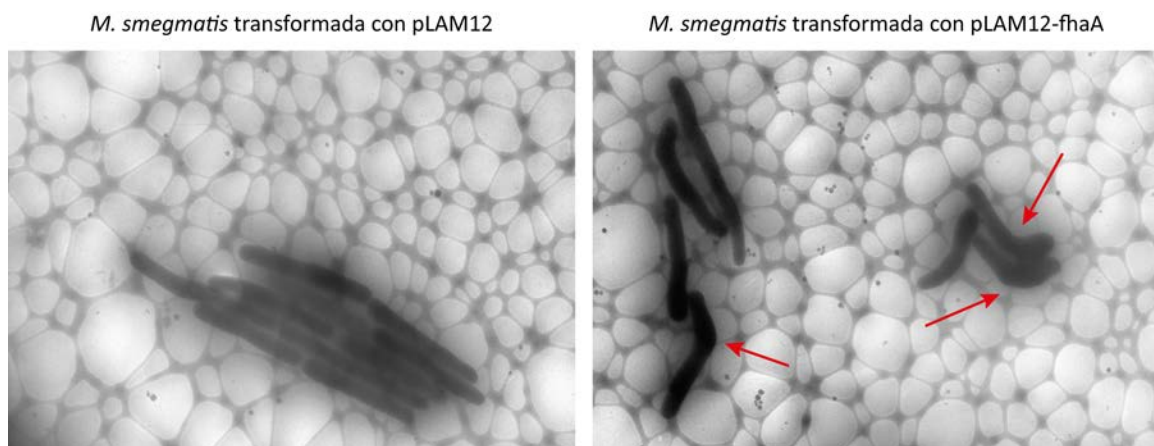
Figura 15. Imágenes de criomicroscopía electrónica de transmisión de *M. smegmatis* MC^{2}155. Izquierda: cepa de *M. smegmatis* transformada con el plásmido pLAM12. Derecha: cepa de *M. smegmatis* transformada con el plásmido pLAM12-fhaA. Se inició un cultivo líquido en 7H9 suplementado con ADC, Tween® 80, kanamicina a una OD_{600nm} de 0.05 y se indujo la expresión de FhaA con acetamida 0.2%. Las células se colectaron luego de 18 horas (pLAM12 OD_{600nm} = 0.84 y pLAM12-fhaA OD_{600nm} = 0.45) y se congelaron en presencia de bolitas de oro.

Las imágenes fueron adquiridas en un microscopio electrónico de transmisión JEOL-2200FS. Las condiciones de cultivo y adquisición de imágenes fueron las mismas para ambas condiciones.

Tomados en conjunto nuestros datos sugieren que las vías de señalización mediadas por PknG afectarían una amplia variedad de procesos celulares, lo que condice con el hecho de que las quinasas de proteínas en serinas y treoninas bacterianas funcionan como nodos de integración de señales y coordinación de diferentes procesos celulares [147]. Los resultados presentados aquí sugieren que PknG podría regular el metabolismo micobacteriano actuando en la encrucijada entre el metabolismo del carbono y del nitrógeno y la producción de energía (esquema 2 en artículo 3). Más aún, postulamos que la fosforilación mediada por PknG constituye un nuevo mecanismo de control de la glutamina sintetasa y que PknG sería la enzima encargada de la fina regulación tendiente a mantener el balance entre la síntesis y la degradación de intermediarios nitrogenados y por tanto asegurar la supervivencia en condiciones de deficiencia de nutrientes como las que encuentra la bacteria en el hospedero (Fig. 16).
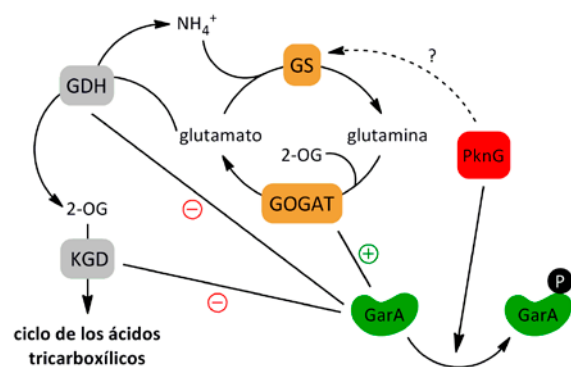
Figura 16. Regulación del metabolismo micobacteriano del nitrógeno por PknG. La glutamina sintetasa (GS) es la enzima responsable de la condensación de amonio y L-glutamato para formar L-glutamina. La glutamina es utilizada, junto al 2-oxoglutarato (2-OG), por GOGAT para sintetizar glutamato. Por otro lado, la glutamato deshidrogenasa (GDH) cataliza la desaminación oxidativa del glutamato, dando como producto 2-oxoglutarato que puede entrar en el ciclo de Krebs. Distintos mecanismos regulan el balance entre la asimilación de nitrógeno y la degradación de glutamato. En condiciones de exceso de nitrógeno, una adenililtransferasa (GlnE) transfiere un AMP a un residuo de Tyr específico de la GS e inhibe su actividad. Por otro lado, en condiciones de privación de nitrógeno, GlnE cataliza la desadenililación de GS. Simultáneamente, PknG regula esta vía actuando a través de la fosforilación de GarA. GarA desfosforilada interactúa con varias enzimas involucradas en esta vía regulando la actividad de las mismas. Por un lado, inhibe a la GDH y KGD y por otro activa a GOGAT. La fosforilación de GarA por PknG lleva a la supresión de las interacciones moleculares con estas enzimas y los efectos sobre la actividad de las mismas. Los resultados de nuestro interactoma apuntan a que PknG también regularía a la GS. Adaptado de [148, 149].

Por otra parte, proponemos un modelo de interacción entre PknG y FhaA que involucra dos regiones de cada una de las proteínas e hipotetizamos que FhaA sería una proteína de andamiaje que mediaría la asociación dependiente de fosforilación de nuevas proteínas (esquema 3 en artículo 3). A fin de validar FhaA como sustrato fisiológico de PknG y de identificar las interacciones mediadas por FhaA, nuestro grupo de investigación está llevando adelante estudios de proteómica comparativa utilizando una cepa deficiente en *pknG* así como experimentos de entrecruzamiento *in vivo*.

IV. CONCLUSIONES Y PERSPECTIVAS

Esta tesis se centra en el estudio de PknG, una quinasa de proteínas que cumple un rol fundamental en la virulencia de *Mycobacterium tuberculosis* posibilitando la supervivencia de la bacteria en el hospedero. Si bien el papel de PknG en la inhibición de la fusión fagolisosomal ha sido reportado hace ya más de diez años, todavía no han logrado esclarecerse los mecanismos que regulan la actividad de esta enzima, las señales que la misma detecta (sea en la bacteria o en el hospedero) y los mecanismos moleculares subyacentes a su acción. Los resultados obtenidos en este trabajo nos han permitido comenzar a responder a algunas de estas interrogantes acerca de PknG.

Por un lado, en este trabajo describimos la modulación de la actividad quinasa de PknG por el dominio rubredoxina extracatalítico. Esta proteína multidominio, con características estructurales muy particulares, no es capaz de regular su actividad por autofosforilación del bucle de activación tal como ha sido reportado para otras STPKs micobacterianas [63, 66, 119]. Por otro lado, la ocurrencia simultánea de los dominios quinasa y rubredoxina es una peculiaridad que se restringe a las quinasas tipo PknG de unos pocos *Actinomycetales*. En nuestro trabajo demostramos que la modificación covalente de las cisteínas del dominio rubredoxina por un ácido graso nitrado provoca la inhibición de la actividad quinasa de PknG. Asimismo establecimos que esta pérdida de actividad se debe a la salida del hierro de su posición de coordinación aunque sin alteración de la estructura global de la quinasa. En el mismo sentido probamos que la ausencia del dominio Rbx da origen a una quinasa con mayor actividad frente a su sustrato, lo que claramente indica que la presencia de este dominio inhibe la actividad catalítica de PknG. Los datos estructurales generados en este trabajo nos permitieron postular que el movimiento tipo bisagra del dominio rubredoxina, respecto al dominio catalítico, modula el acceso del sustrato al sitio activo.

Tomando en consideración los resultados de ambos trabajos [150, 151] podemos plantear que **el dominio rubredoxina es un regulador de la actividad quinasa de PknG**. Una pregunta que surge es cuál(es) son las señales que provocan la alternancia entre la conformación "activa" e "inactiva" de PknG. Una posible hipótesis es que el movimiento del Rbx sea gatillado por la unión de pequeñas moléculas o incluso proteínas a regiones próximas o no a este dominio. Para validar nuestra hipótesis sería necesario evaluar la actividad catalítica de PknG en presencia de distintos mediadores del metabolismo intermediario de la bacteria, la selección de estas moléculas se sustenta en el rol regulador que ejerce PknG sobre el metabolismo energético y del nitrógeno.

La utilización del ácido nitro-oleico nos proporcionó información a dos niveles, por una parte pudimos comenzar a elucidar el rol del dominio Rbx y por otra empezamos a caracterizar un posible inhibidor de PknG con potenciales fines terapéuticos. Ya mencionamos que el plegamiento global y la

maquinaria catalítica están muy conservados entre las STPKs eucariotas y procariotas [30]. Esto provoca cierta inespecificidad en la mayoría de los inhibidores de quinasas que se dirigen al sitio activo. Por el contrario, el OA-NO$_2$ modifica específicamente un dominio no catalítico y esto es lo que origina la pérdida de actividad. Si bien estas moléculas constituyen candidatos interesantes como inhibidores específicos de PknG con potencial acción sobre las micobacterias, debe tomarse en consideración que el OA-NO$_2$ es una molécula electrofílica capaz de reaccionar con Cys e His en múltiples proteínas blanco [135]. Lo último condiciona el potencial del OA-NO$_2$ como droga antituberculosa.

Para sortear la dificultad mencionada en el párrafo previo es preciso dirigir el fármaco a las células infectadas por *M. tuberculosis*. En este sentido, en nuestro laboratorio está en curso el desarrollo de una nueva generación de compuestos híbridos miméticos del α-tocoferol a los que se les adicionó el grupo funcional nitroalqueno. La estrategia farmacológica que plantea el grupo dirigido por el Dr. Carlos Batthyány se basa en la utilización de lipoproteínas como transportadores del compuesto híbrido. Lo anterior se sustenta en que el α-tocoferol es reconocido selectivamente por la proteína de transferencia de α-tocoferol en el hígado que carga a las LDL de la vitamina durante su metabolismo [152]. Una vez en la LDL, el compuesto híbrido es transportado a todo el organismo incluyendo los bordes de los granulomas donde se acumulan células espumosas capaces de reconocer a las LDL oxidadas [153]. De esta forma se lograría que el compuesto ejerza su acción inhibitoria *in situ*. Esta estrategia ya ha sido patentada para el tratamiento de la aterosclerosis y esperamos poder evaluarla prontamente en cultivos celulares de macrófagos infectados con *M. tuberculosis* (Nitroalkenes tocopherols and analogs thereof for use in the treatment and prevention of inflammation related conditions; C. Batthyány & G.V. López inventors; U.S. PCT Application 2014 No 61/903, 068).

Por otro lado, los resultados obtenidos en el marco de este trabajo nos han permitido identificar nuevos procesos regulados por PknG y nuevos sustratos que podrían estar mediando los mismos. En los últimos cinco años han aparecido múltiples estudios fosfoproteómicos en micobacterias. Se demostró que más de 500 proteínas aparecen fosforiladas en aminoácidos hidroxilados y que los cambios a nivel metabólico se asocian a cambios del patrón de fosforilación. Sin embargo continúa sin saberse cuáles son las quinasas responsables de tales eventos de fosforilación. Nuestro abordaje experimental reveló nueve sustratos de PknG que no habían sido previamente reportados, entre ellos es de particular relevancia la glutamina sintasa. A pesar de que aún desconocemos el efecto que pueda tener PknG sobre esta enzima, es muy llamativo que cuatro enzimas cruciales en el metabolismo del nitrógeno estén directa o indirectamente reguladas por PknG.

En bacterias está establecido que la disminución de los niveles de amonio induce la activación de GS y de GOGAT y la inhibición de GDH, este conjunto de ajustes metabólicos es similar a lo que ocurre con GarA en ausencia de PknG [93, 120, 154] (Fig. 16). Aunque estamos lejos de comprender el rol de la fosforilación de estas proteínas en la adaptación de *M. tuberculosis* al medio intracelular podemos sugerir que **PknG cumpliría un papel en la fina regulación destinada a mantener el balance entre la síntesis y la degradación de intermediarios nitrogenados**. La comprensión de la interacción PknG-GS es fundamental para entender cabalmente cómo se regula el metabolismo del nitrógeno en micobacterias.

Entre los interactores de PknG seleccionamos uno para su validación posterior: la proteína FhaA; la elección se basó en que esta proteína posee un dominio de reconocimiento de fosfotreoninas y en su posible rol en el mantenimiento de la morfología celular. Nuestros resultados indican que FhaA interacciona con PknG a través de los residuos fosforilados del segmento N-terminal de la quinasa y que es asimismo sustrato de la misma. . Por otro lado, la cepa productora de FhaA presenta alteraciones morfológicas que podrían deberse a defectos en la síntesis de la pared celular o en la división así como una disminución de la velocidad de crecimiento. En este trabajo determinamos que la presencia de los residuos fosforilables en FhaA es importante para el fenotipo de crecimiento, lo que apunta a un rol fisiológico de estos eventos de fosforilación.

FhaA posee además del dominio C-terminal de unión a fosfotreoninas un dominio N-terminal de función desconocida y donde ocurren los eventos de fosforilación. Nuestra hipótesis actual de trabajo es que **FhaA podría constituir una proteína de andamiaje que medie la asociación dependiente de fosforilación de nuevas proteínas y participe en la transmisión de información a través de una vía de señalización mediada por PknG**. En este sentido, la identificación de los interactores *downstream* a FhaA y el mapeo de los mismos a cada uno de sus dominios permitirá comenzar a dilucidar la vía en la que esta proteína participa. Los resultados obtenidos en este trabajo han dado lugar a un nuevo proyecto del laboratorio donde nos proponemos caracterizar los complejos funcionales de señalización en los que participa FhaA formados *in vivo* en micobacterias. Para ello estamos utilizando una estrategia conocida como SPINE (*Strep-Protein Interaction Experiment*) que combina las ventajas de una purificación específica de proteínas conteniendo etiquetas (*Strep-tag*) con el entrecruzamiento reversible para obtener una "instantánea" de las interacciones proteína-proteína *in vivo.*

En suma, al día de hoy existen dos sustratos de PknG conteniendo dominios FHA de reconocimiento específico de fosfoThr: las proteínas GarA y FhaA que han sido relacionadas al metabolismo

intermediario y la síntesis de la pared celular respectivamente [86, 93]. Nuestros resultados indicarían que, en micobacterias, las proteínas con dominios FHA participan en la organización de vías de transducción de señales mediadas por PknG. Esta participación ocurre a través del reconocimiento de proteínas fosforiladas incluyendo las propias quinasas, y a su vez siendo ellas mismas sustratos fosforilables de estas enzimas. Por otro lado se demostró que FhaA y GarA son esenciales para la sobrevida de la bacteria en el hospedero. Mientras que la depleción de FhaA no afecta el crecimiento de las micobacterias en cultivo (aunque si su morfología), se demostró que FhaA es necesaria para la supervivencia de la bacteria en modelos de macrófagos murinos [86, 155]. Asimismo, micobacterias deficientes en GarA son rápidamente destruidas por el macrófago y tienen dificultad para crecer *in vitro* en medios pobres en aminoácidos; condiciones nutricionales a las que también debe enfrentarse la bacteria dentro del fagosoma [130, 156]. La capacidad de *M. tuberculosis* de inhibir la maduración fagolisosomal y de adaptarse al ambiente intracelular del hospedero mediante la reprogramación de vías metabólicas para acceder a nutrientes, minimizar las consecuencias del estrés oxidativo y mantener la infección son elementos centrales en su patogenicidad [157-159]. Nuestros resultados podrían indicar que PknG está jugando un papel clave en la sobrevida de la bacteria en el hospedero a través de la regulación de la capacidad de la misma de adaptarse a las condiciones nutricionales del ambiente intracelular.

