Estructura y función de las adhesinas de micoplasma del grupo *Pneumoniae*

David Vizarraga Revuelto

UNIVERSITAT DE BARCELONA

UNIVERSIDAD DE BARCELONA
FACULTAD DE FARMACIA Y CIENCIAS DE LA ALIMENTACIÓN
DEPARTAMENTO DE BIOQUÍMICA Y FISIOLOGÍA

PROGRAMA DE DOCTORADO EN BIOTECNOLOGÍA

ESTRUCTURA Y FUNCIÓN DE LAS ADHESINAS DE MICOPLASMA DEL GRUPO *PNEUMONIAE*

David Vizarraga Revuelto

Barcelona, 2020

UNIVERSITAT DE BARCELONA

UNIVERSIDAD DE BARCELONA
FACULTAD DE FARMACIA Y CIENCIAS DE LA ALIMENTACIÓN
DEPARTAMENTO DE BIOQUÍMICA Y FISIOLOGÍA

Memoria presentada por **David Vizarraga Revuelto** para optar al grado de Doctor por la Universidad de Barcelona

Esta Tesis Doctoral ha sido realizada bajo la dirección del **Dr. David Aparicio** y el **Dr. Ignacio Fita**, en el Instituto de Biología Molecular de Barcelona (CSIC). Tutora de la tesis, la **Dra. Josefa Badía**

Barcelona, 2020

Agradecimientos

La vida es una sucesión de elecciones, tanto buenas como malas. Durante este tiempo han ido sucediendo constantes elecciones que me han llevado a ser la persona que soy hoy en día. Sin ir más lejos, la elección de estudiar la carrera de biotecnología en la universidad de Zaragoza fue de rebote, yo quería ser matemático, pero mi padre me convenció para que fuera biotecnólogo. Visto a día de hoy, le tengo que agradecer esas palabras en su día "Hijo, ¿Por qué no te apuntas a biotecnología? Es el futuro". Gracias a él, he estudiado algo que realmente me ha llegado a apasionar en algunos aspectos y que me ha permitido conocer a personas INCREÍBLES tanto en el grado, como en el máster y ahora en el doctorado. Aunque muy dentro de mí, sigan esas ganas por aprender matemáticas.

En los tiempos que corren ahora mismo, vas por la calle y escuchas comentarios como "¿Cuándo van a sacar la vacuna? A este paso nos morimos todos" o "¿Para qué tanta investigación si luego tardan 1 año en sacar una vacuna?". Me hacen mucha gracia este tipo de comentarios, porque realmente cuando estás dentro de cualquier centro de investigación te das cuenta de que hay personas con gran inteligencia que realmente aman su trabajo, se dedican a tiempo completo para él y tienen esa capacidad "especial" de reconducir aquellos fracasos en futuros éxitos. En esta profesión, abundan los fracasos y escasean los éxitos y para que un resultado positivo se de, han tenido que pasar horas y horas de pensamientos en tu casa, días y días de trabajos en el laboratorio y meses y meses de análisis exhaustivos. Por lo que, espero que este párrafo sirva para que aquel que no pertenezca al mundo de la ciencia, se de cuenta del sacrificio que desempeña.

Hace ya 5 años que entré por primera vez por las puertas del Instituto de biología molecular de Barcelona (IBMB). Me acuerdo de que en esa época entraba un chico con miedos e inseguridades a partir de ese momento, ya que me enfrentaba a un reto muy

complejo. Por ello, quiero agradecer a todas y cada una de las personas que a lo largo de este periodo de tiempo me han ayudado tanto laboral como personalmente.

Para empezar, quería agradecer a mi director de tesis el **Dr. Ignacio Fita**. Verdaderamente ha sido un honor trabajar con usted y ya sabe que ojalá podamos encontrarnos en un futuro no muy lejano. Aunque todo no fue como se dice "un paseo de rosas", tengo que agradecerle muchas cosas, pero las que más destaco es que gracias a usted entré siendo un mero estudiante y he salido con esa capacidad de pensar y ser un mejor científico. También, gracias por las numerosas charlas/emails acerca de los micoplasmas, cristalografía y sobre el mundo científico, las cuales me han servido para adquirir muchos conocimientos. Gracias por aguantarme y tranquilizarme aquellos días que podía estar más nervioso o impaciente. Por último, quiero decirle que no me llevo un jefe, sino que me llevo a un amigo al cual admiro y que aprendí muchísimo de usted. También, quería agradecer a mi otro director de tesis y compañero **Dr. David Aparicio** por su ayuda al iniciar este camino de la tesis doctoral, por guiarme a lo largo del recorrido y por todas esas charlas que mantuvimos acerca del micoplasma. Espero que nos veamos pronto y nunca olvidaré ese confinamiento preparando el artículo a contrarreloj.

Agradecer a mi compañera de "lab" **Rosa Pérez** por la ayuda en cada momento que se lo pides. Gracias a ti hemos podido sacar muchos de los proyectos y sobre todo a pescar aquellos cristales que eran más delicados. También, quería agradecer a otro compañero que, aunque solamente estuvo con nosotros 1 año, fue más que suficiente para que dejara huella en el "lab", muchas gracias, **Ramiro Illanes**.

Dar las gracias al **Dr. Oscar Quijada** y la **Dra. María Lluch** que, aunque directamente no sean mis compañeros, para mí es como si lo fuesen. Agradeceros toda la ayuda a lo largo del proyecto de *M. pneumoniae*. En particular a ti Oscar, las amplias charlas y los audios de WhatsApp de 5 minutos o más comentando resultados o teorías acerca de la biología del parásito.

Por último, para finalizar con las personas que pertenecen al IBMB, agradecer a todo el personal que hace posible que podamos disfrutar de un trabajo tan maravilloso.

Fuera de la vida científica, quería agradecer a tres personas en especial, tres personas a los que considero amigos y si dios quiere espero estar mucho tiempo junto a ellos. Primeramente, agradecer a **Sergio Garnica**, porque eres mi "enano" querido y gracias a esas charlas y conversaciones de horas y horas me pude evadir del trabajo en muchas ocasiones. También al **Dr. Daniel Fuentes**, porque sí, ¡es doctor a pesar de todo! Gracias por el apoyo en todo momento y los consejos que me has dado. Aunque te pareciese que no siempre te hiciera caso, te tengo en gran admiración. Por último, pero no menos importante a mi *sensei* **Andrea Ardid**, gracias a ti pude tomar el camino que estoy llevando y en parte esto también es tuyo.

Agradecer a toda mi familia, en especial a **mis padres** por darme la educación y la guía para ser quien soy en este momento. A **mi tía** por apoyarme y animarme a lo largo de este tiempo. Y a **mis abuelos**, en particular a mi abuelo **Paco**, que estés donde estés sé que me seguirás cuidando como cuando tenía 4 años.

Para el final, he querido dejar a las dos personas que más especiales me llevo de estos 5 años y que seguro que permanecerán en mi vida. Primero, a mi compañero, el cual considero un muy buen amigo, **Jesús Martín Romero**. Gracias por todas esas conversaciones acerca de animales "raros" de los viernes, por esos "cardios", por aquellas charlas sobre videojuegos o de cómo transcurre la vida. Ten por seguro que el poco tiempo que llevas en el "lab" me ha sido suficiente para tenerte ese cariño "especial". ¡Espero y deseo que la vida te de muchas alegrías! Por último, gracias a la mejor persona que entró en mi vida un 6 de septiembre de 2017, **Anna Giménez Sánchez**, mi chiquita. Gracias por apoyarme en cada momento, por ayudarme hasta en lo que tú no entiendes y haces por comprender. Te debo tanto que me faltarían años para devolverlo, por ello espero estar siempre contigo. Te quiero mucho.