En este sentido, debemos recordar que la tuberculosis continúa siendo hoy en día un problema mayor en salud pública a nivel mundial. Un aporte desde la investigación básica que permita profundizar en el conocimiento de la fisiopatología del agente causante de la enfermedad así como identificar blancos para el desarrollo de nuevos fármacos podría ser de crucial importancia en el desarrollo de estrategias para el futuro control de la misma.

V. REFERENCIAS BIBLIOGRÁFICAS

1.   (2015) Global tuberculosis report, World Health Organization, Geneva, Switzerland.

2.   (2014) Informe Epidemiológico 2014, Comisión Honoraria para la lucha antituberculosa y enfermedades prevalentes, Uruguay.

3.   (2014) Tuberculosis in the Americas, Pan American Health Organization.

4.   Comas, I., Coscolla, M., Luo, T., Borrell, S., Holt, K. E., Kato-Maeda, M., Parkhill, J., Malla, B., Berg, S., Thwaites, G., Yeboah-Manu, D., Bothamley, G., Mei, J., Wei, L., Bentley, S., Harris, S. R., Niemann, S., Diel, R., Aseffa, A., Gao, Q., Young, D., and Gagneux, S. (2013) Out-of-Africa migration and Neolithic coexpansion of Mycobacterium tuberculosis with modern humans, *Nat Genet 45*, 1176-1182.

5.   Cambau, E., and Drancourt, M. (2014) Steps towards the discovery of Mycobacterium tuberculosis by Robert Koch, 1882, *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases 20*, 196-201.

6.   Cambier, C. J., Falkow, S., and Ramakrishnan, L. (2014) Host evasion and exploitation schemes of Mycobacterium tuberculosis, *Cell 159*, 1497-1509.

7.   Madigan, M. T., Martinko, J. M., and Parker, J. (1998) *Brock Biología de los microorganismos*, 8va ed., Prentice Hall Iberia, Madrid.

8.   Gram, H. C. (1884) Über die isolierte Färbung der Schizomyceten in Schnitt- und Trockenpräparaten, *Fortschritte der Medizin 2*, 185-189.

9.   Hett, E. C., and Rubin, E. J. (2008) Bacterial growth and cell division: a mycobacterial perspective, *Microbiology and molecular biology reviews : MMBR 72*, 126-156.

10.   Kieser, K. J., and Rubin, E. J. (2014) How sisters grow apart: mycobacterial growth and division, *Nat Rev Microbiol 12*, 550-562.

11.   Jankute, M., Cox, J. A., Harrison, J., and Besra, G. S. (2015) Assembly of the Mycobacterial Cell Wall, *Annual review of microbiology 69*, 405-423.

12.   Nataraj, V., Varela, C., Javid, A., Singh, A., Besra, G. S., and Bhatt, A. (2015) Mycolic acids: deciphering and targeting the Achilles' heel of the tubercle bacillus, *Mol Microbiol 98*, 7-16.

13.   Smith, I. (2003) Mycobacterium tuberculosis pathogenesis and molecular determinants of virulence, *Clinical microbiology reviews 16*, 463-496.

14.   van Crevel, R., Ottenhoff, T. H., and van der Meer, J. W. (2002) Innate immunity to Mycobacterium tuberculosis, *Clinical microbiology reviews 15*, 294-309.

15.   Belkaid, Y., and Hand, T. W. (2014) Role of the microbiota in immunity and inflammation, *Cell 157*, 121-141.

16.   Niedergang, F., and Chavrier, P. (2004) Signaling and membrane dynamics during phagocytosis: many roads lead to the phagos(R)ome, *Current opinion in cell biology 16*, 422-428.

17.   Rohde, K., Yates, R. M., Purdy, G. E., and Russell, D. G. (2007) Mycobacterium tuberculosis and the environment within the phagosome, *Immunological reviews 219*, 37-54.

18.   Ramakrishnan, L. (2012) Revisiting the role of the granuloma in tuberculosis, *Nature reviews. Immunology 12*, 352-366.

19.   Taylor, J. L., Hattle, J. M., Dreitz, S. A., Troudt, J. M., Izzo, L. S., Basaraba, R. J., Orme, I. M., Matrisian, L. M., and Izzo, A. A. (2006) Role for matrix metalloproteinase 9 in granuloma formation during pulmonary Mycobacterium tuberculosis infection, *Infect Immun 74*, 6135-6144.

20.   Volkman, H. E., Pozos, T. C., Zheng, J., Davis, J. M., Rawls, J. F., and Ramakrishnan, L. (2010) Tuberculous granuloma induction via interaction of a bacterial secreted protein with host epithelium, *Science 327*, 466-469.

21.   Cosma, C. L., Sherman, D. R., and Ramakrishnan, L. (2003) The secret lives of the pathogenic mycobacteria, *Annual review of microbiology 57*, 641-676.

22.   Calmette, A., Guérin, C., and Weil-Hallé, V. (1924) Essais d'immunisation contre l'infection tuberculeuse, *Revue de la tuberculose V*, 481-491.

23.   Liu, J., Tran, V., Leung, A. S., Alexander, D. C., and Zhu, B. (2009) BCG vaccines: their mechanisms of attenuation and impact on safety and protective efficacy, *Human vaccines 5*, 70-78.

24.   Montagnani, C., Chiappini, E., Galli, L., and de Martino, M. (2014) Vaccine against tuberculosis: what's new?, *BMC infectious diseases 14 Suppl 1*, S2.

25.   (2010) Guidelines for treatment of tuberculosis, World Health Organization, Geneva, Switzerland.

26.   Tratamiento de la tuberculosis, Comisión Honoraria para la lucha antituberculosa y enfermedades prevalentes, Uruguay.

27.   (2011) Guidelines for the programmatic management of drug-resistant tuberculosis, World Health Organization, Geneva, Switzerland.

28.   Dorman, S. E., and Chaisson, R. E. (2007) From magic bullets back to the magic mountain: the rise of extensively drug-resistant tuberculosis, *Nature medicine 13*, 295-298.

29.   Szekely, R., Waczek, F., Szabadkai, I., Nemeth, G., Hegymegi-Barakonyi, B., Eros, D., Szokol, B., Pato, J., Hafenbradl, D., Satchell, J., Saint-Joanis, B., Cole, S. T., Orfi, L., Klebl, B. M., and Keri, G. (2008) A novel drug discovery concept for tuberculosis: inhibition of bacterial and host cell signalling, *Immunology letters 116*, 225-231.

30.   Wehenkel, A., Bellinzoni, M., Grana, M., Duran, R., Villarino, A., Fernandez, P., Andre-Leroux, G., England, P., Takiff, H., Cervenansky, C., Cole, S. T., and Alzari, P. M. (2008) Mycobacterial Ser/Thr protein kinases and phosphatases: physiological roles and therapeutic potential, *Biochim Biophys Acta 1784*, 193-202.

31.   Berg, J. M., Tymoczko, J. L., and Stryer, L. (2002) *Biochemistry*, 5th ed., Macmillan, New York.

32.   Cohen, P. (2002) The origins of protein phosphorylation, *Nature cell biology 4*, E127-130.

33.   Ciesla, J., Fraczyk, T., and Rode, W. (2011) Phosphorylation of basic amino acid residues in proteins: important but easily missed, *Acta biochimica Polonica 58*, 137-148.

34.   Hunter, T. (2012) Why nature chose phosphate to modify proteins, *Philosophical transactions of the Royal Society of London. Series B, Biological sciences 367*, 2513-2516.

35.   Sickmann, A., and Meyer, H. E. (2001) Phosphoamino acid analysis, *Proteomics 1*, 200-206.

36.   Stock, A. M., Robinson, V. L., and Goudreau, P. N. (2000) Two-component signal transduction, *Annual review of biochemistry 69*, 183-215.

37.   Av-Gay, Y., and Deretic, V. (2005) Two-Component Systems, Protein Kinases, and Signal Transduction in *Mycobacterium tuberculosis*, In *Tuberculosis and the Tubercle Bacillus* (Cole, S. T., Ed.), ASM Press, Washington, D.C.

38.   Laub, M. T., and Goulian, M. (2007) Specificity in two-component signal transduction pathways, *Annual review of genetics 41*, 121-145.

39.   Haydel, S. E., Malhotra, V., Cornelison, G. L., and Clark-Curtiss, J. E. (2012) The prrAB two-component system is essential for Mycobacterium tuberculosis viability and is induced under nitrogen-limiting conditions, *J Bacteriol 194*, 354-361.

40.   Zahrt, T. C., and Deretic, V. (2000) An essential two-component signal transduction system in Mycobacterium tuberculosis, *J Bacteriol 182*, 3832-3838.

41.   Wang, J. Y., and Koshland, D. E., Jr. (1978) Evidence for protein kinase activities in the prokaryote Salmonella typhimurium, *J Biol Chem 253*, 7605-7608.

42.   Sherman, D. R., and Grundner, C. (2014) Agents of change - concepts in Mycobacterium tuberculosis Ser/Thr/Tyr phosphosignalling, *Mol Microbiol 94*, 231-241.

43.   Leonard, C. J., Aravind, L., and Koonin, E. V. (1998) Novel families of putative protein kinases in bacteria and archaea: evolution of the "eukaryotic" protein kinase superfamily, *Genome research 8*, 1038-1047.

44.   Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S., Barry, C. E., 3rd, Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, K., Osborne, J., Quail, M. A., Rajandream, M. A., Rogers, J., Rutter, S., Seeger, K., Skelton, J., Squares, R., Squares, S., Sulston, J. E., Taylor, K., Whitehead, S., and Barrell, B. G. (1998) Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence, *Nature 393*, 537-544.

45.   Hanks, S. K., and Hunter, T. (1995) Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification, *FASEB journal : official publication of the Federation of American Societies for Experimental Biology 9*, 576-596.

46.   Kusebauch, U., Ortega, C., Ollodart, A., Rogers, R. S., Sherman, D. R., Moritz, R. L., and Grundner, C. (2014) Mycobacterium tuberculosis supports protein tyrosine phosphorylation, *Proc Natl Acad Sci U S A 111*, 9265-9270.

47.   Prisic, S., Dankwa, S., Schwartz, D., Chou, M. F., Locasale, J. W., Kang, C. M., Bemis, G., Church, G. M., Steen, H., and Husson, R. N. (2010) Extensive phosphorylation with overlapping specificity by Mycobacterium tuberculosis serine/threonine protein kinases, *Proc Natl Acad Sci U S A 107*, 7521-7526.

48.   Koul, A., Choidas, A., Treder, M., Tyagi, A. K., Drlica, K., Singh, Y., and Ullrich, A. (2000) Cloning and characterization of secretory tyrosine phosphatases of Mycobacterium tuberculosis, *J Bacteriol 182*, 5425-5432.

49.   Kalume, D. E., Molina, H., and Pandey, A. (2003) Tackling the phosphoproteome: tools and strategies, *Current opinion in chemical biology 7*, 64-69.

50.   Aebersold, R., and Goodlett, D. R. (2001) Mass spectrometry in proteomics, *Chemical reviews 101*, 269-295.

51.   Fortuin, S., Tomazella, G. G., Nagaraj, N., Sampson, S. L., Gey van Pittius, N. C., Soares, N. C., Wiker, H. G., de Souza, G. A., and Warren, R. M. (2015) Phosphoproteomics analysis of a clinical Mycobacterium tuberculosis Beijing isolate: expanding the mycobacterial phosphoproteome catalog, *Front Microbiol 6*, 6.

52.   Nakedi, K. C., Nel, A. J., Garnett, S., Blackburn, J. M., and Soares, N. C. (2015) Comparative Ser/Thr/Tyr phosphoproteomics between two mycobacterial species: the fast growing Mycobacterium smegmatis and the slow growing Mycobacterium bovis BCG, *Front Microbiol 6*, 237.

53.   Zheng, J., Liu, L., Liu, B., and Jin, Q. (2015) Phosphoproteomic analysis of bacillus Calmette-Guerin using gel-based and gel-free approaches, *J Proteomics 126*, 189-199.

54.   Parandhaman, D. K., Sharma, P., Bisht, D., and Narayanan, S. (2014) Proteome and phosphoproteome analysis of the serine/threonine protein kinase E mutant of Mycobacterium tuberculosis, *Life Sci 109*, 116-126.

55.   Calder, B., Albeldas, C., Blackburn, J. M., and Soares, N. C. (2016) Mass Spectrometry Offers Insight into the Role of Ser/Thr/Tyr Phosphorylation in the Mycobacteria, *Front Microbiol 7*, 141.

56.   Baer, C. E., Iavarone, A. T., Alber, T., and Sassetti, C. M. (2014) Biochemical and spatial coincidence in the provisional Ser/Thr protein kinase interaction network of Mycobacterium tuberculosis, *J Biol Chem 289*, 20422-20433.

57.   Prisic, S., and Husson, R. N. (2014) Mycobacterium tuberculosis Serine/Threonine Protein Kinases, *Microbiol Spectr 2*.

58.   Manning, G., Whyte, D. B., Martinez, R., Hunter, T., and Sudarsanam, S. (2002) The protein kinase complement of the human genome, *Science 298*, 1912-1934.

59.   Av-Gay, Y., and Everett, M. (2000) The eukaryotic-like Ser/Thr protein kinases of Mycobacterium tuberculosis, *Trends in microbiology 8*, 238-244.

60.   Ortiz-Lombardia, M., Pompeo, F., Boitel, B., and Alzari, P. M. (2003) Crystal structure of the catalytic domain of the PknB serine/threonine kinase from Mycobacterium tuberculosis, *J Biol Chem 278*, 13094-13100.

61.   Young, T. A., Delagoutte, B., Endrizzi, J. A., Falick, A. M., and Alber, T. (2003) Structure of Mycobacterium tuberculosis PknB supports a universal activation mechanism for Ser/Thr protein kinases, *Nature structural biology 10*, 168-174.

62.   Gay, L. M., Ng, H. L., and Alber, T. (2006) A conserved dimer and global conformational changes in the structure of apo-PknE Ser/Thr protein kinase from Mycobacterium tuberculosis, *Journal of molecular biology 360*, 409-420.

63.   Scherr, N., Honnappa, S., Kunz, G., Mueller, P., Jayachandran, R., Winkler, F., Pieters, J., and Steinmetz, M. O. (2007) Structural basis for the specific inhibition of protein kinase G, a virulence factor of Mycobacterium tuberculosis, *Proc Natl Acad Sci U S A 104*, 12151-12156.

64.   Huse, M., and Kuriyan, J. (2002) The conformational plasticity of protein kinases, *Cell 109*, 275-282.

65.   Johnson, L. N., Noble, M. E., and Owen, D. J. (1996) Active and inactive protein kinases: structural basis for regulation, *Cell 85*, 149-158.

66.   Boitel, B., Ortiz-Lombardia, M., Duran, R., Pompeo, F., Cole, S. T., Cervenansky, C., and Alzari, P. M. (2003) PknB kinase activity is regulated by phosphorylation in two Thr residues and dephosphorylation by PstP, the cognate phospho-Ser/Thr phosphatase, in Mycobacterium tuberculosis, *Mol Microbiol 49*, 1493-1508.

67.   Greenstein, A. E., Echols, N., Lombana, T. N., King, D. S., and Alber, T. (2007) Allosteric activation by dimerization of the PknD receptor Ser/Thr protein kinase from Mycobacterium tuberculosis, *J Biol Chem 282*, 11427-11435.

68.   Wehenkel, A., Fernandez, P., Bellinzoni, M., Catherinot, V., Barilone, N., Labesse, G., Jackson, M., and Alzari, P. M. (2006) The structure of PknB in complex with mitoxantrone, an ATP-competitive inhibitor, suggests a mode of protein kinase regulation in mycobacteria, *FEBS letters 580*, 3018-3022.

69.   Amano, M., Hamaguchi, T., Shohag, M. H., Kozawa, K., Kato, K., Zhang, X., Yura, Y., Matsuura, Y., Kataoka, C., Nishioka, T., and Kaibuchi, K. (2015) Kinase-interacting substrate screening is a novel method to identify kinase substrates, *The Journal of cell biology 209*, 895-912.

70.   Mok, J., Im, H., and Snyder, M. (2009) Global identification of protein kinase substrates by protein microarray analysis, *Nature protocols 4*, 1820-1827.

71.   Berwick, D. C., and Tavare, J. M. (2004) Identifying protein kinase substrates: hunting for the organ-grinder's monkeys, *Trends in biochemical sciences 29*, 227-232.

72.   Hodgson, D. R., and Schroder, M. (2011) Chemical approaches towards unravelling kinase-mediated signalling pathways, *Chemical Society reviews 40*, 1211-1223.

73.   Chao, J., Wong, D., Zheng, X., Poirier, V., Bach, H., Hmama, Z., and Av-Gay, Y. (2010) Protein kinase and phosphatase signaling in Mycobacterium tuberculosis physiology and pathogenesis, *Biochim Biophys Acta 1804*, 620-627.

74.   Sharma, K., Gupta, M., Krupa, A., Srinivasan, N., and Singh, Y. (2006) EmbR, a regulatory protein with ATPase activity, is a substrate of multiple serine/threonine kinases and phosphatase in Mycobacterium tuberculosis, *The FEBS journal 273*, 2711-2721.

75.   Molle, V., Brown, A. K., Besra, G. S., Cozzone, A. J., and Kremer, L. (2006) The condensing activities of the Mycobacterium tuberculosis type II fatty acid synthase are differentially regulated by phosphorylation, *J Biol Chem 281*, 30094-30103.

76.   Veyron-Churlet, R., Molle, V., Taylor, R. C., Brown, A. K., Besra, G. S., Zanella-Cleon, I., Futterer, K., and Kremer, L. (2009) The Mycobacterium tuberculosis beta-ketoacyl-acyl carrier protein synthase III activity is inhibited by phosphorylation on a single threonine residue, *J Biol Chem 284*, 6414-6424.

77.   Thakur, M., and Chakraborti, P. K. (2006) GTPase activity of mycobacterial FtsZ is impaired due to its transphosphorylation by the eukaryotic-type Ser/Thr kinase, PknA, *J Biol Chem 281*, 40107-40113.

78.   Parikh, A., Verma, S. K., Khan, S., Prakash, B., and Nandicoori, V. K. (2009) PknB-mediated phosphorylation of a novel substrate, N-acetylglucosamine-1-phosphate uridyltransferase, modulates its acetyltransferase activity, *Journal of molecular biology 386*, 451-464.

79.   Canova, M. J., Kremer, L., and Molle, V. (2009) The Mycobacterium tuberculosis GroEL1 chaperone is a substrate of Ser/Thr protein kinases, *J Bacteriol 191*, 2876-2883.

80.   Thakur, M., and Chakraborti, P. K. (2008) Ability of PknA, a mycobacterial eukaryotic-type serine/threonine kinase, to transphosphorylate MurD, a ligase involved in the process of peptidoglycan biosynthesis, *The Biochemical journal 415*, 27-33.

81.   Kang, C. M., Abbott, D. W., Park, S. T., Dascher, C. C., Cantley, L. C., and Husson, R. N. (2005) The Mycobacterium tuberculosis serine/threonine kinases PknA and PknB: substrate identification and regulation of cell shape, *Genes & development 19*, 1692-1704.

82.   Villarino, A., Duran, R., Wehenkel, A., Fernandez, P., England, P., Brodin, P., Cole, S. T., Zimny-Arndt, U., Jungblut, P. R., Cervenansky, C., and Alzari, P. M. (2005) Proteomic identification of M. tuberculosis protein kinase substrates: PknB recruits GarA, a FHA domain-containing protein, through activation loop-mediated interactions, *Journal of molecular biology 350*, 953-963.

83.   Grundner, C., Gay, L. M., and Alber, T. (2005) Mycobacterium tuberculosis serine/threonine kinases PknB, PknD, PknE, and PknF phosphorylate multiple FHA domains, *Protein Sci 14*, 1918-1921.

84.   Roumestand, C., Leiba, J., Galophe, N., Margeat, E., Padilla, A., Bessin, Y., Barthe, P., Molle, V., and Cohen-Gonsaud, M. (2011) Structural insight into the Mycobacterium tuberculosis Rv0020c protein and its interaction with the PknB kinase, *Structure 19*, 1525-1534.

85.   Gupta, M., Sajid, A., Arora, G., Tandon, V., and Singh, Y. (2009) Forkhead-associated domain-containing protein Rv0019c and polyketide-associated protein PapA5, from substrates of serine/threonine protein kinase PknB to interacting proteins of Mycobacterium tuberculosis, *J Biol Chem 284*, 34723-34734.

86.   Gee, C. L., Papavinasasundaram, K. G., Blair, S. R., Baer, C. E., Falick, A. M., King, D. S., Griffin, J. E., Venghatakrishnan, H., Zukauskas, A., Wei, J. R., Dhiman, R. K., Crick, D. C., Rubin, E. J., Sassetti, C. M., and Alber, T. (2012) A phosphorylated pseudokinase complex controls cell wall synthesis in mycobacteria, *Sci Signal 5*, ra7.

87.   Greenstein, A. E., MacGurn, J. A., Baer, C. E., Falick, A. M., Cox, J. S., and Alber, T. (2007) M. tuberculosis Ser/Thr protein kinase D phosphorylates an anti-anti-sigma factor homolog, *PLoS Pathog 3*, e49.

88.   Hatzios, S. K., Baer, C. E., Rustad, T. R., Siegrist, M. S., Pang, J. M., Ortega, C., Alber, T., Grundner, C., Sherman, D. R., and Bertozzi, C. R. (2013) Osmosensory signaling in Mycobacterium tuberculosis mediated by a eukaryotic-like Ser/Thr protein kinase, *Proc Natl Acad Sci U S A 110*, E5069-5077.

89.   Dasgupta, A., Datta, P., Kundu, M., and Basu, J. (2006) The serine/threonine kinase PknB of Mycobacterium tuberculosis phosphorylates PBPA, a penicillin-binding protein required for cell division, *Microbiology 152*, 493-504.

90.   Park, S. T., Kang, C. M., and Husson, R. N. (2008) Regulation of the SigH stress response regulon by an essential protein kinase in Mycobacterium tuberculosis, *Proc Natl Acad Sci U S A 105*, 13105-13110.

91.   Perez, J., Garcia, R., Bach, H., de Waard, J. H., Jacobs, W. R., Jr., Av-Gay, Y., Bubis, J., and Takiff, H. E. (2006) Mycobacterium tuberculosis transporter MmpL7 is a potential substrate for kinase PknD, *Biochemical and biophysical research communications 348*, 6-12.

92.   Molle, V., Soulat, D., Jault, J. M., Grangeasse, C., Cozzone, A. J., and Prost, J. F. (2004) Two FHA domains on an ABC transporter, Rv1747, mediate its phosphorylation by PknF, a Ser/Thr protein kinase from Mycobacterium tuberculosis, *FEMS Microbiol Lett 234*, 215-223.

93.   O'Hare, H. M., Duran, R., Cervenansky, C., Bellinzoni, M., Wehenkel, A. M., Pritsch, O., Obal, G., Baumgartner, J., Vialaret, J., Johnsson, K., and Alzari, P. M. (2008) Regulation of glutamate metabolism by protein kinases in mycobacteria, *Mol Microbiol 70*, 1408-1423.

94.   Zheng, X., Papavinasasundaram, K. G., and Av-Gay, Y. (2007) Novel substrates of Mycobacterium tuberculosis PknH Ser/Thr kinase, *Biochemical and biophysical research communications 355*, 162-168.

95.   Chao, J. D., Papavinasasundaram, K. G., Zheng, X., Chavez-Steenbock, A., Wang, X., Lee, G. Q., and Av-Gay, Y. (2010) Convergence of Ser/Thr and two-component signaling to coordinate expression of the dormancy regulon in Mycobacterium tuberculosis, *J Biol Chem 285*, 29239-29246.

96.   Molle, V., Kremer, L., Girard-Blanc, C., Besra, G. S., Cozzone, A. J., and Prost, J. F. (2003) An FHA phosphoprotein recognition domain mediates protein EmbR phosphorylation by PknH, a Ser/Thr protein kinase from Mycobacterium tuberculosis, *Biochemistry 42*, 15300-15309.

97.   Arora, G., Sajid, A., Gupta, M., Bhaduri, A., Kumar, P., Basu-Modak, S., and Singh, Y. (2010) Understanding the role of PknJ in Mycobacterium tuberculosis: biochemical characterization and identification of novel substrate pyruvate kinase A, *PLoS One 5*, e10772.

98.   Kumar, P., Kumar, D., Parikh, A., Rananaware, D., Gupta, M., Singh, Y., and Nandicoori, V. K. (2009) The Mycobacterium tuberculosis protein kinase K modulates activation of transcription from the promoter of mycobacterial monooxygenase operon through phosphorylation of the transcriptional regulator VirS, *J Biol Chem 284*, 11090-11099.

99.   Canova, M. J., Veyron-Churlet, R., Zanella-Cleon, I., Cohen-Gonsaud, M., Cozzone, A. J., Becchi, M., Kremer, L., and Molle, V. (2008) The Mycobacterium tuberculosis serine/threonine kinase PknL phosphorylates Rv2175c: mass spectrometric profiling of the activation loop phosphorylation sites and their role in the recruitment of Rv2175c, *Proteomics 8*, 521-533.

100.   Fernandez, P., Saint-Joanis, B., Barilone, N., Jackson, M., Gicquel, B., Cole, S. T., and Alzari, P. M. (2006) The Ser/Thr protein kinase PknB is essential for sustaining mycobacterial growth, *J Bacteriol 188*, 7778-7784.

101.   Betts, J. C., Lukey, P. T., Robb, L. C., McAdam, R. A., and Duncan, K. (2002) Evaluation of a nutrient starvation model of Mycobacterium tuberculosis persistence by gene and protein expression profiling, *Mol Microbiol 43*, 717-731.

102.   Durocher, D., and Jackson, S. P. (2002) The FHA domain, *FEBS letters 513*, 58-66.

103.   Niebisch, A., Kabus, A., Schultz, C., Weil, B., and Bott, M. (2006) Corynebacterial protein kinase G controls 2-oxoglutarate dehydrogenase activity via the phosphorylation status of the OdhI protein, *J Biol Chem 281*, 12300-12307.

104.   Molle, V., and Kremer, L. (2010) Division and cell envelope regulation by Ser/Thr phosphorylation: Mycobacterium shows the way, *Mol Microbiol 75*, 1064-1077.

105.   Jayakumar, D., Jacobs, W. R., Jr., and Narayanan, S. (2008) Protein kinase E of Mycobacterium tuberculosis has a role in the nitric oxide stress response and apoptosis in a human macrophage model of infection, *Cellular microbiology 10*, 365-374.

106.   Spivey, V. L., Molle, V., Whalan, R. H., Rodgers, A., Leiba, J., Stach, L., Walker, K. B., Smerdon, S. J., and Buxton, R. S. (2011) Forkhead-associated (FHA) domain containing ABC transporter Rv1747 is positively regulated by Ser/Thr phosphorylation in Mycobacterium tuberculosis, *J Biol Chem 286*, 26198-26209.

107.   Deol, P., Vohra, R., Saini, A. K., Singh, A., Chandra, H., Chopra, P., Das, T. K., Tyagi, A. K., and Singh, Y. (2005) Role of Mycobacterium tuberculosis Ser/Thr kinase PknF: implications in glucose transport and cell division, *J Bacteriol 187*, 3415-3420.

108.   Alderwick, L. J., Molle, V., Kremer, L., Cozzone, A. J., Dafforn, T. R., Besra, G. S., and Futterer, K. (2006) Molecular structure of EmbR, a response element of Ser/Thr kinase signaling in Mycobacterium tuberculosis, *Proc Natl Acad Sci U S A 103*, 2558-2563.

109.   Gomez-Velasco, A., Bach, H., Rana, A. K., Cox, L. R., Bhatt, A., Besra, G. S., and Av-Gay, Y. (2013) Disruption of the serine/threonine protein kinase H affects phthiocerol dimycocerosates synthesis in Mycobacterium tuberculosis, *Microbiology 159*, 726-736.

110.   Gopalaswamy, R., Narayanan, S., Chen, B., Jacobs, W. R., and Av-Gay, Y. (2009) The serine/threonine protein kinase PknI controls the growth of Mycobacterium tuberculosis upon infection, *FEMS Microbiol Lett 295*, 23-29.

111.   Cohen-Gonsaud, M., Barthe, P., Canova, M. J., Stagier-Simon, C., Kremer, L., Roumestand, C., and Molle, V. (2009) The Mycobacterium tuberculosis Ser/Thr kinase substrate Rv2175c is a DNA-binding protein regulated by phosphorylation, *J Biol Chem 284*, 19290-19300.

112.   Kruh, N. A., Troudt, J., Izzo, A., Prenni, J., and Dobos, K. M. (2010) Portrait of a pathogen: the Mycobacterium tuberculosis proteome in vivo, *PLoS One 5*, e13938.

113.   Cole, S. T., Eiglmeier, K., Parkhill, J., James, K. D., Thomson, N. R., Wheeler, P. R., Honore, N., Garnier, T., Churcher, C., Harris, D., Mungall, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R. M., Devlin, K., Duthoy, S., Feltwell, T., Fraser, A., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Lacroix, C., Maclean, J., Moule, S., Murphy, L., Oliver, K., Quail, M. A., Rajandream, M. A., Rutherford, K. M., Rutter, S., Seeger, K., Simon, S., Simmonds, M., Skelton, J., Squares, R., Squares, S., Stevens, K., Taylor, K., Whitehead, S., Woodward, J. R., and Barrell, B. G. (2001) Massive gene decay in the leprosy bacillus, *Nature 409*, 1007-1011.

114.   Sieker, L. C., Stenkamp, R. E., and LeGall, J. (1994) Rubredoxin in crystalline state, *Methods in enzymology 243*, 203-216.