Abreviaturas

2D	Dos dimensiones
3D	Tres dimensiones
Å	Angstrom
aa	Aminoácido
ADN	Ácido desoxirribonucleico
AIM2	Ausente en melanoma 2
ARN	Ácido ribonucleico
ASC	Proteína tipo punto asociada a apoptosis
ATP	Adenosín trifosfato
^{o}C	Grado centígrado
CARDS	Síndrome de dificultad respiratoria adquirida en la comunidad
CD4	Cúmulo de diferenciación 4
CD8	Cúmulo de diferenciación 8
Crio-EM	Crio-microscopía electrónica
Crio-ET o CET	Crio-tomografía electrónica
CTF	Función de transferencia de contraste
dRI	Detector de índice de refracción
e^{-}	Electrón
EFTEM	Energía de microscopía electrónica de transmisión filtrada
ELISA	Ensayo de inmunoabsorción ligado a enzima
FSC	Correlación de capa de Fourier
g	Fuerza g o fuerza centrífuga relativa
g	Gramo
Gal	Galactosa
gDNA	Ácido desoxirribonucleico genómico
Glc	Glucosa
GO	Óxido de grafeno.
H_2O_2	Peróxido de hidrógeno
ICP-MS	Espectrometría de masas de plasma acoplado inductivamente
IgA	Inmunoglobulina tipo A
IgE	Inmunoglobulina tipo E
IgG	Inmunoglobulina tipo G
IgM	Inmunoglobulina tipo M
IL-1β	Interleucina-1-β
IL-8	Interleucina-8
Kb	Kilobase
K_d	Constante de disociación
KDa	Kilodalton
M	Molar
MALS	Dispersión de luz multiangular
min	Minuto
MgPar	Elementos repetitivos de *Mycoplasma genitalium*
ml	Mililitro

mM	Milimolar
Nap	Complejo de adhesión
Neu5Ac	Ácido N-Acetilneuramínico
NF-κB	Factor nuclear κβ
nl	Nanolitro
nm	Nanomolar
NOD	Dominio de oligomerización de unión a nucleótidos
ORF	Marco abierto de lectura
OT	*Organela terminal*
P110N	Dominio N de la proteína P110
P40/P90N	Dominio N de la proteína P40/P90
PAA	Ácido fosfonoacético
PCR	Reacción en cadena de la polimerasa
PPI	Suero pre-inmunización
qPCR	Reacción cuantitativa en cadena de la polimerasa
RepMps	Elementos repetitivos de *Mycoplasma pneumoniae*
RMSD	Desviación cuadrática media
ROS	Especies reactivas de oxígeno
s	Segundo
S1	Inserción 1 P40/P90 (residuos 297-368)
S2	Inserción 2 P40/P90 (residuos 395-462)
SAD	Difracción anómala de una longitud de onda
SARS	Síndrome respiratorio agudo grave
SEC	Cromatografía por exclusión de tamaño
SeMet	Seleniometionina.
SDS-PAGE	Electroforesis en gel de poliacrilamida con dodecilsulfato sódico
SPR	Resonancia de plasmón superficial
Th1	Linfocitos T *helper* 1
Th2	Linfocitos T *helper* 2
TNF-α	Factor de necrosis tumoral α
µg	Microgramos
µl	Microlitro
µm	Micrómetro
VIH	Virus inmunodeficiencia humana
WT	Tipo natural

"El valor de una idea radica en el uso de la misma"

(Thomas Alva Edison)

Para mis padres, mi chiquita y Bú

Índice

Resumen

RESUMEN

Mycoplasma pneumoniae y *Mycoplasma genitalium* son dos patógenos humanos que infectan tejido pulmonar y urogenital respectivamente. El mecanismo principal para llevar a cabo la infección es a través de la adhesión a las células huésped por una pequeña protuberancia llamada *Organela terminal* (OT). En la membrana de esta OT se encuentra un complejo multiproteico, denominado Nap, compuesto por un dímero de heterodímeros de las proteínas P1 y P40/P90 en *M. pneumoniae,* o P140 y P110 en *M. genitalium*. P1 ha sido el foco de numerosos estudios, identificándola como la adhesina principal del complejo Nap en *M. pneumoniae*. Además de la función de adhesión, las proteínas de los Naps también participan en otros procesos esenciales para la patogenicidad, como son la motilidad y la evasión del sistema inmune.

En nuestro laboratorio se determinó hace unos años la estructura del ectodominio de la proteína P110 de *M. genitalium* (Aparicio *et al.*, Nature comm. 2018). Continuando la investigación sobre las proteínas que componen el Nap, el objetivo principal de esta Tesis doctoral ha sido determinar la estructura tridimensional de P140 (*M. genitalium*) y de P1 y P40/P90 (*M. pneumoniae*), estudiando también el papel que desempeñan estas proteínas en adhesión, motilidad e inmunogenicidad.

El primer capítulo presenta la obtención, mediante cristalografía de rayos X, de la estructura del dominio C de P140. La estructura se resolvió en dos grupos espaciales cristalográficos diferentes: I) tetragonal I422, a 1.43 Å de resolución, con una molécula en la unidad asimétrica y II) monoclínico C2, a 1.92 Å, con ocho moléculas en la unidad asimétrica. El análisis estructural ha proporcionado dos resultados que cabe destacar aquí: I) El dominio C de P140 puede formar oligómeros (octámeros) compactos; II) Por otro lado, las grandes diferencias entre las diversas conformaciones halladas del dominio C, cuando *a posteriori* se pudo disponer de la estructura del ectodominio completo de P140, indican la gran plasticidad de este dominio y sugieren un papel importante de esa plasticidad en el funcionamiento del Nap.

El segundo capítulo corresponde a la determinación de la estructura de P140 y del complejo P110-P140. P140 se resolvió a través de un procedimiento en el cual fue

crítica la información proporcionada por crio-tomografía electrónica (crio-ET) del Nap y por diversos cristales tanto de P140 como de P1. En este trabajo también se pudieron modelar diversas conformaciones del Nap a través del "docking" de las estructuras de P140 y P110 en los correspondientes mapas de crio-ET y de crio-microscopía electrónica (crio-EM). Esta información permitió proponer un primer mecanismo de unión y liberación del Nap a los receptores celulares.

El tercer capítulo corresponde a la determinación de las estructuras de P1 y P40/P90N y al análisis de sus propiedades inmunogénicas. La estructura de ambas proteínas se resolvió mediante cristalografía de rayos X. La estructura de P1 se resolvió también (2.8 Å de resolución) mediante crio-EM. Las estructuras de P1 y P40/P90 presentan una elevada similitud con P140 y P110, respectivamente. No obstante, las proteínas de *M. pneumoniae*, en particular P1, presentan regiones de alta variabilidad estructural en su superficie que no existen en las proteínas de *M. genitalium*. Esta variabilidad es particularmente acusada en el entorno del sitio de unión al receptor celular, lo que sugiere un mecanismo de protección frente a la respuesta inmunológica del huésped. Se realizaron ensayos inmunológicos con distintas construcciones de P1 y de P40/P90 frente a sueros de pacientes infectados por *M. pneumoniae* que revelaron que los dominios C de P1 son reconocidos exclusivamente por sueros con alto contenido de inmunoglobulinas del tipo G (IgG). Además, se verificó que anticuerpos policlonales dirigidos contra el dominio C de P1 producían una importante inhibición de la adhesión. Para P40/P90, de la cual prácticamente no existen estudios inmunológicos, se observó que también es reconocida por la mayoría de los sueros de pacientes. Por último, se demostró estructural y bioquímicamente, que el sitio de unión a ácidos siálicos está en P40/P90 y no en P1, como se había propuesto hasta ahora.