115.   Hagelueken, G., Wiehlmann, L., Adams, T. M., Kolmar, H., Heinz, D. W., Tummler, B., and Schubert, W. D. (2007) Crystal structure of the electron transfer complex rubredoxin rubredoxin reductase of Pseudomonas aeruginosa, *Proc Natl Acad Sci U S A 104*, 12276-12281.

116.   Tiwari, D., Singh, R. K., Goswami, K., Verma, S. K., Prakash, B., and Nandicoori, V. K. (2009) Key residues in Mycobacterium tuberculosis protein kinase G play a role in regulating kinase activity and survival in the host, *J Biol Chem 284*, 27467-27479.

117.   Zeytuni, N., and Zarivach, R. (2012) Structural and functional discussion of the tetra-trico-peptide repeat, a protein interaction module, *Structure 20*, 397-405.

118.   Mieczkowski, C., Iavarone, A. T., and Alber, T. (2008) Auto-activation mechanism of the Mycobacterium tuberculosis PknB receptor Ser/Thr kinase, *The EMBO journal 27*, 3186-3197.

119.   Duran, R., Villarino, A., Bellinzoni, M., Wehenkel, A., Fernandez, P., Boitel, B., Cole, S. T., Alzari, P. M., and Cervenansky, C. (2005) Conserved autophosphorylation pattern in activation loops and juxtamembrane regions of Mycobacterium tuberculosis Ser/Thr protein kinases, *Biochemical and biophysical research communications 333*, 858-867.

120.   Cowley, S., Ko, M., Pick, N., Chow, R., Downing, K. J., Gordhan, B. G., Betts, J. C., Mizrahi, V., Smith, D. A., Stokes, R. W., and Av-Gay, Y. (2004) The Mycobacterium tuberculosis protein serine/threonine kinase PknG is linked to cellular glutamate/glutamine levels and is important for growth in vivo, *Mol Microbiol 52*, 1691-1702.

121.   Walburger, A., Koul, A., Ferrari, G., Nguyen, L., Prescianotto-Baschong, C., Huygen, K., Klebl, B., Thompson, C., Bacher, G., and Pieters, J. (2004) Protein kinase G from pathogenic mycobacteria promotes survival within macrophages, *Science 304*, 1800-1804.

122.   Koul, A., Choidas, A., Tyagi, A. K., Drlica, K., Singh, Y., and Ullrich, A. (2001) Serine/threonine protein kinases PknF and PknG of Mycobacterium tuberculosis: characterization and localization, *Microbiology 147*, 2307-2314.

123.   Scherr, N., Muller, P., Perisa, D., Combaluzier, B., Jeno, P., and Pieters, J. (2009) Survival of pathogenic mycobacteria in macrophages is mediated through autophosphorylation of protein kinase G, *J Bacteriol 191*, 4546-4554.

124.   van der Woude, A. D., Stoop, E. J., Stiess, M., Wang, S., Ummels, R., van Stempvoort, G., Piersma, S. R., Cascioferro, A., Jimenez, C. R., Houben, E. N., Luirink, J., Pieters, J., van der Sar, A. M., and Bitter, W. (2014) Analysis of SecA2-dependent substrates in Mycobacterium marinum identifies protein kinase G (PknG) as a virulence effector, *Cellular microbiology 16*, 280-295.

125.   Braibant, M., Gilot, P., and Content, J. (2000) The ATP binding cassette (ABC) transport systems of Mycobacterium tuberculosis, *FEMS microbiology reviews 24*, 449-467.

126.   Nguyen, L., Walburger, A., Houben, E., Koul, A., Muller, S., Morbitzer, M., Klebl, B., Ferrari, G., and Pieters, J. (2005) Role of protein kinase G in growth and glutamine metabolism of Mycobacterium bovis BCG, *J Bacteriol 187*, 5852-5856.

127.   Belanger, A. E., and Hatfull, G. F. (1999) Exponential-phase glycogen recycling is essential for growth of Mycobacterium smegmatis, *J Bacteriol 181*, 6670-6678.

128.   England, P., Wehenkel, A., Martins, S., Hoos, S., Andre-Leroux, G., Villarino, A., and Alzari, P. M. (2009) The FHA-containing protein GarA acts as a phosphorylation-dependent molecular switch in mycobacterial signaling, *FEBS letters 583*, 301-307.

129.   Nott, T. J., Kelly, G., Stach, L., Li, J., Westcott, S., Patel, D., Hunt, D. M., Howell, S., Buxton, R. S., O'Hare, H. M., and Smerdon, S. J. (2009) An intramolecular switch regulates phosphoindependent FHA domain interactions in Mycobacterium tuberculosis, *Sci Signal 2*, ra12.

130.   Ventura, M., Rieck, B., Boldrin, F., Degiacomi, G., Bellinzoni, M., Barilone, N., Alzaidi, F., Alzari, P. M., Manganelli, R., and O'Hare, H. M. (2013) GarA is an essential regulator of metabolism in Mycobacterium tuberculosis, *Mol Microbiol 90*, 356-366.

131.   Anand, N., Singh, P., Sharma, A., Tiwari, S., Singh, V., Singh, D. K., Srivastava, K. K., Singh, B. N., and Tripathi, R. P. (2012) Synthesis and evaluation of small libraries of triazolylmethoxy chalcones, flavanones and 2-aminopyrimidines as inhibitors of mycobacterial FAS-II and PknG, *Bioorganic & medicinal chemistry 20*, 5150-5163.

132.   Singh, N., Tiwari, S., Srivastava, K. K., and Siddiqi, M. I. (2015) Identification of Novel Inhibitors of Mycobacterium tuberculosis PknG Using Pharmacophore Based Virtual Screening, Docking, Molecular Dynamics Simulation, and Their Biological Evaluation, *Journal of chemical information and modeling 55*, 1120-1129.

133. Fiuza, M., Canova, M. J., Zanella-Cleon, I., Becchi, M., Cozzone, A. J., Mateos, L. M., Kremer, L., Gil, J. A., and Molle, V. (2008) From the characterization of the four serine/threonine protein kinases (PknA/B/G/L) of Corynebacterium glutamicum toward the role of PknA and PknB in cell division, *J Biol Chem 283*, 18099-18112.

134. Narayan, A., Sachdeva, P., Sharma, K., Saini, A. K., Tyagi, A. K., and Singh, Y. (2007) Serine threonine protein kinases of mycobacterial genus: phylogeny to function, *Physiological genomics 29*, 66-75.

135. Batthyany, C., Schopfer, F. J., Baker, P. R., Duran, R., Baker, L. M., Huang, Y., Cervenansky, C., Branchaud, B. P., and Freeman, B. A. (2006) Reversible post-translational modification of proteins by nitrated fatty acids in vivo, *J Biol Chem 281*, 20450-20463.

136. Hatfull, G. F. (2012) The secret lives of mycobacteriophages, *Advances in virus research 82*, 179-288.

137. Piuri, M., Jacobs, W. R., Jr., and Hatfull, G. F. (2009) Fluoromycobacteriophages for rapid, specific, and sensitive antibiotic susceptibility testing of Mycobacterium tuberculosis, *PLoS One 4*, e4870.

138. Piuri, M., Rondon, L., Urdaniz, E., and Hatfull, G. F. (2013) Generation of affinity-tagged fluoromycobacteriophages by mixed assembly of phage capsids, *Appl Environ Microbiol 79*, 5608-5615.

139. Rondon, L., Piuri, M., Jacobs, W. R., Jr., de Waard, J., Hatfull, G. F., and Takiff, H. E. (2011) Evaluation of fluoromycobacteriophages for detecting drug resistance in Mycobacterium tuberculosis, *Journal of clinical microbiology 49*, 1838-1842.

140. Sambandamurthy, V. K., Derrick, S. C., Hsu, T., Chen, B., Larsen, M. H., Jalapathy, K. V., Chen, M., Kim, J., Porcelli, S. A., Chan, J., Morris, S. L., and Jacobs, W. R., Jr. (2006) Mycobacterium tuberculosis DeltaRD1 DeltapanCD: a safe and limited replicating mutant strain that protects immunocompetent and immunocompromised mice against experimental tuberculosis, *Vaccine 24*, 6309-6320.

141. Adams, J. A. (2001) Kinetic and catalytic mechanisms of protein kinases, *Chemical reviews 101*, 2271-2290.

142. Hunter, T. (2000) Signaling--2000 and beyond, *Cell 100*, 113-127.

143. Krupa, A., and Srinivasan, N. (2005) Diversity in domain architectures of Ser/Thr kinases and their homologues in prokaryotes, *BMC genomics 6*, 129.

144. Reckel, S., and Hantschel, O. (2015) Kinase Regulation in Mycobacterium tuberculosis: Variations on a Theme, *Structure 23*, 975-976.

145. Trinkle-Mulcahy, L., Boulon, S., Lam, Y. W., Urcia, R., Boisvert, F. M., Vandermoere, F., Morrice, N. A., Swift, S., Rothbauer, U., Leonhardt, H., and Lamond, A. (2008) Identifying specific protein interaction partners using quantitative mass spectrometry and bead proteomes, *The Journal of cell biology 183*, 223-239.

146. Carvalho, P. C., Lima, D. B., Leprevost, F. V., Santos, M. D., Fischer, J. S., Aquino, P. F., Moresco, J. J., Yates, J. R., 3rd, and Barbosa, V. C. (2016) Integrated analysis of shotgun proteomic data with PatternLab for proteomics 4.0, *Nature protocols 11*, 102-117.

147. Cousin, C., Derouiche, A., Shi, L., Pagot, Y., Poncet, S., and Mijakovic, I. (2013) Protein-serine/threonine/tyrosine kinases in bacterial signaling and regulation, *FEMS Microbiol Lett 346*, 11-19.

148. Gouzy, A., Poquet, Y., and Neyrolles, O. (2014) Nitrogen metabolism in Mycobacterium tuberculosis physiology and virulence, *Nat Rev Microbiol 12*, 729-737.

149. Leigh, J. A., and Dodsworth, J. A. (2007) Nitrogen regulation in bacteria and archaea, *Annual review of microbiology 61*, 349-377.

150. Gil, M., Grana, M., Schopfer, F. J., Wagner, T., Denicola, A., Freeman, B. A., Alzari, P. M., Batthyany, C., and Duran, R. (2013) Inhibition of Mycobacterium tuberculosis PknG by non-catalytic rubredoxin domain specific modification: reaction of an electrophilic nitro-fatty acid with the Fe-S center, *Free Radic Biol Med 65*, 150-161.

151.   Lisa, M. N., Gil, M., Andre-Leroux, G., Barilone, N., Duran, R., Biondi, R. M., and Alzari, P. M. (2015) Molecular Basis of the Activity and the Regulation of the Eukaryotic-like S/T Protein Kinase PknG from Mycobacterium tuberculosis, *Structure 23*, 1039-1048.

152.   Stocker, A. (2004) Molecular mechanisms of vitamin E transport, *Annals of the New York Academy of Sciences 1031*, 44-59.

153.   Caceres, N., Tapia, G., Ojanguren, I., Altare, F., Gil, O., Pinto, S., Vilaplana, C., and Cardona, P. J. (2009) Evolution of foamy macrophages in the pulmonary granulomas of experimental tuberculosis models, *Tuberculosis (Edinb) 89*, 175-182.

154.   Tempest, D. W., Meers, J. L., and Brown, C. M. (1970) Synthesis of glutamate in Aerobacter aerogenes by a hitherto unknown route, *The Biochemical journal 117*, 405-407.

155.   Sassetti, C. M., Boyd, D. H., and Rubin, E. J. (2003) Genes required for mycobacterial growth defined by high density mutagenesis, *Mol Microbiol 48*, 77-84.

156.   Tullius, M. V., Harth, G., and Horwitz, M. A. (2003) Glutamine synthetase GlnA1 is essential for growth of Mycobacterium tuberculosis in human THP-1 macrophages and guinea pigs, *Infect Immun 71*, 3927-3936.

157.   Deretic, V., Singh, S., Master, S., Harris, J., Roberts, E., Kyei, G., Davis, A., de Haro, S., Naylor, J., Lee, H. H., and Vergne, I. (2006) Mycobacterium tuberculosis inhibition of phagolysosome biogenesis and autophagy as a host defence mechanism, *Cellular microbiology 8*, 719-727.

158.   Lee, W., VanderVen, B. C., Fahey, R. J., and Russell, D. G. (2013) Intracellular Mycobacterium tuberculosis exploits host-derived fatty acids to limit metabolic stress, *J Biol Chem 288*, 6788-6800.

159.   Russell, D. G., VanderVen, B. C., Lee, W., Abramovitch, R. B., Kim, M. J., Homolka, S., Niemann, S., and Rohde, K. H. (2010) Mycobacterium tuberculosis wears what it eats, *Cell host & microbe 8*, 68-76.

VI. ANEXO

Tablas suplementarias del manuscrito "*Mycobacterium tuberculosis* PknG interactome reveals new substrate of this kinase related to cell wall synthesis"

Table S1. Oligonucleotides used in this work

Oligo name	Plasmid	Base pair sequence (5' - 3')
FhaA F	pLAM12-fhaA	AAATTCATATGTGGAGCCACCCGCAGTTCGAAAAAAGCGCTGGTAGCCAGAAAAGGCTGGTTC
FhaA R	pLAM12-fhaA	ATATTGAATTCTCAGTGCATGCGGACGATGATC
FhaA_T18A F	pLAM12-fhaA_T_{18}A	CTCGAGCAGGCGGTTGGCGATG
FhaA_T18A R	pLAM12-fhaA_T_{18}A	CATCGCCAACCGCCTGCTCGAG
FhaA_T116A F	pLAM12-fhaA_T_{116}A	GAACCTGCATGCCGGCCAGTTCC
FhaA_T116A R	pLAM12-fhaA_T_{116}A	GGAACTGGCCGGCATGCAGGTTC