Los resultados obtenidos en esta Tesis permiten una mejor comprensión de los mecanismos de adhesión, motilidad y evasión del sistema inmune en *M. pneumoniae* y *M. genitalium* y posibilitan el desarrollo de antígenos para la obtención de vacunas frente a estos microorganismos.

Introducción

1. Características generales de los micoplasmas

Los micoplasmas son bacterias gram positivas que, basándose en un análisis de la subunidad 16s del ribosoma, se clasifican en el filo *Firmicutes* donde también se encuentran otros géneros bacterianos como *Clostridium*, *Lactobacillus* y *Streptococcus* [1]. Taxonómicamente, los micoplasmas pertenecen al dominio de las bacterias de la clase *Mollicutes*, familia *Mycoplasmataceae* y género *Mycoplasma*. Dentro del género *Mycoplasma* hay más de 200 especies de micoplasmas conocidos, entre los que cabe destacar a *Mycoplasma pneumoniae* como patógeno humano importante de las vías respiratorias que produce principalmente la denominada neumonía atípica [2]. Otros micoplasmas importantes que afectan a humanos son *Mycoplasma genitalium*, patógeno de transmisión sexual de las vías urogenitales [3] y *Mycoplasma hominis*, patógeno de las trompas de Falopio que produce infecciones en el tracto genitourinario, en especial en mujeres, e infecciones sistémicas en neonatos [4].

Los micoplasmas representan a las bacterias de menor tamaño (0,5-2 µm) que pueden vivir y replicarse en la naturaleza. A diferencia del resto de bacterias gram positivas, los micoplasmas no poseen pared celular [5], lo que los hace sensibles a sustancias como alcoholes, detergentes o lisis por choque osmótico pero, a su vez, les otorga resistencia frente a los antibióticos de la familia β-lactámicos (inhiben la síntesis de la barrera de peptidoglicanos de la pared celular). La ausencia de pared hace que la membrana de estos microorganismos esté recubierta de lipoproteínas, glicolípidos, lipoglucanos, colesterol y de una serie de proteínas ancladas a membrana involucradas en la adhesión, variación antigénica, movimiento y transporte de metabolitos [6].

Los micoplasmas son patógenos ubicuos que pueden encontrarse en plantas, en animales y seres humanos [7]. El inicio de la infección implica la adherencia a las células de un tejido específico del huésped donde los micoplasmas se multiplican y pueden persistir durante largos periodos de tiempo [8]. Una vez llevado a cabo el proceso de adhesión, se desencadenan una serie de procesos patogénicos que pueden ser muy

variados como la producción de radicales libres de oxígeno (ROS), la ruptura de los cilios (ciliostasis) y/o procesos proinflamatorios entre otros [9,10].

El genoma de los micoplasmas presenta varias peculiaridades:

I) **Tamaño genómico reducido**, con genomas que oscilan entre las 500 y las 1400 kb dependiendo de la especie de micoplasma [11]. La reducción del genoma puede ser debida a un proceso de evolución degenerativa donde una bacteria ancestral gram positiva acorta el tamaño de su genoma dejando el mínimo de genes necesarios para un organismo vivo [12]. Existe una relación entre el tamaño del organismo y la reducción del genoma, ya que la mayoría de los genes (alrededor del 80%) están dedicados a la obtención de energía, mientras que una menor parte se destinan al crecimiento celular.

II) ADN con bajo contenido en citosinas y guaninas, con valores de C+G que oscilan entre el 18 y el 40%. Esta es una característica que los micoplasmas comparten con todos los miembros del filo *Firmicutes* [11].

III) Codón UGA con traducción alternativa al código genético estándar. Los codones UGA se traducen normalmente como triptófanos, sin embargo, en el género *Mycoplasma* se interpretan como codones STOP [11]. Esta variación en el código genético limitó durante años los estudios de proteínas recombinantes antes del descubrimiento de la mutagénesis dirigida y/o de la síntesis sintética de ADN.

IV) Elevada frecuencia de mutaciones genéticas con respecto a otras bacterias. La reducción del genoma conlleva la eliminación de sistemas reparadores, por ejemplo, la actividad exonucleasa 3' a 5' de la ADN polimerasa está ausente en micoplasmas, lo que da lugar a su vez a la pérdida de funciones enzimáticas [12].

La producción de la mayoría de las moléculas esenciales en micoplasmas, como los aminoácidos o los cofactores, es limitada debido a la reducción del tamaño del genoma y a la pérdida de función de algunas enzimas por el incremento de mutaciones. Por consiguiente, esas moléculas esenciales las deben obtener los micoplasmas del

entorno ya sea de las células del huésped que parasita, o bien de un medio de cultivo enriquecido. Las bases púricas y pirimidínicas para la síntesis de ácido desoxirribonucleico (ADN) o ácido ribonucleico (ARN) se obtienen a partir de la degradación, mediante nucleasas, del ADN de las células del huésped [13]. El papel que juegan los lípidos es esencial para los micoplasmas, como el colesterol en la estabilización de la membrana [14], pero las vías para obtenerlo están limitadas. Los micoplasmas son capaces de sintetizar fosfolípidos y glicolípidos, aunque esto representa un gasto elevado de energía, por lo que en general los lípidos también se obtienen exógenamente. La mayoría del genoma del micoplasma tiene como finalidad la producción de adenosín trifosfato (ATP) por la fermentación principalmente de glucosa a ácido láctico y acético [15] o, alternativamente, a partir de pequeños carbohidratos o glicerol [16].

2. Micoplasmas del *cluster Pneumoniae*

2.1 *Mycoplasma pneumoniae*

M. pneumoniae, un importante patógeno humano, es el representante más característico del grupo (*cluster*) *Pneumoniae* de micoplasmas, donde también se encuentran *M. genitalium* y *Mycoplasma gallisepticum* [11]. *M. pneumoniae* causa principalmente la denominada neumonía atípica cuyos síntomas son, en general, más leves que la neumonía típica provocada por bacterias como *Streptococcus pneumoniae* [2]. El tamaño de *M. pneumoniae* es de 1-2 µm de largo y 0.1-0.2 µm de ancho, con una morfología en la que pueden diferenciarse tres partes: cola, cuerpo y cuello (**Figura 1a**). El genoma de *M. pneumoniae* consta de 816 kb, codificando un total de 687 genes, de los cuales el 57% son esenciales [16].

M. pneumoniae se denominó inicialmente "agente Eaton", ya que fue el Dr. D.M. Eaton quien logró por primera vez, en 1944, el cultivo del "agente" causante de la neumonía atípica. Inicialmente, el "agente Eaton" fue considerado un virus ya que al pasarlo por filtros de exclusión bacteriana continuaba filtrándose, no se podía tratar por antibióticos como la penicilina (β-lactámicos) y no se observaba por microscopía [17,18].