Table S2a. Proteins identified in biological replicate n°1

Proteins highlighted in grey are found in mock experiments

UniProt accesion number	#Unique peptides	Sequence count	Spectrum count	NSAF	Coverage	Description
A0QS97	5	5	39	0.069140851	0.4744	30S ribosomal protein S7 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpsG PE=1 SV=1
A0R7F9	4	4	20	0.061458534	0.5556	30S ribosomal protein S6 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpsF PE=1 SV=9
A0R5R3	5	5	25	0.047683345	0.5724	KanY protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_6282 PE=4 SV=1
Q3I5Q7	7	7	30	0.035762509	0.3448	HBHA-like protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_0919 PE=4 SV=1
A0QQU5	13	13	67	0.034250921	0.3179	60 kDa chaperonin 1 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=groL1 PE=1 SV=1
A0R3B8	9	9	45	0.029146026	0.3396	Enolase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=eno PE=1 SV=1
A0QSD9	3	3	8	0.02873386	0.3506	50S ribosomal protein L29 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpmC PE=3 SV=1
A0QS98	4	4	40	0.027935697	0.1439	Elongation factor Tu OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=tuf PE=1 SV=1
A0QPE7	10	10	44	0.027041755	0.3022	3-oxoacyl-acyl-carrier protein reductase FabG4 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=fabG4 PE=1 SV=1
A0QWS8	2	2	10	0.026339372	0.2286	Integration host factor OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=mihF PE=4 SV=1
A0QXT5	6	6	29	0.025063558	0.2656	Uncharacterized protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_3419 PE=4 SV=1
A0R5H1	6	6	18	0.022223845	0.3571	Crp/Fnr familytranscriptional regulator OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_6189 PE=4 SV=1
A0QTS9	5	5	17	0.021370808	0.2	Conserved domain protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_1951 PE=4 SV=1
A0R7F6	2	2	11	0.020147003	0.2384	50S ribosomal protein L9 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rplI PE=1 SV=1
A0R1B5	4	4	11	0.019377054	0.4459	Uncharacterized protein MSMEG_4692/MSMEI_4575 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_4692 PE=1 SV=1
A0QS63	3	3	9	0.019146697	0.2846	50S ribosomal protein L7/L12 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rplL PE=3 SV=1
A0QSG3	1	3	9	0.018856596	0.2424	30S ribosomal protein S8 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpsH PE=1 SV=1; Additional IDs concatenated into MaxParsimony group: P9WH27
A0QS46	5	5	15	0.017652983	0.3106	50S ribosomal protein L1 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rplA PE=1 SV=1
A0QX24	4	4	22	0.017483893	0.2328	ATPase, MoxR family protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=moxR PE=4 SV=1
P9WI73	12	12	47	0.017331307	0.2227	Serine/threonine-protein kinase PknG OS=Mycobacterium tuberculosis (strain ATCC 25618 / H37Rv) GN=pknG PE=1 SV=1
A0R201	4	4	17	0.015314586	0.2182	ATP synthase gamma chain OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=atpG PE=1 SV=1
A0R203	6	6	23	0.014294288	0.1933	ATP synthase subunit b-delta OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=atpFH PE=1 SV=1
A0QS62	5	5	9	0.014223261	0.4229	50S ribosomal protein L10 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rplJ PE=1 SV=1
A0QQJ4	2	2	17	0.013992791	0.0774	F420-dependent glucose-6-phosphate dehydrogenase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=fgd PE=1 SV=2
A0R7G5	3	3	9	0.013527558	0.1739	Transcriptional regulator, PadR family protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_6903 PE=4 SV=1
A0QSS4	9	9	26	0.013316016	0.2722	60 kDa chaperonin 2 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=groL2 PE=1 SV=1
A0QVU2	5	5	13	0.012886467	0.2509	35 kDa protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_2695 PE=4 SV=1
A0R729	6	6	23	0.012595957	0.1386	Glycerol kinase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=glpK PE=3 SV=1
A0QXC0	7	7	17	0.01214878	0.2894	Branched-chain amino acid ABC transporter substrate-binding protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_3247 PE=4 SV=1
P9WLR5	1	2	7	0.011804535	0.1646	Uncharacterized protein Rv1829 OS=Mycobacterium tuberculosis (strain ATCC 25618 / H37Rv) GN=Rv1829 PE=1 SV=1; Additional IDs concatenated into MaxParsimony group: A0QYG0
A0R1Z7	2	2	8	0.011644775	0.1842	ATP:cob(I)alamin adenosyltransferase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_4934 PE=4 SV=1
A0R200	6	6	19	0.011062536	0.1811	ATP synthase subunit beta OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=atpD PE=1 SV=1
A0QPE8	6	6	17	0.010883282	0.1667	3-ketoacyl-CoA thiolase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=fadA2 PE=3 SV=1
A0QUX7	2	2	6	0.009761061	0.1353	Acetolactate synthase small subunit OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=ilvH PE=1 SV=1
A0R656	3	3	5	0.009738148	0.2817	Uncharacterized protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_6431 PE=4 SV=1
A0R0W7	6	6	12	0.009509343	0.1805	Sulfate-binding lipoprotein SubI OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_4533 PE=4 SV=1

Accession						Description
A0R157	6	6	14	0.009307422	0.1899	Saccharopine dehydrogenase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_4632 PE=4 SV=1
A0R1Z9	2	2	4	0.009142592	0.2397	ATP synthase epsilon chain OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=atpC PE=1 SV=1
A0QNG7	4	4	16	0.009086272	0.1786	FHA domain protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_0035 PE=4 SV=1
A0R202	5	5	18	0.009084199	0.1259	ATP synthase subunit alpha OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=atpA PE=1 SV=1
A0QSG6	3	3	7	0.009046466	0.1963	30S ribosomal protein S5 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpsE PE=1 SV=1
A0R0B3	2	2	3	0.008380709	0.2424	Meromycolate extension acyl carrier protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=acpM PE=1 SV=1
A0R4L2	3	3	8	0.008134218	0.1618	Conserved membrane protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_5873 PE=4 SV=1
A0R204	2	2	5	0.008134218	0.1588	ATP synthase subunit b OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=atpF PE=3 SV=1
A0QYY6	5	5	14	0.008083273	0.1608	30S ribosomal protein S1 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpsA PE=3 SV=1
A0QTT5	6	6	13	0.007665937	0.2004	Uncharacterized protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_1957 PE=4 SV=1
A0QYG2	2	2	4	0.007001605	0.2975	Glycogen accumulation regulator GarA OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=garA PE=1 SV=2
A0QWT6	2	2	7	0.006938867	0.1147	Lipoprotein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_3058 PE=3 SV=1
A0R042	5	5	15	0.006925628	0.1252	AMP-binding enzyme OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_4254 PE=4 SV=1
A0QU53	4	4	10	0.006811906	0.1108	Acyl-CoA dehydrogenase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=fadE23 PE=3 SV=1
A0QQC8	4	4	15	0.006669535	0.1286	Chaperone protein DnaK OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=dnaK PE=1 SV=8
A0QXB0	3	3	7	0.006562521	0.1763	ABC-type amino acid transport system, secreted component OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_3235 PE=3 SV=1
A0QSH0	5	5	13	0.006052734	0.1364	Putative protease IV Sppa OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=sppA PE=4 SV=1
A0QYH8	2	2	10	0.005871834	0.0594	Oxidoreductase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_3663 PE=4 SV=1
A0R730	6	6	12	0.005842889	0.1109	Glycerol-3-phosphate dehydrogenase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_6761 PE=3 SV=1
A0QS45	2	2	3	0.005842889	0.1479	50S ribosomal protein L11 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rplK PE=1 SV=1
A0R4C3	3	3	7	0.005121544	0.1587	Phosphate-binding protein PstS OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=pstS PE=3 SV=1
A0QVX3	2	2	5	0.005065264	0.1392	Glutamate binding periplasmic protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_2727 PE=3 SV=1
A0QV12	2	2	4	0.004571296	0.1694	Putative conserved membrane protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_2410 PE=4 SV=1
P71534	2	2	4	0.004338249	0.1333	3-oxoacyl-[acyl-carrier-protein] reductase FabG OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=fabG PE=3 SV=2
A0QQW8	4	4	7	0.004172293	0.1466	Dihydrolipoyl dehydrogenase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=lpdA PE=3 SV=1
A0R0B4	3	3	6	0.003988895	0.1538	3-oxoacyl-(Acyl-carrier-protein) synthase 1 KasA OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=kasA PE=3 SV=1
A0R5M3	3	3	6	0.003922885	0.0804	Alcohol dehydrogenase, iron-containing OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_6242 PE=4 SV=1
A0QP15	2	2	2	0.003545685	0.1667	CoA binding domain protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_0238 PE=4 SV=1
A0QT07	3	3	3	0.00317889	0.1456	Succinate dehydrogenase iron-sulfur protein subunit SdhB OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=sdhB PE=4 SV=1
A0R5R5	4	4	5	0.003164341	0.1373	Cyclopropane fatty-acyl-phospholipid synthase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_6284 PE=4 SV=1
A0QU54	2	2	5	0.003128545	0.086	Putative acyl-CoA dehydrogenase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_2081 PE=3 SV=1
A0QQ65	4	4	6	0.00301158	0.0563	ABC transporter solute-binding protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_0643 PE=4 SV=1
contaminant_KERATIN03	2	2	6	0.002798281	0.0607	
A0R3L1	2	2	4	0.002399682	0.0738	Magnesium chelatase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_5512 PE=4 SV=1
A0R7J3	2	2	3	0.00236379	0.0684	ParB-like partition proteins OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=parB PE=4 SV=1
A0QUX6	5	5	5	0.002237568	0.0874	Acetolactate synthase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=ilvB PE=3 SV=1
A0R3D1	2	2	2	0.001907334	0.0793	Retinol dehydrogenase 13 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_5430 PE=3 SV=1
A0R072	2	2	4	0.001891032	0.0735	2-oxoglutarate dehydrogenase, E2 component, dihydrolipoamide succinyltransferase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=sucB PE=3 SV=1
A0QPX3	2	2	2	0.001717785	0.0963	Aliphatic sulfonates family ABC transporter, periplasmic ligand-binding protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_0550 PE=4 SV=1
contaminant_KERATIN22	4	4	4	0.001715122	0.0992	
A0QXA3	2	2	2	0.001171879	0.0551	Pyruvate kinase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=pyk PE=3 SV=1
A0R465	2	2	3	0.001152348	0.0708	Enoyl-CoA hydratase/3-hydroxyacyl-CoA dehydrogenase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=fadB PE=4 SV=1

146

Table S2b. Proteins identified in biological replicate n°2

Proteins highlighted in grey are found in mock experiments

UniProt accesion number	#Unique peptides	Sequence count	Spectrum count	NSAF	Coverage	Description
A0R7F9	6	6	25	0.085667014	0.6667	30S ribosomal protein S6 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpsF PE=1 SV=1
A0QSG3	1	3	24	0.056072955	0.2424	30S ribosomal protein S8 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpsH PE=1 SV=1; Additional IDs concatenated into MaxParsimony group: P9WH27
A0QS97	0	4	19	0.037561691	0.4038	30S ribosomal protein S7 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpsG PE=1 SV=1; Additional IDs concatenated into MaxParsimony group: P9WH29
A0QS62	3	3	21	0.03700815	0.28	50S ribosomal protein L10 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rplJ PE=1 SV=1
A0QSD9	4	4	9	0.0360469	0.4805	50S ribosomal protein L29 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpmC PE=3 SV=1
Q3I5Q7	7	7	27	0.035891525	0.3103	HBHA-like protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_0919 PE=4 SV=1
A0QWS8	2	2	12	0.035245857	0.2286	Integration host factor OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=mihF PE=4 SV=1
A0R3B8	7	7	39	0.028167796	0.3162	Enolase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=eno PE=1 SV=1
A0R7F6	3	3	13	0.026551101	0.3046	50S ribosomal protein L9 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rplI PE=1 SV=1
A0R5R3	3	3	11	0.023395957	0.3103	KanY protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_6282 PE=4 SV=1
A0QPE7	11	11	34	0.023301428	0.3444	3-oxoacyl-acyl-carrier protein reductase FabG4 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=fabG4 PE=1 SV=1
A0QVU2	6	6	21	0.023212997	0.3441	35 kDa protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_2695 PE=4 SV=1
A0R200	10	10	33	0.021425771	0.3453	ATP synthase subunit beta OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=atpD PE=1 SV=1
A0QXT5	8	8	22	0.021202586	0.2969	Uncharacterized protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_3419 PE=4 SV=1
A0R203	8	8	26	0.01801895	0.3079	ATP synthase subunit b-delta OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=atpFH PE=1 SV=1
A0R0W7	8	8	20	0.017673424	0.3811	Sulfate-binding lipoprotein SubI OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_4533 PE=4 SV=1
A0QQU5	12	12	30	0.017101733	0.3179	60 kDa chaperonin 1 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=groL1 PE=1 SV=1
A0QS63	3	3	7	0.016606221	0.2846	50S ribosomal protein L7/L12 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rplL PE=3 SV=1
P9WI73	13	13	40	0.016448067	0.2293	Serine/threonine-protein kinase PknG OS=Mycobacterium tuberculosis (strain ATCC 25618 / H37Rv) GN=pknG PE=1 SV=1
A0R1B5	4	4	8	0.015714713	0.2611	Uncharacterized protein MSMEG_4692/MSMEI_4575 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_4692 PE=1 SV=1
A0R5H1	4	4	11	0.015144704	0.2411	Crp/Fnr familytranscriptional regulator OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_6189 PE=4 SV=1
A0R4L2	5	5	13	0.014739766	0.2757	Conserved membrane protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_5873 PE=4 SV=1
A0R150	2	2	4	0.014018239	0.3523	50S ribosomal protein L27 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpmA PE=1 SV=1
A0QTS9	5	5	10	0.014018239	0.2545	Conserved domain protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_1951 PE=4 SV=1
A0R201	4	4	13	0.013059336	0.1857	ATP synthase gamma chain OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=atpG PE=1 SV=1
A0R197	2	2	9	0.012732162	0.2523	ATP-dependent Clp protease proteolytic subunit OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=clpP PE=3 SV=1
A0R204	3	3	7	0.012698875	0.2353	ATP synthase subunit b OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=atpF PE=3 SV=1
A0QWT6	4	4	10	0.011053808	0.2366	Lipoprotein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_3058 PE=3 SV=1
A0R656	3	3	5	0.010859199	0.2183	Uncharacterized protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_6431 PE=4 SV=1
A0QZ11	2	2	4	0.010821097	0.3684	RNA polymerase-binding protein RbpA OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rbpA PE=1 SV=2
A0QX24	5	5	12	0.010634526	0.2874	ATPase, MoxR family protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=moxR PE=4 SV=1
A0QXB0	4	4	10	0.01045428	0.2644	ABC-type amino acid transport system, secreted component OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_3235 PE=3 SV=1
A0R1Z7	3	3	6	0.009738987	0.2368	ATP:cob(I)alamin adenosyltransferase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_4934 PE=4 SV=1
A0QPX3	5	5	10	0.009577679	0.2671	Aliphatic sulfonates family ABC transporter, periplasmic ligand-binding protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_0550 PE=4 SV=1
A0QXC0	5	5	12	0.00956283	0.2661	Branched-chain amino acid ABC transporter substrate-binding protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_3247 PE=4 SV=1