Todo esto, junto a la imposibilidad de cultivar micoplasmas en aquella época, llevó a considerarlo como un virus. Fue a principios de los años 60 cuando el Dr. L.Hayflick, a través de una fórmula especial de medio de cultivo, logró aislar el "agente Eaton" de la yema de huevo, comprobando que no se trataba de un virus sino de un micoplasma. Inicialmente, iba a ser denominado con el apellido de su descubridor, *Mycoplasma hayflickiae*, pero finalmente se denominó con el nombre actual de *Mycoplasma pneumoniae* [19].

2.2 *Mycoplasma genitalium*

M. genitalium pertenece, como *M. pneumoniae,* al *cluster Pneumoniae* de micoplasmas y es un patógeno humano de las vías urogenitales, considerado por la OMS como un patógeno de transmisión sexual emergente. *M. genitalium*, con una morfología que se describe como la de un matraz (del inglés "flask-like"), es más ancho (0.3-0.4 µm) pero de menor longitud (0.6-0.7 µm) que *M. pneumoniae* (**Figura 1b**). La presencia de cola no es aparente en *M. genitalium*, pero sí que es muy evidente la presencia de la protuberancia correspondiente a la *Organela terminal* (OT) [20]. El genoma de *M. genitalium* posee un tamaño de 580 kb, codificando un total de 482 genes, de los cuales el 80% son esenciales [21].

M. genitalium fue aislado en los años 80 a partir de muestras de pacientes que presentaban uretritis no gonocócica, en 1981 el Dr. J.G.Tully consiguió observar células de *M. genitalium* mediante microscopía electrónica determinando su morfología y clasificándolo como un micoplasma [21].

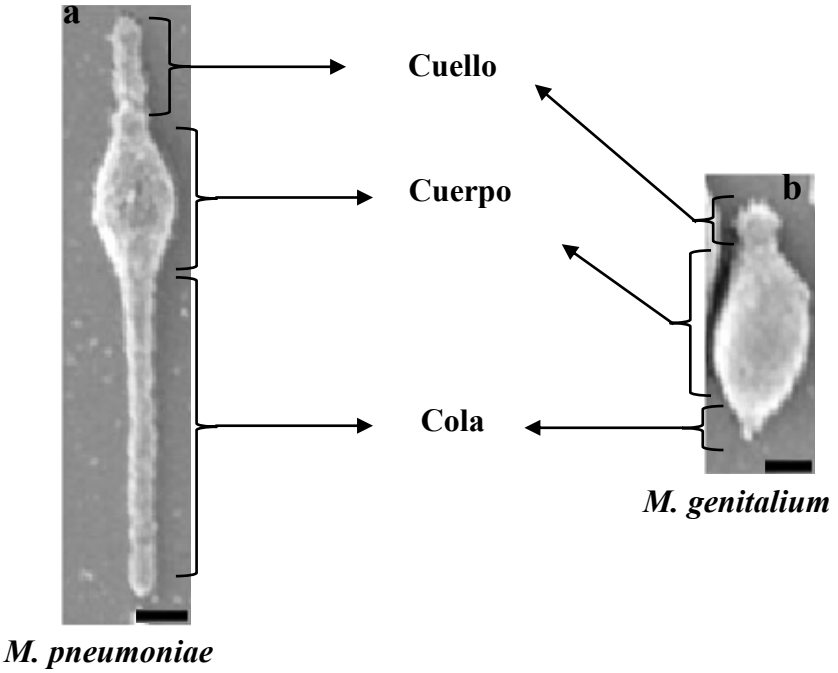


Fig. 1. Microscopía electrónica de barrido de células de **a)** *M. pneumoniae* y **b)** *M. genitalium* unidas a cubreobjetos. Imagen adaptada de J.M.Hatchel *et al*. Microbiology, 2006. Escala 250 nm

3 Propiedades de *M. pneumoniae* y *M. genitalium*

3.1 Adhesión en el *cluster Pneumoniae*

M. pneumoniae y *M. genitalium* presentan la protuberancia citada anteriormente como OT [22], con un papel fundamental en la adhesión a las células del huésped, en la división celular del micoplasma y en la motilidad. Los elementos estructurales que componen la OT de *M. pneumoniae* se caracterizaron en el año 2001 tras realizar una extracción con detergente (Tritón X-100 al 0.5-2%) obteniendo una estructura compacta y soluble, donde se identificaron mediante un gel de plata alrededor de 100 proteínas, mientras que en la tinción con Coomassie se determinaron 50 proteínas de las cuales 41 de ellas se pudo caracterizar su secuencia y peso molecular [23]. Por otra parte, mediante crio-microscopía electrónica (crio-EM) se identificaron tres regiones en la OT (**Figura 2**) [24]:

I) Botón terminal: Zona más próxima a la membrana y que sugiere ser la encargada de transmitir la información del exterior del micoplasma. Se encuentran proteínas como la P65 y HMW3 [25].

II) Zona electrodensa: Zona central alargada, de unos 300 nm de longitud, encargada de conectar las otras dos partes de la OT. Su función se considera principalmente estructural, aunque en ella se han encontrado proteínas (HMW1 y HMW2) con motivos de unión a ATP relacionadas con adhesión y movimiento [25].

III) Complejo rueda (*the wheel or the bowl complex*): Zona próxima al cuerpo del micoplasma y posiblemente la encargada de transmitir la información entre el cuerpo y la OT. Se encuentran proteínas como la P200, P41… [25] Además de esas tres regiones, anclada a la membrana de la OT del micoplasma existe un complejo denominado Nap [26], formado principalmente por P1 y P40/P90 y por otras proteínas como P30 [27].

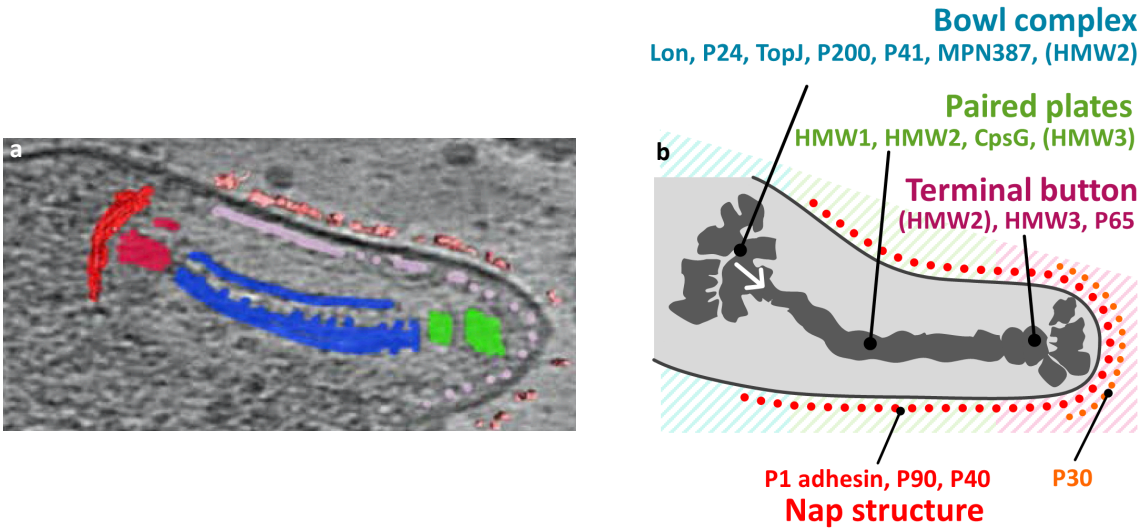


Fig. 2. Esquema de la arquitectura de la OT de *M. pneumoniae*. **a)** Tomograma de la OT marcando en rojo, azul, verde y rosa el complejo rueda, la zona electrodensa, el botón terminal y el Nap, respectivamente. **b)** Representación de las proteínas que componen la OT. Imágenes adaptadas de D.C.Krause et al, Molecular microbiology, 2018 y D.Nakane et al, PLOS pathogens, 2015.

Estudios *in vivo* demostraron que *M. pneumoniae* se unía a la superficie de las células mediante la unión a moléculas de ácidos siálicos, ya que al eliminarlos mediante neuramidasas la adhesión del micoplasma se reducía muy significativamente [28-30]. Los

ácidos siálicos son oligosacáridos hidrofílicos cargados negativamente y uno de los principales constituyentes de muchos de los oligosacáridos, glicoproteínas y glicolípidos de las membranas de las células eucariotas y de las mucosas [31,32]. Los dos compuestos sialidados más abundantes en las membranas de las células son los oligosacáridos 3SL (α-Neu5Ac-2,3-β-Gal-1,4-β–Glc) y 6SL (α-Neu5Ac-2,6-β-Gal-1,4-β–Glc), en los ensayos realizados muestran afinidades diferentes en la unión por *M. pneumoniae*, con una afinidad por el 3SL ligeramente más elevada que el 6SL. La variación en la ratio 3SL:6SL a favor del 6SL hacía disminuir la frecuencia de movimiento llegando a ser 0% en los casos con un porcentaje nulo en 3SL [33]. Desde 1970 se han realizado numerosos estudios que han permitido definir y caracterizar, por lo menos en parte, una serie de proteínas que están relacionadas con la adhesión de *M. pneumoniae*/*M. genitalium***.**

3.1.1 Proteínas involucradas en el proceso de adhesión

I) P1. Desde que comenzaron los estudios de adhesión, P1 (con 170 KDa de peso molecular) ha sido considerada la adhesina principal. Según el análisis de la secuencia de aminoácidos, los residuos 1 al 59 corresponden al péptido señal, quien permite su translocación y transición hacia la membrana para posteriormente escindirse [34,35]. Desde el residuo 60 al 1526, se encuentra la parte extracelular de la proteína que se ancla a la membrana del micoplasma a través de una hélice transmembrana localizada en los residuos 1527 al 1547 [35]. Por último, desde el residuo 1548 al 1627 se localiza la zona citoplasmática [35], la cual conectaría con el interior de la célula. P1 es codificada por el gen *p1* que se encuentra en el mismo operón que los genes ORF4 y ORF6. ORF4 codifica una proteína bifuncional oligoribonucleasa y fosfatasa, mientras que ORF6 codifica P40/P90, la otra adhesina que compone el Nap [11]. La proteína ortóloga de P1 en *M. genitalium* es P140, de 1444 residuos y un 44% de identidad de secuencia (**Tabla 1**). La diferencia más importante entre ambas proteínas se encuentra en varias inserciones repartidas a lo largo del dominio extracelular, conservando una alta homología en el dominio citoplasmático. De P1 existen varios estudios que la relacionan en los procesos de adhesión y movimiento. Uno de los primeros estudios fue el del Dr. S.F.Dallo en 1988, donde se concretó una secuencia entre los residuos 1384-1394, en la que se unía un anticuerpo monoclonal que producía inhibición en la adhesión del micoplasma a

plástico. Esta secuencia fue denominada como "péptido de citoadherencia" [36]. Posteriormente, en otros ensayos se usaron anticuerpos contra un fragmento de 360 residuos del dominio C de P1 (1160-1520). Los anticuerpos generados produjeron una reducción tanto en la adhesión del micoplasma a plástico y a células como al movimiento [37,38]. Los resultados sugieren que este fragmento de 360 residuos debería contener el sitio de unión a los ácidos siálicos. Además, el estudio realizado por el Dr. S.Seto en el año 2005, indicaba que P1 debía experimentar cambios conformacionales durante el movimiento, que explican porqué los anticuerpos generados para el fragmento de 360 residuos inhibían solamente la adhesión de aquellos micoplasmas que estaban en movimiento. En otras palabras, P1 durante el movimiento debía tener una conformación diferente, haciéndola así más accesible al anticuerpo en comparación a cuando estaba en estático [37]. Otra característica importante de P1 es la alta inmunogenicidad que presenta según ensayos realizados con el fragmento de 360 residuos en sueros de pacientes [39]. En conjunto, las propiedades que presenta el fragmento del dominio C lo convierten en un potencial antígeno para el desarrollo de vacunas, como las realizadas para otros patógenos productores de infecciones respiratorias como *S. pneumoniae* y *Haemophilus influenzae*. Sin embargo, hasta la fecha no se ha tenido éxito en este campo, pero aprovechando la alta inmunogenicidad de este fragmento, se utiliza como método de diagnóstico por ensayo de inmunoabsorción ligado a enzima (ELISA) detectando infecciones por *M. pneumoniae* (DiaSorin). A parte de los estudios sobre la importancia de P1 en la infectividad del micoplasma, también se han realizado otro tipo de investigaciones como el análisis de las regiones repetitivas del genoma (RepMPs) [40] o el reconocimiento de inmunoglobulinas de tipo E (IgE) sobre P1 y la implicación de esta proteína en el desarrollo de alergias y asma [41].

II) P40/P90. Esta proteína, de 130 KDa de peso molecular, está codificada por el gen ORF6 y presenta una organización de dominios similar a P1 teniendo un péptido señal entre los residuos 1 y 25, una región extracelular del 26 al 1120, una hélice transmembrana entre 1121 y 1141 y un dominio citoplasmático del 1142 al 1218 [35]. Su ortóloga en *M. genitalium* es P110, de 1053 residuos, cuya estructura ha sido recientemente resuelta en un trabajo en el que también se ha demostrado que es capaz de unirse a los ácidos siálicos [42]. La identidad secuencial entre ambas proteínas es de un

46% y, como también ocurre en P1, el dominio citoplasmático es el más conservado (**Tabla 1**). La gran diferencia a nivel secuencial entre P40/P90 y P110 consiste en dos grandes inserciones en la zona extracelular de P40/P90 que denominaremos S1 y S2. Históricamente, P40/P90 ha tenido menos impacto que P1, de modo que no se encuentra mucha información acerca de ella. Los estudios que tratan sobre esta proteína se centran en el punto de corte que da lugar a las dos subunidades P40 y P90 y sobre la formación del complejo P1-P40/P90. Existen dos artículos sobre el punto de corte, el primero de ellos, de 1998, localizaba el punto de corte antes de la Arg_{455} [43], mientras que, en el segundo (muy reciente) se concluye que el corte se produce 10 aminoácidos antes, entre la Arg_{445} y la Ala_{446} [44]. Ambos estudios reflejan ambigüedad sobre el lugar exacto de escisión, generando interés respecto a la producción de dos dominios independientes denominados P40 y P90 a partir de la traducción del gen ORF6. Sobre la formación de la partícula Nap, hay un estudio que afirma que el complejo que obtuvieron directamente de *M. pneumoniae* fue el de P1 con la subunidad P90, quedando P40 libre [45].