A0QQC8	6	6	18	0.008924795	0.1624	Chaperone protein DnaK OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=dnaK PE=1 SV=1
A0QVB8	2	2	8	0.008906895	0.1733	30S ribosomal protein S2 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpsB PE=3 SV=2
A0QTT5	5	5	13	0.008548436	0.1578	Uncharacterized protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_1957 PE=4 SV=1
A0QYY6	4	4	13	0.008369971	0.1441	30S ribosomal protein S1 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpsA PE=3 SV=1
A0QWN0	4	4	8	0.008307104	0.1886	Conserved membrane glycine rich protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_3000 PE=4 SV=1
A0R4C3	5	5	10	0.008158763	0.2381	Phosphate-binding protein PstS OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=pstS PE=3 SV=1
A0R157	5	5	11	0.008154841	0.1298	Saccharopine dehydrogenase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_4632 PE=4 SV=1
A0QU53	2	2	10	0.00759609	0.0911	Acyl-CoA dehydrogenase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=fadE23 PE=3 SV=1
contaminant_KERATIN03	8	8	14	0.007280974	0.1686	
contaminant_KERATIN02	7	7	14	0.006941507	0.1897	
A0QSS4	6	6	12	0.006853361	0.1648	60 kDa chaperonin 2 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=groL2 PE=1 SV=1
A0R729	4	4	11	0.006717651	0.103	Glycerol kinase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=glpK PE=3 SV=1
A0R7G5	2	2	4	0.006704375	0.1196	Transcriptional regulator, PadR family protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_6903 PE=4 SV=1
A0QS98	3	3	8	0.006230328	0.1136	Elongation factor Tu OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=tuf PE=1 SV=1
A0R202	3	3	11	0.006190536	0.0821	ATP synthase subunit alpha OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=atpA PE=1 SV=1
A0QP15	2	2	3	0.005930793	0.2436	CoA binding domain protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_0238 PE=4 SV=1
A0R0B5	3	3	8	0.005916571	0.1127	3-oxoacyl-(Acyl-carrier-protein) synthase 1 KasA OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=kasB PE=3 SV=1
A0QSG6	2	2	4	0.005764509	0.1449	30S ribosomal protein S5 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpsE PE=1 SV=1
A0QSH0	7	7	11	0.005711134	0.1869	Putative protease IV Sppa OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=sppA PE=4 SV=1
A0QPE8	3	3	8	0.005711134	0.1343	3-ketoacyl-CoA thiolase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=fadA2 PE=3 SV=1
A0R5N7	2	2	6	0.005347999	0.1012	Aspartate-semialdehyde dehydrogenase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=asd PE=3 SV=1
A0QQW8	4	4	8	0.005317263	0.1401	Dihydrolipoyl dehydrogenase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=lpdA PE=3 SV=1
A0QS46	2	2	4	0.005249383	0.166	50S ribosomal protein L1 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rplA PE=1 SV=1
A0QZ46	2	2	4	0.005014654	0.1301	Proteasome subunit alpha OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=prcA PE=1 SV=1
A0QYU6	2	2	2	0.004781415	0.2326	50S ribosomal protein L20 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rplT PE=1 SV=1
A0QUV6	3	3	4	0.004690513	0.1711	Electron transfer flavoprotein beta subunit OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=etfB PE=4 SV=1
A0QNG7	5	5	7	0.004432872	0.1951	FHA domain protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_0035 PE=4 SV=1
A0R7J3	3	3	5	0.00439318	0.1681	ParB-like partition proteins OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=parB PE=4 SV=1
A0R648	2	2	4	0.004390053	0.0925	Glycerophosphoryl diester phosphodiesterase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=glpQ1 PE=4 SV=1
A0R465	6	6	10	0.004283351	0.1722	Enoyl-CoA hydratase/3-hydroxyacyl-CoA dehydrogenase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=fadB PE=4 SV=1
A0R042	4	4	8	0.004118881	0.0985	AMP-binding enzyme OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_4254 PE=4 SV=1
A0QT07	3	3	3	0.003544842	0.1034	Succinate dehydrogenase iron-sulfur protein subunit SdhB OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=sdhB PE=4 SV=1
A0QR33	3	3	5	0.003419083	0.1086	Glutamate-1-semialdehyde 2,1-aminomutase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=hemL PE=3 SV=1
A0QUV7	2	2	3	0.002955923	0.1406	Electron transfer flavoprotein, alpha subunit OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=etfA PE=4 SV=1
contaminant_KERATIN13	3	3	6	0.002877772	0.042	
A0QQ65	3	3	5	0.002798559	0.0817	ABC transporter solute-binding protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_0643 PE=4 SV=1
contaminant_KERATIN22	3	3	5	0.002390707	0.0651	
A0QQF9	2	2	2	0.002042392	0.1093	Uncharacterized protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_0741 PE=4 SV=1
A0R3L1	3	3	3	0.00200695	0.1497	Magnesium chelatase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_5512 PE=4 SV=1
A0QWX3	2	2	2	0.001983288	0.1479	D-ribose-binding periplasmic protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_3095 PE=4 SV=1
A0QYH8	3	3	3	0.001964339	0.0743	Oxidoreductase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_3663 PE=4 SV=1
A0R730	3	3	3	0.00162888	0.0651	Glycerol-3-phosphate dehydrogenase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_6761 PE=3 SV=1
Q3L887	2	2	3	0.001514245	0.0507	Acyl-CoA dehydrogenase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=fadE5 PE=3 SV=1
A0R5R5	2	2	2	0.001411447	0.0709	Cyclopropane fatty-acyl-phospholipid synthase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_6284 PE=4 SV=1

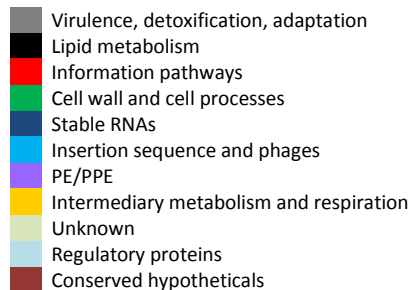| A0QU54 | 2 | 2 | 2 | 0.001395481 | 0.0905 | Putative acyl-CoA dehydrogenase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_2081 PE=3 SV=1 |
| P71533 | 2 | 2 | 3 | 0.000970833 | 0.0346 | Protein translocase subunit SecA 1 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=secA1 PE=1 SV=2 |
| A0R574 | 2 | 2 | 2 | 0.000727361 | 0.0366 | ATP-dependent Clp protease ATP-binding subunit ClpC1 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=clpC1 PE=3 SV=1 |

Table S2c. Proteins identified in biological replicate n°3

Proteins highlighted in grey are found in mock experiments

UniProt accesion number	#Unique peptides	Sequence count	Spectrum count	NSAF	Coverage	Description
A0QS97	1	6	35	0.063500905	0.4103	30S ribosomal protein S7 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpsG PE=1 SV=1
Q3I5Q7	8	8	42	0.051238661	0.2759	HBHA-like protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_0919 PE=4 SV=1
A0R5R3	5	5	26	0.050750674	0.4966	KanY protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_6282 PE=4 SV=1
A0R7F9	4	4	15	0.047172101	0.5556	30S ribosomal protein S6 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpsF PE=1 SV=1
A0QTS9	5	5	29	0.037308843	0.2364	Conserved domain protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_1951 PE=4 SV=1
A0QSD9	4	4	9	0.033081733	0.4805	50S ribosomal protein L29 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpmC PE=3 SV=1
A0QPE7	9	9	49	0.030819106	0.3044	3-oxoacyl-acyl-carrier protein reductase FabG4 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=fabG4 PE=1 SV=1
A0QZ11	2	2	12	0.029792906	0.2281	RNA polymerase-binding protein RbpA OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rbpA PE=1 SV=2
A0R3B8	10	10	40	0.026513593	0.3489	Enolase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=eno PE=1 SV=1
A0QS98	5	5	35	0.025015508	0.1768	Elongation factor Tu OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=tuf PE=1 SV=1
A0QRD4	3	3	8	0.024881987	0.2747	Uncharacterized protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_1076 PE=4 SV=1
A0QS62	4	4	15	0.024259937	0.3486	50S ribosomal protein L10 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rplJ PE=1 SV=1
A0QS46	6	6	20	0.024087881	0.3106	50S ribosomal protein L1 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rplA PE=1 SV=1
A0QQU5	12	12	45	0.023542453	0.2865	60 kDa chaperonin 1 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=groL1 PE=1 SV=1
A0R200	6	6	36	0.021450892	0.1263	ATP synthase subunit beta OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=atpD PE=1 SV=1
A0R5H1	4	4	16	0.020216615	0.2054	Crp/Fnr familytranscriptional regulator OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_6189 PE=4 SV=1
A0R204	2	2	12	0.019978772	0.1588	ATP synthase subunit b OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=atpF PE=3 SV=1
A0QS63	3	3	9	0.019594565	0.2846	50S ribosomal protein L7/L12 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rplL PE=3 SV=1
A0QXB0	4	4	19	0.018229219	0.2305	ABC-type amino acid transport system, secreted component OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_3235 PE=3 SV=1
A0QXT5	6	6	19	0.016805061	0.2188	Uncharacterized protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_3419 PE=4 SV=1
A0QVU2	5	5	15	0.015216807	0.2509	35 kDa protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_2695 PE=4 SV=1
A0R7F6	2	2	8	0.014995105	0.2053	50S ribosomal protein L9 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rplI PE=1 SV=1
P9WI73	12	12	39	0.014717695	0.1787	Serine/threonine-protein kinase PknG OS=Mycobacterium tuberculosis (strain ATCC 25618 / H37Rv) GN=pknG PE=1 SV=1
P71534	2	2	10	0.011099318	0.102	3-oxoacyl-[acyl-carrier-protein] reductase FabG OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=fabG PE=3 SV=2
contaminant_KERATIN13	6	6	25	0.011004378	0.0824	
A0R0B5	4	4	16	0.010859764	0.1127	3-oxoacyl-(Acyl-carrier-protein) synthase 1 KasA OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=kasB PE=3 SV=1
A0R7G5	2	2	7	0.010767545	0.1196	Transcriptional regulator, PadR family protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_6903 PE=4 SV=1
A0R729	5	5	19	0.010648751	0.1327	Glycerol kinase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=glpK PE=3 SV=1
A0R0W7	6	6	13	0.010542762	0.1519	Sulfate-binding lipoprotein SubI OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_4533 PE=4 SV=1
A0R157	6	6	15	0.010205503	0.1562	Saccharopine dehydrogenase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_4632 PE=4 SV=1
A0QQJ4	3	3	12	0.010108307	0.122	F420-dependent glucose-6-phosphate dehydrogenase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=fgd PE=1 SV=2
A0QNG7	4	4	16	0.009298812	0.1766	FHA domain protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_0035 PE=4 SV=1
A0QTT5	4	4	14	0.008448734	0.1258	Uncharacterized protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_1957 PE=4 SV=1
A0R4L2	4	4	8	0.008324488	0.0809	Conserved membrane protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_5873 PE=4 SV=1
A0R201	5	5	9	0.008297373	0.1661	ATP synthase gamma chain OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=atpG PE=1 SV=1
A0QVX3	4	4	8	0.008293996	0.1099	Glutamate binding periplasmic protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_2727 PE=3 SV=1
A0QSG6	3	3	6	0.007935494	0.2103	30S ribosomal protein S5 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpsE PE=1 SV=1
A0QTP2	0	2	6	0.007898584	0.1256	ECF RNA polymerase sigma factor SigH OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=sigH PE=1 SV=2
A0R1Z7	2	2	5	0.007448226	0.1842	ATP:cob(I)alamin adenosyltransferase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_4934 PE=4 SV=1

Accession						Description
A0QSS4	5	5	14	0.007337882	0.1463	60 kDa chaperonin 2 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=groL2 PE=1 SV=1
A0QYY6	3	3	12	0.007090587	0.0981	30S ribosomal protein S1 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpsA PE=3 SV=1
A0QYG0	0	2	4	0.006903234	0.1646	Uncharacterized protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_3645 PE=4 SV=1
A0QWT6	4	4	6	0.006086723	0.0932	Lipoprotein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_3058 PE=3 SV=1
A0QXC0	4	4	8	0.005850803	0.1654	Branched-chain amino acid ABC transporter substrate-binding protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_3247 PE=4 SV=1
A0R203	5	5	9	0.005724255	0.0899	ATP synthase subunit b-delta OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=atpFH PE=1 SV=1
A0QX24	4	4	7	0.005693185	0.1178	ATPase, MoxR family protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=moxR PE=4 SV=1
A0R202	3	3	11	0.005681311	0.073	ATP synthase subunit alpha OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=atpA PE=1 SV=1
A0R1H3	2	2	3	0.005408266	0.242	Antioxidant, AhpC/TSA family protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_4753 PE=4 SV=1
A0QSH0	5	5	11	0.005241345	0.1162	Putative protease IV Sppa OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=sppA PE=4 SV=1
A0R4C3	4	4	7	0.005241345	0.1402	Phosphate-binding protein PstS OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=pstS PE=3 SV=1
A0QU53	3	3	7	0.004879872	0.0665	Acyl-CoA dehydrogenase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=fadE23 PE=3 SV=1
A0QUV6	3	3	4	0.004304678	0.1711	Electron transfer flavoprotein beta subunit OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=etfB PE=4 SV=1
A0QQW8	4	4	7	0.004269888	0.1336	Dihydrolipoyl dehydrogenase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=lpdA PE=3 SV=1
A0QUV7	2	2	4	0.00361703	0.115	Electron transfer flavoprotein, alpha subunit OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=etfA PE=4 SV=1
A0R0B4	2	2	5	0.003401834	0.0649	3-oxoacyl-(Acyl-carrier-protein) synthase 1 KasA OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=kasA PE=3 SV=1
contaminant_KERATIN03	3	3	7	0.003341026	0.0742	
A0R042	3	3	7	0.00330756	0.0684	AMP-binding enzyme OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_4254 PE=4 SV=1
A0QPE8	4	4	5	0.00327584	0.0995	3-ketoacyl-CoA thiolase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=fadA2 PE=3 SV=1
A0QUX6	2	2	7	0.003205871	0.0518	Acetolactate synthase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=ilvB PE=3 SV=1
A0QQC8	3	3	7	0.003185254	0.0932	Chaperone protein DnaK OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=dnaK PE=1 SV=1
contaminant_KERATIN22	4	4	7	0.003071672	0.107	
A0R730	2	2	6	0.002989781	0.0458	Glycerol-3-phosphate dehydrogenase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_6761 PE=3 SV=1
A0QU54	3	3	4	0.002561381	0.1222	Putative acyl-CoA dehydrogenase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_2081 PE=3 SV=1
A0R5R5	2	2	2	0.001295344	0.0686	Cyclopropane fatty-acyl-phospholipid synthase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_6284 PE=4 SV=1
A0R465	2	2	2	0.000786202	0.05	Enoyl-CoA hydratase/3-hydroxyacyl-CoA dehydrogenase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=fadB PE=4 SV=1

Table S3a. PknG protein interactors recovered in one-step elution experiment

- ◼ (grey) Virulence, detoxification, adaptation
- ◼ (black) Lipid metabolism
- ◼ (red) Information pathways
- ◼ (green) Cell wall and cell processes
- ◼ (dark blue) Stable RNAs
- ◼ (light blue) Insertion sequence and phages
- ◼ (purple) PE/PPE
- ◼ (yellow) Intermediary metabolism and respiration
- ◼ (pale) Unknown
- ◼ (light blue) Regulatory proteins
- ◼ (brown) Conserved hypotheticals

UniProt accesion number	Description	Replicate count	Gene name in *M. smegmatis*	Gene ortholog name in *M. tuberculosis*	Functional classification in *M. tuberculosis*										
A0QQC8	Chaperone protein DnaK OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=dnaK PE=1 SV=1	3	MSMEG_0709	Rv0350	X										
A0QPE8	3-ketoacyl-CoA thiolase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=fadA2 PE=3 SV=1	3	MSMEG_0373	Rv0243		X									
A0R465	Enoyl-CoA hydratase/3-hydroxyacyl-CoA dehydrogenase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=fadB PE=4 SV=1	3	MSMEG_5720	Rv0860		X									
A0QU53	Acyl-CoA dehydrogenase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=fadE23 PE=3 SV=1	3	MSMEG_2080	Rv3140		X									
A0QPE7	3-oxoacyl-acyl-carrier protein reductase FabG4 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=fabG4 PE=1 SV=1	3	MSMEG_0372	Rv0242c		X									
A0QU54	Putative acyl-CoA dehydrogenase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_2081 PE=3 SV=1	3	MSMEG_2081	Rv3139		X									
A0R042	AMP-binding enzyme OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_4254 PE=4 SV=1	3	MSMEG_4254	Rv2187		X									
A0R5R5	Cyclopropane fatty-acyl-phospholipid synthase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_6284 PE=4 SV=1	3	MSMEG_6284	Rv3720		X									
A0R0B5	3-oxoacyl-(Acyl-carrier-protein) synthase 1 KasA OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=kasB PE=3 SV=1	2	MSMEG_4328	Rv2246		X									
P71534	3-oxoacyl-[acyl-carrier-protein] reductase FabG OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=fabG PE=3 SV=2	2	MSMEG_3150	Rv1483		X									
A0R0B4	3-oxoacyl-(Acyl-carrier-protein) synthase 1 KasA OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=kasA PE=3 SV=1	2	MSMEG_4327	Rv2245		X									
A0QYY6	30S ribosomal protein S1 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpsA PE=3 SV=1	3	MSMEG_3833	Rv1630			X								
A0QS46	50S ribosomal protein L1 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rplA PE=1 SV=1	3	MSMEG_1347	Rv0641			X								
A0QS63	50S ribosomal protein L7/L12 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rplL PE=3 SV=1	3	MSMEG_1365	Rv0652			X								