III) P30. Esta proteína, de 30 KDa de peso molecular y codificada por el gen *p30,* ha sido históricamente considerada como la segunda adhesina junto a P1. La organización estructural comprende 3 dominios y 2 hélices transmembranas (en los residuos 11-31 y 73-99) [46]. El dominio I (32-72) se encontraría en el interior mientras que los dominios II y III estarían en el exterior. El dominio II (100-177) está altamente conservado entre las otras especies del género *Pneumoniae*: *M. pneumoniae*, *M. genitalium* y *M. gallisepticum* [35]. Por último, el dominio III (178-274) está constituido por repeticiones ricas en prolinas que varían en secuencia entre las diferentes especies [35]. Un estudio sobre esta proteína muestra que las mutaciones sobre diversos residuos producen un efecto negativo en la adhesión y el movimiento, otorgando un papel fundamental en ambos aspectos [47]. Sobre las interacciones de esta proteína con otras que conforman la OT, un estudio llevado a cabo en *M. genitalium* con P32, la ortóloga a P30 (**Tabla 1**), demuestra una relación estrecha con P1 en la parte extracelular y P65 en la parte intracelular [48], lo que sugiere que P30 pudiera participar en el movimiento y adhesión llevado a cabo por P1 y/o un papel en la transmisión de la información entre el exterior y el interior del micoplasma.

IV) HMW1 y HMW2. Otras proteínas relacionadas con la adhesión son la HMW1 y la HMW2 de 112 y 215 KDa, respectivamente. En 1995 se realizó un *knock-out* para estas proteínas que revelaron inhibición en la adhesión [49]. Sin embargo, posteriormente se demostró que esta inhibición se debía a una incorrecta organización de la OT ya que afectaba a otras proteínas, en particular las proteínas del Nap y, por consiguiente, no se puede asegurar que la inhibición de la adhesión sea producida directamente por HMW1 y HMW2 [50]. No obstante, es incuestionable que HMW1 y HMW2 son fundamentales para un correcto ensamblaje de la OT.

Todas estas proteínas tienen un papel dentro del proceso de adhesión del microorganismo a la célula huésped. Sin embargo, los estudios realizados hasta la fecha muestran a P1 como la principal en la adhesión, el resto de las proteínas son componentes accesorios de adhesión.

Tabla 1. Proteínas implicadas en la *Organela terminal* de *M. pneumoniae* y *M. genitalium*.

M. pneumoniae		***M. genitalium***		**Sequence identity (%)**
Gene name	**Protein name**	**Gene name**	**Protein name**	
mpn141	P1	MgpB	P140	44
mpn142	P40/P90	MgpC	P110	46
mpn453	P30	MG318	P32	43
mpn309	P65	MG217	P65	40
mpn447	HMW1	MG312	HMW1	33
mpn310	HMW2	MG218	HMW2	57
mpn452	HMW3	MG317	HMW3	33
mpn567	P200	MG386	P200	29
mpn311	P41	MG491	MG491	53
mpn312	P24	MG219	MG219	17
mpn119	TopJ	MG200	DNAJ	35
mpn213	P116	MG075	MG075	51

3.2 Motilidad

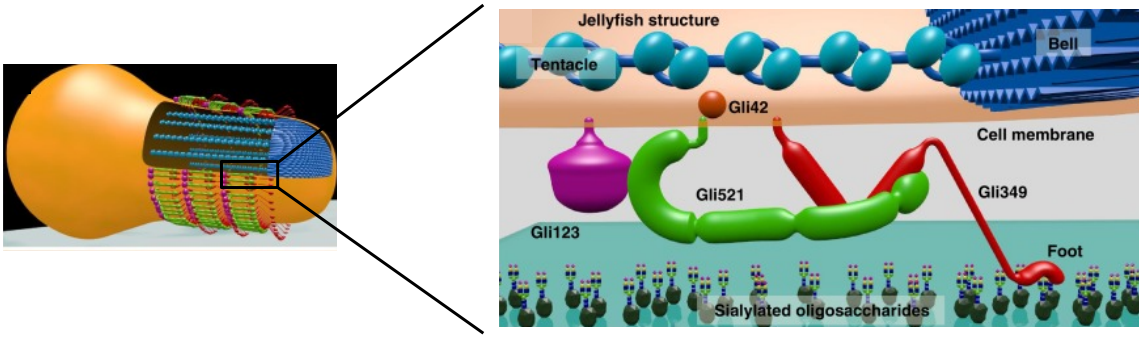
Ciertas especies de micoplasmas, en particular todas las pertenecientes al *cluster Pneumoniae*, son capaces de desplazarse. No obstante, los micoplasmas no poseen motores moleculares dedicados al movimiento como pueden ser los flagelos, los cilios y otros dispositivos presentes en muchas bacterias, por lo que debe tratarse de otro componente el que realice dicha función [51]. En *M. pneumoniae*, se ha demostrado que tanto P1 como P30 están involucradas en el movimiento del micoplasma. Así pues, la OT aparte de ser importante en el proceso de adhesión también debe tener un papel en el movimiento [37,52]. El mecanismo por el que se mueven se denomina *gliding* (del inglés, deslizamiento) y permite a los micoplasmas "deslizarse" sobre superficies sólidas. La direccionalidad del movimiento viene dada por la protuberancia que presentan algunos micoplasmas como las que encontramos en *M. pneumoniae* o *Mycoplasma mobile* [53,54]. Con respecto al *gliding*, en *M. pneumoniae* y *M. gallisepticum* se ha podido observar la importancia del ATP [55] y del sistema fosfatasa/kinasa (PrpC/PrkC) [56] en la frecuencia del movimiento, aportando protagonismo de la producción de ATP y el proceso de fosforilación/desfosforilación sobre tres proteínas: P1, HMW1 y HMW2. Aunque todavía se desconoce el motivo de este proceso, estudios reportan que mutantes defectivos en PrpC o PrkC afectan a la frecuencia de *gliding,* pudiendo tener un papel en la transmisión del movimiento a lo largo de la OT [56,57]. Actualmente, no se conoce con certeza cómo se mueve *M. pneumoniae* y *M. genitalium,* pero sí hay dos propuestas de modelos de movimiento.

3.2.1 Modelos de movimiento

I) Modelo cien pies: Modelo descrito para *M. mobile* [58], parásito que se encuentra en las branquias de los peces y del mismo modo que ocurre en *M. pneumoniae* y *M. genitalium*, se divide en tres partes: cabeza, cuerpo y cuello [59]. En la cabeza se encuentra una estructura denominada medusa que se compone de una campana ovalada de 155 nm de largo, llena de un enrejado hexagonal de 12 nm y conectada con varios tentáculos flexibles compuestos por diversas partículas de 20 nm de diámetro (**Figura 3a**) [59]. En la superficie de la cabeza se encuentran las denominadas

patas, formadas por 4 proteínas principales: **I**) **Gli349** de 349 KDa, es responsable de la unión a los ácidos siálicos y realiza una función de "pata", tirando de ellos utilizando ATP [60]. **II) Gli521** de 521 KDa, podría actuar como una especie de manivela a la hora de transmitir la fuerza [61]. **III) Gli123** de 123 KDa, tiene la función de determinar la localización de todas aquellas proteínas relacionadas con el movimiento. Junto con gli349 y gli521 se encuentra en el mismo operón [62]. **IV) P42** de 42 KDa, tiene actividad ATPasa y es la encargada de proporcionar ATP como fuente de energía para el funcionamiento del mecanismo [63] (**Figura 3a**).

a



b

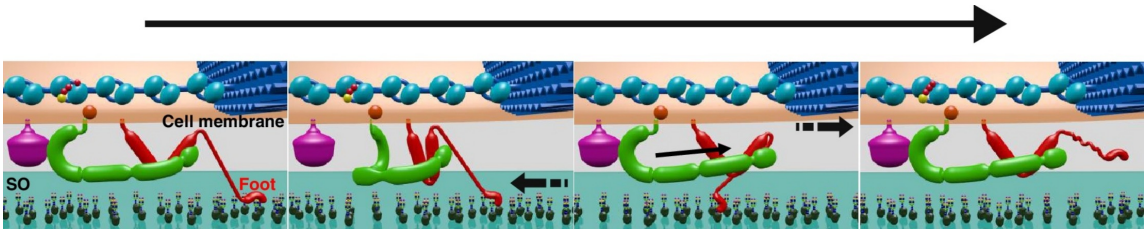


Fig. 3. Componentes de la estructura de *M.mobile* para llevar a cabo el movimiento. **a)** Representación de la maquinaria de *M.mobile* con una ampliación de la unidad de superficie y parte de la estructura interna. En el interior se encuentra la denominada estructura de medusa, compuesta por tentáculos y una campana. En la superficie se encuentran las cuatro proteínas fundamentales para la adhesión y movimiento, gli349 (rojo), gli521 (verde), gli123 (rosa) y gli42 (naranja). **b)** Representación esquemática del modelo de *gliding* denominado cien pies para *M.mobile*. Imágenes adaptadas de M.Miyata et al, Current Opinion in Microbiology, 2015 y M.S.Nishikawa et al, bioRxiv, 2019.

La velocidad con la que se mueve *M. mobile* es una de las más altas dentro del género de los micoplasmas, 2.0-4.5 μms^{-1}. El modelo de cien pies se compone de varias etapas: Inicialmente, ocurre una unión del "pie" de gli349 a los ácidos siálicos, tras esta unión se produce un tirón o golpe mediado por ATP de la parte superior de gli349, el cual impulsa hacia adelante al micoplasma. Finalmente, la tensión aplicada por la etapa

anterior disminuye, produciendo así una desunión del "pie" de los ácidos siálicos consumiendo una molécula de ATP, para devolverlo a una posición inicial [58] (**Figura 3b**). Se han realizado varios ensayos de inhibición y mutaciones sobre las proteínas gli349, gli521 y gli123 produciendo efectos negativos en la adherencia y movimiento [64,62].

II) Modelo oruga: Modelo propuesto para *M. pneumoniae* [65,66] que en comparación con el anterior es más lento, 0.3-0.5 μms^{-1}. Este modelo consta de una primera fase de adhesión del Nap a los azúcares y una vez adherido, la zona electrodensa se comprime como si de un muelle se tratara. Una vez totalmente comprimida se estira produciendo un tirón, similar al que se da en el modelo de cien pies, que empuja al micoplasma hacia adelante produciendo un movimiento lineal (**Figura 4**). De este modelo existen dos vertientes, ambas comparten el mismo fundamento, pero difieren en el reordenamiento de los Naps. En el primero, el movimiento se da mientras los Naps permanecen unidos, por lo que expone que debe existir una recircularización de los Naps a través de la membrana del micoplasma (**Figura 4, panel izquierdo**), mientras que en el segundo, esta recircularización no sucede ya que el Nap se desprende justo antes del tirón (**Figura 4, panel derecho**).

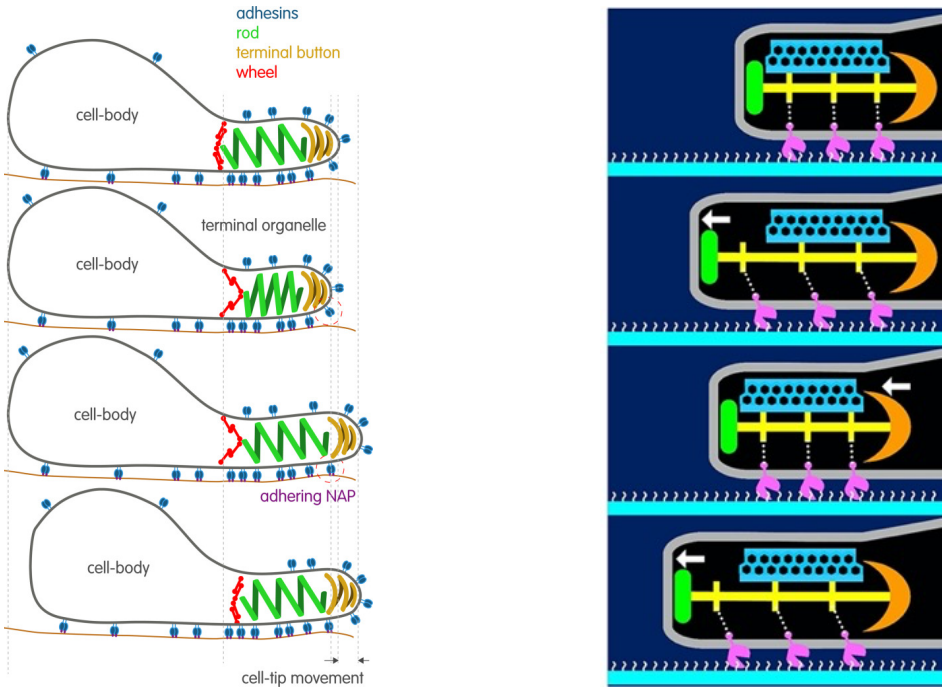


Fig. 4. Representación esquemática del modelo oruga para *M. pneumoniae*. En el panel de la izquierda el modelo propuesto en A. Seybert et al, 2018 y en el de la derecha el modelo propuesto en A.Kawamoto et al, 2016.

En la actualidad, se desconoce el motor molecular encargado del movimiento en *M. pneumoniae* y *M. genitalium*, sin embargo, hay varias proteínas candidatas que se tienen en consideración como el sistema de fosforilación PrpC/PrkC o la HMW1. El sistema PrpC/PrkC puede producir un movimiento a través del cambio conformacional que se produce en el proceso de fosforilación y desfosforilación. La HMW1 posee unos dominios característicos de las ATPasas (Walker A y Walker B) de unión e hidrólisis de ATP, que la convierte en una buena candidata para ser el motor celular. No obstante, un estudio reciente demostró que sigue existiendo un pequeño porcentaje de movimiento eliminando tanto HMW1 como HMW2 en *M. genitalium* [67].

3.3 Evasión del sistema inmune

Los micoplasmas han desarrollado varios métodos para evadir el sistema inmune del huésped y así asegurar y prolongar su existencia dentro de este [68]:

I) Mimetismo molecular y variabilidad antigénica. Son mecanismos que las bacterias y otros microorganismos adoptan para su supervivencia [69]. La mímica molecular se basa en la generación de posibles epítopos antigénicos en el micoplasma que también existen en las células del huésped (relacionadas con enfermedades autoinmunes), esta reactividad cruzada es importante para evitar el sistema inmune adaptativo [69]. Un ejemplo de este modelo de evasión lo encontramos en P1, ya que anticuerpos monoclonales contra dos octapéptidos de su zona extracelular, los cuales inhibían la adherencia, presentaban una actividad entrecruzada con dos moléculas en las células eucariotas: la gliceraldehido-3-fosfato deshidrogenasa y 2-fosfo-D-glicerato hidrolasa [70].