A0R7F6	50S ribosomal protein L9 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rplI PE=1 SV=1	3	MSMEG_6894	Rv0056	**X**	
A0QSG6	30S ribosomal protein S5 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpsE PE=1 SV=1	3	MSMEG_1472	Rv0721	**X**	
A0R7F9	30S ribosomal protein S6 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpsF PE=1 SV=1	3	MSMEG_6897	Rv0053	**X**	
A0QSG3	30S ribosomal protein S8 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpsH PE=1 SV=1; Additional IDs concatenated into MaxParsimony group: P9WH27	2	MSMEG_1469	Rv0718	**X**	
A0QWS8	Integration host factor OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=mihF PE=4 SV=1	2	MSMEG_3050	Rv1388	**X**	
P9WH29	30S ribosomal protein S7 OS=Mycobacterium tuberculosis (strain ATCC 25618 / H37Rv) GN=rpsG PE=1 SV=1	2	Rv0683	Rv0683	**X**	
Q3I5Q7	HBHA-like protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_0919 PE=4 SV=1	3	MSMEG_0919	Rv0475		**X**
A0R4C3	Phosphate-binding protein PstS OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=pstS PE=3 SV=1	3	MSMEG_5782	Rv0928 [1]		**X**
A0QSH0	Putative protease IV Sppa OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=sppA PE=4 SV=1	3	MSMEG_1476	Rv0724A		**X**
A0R0W7	Sulfate-binding lipoprotein SubI OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_4533 PE=4 SV=1	3	MSMEG_4533	Rv2400c		**X**
A0QVX3	Glutamate binding periplasmic protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_2727 PE=3 SV=1	2	MSMEG_2727	---		**X**
A0R7J3	ParB-like partition proteins OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=parB PE=4 SV=1	2	MSMEG_6938	Rv3917c		**X**
A0R729	Glycerol kinase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=glpK PE=3 SV=1	3	MSMEG_6759	Rv3696c [2]			**X**
A0R203	ATP synthase subunit b-delta OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=atpFH PE=1 SV=1	3	MSMEG_4939	Rv1307			**X**
A0R201	ATP synthase gamma chain OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=atpG PE=1 SV=1	3	MSMEG_4937	Rv1309			**X**
A0QQW8	Dihydrolipoyl dehydrogenase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=lpdA PE=3 SV=1	3	MSMEG_0903	Rv0462			**X**
A0R204	ATP synthase subunit b OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=atpF PE=3 SV=1	3	MSMEG_4940	Rv1306			**X**
A0R202	ATP synthase subunit alpha OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=atpA PE=1 SV=1	3	MSMEG_4938	Rv1308			**X**
A0R730	Glycerol-3-phosphate dehydrogenase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_6761 PE=3 SV=1	3	MSMEG_6761	Rv3302c [3]			**X**
A0R3B8	Enolase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=eno PE=1 SV=1	3	MSMEG_5415	Rv1023			**X**
A0R200	ATP synthase subunit beta OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=atpD PE=1 SV=1	3	MSMEG_4936	Rv1310			**X**
A0QUX6	Acetolactate synthase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=ilvB PE=3 SV=1	2	MSMEG_2372	Rv3003c			**X**
A0QUV6	Electron transfer flavoprotein beta subunit OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=etfB PE=4 SV=1	2	MSMEG_2351	Rv3029c			**X**
A0QQJ4	F420-dependent glucose-6-phosphate dehydrogenase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=fgd PE=1 SV=2	2	MSMEG_0777	Rv0407			**X**
A0R3L1	Magnesium chelatase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_5512 PE=4 SV=1	2	MSMEG_5512	Rv0958			**X**

153

UniProt	Description		MSMEG	Rv		
A0QUV7	Electron transfer flavoprotein alpha subunit OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=etfA PE=4 SV=1	2	MSMEG_2352	Rv3028c	X	
A0QT07	Succinate dehydrogenase iron-sulfur protein subunit SdhB OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=sdhB PE=4 SV=1	2	MSMEG_1669	Rv3319	X	
A0QXC0	Branched-chain amino acid ABC transporter substrate-binding protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_3247 PE=4 SV=1	3	MSMEG_3247	---	X	
A0QTS9	Conserved domain protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_1951 PE=4 SV=1	3	MSMEG_1951	---	X	
A0QXB0	ABC-type amino acid transport system, secreted component OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_3235 PE=3 SV=1	3	MSMEG_3235	---	X	
A0QWT6	Lipoprotein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_3058 PE=3 SV=1	3	MSMEG_3058	---	X	
A0QXT5	Uncharacterized protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_3419 PE=4 SV=1	3	MSMEG_3419	---	X	
A0QYH8	Oxidoreductase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_3663 PE=4 SV=1	2	MSMEG_3663	---	X	
A0QPX3	Aliphatic sulfonates family ABC transporter, periplasmic ligand-binding protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_0550 PE=4 SV=1	2	MSMEG_0550	---	X	
A0QP15	CoA binding domain protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_0238 PE=4 SV=1	2	MSMEG_0238	---	X	
A0QQ65	ABC transporter solute-binding protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_0643 PE=4 SV=1	2	MSMEG_0643	---	X	
A0QX24	ATPase, MoxR family protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=moxR PE=4 SV=1	3	MSMEG_3147	Rv1479		X
A0R5H1	Crp/Fnr familytranscriptional regulator OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_6189 PE=4 SV=1	3	MSMEG_6189	Rv3676		X
A0QNG7	FHA domain protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_0035 PE=4 SV=1	3	MSMEG_0035	Rv0020c		X
A0R656	Uncharacterized protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_6431 PE=4 SV=1	2	MSMEG_6431	Rv3849		X
A0QVU2	35 kDa protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_2695 PE=4 SV=1	3	MSMEG_2695	Rv2744c		X
A0R1Z7	ATP:cob(I)alamin adenosyltransferase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_4934 PE=4 SV=1	3	MSMEG_4934	Rv1314c		X
A0R157	Saccharopine dehydrogenase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_4632 PE=4 SV=1	3	MSMEG_4632	Rv2449c		X
A0R7G5	Transcriptional regulator, PadR family protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_6903 PE=4 SV=1	3	MSMEG_6903	Rv0047c		X
A0R4L2	Conserved membrane protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_5873 PE=4 SV=1	3	MSMEG_5873	Rv0756c		X
A0QTT5	Uncharacterized protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_1957 PE=4 SV=1	3	MSMEG_1957	Rv3195		X
A0R5R3	KanY protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_6282 PE=4 SV=1	3	MSMEG_6282	Rv3718c		X
A0R1B5	Uncharacterized protein MSMEG_4692/MSMEI_4575 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_4692 PE=1 SV=1	2	MSMEG_4692	Rv2468c		X
A0QZ11	RNA polymerase-binding protein RbpA OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rbpA PE=1 SV=2	2	MSMEG_3858	Rv2050		X

Rv0928 [1]   50% identity
Rv 3696c [2] 77% identity
Rv3302c [3] 60% identity

Table S3b. Operon ID of proteins recovered in one-step elution experiment

Highlighted genes belongs to the same operon according to DOOR[2]

Gene name in *M. smegmatis*	Operon ID	Description
MSMEG_0035	1917196	
MSMEG_0238	242478	
MSMEG_0372	242509	
MSMEG_0373	1915994	
MSMEG_0550	242543	
MSMEG_0643	242562	
MSMEG_0709	242576	
MSMEG_0777	1917567	
MSMEG_0903	242616	
MSMEG_0919	1916099	
MSMEG_1347	242704	
MSMEG_1365	1916150	
MSMEG_1469	242729	30S ribosomal protein S8 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpsH PE=1 SV=1; Additional IDs concatenated into MaxParsimony group: P9WH27
MSMEG_1472	242729	30S ribosomal protein S5 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpsE PE=1 SV=1
MSMEG_1476	1915803	
MSMEG_1669	242777	
MSMEG_1951	242841	
MSMEG_1957	1916180	
MSMEG_2080	242864	Acyl-CoA dehydrogenase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=fadE23 PE=3 SV=1
MSMEG_2081	242864	Putative acyl-CoA dehydrogenase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_2081 PE=3 SV=1
MSMEG_2351	242928	Electron transfer flavoprotein beta subunit OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=etfB PE=4 SV=1
MSMEG_2352	242928	Electron transfer flavoprotein, alpha subunit OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=etfA PE=4 SV=1
MSMEG_2372	242933	
MSMEG_2695	243016	
MSMEG_2727	243025	
MSMEG_3050	243091	
MSMEG_3058	243093	
MSMEG_3147	243113	
MSMEG_3150	243114	
MSMEG_3235	243134	
MSMEG_3247	243136	
MSMEG_3419	1917030	
MSMEG_3663	243211	
MSMEG_3833	1917541	
MSMEG_3858	1917238	
MSMEG_4254	1916905	
MSMEG_4327	243342	3-oxoacyl-(Acyl-carrier-protein) synthase 1 KasA OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=kasA PE=3 SV=1
MSMEG_4328	243342	3-oxoacyl-(Acyl-carrier-protein) synthase 1 KasA OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=kasB PE=3 SV=1
MSMEG_4533	243382	
MSMEG_4632	243401	
MSMEG_4692	243413	
MSMEG_4934	1915822	

MSMEG_4936	243468	ATP synthase subunit beta OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=atpD PE=1 SV=1
MSMEG_4937	243468	ATP synthase gamma chain OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=atpG PE=1 SV=1
MSMEG_4938	243468	ATP synthase subunit alpha OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=atpA PE=1 SV=1
MSMEG_4939	243468	ATP synthase subunit b-delta OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=atpFH PE=1 SV=1
MSMEG_4940	243468	ATP synthase subunit b OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=atpF PE=3 SV=1
MSMEG_5415	243564	
MSMEG_5512	243582	
MSMEG_5720	243632	
MSMEG_5782	243639	
MSMEG_5873	1917521	
MSMEG_6189	1916303	
MSMEG_6282	1916959	
MSMEG_6284	243744	
MSMEG_6431	1917910	
MSMEG_6759	243851	Glycerol kinase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=glpK PE=3 SV=1
MSMEG_6761	243851	Glycerol-3-phosphate dehydrogenase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_6761 PE=3 SV=1
MSMEG_6894	243880	
MSMEG_6897	243881	
MSMEG_6903	1917029	
MSMEG_6938	243889	
Rv0683		

UniProt accesion number	Gene name in *M. smegmatis*	Phosphopeptides (unambigously phosphorylated residues are indicated in bold red)	Site(s) of phosphorylation	Gene ortholog name in *M. tuberculosis*	Phosphopeptides (unambigously phosphorylated residues are indicated in bold red)	Site(s) of phosphorylation		
			M. smegmatis			*M. bovis* BCG	*M. tuberculosis* H37Rv	*M. tuberculosis* SAW5527 (clinical isolate)
A0QSG6	MSMEG_1472	ESEALAAAAAREG**S**A *	S213 *					
A0QNG7	MSMEG_0035	FEQSPNLH**T**GQFR *	T118 *	Rv0020c	FEQ**S**SNLH**T**GQFR	T116 **	T116 **	S111, T116 **
					GGQGQGRPDE**Y**YDD	Y174 **		Y174 **
					GGYPPETGGYPPQPG**Y**PRPR	Y215 **		Y215 **
					HPGQGD**Y**PEQIG**Y**PDQGGYPEQR	Y232 **		Y226 **
A0QQC8	MSMEG_0709			Rv0350	AALGG**S**DI**S**AIK	S558 **		S561 **
A0R042	MSMEG_4254			Rv2187	NEQATTEAF**T**DGWFK		T437 **	T437 **
A0R656	MSMEG_6431			Rv3849	AHGLP**S**AAQQK	S111 **		S111 **
A0QVU2	MSMEG_2695			Rv2744c	YANAIGSAELAE**S**SVQGR	S209 **		S209 **

* extracted from Nakedi, K. *et al.* (2015) Front Microbiol. 6, 237
** extracted from Calder, B. et al. (2016) Front Microbiol 7, 141

Table S3d. PANTHER classification of putative PknG interactors

	Category name	# genes	Percent of gene hit against total # genes	Percent of gene hit against total # Protein Class hits
1	transporter (PC00227)	3	6.30%	6.80%
2	hydrolase (PC00121)	3	6.30%	6.80%
3	oxidoreductase (PC00176)	13	27.10%	29.50%
4	lyase (PC00144)	2	4.20%	4.50%
5	transferase (PC00220)	9	18.80%	20.50%
6	ligase (PC00142)	2	4.20%	4.50%
7	nucleic acid binding (PC00171)	8	16.70%	18.20%
8	receptor (PC00197)	2	4.20%	4.50%
9	kinase (PC00137)	1	2.10%	2.30%
10	isomerase (PC00135)	1	2.10%	2.30%

Table S4a. Proteins identified as exclusive interactors of full-length PknG recovered under phosphorylation conditions of elution (*E1*)

- ■ Virulence, detoxification, adaptation
- ■ Lipid metabolism
- ■ Information pathways
- ■ Cell wall and cell processes
- ■ Stable RNAs
- ■ Insertion sequence and phages
- ■ PE/PPE
- ■ Intermediary metabolism and respiration
- ■ Unknown
- ■ Regulatory proteins
- ■ Conserved hypotheticals
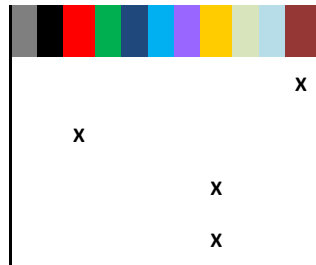
UniProt accesion number	Description	Replicate count	Gene name in *M. smegmatis*	Gene ortholog name in *M. tuberculosis*											
															Functional classification in *M. tuberculosis*
A0QYG2	Glycogen accumulation regulator GarA OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=garA PE=1 SV=2	3	MSMEG_3647	Rv1827											X
A0QWS8	Integration host factor OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=mihF PE=4 SV=1	2	MSMEG_3050	Rv1388			X								
A0QUV6	Electron transfer flavoprotein beta subunit OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=etfB PE=4 SV=1	2	MSMEG_2351	Rv3029c								X			
A0R079	Glutamine synthetase 1 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=glnA PE=1 SV=1	2	MSMEG_4290	Rv2220								X			

Table S4b. Proteins identified as common interactors of PknG and PknG$_{\Delta73}$ recovered under phosphorylation conditions of elution (*E1*)

Functional classification colour key:

- ▉ Virulence, detoxification, adaptation
- ▉ Lipid metabolism
- ▉ Information pathways
- ▉ Cell wall and cell processes
- ▉ Stable RNAs
- ▉ Insertion sequence and phages
- ▉ PE/PPE
- ▉ Intermediary metabolism and respiration
- ▉ Unknown
- ▉ Regulatory proteins
- ▉ Conserved hypotheticals

UniProt accesion number	Description	Replicate count	Gene name in *M. smegmatis*	Gene ortholog name in *M. tuberculosis*	Virulence, detoxification, adaptation	Lipid metabolism	Information pathways	Cell wall and cell processes	Stable RNAs	Insertion sequence and phages	PE/PPE	Intermediary metabolism and respiration	Unknown	Regulatory proteins	Conserved hypotheticals
A0R2I1	4Fe-4S ferredoxin, iron-sulfur binding protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=fdxC PE=4 SV=1	5	MSMEG_5122	Rv1177								X			
A0R597	Inorganic pyrophosphatase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=ppa PE=3 SV=1	6	MSMEG_6114	Rv3628								X			
A0R5M3	Alcohol dehydrogenase, iron-containing OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_6242 PE=4 SV=1	6	MSMEG_6242	---											X
A0QQC8	Chaperone protein DnaK OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=dnaK PE=1 SV=1	6	MSMEG_0709	Rv0350	X										
A0QQU5	60 kDa chaperonin 1 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=groL1 PE=1 SV=1	5	MSMEG_0880	Rv0440	X										
P9WI73	Serine/threonine-protein kinase PknG OS=Mycobacterium tuberculosis (strain ATCC 25618 / H37Rv) GN=pknG PE=1 SV=1	5	Rv0410c	Rv0410c										X	
A0QTE1	Acetyl-/propionyl-coenzyme A carboxylase alpha chain OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=accA3 PE=4 SV=1	5	MSMEG_1807	Rv3285		X									

Table S4c. Reported phosphosites for proteins identified as putative PknG substrates in *M. smegmatis* and other mycobacterial phosphoproteomes

UniProt accesion number	Gene name in *M. smegmatis*	Phosphopeptides (unambigously phosphorylated residues are indicated in bold red)	Site(s) of phosphorylation	Gene ortholog name in *M. tuberculosis*	Phosphopeptides (unambigously phosphorylated residues are indicated in bold red)	Site(s) of phosphorylation		
			M. smegmatis			*M. bovis* BCG	*M. tuberculosis* H37Rv	*M. tuberculosis* SAW5527 (clinical isolate)
A0R2I1	MSMEG_5122			Rv1177				
A0R597	MSMEG_6114			Rv3628	HFFVH**Y**K			Y126 **
A0R5M3	MSMEG_6242			---				
A0QQC8	MSMEG_0709			Rv0350	AALGG**S**DI**S**AIK	S558 *		S561 **
A0QQU5	MSMEG_0880			Rv0440	AMLQDMAIL**T**GGQVI**S**EEVGLTLENADLSLLGK	T294, S300 *	T294, S300 ***	
					AVEKV**T**ETLLK		T127 ***	
					EQIAA**T**AAI**S**AGDQSIGDLIAEAMDK	T146 *	T146, S150 ***	
					KWGAP**TIT**NDGV**S**IAK	T47, T49, S54 *		
					WGAPTI**T**NDGV**S**IAK	S54 *		T49 **
P9WI73	Rv0410c			Rv0410c				
A0QTE1	MSMEG_1807			Rv3285				
A0QYG2	MSMEG_3647	ADFLNELDAPAAAG**T**EGAVSGVEGLPSGSALLVVK	T40 *	Rv1827	DQ**T**SDEV**T**VE**TT**SVFR		T21 ***	T13, T18, T21, T22 **
					TDMNPDIEKDQTSDEVTVE**T**TSVFR	T21 *	T21 ***	
A0QWS8	MSMEG_3050			Rv1388	AQEIMTELEIAP**T**RR		T169 **	
A0QUV6	MSMEG_2351			Rv3029c				
A0R079	MSMEG_4290			Rv2220	VN**SY**KRLVPGYEAPINLVYSQR		S326, Y327 ****	

* extracted from Nakedi, K. *et al.* (2015) Front Microbiol. 6, 237
** extracted from Fortuin, S. et al. (2015) Front Microbiol. 6, 6
*** extracted from Prisic, S. et al. (2010) Proc Natl Acad Sci U S A. 107, 7521-752
**** extracted from Kusebauch, U. et al. (2014) Proc Natl Acad Sci U S A. 111, 9265-9270

Table S5a. Proteins identified as exclusive interactors of full-length PknG recovered under dephosphorylation conditions of elution (*E2*)

■ Virulence, detoxification, adaptation
■ Lipid metabolism
■ Information pathways
■ Cell wall and cell processes
■ Stable RNAs
■ Insertion sequence and phages
■ PE/PPE
■ Intermediary metabolism and respiration
■ Unknown
■ Regulatory proteins
■ Conserved hypotheticals
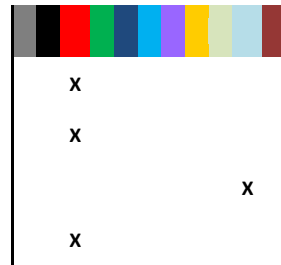
UniProt accesion number	Description	Replicate count	Gene name in *M. smegmatis*	Gene ortholog name in *M. tuberculosis*	Functional classification in *M. tuberculosis*										
A0QSE0	30S ribosomal protein S17 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpsQ PE=1 SV=1	3	MSMEG_1445	Rv0710			X								
A0QWS8	Integration host factor OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=mihF PE=4 SV=1	2	MSMEG_3050	Rv1388			X								
A0QNG7	FHA domain protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_0035 PE=4 SV=1	3	MSMEG_0035	Rv0020c											X
A0QSD9	50S ribosomal protein L29 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpmC PE=3 SV=1	2	MSMEG_1444	Rv0709			X								

Table S5b. Proteins identified as exclusive interactors of PknG$_{\Delta73}$ recovered under dephosphorylation conditions of elution (*E2*)
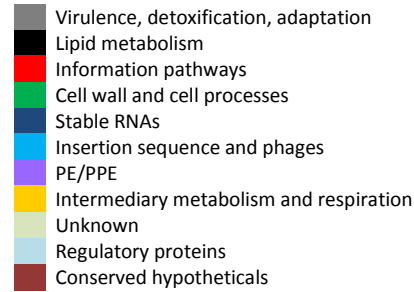
■ Virulence, detoxification, adaptation
■ Lipid metabolism
■ Information pathways
■ Cell wall and cell processes
■ Stable RNAs
■ Insertion sequence and phages
■ PE/PPE
■ Intermediary metabolism and respiration
■ Unknown
■ Regulatory proteins
■ Conserved hypotheticals

UniProt accesion number	Description	Replicate count	Gene name in *M. smegmatis*	Gene ortholog name in *M. tuberculosis*	Functional classification in *M. tuberculosis*										
A0QUV6	Electron transfer flavoprotein beta subunit OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=etfB PE=4 SV=1	2	MSMEG_2351	Rv3029c								X			
A0R7J6	R3H domain-containing protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_6941 PE=4 SV=1	2	MSMEG_6941	Rv3920c											X
A0QU54	Putative acyl-CoA dehydrogenase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_2081 PE=3 SV=1	2	MSMEG_2081	Rv3139		X									
A0QSS4	60 kDa chaperonin 2 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=groL2 PE=1 SV=1	2	MSMEG_1583	Rv3417c	X										
A0QSZ3	Isocitrate dehydrogenase (NADP) Icd2 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=icd2 PE=4 SV=1	2	MSMEG_1654	Rv0066c								X			
A0R0Q9	Acyl-CoA oxidase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_4474 PE=4 SV=1	2	MSMEG_4474	---								X			
A0QW71	Transcriptional accessory protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_2839 PE=4 SV=1	2	MSMEG_2839	---								X			

Color legend:
- ⬛ (grey) Virulence, detoxification, adaptation
- ⬛ (black) Lipid metabolism
- 🟥 Information pathways
- 🟩 Cell wall and cell processes
- 🟦 Stable RNAs
- 🟦 Insertion sequence and phages
- 🟪 PE/PPE
- 🟨 Intermediary metabolism and respiration
- ⬜ Unknown
- 🟦 Regulatory proteins
- 🟫 Conserved hypotheticals

UniProt accesion number	Description	Replicate count	Gene name in *M. smegmatis*	Gene ortholog name in *M. tuberculosis*	Virulence, detoxification, adaptation	Lipid metabolism	Information pathways	Cell wall and cell processes	Stable RNAs	Insertion sequence and phages	PE/PPE	Intermediary metabolism and respiration	Unknown	Regulatory proteins	Conserved hypotheticals
A0QQC8	Chaperone protein DnaK OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=dnaK PE=1 SV=1	5	MSMEG_0709	Rv0350	X										
A0QPE7	3-oxoacyl-acyl-carrier protein reductase FabG4 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=fabG4 PE=1 SV=1	6	MSMEG_0372	Rv0242c		X									
A0R5R5	Cyclopropane fatty-acyl-phospholipid synthase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_6284 PE=4 SV=1	5	MSMEG_6284	Rv3720		X									
A0QTE7	Propionyl-CoA carboxylase beta chain OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=accD5 PE=4 SV=1; Additional IDs concatenated into MaxParsimony group: P9WQH7	4	MSMEG_1813	Rv3280		X									
A0R042	AMP-binding enzyme OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_4254 PE=4 SV=1	4	MSMEG_4254	Rv2187		X									
A0QVQ3	30S ribosomal protein S15 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpsO PE=1 SV=1	6	MSMEG_2654	Rv2785c			X								
A0QYU6	50S ribosomal protein L20 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rplT PE=1 SV=1	6	MSMEG_3791	Rv1643			X								
A0QSG3	30S ribosomal protein S8 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpsH PE=1 SV=1	6	MSMEG_1469	Rv0718			X								
A0QSD1	50S ribosomal protein L3 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rplC PE=1 SV=1	6	MSMEG_1436	Rv0701			X								
A0R7F9	30S ribosomal protein S6 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpsF PE=1 SV=1	5	MSMEG_6897	Rv0053			X								
Q9ZHC5	DNA-binding protein HU homolog OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=hup PE=3 SV=1	6	MSMEG_2389	Rv2986c			X								
A0QSG5	50S ribosomal protein L18 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rplR PE=3 SV=1	6	MSMEG_1471	Rv0720			X								
A0QSD4	50S ribosomal protein L2 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rplB PE=1 SV=1	5	MSMEG_1439	Rv0704			X								
A0QSL5	30S ribosomal protein S13 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155)	5	MSMEG_1521	Rv3460c			X								

Functional classification in *M. tuberculosis*

	GN=rpsM PE=1 SV=1					
A0QSP8	50S ribosomal protein L13 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rplM PE=1 SV=1	6	MSMEG_1556	Rv3443c	X	
A0QS97	30S ribosomal protein S7 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpsG PE=1 SV=1; Additional IDs concatenated into MaxParsimony group: P9WH29	5	MSMEG_1399	Rv0683	X	
A0R102	30S ribosomal protein S20 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpsT PE=1 SV=1	4	MSMEG_4571	Rv2412	X	
A0QSD3	50S ribosomal protein L23 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rplW PE=1 SV=1	4	MSMEG_1438	Rv0703	X	
A0R151	50S ribosomal protein L21 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rplU PE=1 SV=1	4	MSMEG_4625	Rv2442c	X	
A0QYY6	30S ribosomal protein S1 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpsA PE=3 SV=1	5	MSMEG_3833	Rv1630	X	
A0R7F6	50S ribosomal protein L9 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rplI PE=1 SV=1	4	MSMEG_6894	Rv0056	X	
A0QSG0	50S ribosomal protein L24 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rplX PE=1 SV=1; Additional IDs concatenated into MaxParsimony group: P9WHB7	5	MSMEG_1466	Rv0715	X	
A0QV37	30S ribosomal protein S16 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpsP PE=1 SV=1	4	MSMEG_2435	Rv2909c	X	
A0QSG4	50S ribosomal protein L6 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rplF PE=1 SV=1	4	MSMEG_1470	Rv0719	X	
Q9AFI5	Single-stranded DNA-binding protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=ssb PE=1 SV=1	4	MSMEG_6896	Rv0054	X	
A0QSP9	30S ribosomal protein S9 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpsI PE=1 SV=1	4	MSMEG_1557	Rv3442c	X	
A0QSD6	50S ribosomal protein L22 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rplV PE=1 SV=1	4	MSMEG_1441	Rv0706	X	
A0QVB8	30S ribosomal protein S2 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpsB PE=3 SV=2	5	MSMEG_2519	Rv2890c	X	
A0QSD2	50S ribosomal protein L4 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rplD PE=1 SV=1	4	MSMEG_1437	Rv0702	X	
A0QVQ5	Polyribonucleotide nucleotidyltransferase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=pnp PE=1 SV=1	5	MSMEG_2656	Rv2783c	X	
A0QSL7	30S ribosomal protein S4 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpsD PE=1 SV=1	5	MSMEG_1523	Rv3458c	X	
A0QSG8	50S ribosomal protein L15 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rplO PE=3 SV=1	5	MSMEG_1474	Rv0723	X	
A0QS46	50S ribosomal protein L1 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rplA PE=1 SV=1	4	MSMEG_1347	Rv0641	X	
A0R3D2	50S ribosomal protein L25 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rplY PE=3 SV=1	5	MSMEG_5431	Rv1015c	X	
A0QSL8	DNA-directed RNA polymerase subunit alpha OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpoA PE=1 SV=1	4	MSMEG_1524	Rv3457c	X	
A0R5D9	DNA topoisomerase 1 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=topA PE=3 SV=1	4	MSMEG_6157	Rv3646c	X	
A0R1Z9	ATP synthase epsilon chain OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=atpC PE=1 SV=1	6	MSMEG_4935	Rv1311		X
A0R079	Glutamine synthetase 1 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155)	6	MSMEG_4290	Rv2220		X

A0R3B8	GN=glnA PE=1 SV=1 Enolase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=eno PE=1 SV=1	6	MSMEG_5415	Rv1023	X
A0QQW8	Dihydrolipoyl dehydrogenase OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=lpdA PE=3 SV=1	4	MSMEG_0903	Rv0462	X
A0R200	ATP synthase subunit beta OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=atpD PE=1 SV=1; Additional IDs concatenated into MaxParsimony group: P9WPU5	4	MSMEG_4936	Rv1310	X
A0R5M3	Alcohol dehydrogenase, iron-containing OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_6242 PE=4 SV=1	6	MSMEG_6242	---	X
A0QXT5	Uncharacterized protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_3419 PE=4 SV=1	5	MSMEG_3419	---	X
P9WI73	Serine/threonine-protein kinase PknG OS=Mycobacterium tuberculosis (strain ATCC 25618 / H37Rv) GN=pknG PE=1 SV=1	5	Rv0410c	Rv0410c	X
A0R061	HesB/YadR/YfhF family protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_4272 PE=4 SV=1; Additional IDs concatenated into MaxParsimony group: P9WMN5	6	MSMEG_4272	Rv2204c	X
A0R1B5	Uncharacterized protein MSMEG_4692/MSMEI_4575 OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_4692 PE=1 SV=1	4	MSMEG_4692	Rv2468c	X
A0QT22	Uncharacterized protein OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_1684 PE=4 SV=1	4	MSMEG_1684	Rv3311	X
A0QVX6	ATPase involved in DNA repair OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=MSMEG_2731 PE=4 SV=1	4	MSMEG_2731	Rv2731	X

VII. AGRADECIMIENTOS