La variabilidad antigénica se basa en variaciones en el genoma que ocurren con una tasa mayor a la normal. Se ha demostrado que las lipoproteínas cambian en tamaño, combinan sus dominios… [71] Del mismo modo, existen unas secuencias dentro del genoma de *M. pneumoniae* y *M. genitalium* llamadas RepMPs y MgPars, respectivamente [11-12]. Se trata de fragmentos de ciertas proteínas con variaciones en su secuencia y que por recombinación homóloga pueden ser insertadas dentro del gen

codificante para la proteína correspondiente, resultando finalmente una proteína con pequeñas inserciones, deleciones o cambios de aminoácidos en su secuencia. En *M. pneumoniae* se encuentran 4 RepMps diferentes (**Figura 5**):

- **RepMP1**: Afecta a proteínas transcritas por los genes *mpn130*, *mpn137* y *mpn138*.
- **RepMP4**: Afecta a la parte N-terminal-intermedia de la zona extracelular (desde el residuo 1 hasta el 515) de P1. Hay 8 copias repartidas por todo el genoma.
- **RepMP2/3**: Afecta a la parte intermedia-C-terminal de la zona extracelular (desde el residuo 683 hasta el 1367) de P1. Hay 10 copias repartidas por todo el genoma.
- **RepMP5**: Afecta a gran parte del dominio globular de la zona extracelular (desde el residuo 1 hasta el 900) de P40/P90. Hay 8 copias repartidas por todo el genoma.

En *M. genitalium* hay repartidas a lo largo del genoma 9 copias MgPar que afectan a las regiones B (del residuo 192 al 339), EF (del residuo 763 al 959) y G (del residuo 1102 al 1183) del gen *mgpB* (codifica P140) y a las regiones KL (del residuo 42 al 294) y LM (del residuo 462 al 517) del gen *mgpC* (codifica P110) [12].

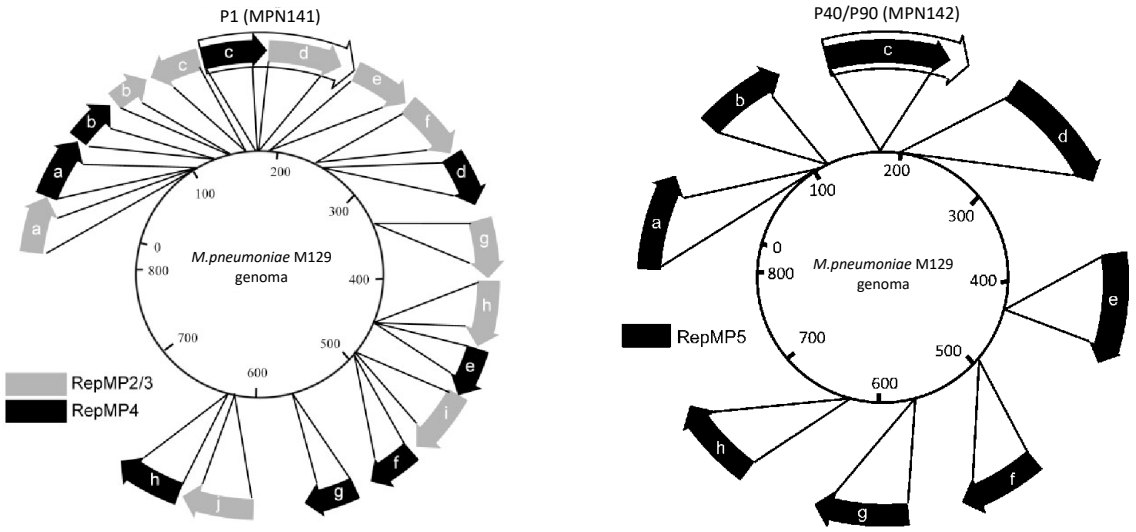


Fig. 5. Genoma de *M. pneumoniae*. En el panel de la derecha están representadas todas las copias de RepMPs de P1 y en el panel de la izquierda los RepMPs de P40/P90. Imágenes adaptadas de E.B.M. Spuesens et al, Microbiology, 2009 y 2011.

La función de todas estas secuencias es conferir al micoplasma una alta variabilidad en diferentes proteínas que le permita evadir al sistema inmune y ser más resistente al ataque de anticuerpos [71]. Una diferencia entre *M. pneumoniae* y *M. genitalium* es que tanto P1 como P40/P90 presentan más copias para variar su secuencia (o genoma) con respecto a sus ortólogas, P140 y P110, lo que hace a *M. pneumoniae* poseer un método de evasión más sofisticado. Además, tanto en *M. pneumoniae* como *M. genitalium* estas secuencias RepMPs/MgPars que afectan a P1/P140 permiten clasificar a la población de micoplasmas de cada especie en tipos. En *M. pneumoniae* existen dos tipos: P1-tipo1 y P1-tipo2, y estas a su vez tienen diferentes subtipos, según se han ido descubriendo más variantes clínicas [72].

II) Defensa frente a moléculas del sistema inmune. Durante el proceso de infección del micoplasma, el sistema inmune del huésped secreta distintas moléculas y células que afectan a su supervivencia. Entre estas moléculas se encuentran los ROS liberados por los macrófagos, para contrarrestarlos el micoplasma posee una serie de moléculas que los degradan [68]. Otro tipo de moléculas son las inmunoglobulinas, quienes llevan a cabo un papel crucial en la neutralización del antígeno. Se ha observado que en algunas especies de micoplasmas existen proteínas capaces de unirse a inmunoglobulinas del tipo A (IgA) humanas y no humanas y escindirlas [73]. Por último, los neutrófilos son leucocitos capaces de eliminar a patógenos, sin embargo, el micoplasma va a secretar nucleasas que entre otras funciones tienen la capacidad de degradar el ADN de los neutrófilos [74].

III) Invasión celular y creación de biofilms. De entrada, se creía que los micoplasmas eran patógenos extracelulares que se quedaban adheridos a la superficie de las células del huésped y desarrollaban un proceso de infección. Sin embargo, el avance de técnicas como la crio-EM han permitido identificar micoplasmas en el interior de estas células. La internalización del micoplasma, le permite escaparse del sistema inmune y de los fármacos, sobreviviendo dentro de la célula del huésped (además de tener una mayor accesibilidad a las moléculas producidas en el interior de las células) [75].

Los biofilms, son una agrupación de uno o varios microorganismos adheridos a una superficie, con características funcionales y una estructura compleja que les confieren protección ante los antibióticos, anticuerpos… [76] Varios estudios reportaron que *M. pneumoniae* puede agruparse formando biofilms en vidrio y que la formación de esta biopelícula se impide al usar inhibidores de la adhesión, como pueden ser anticuerpos contra P1 o cortando los ácidos siálicos de la superficie a través de neuraminidasas [77].

3.4 Patogenicidad

El mecanismo de infección de *M. pneumoniae y M. genitalium* comienza con la adhesión a las células del huésped principalmente mediado por el Nap para posteriormente, desencadenar una serie de acontecimientos (**Figura 6**):

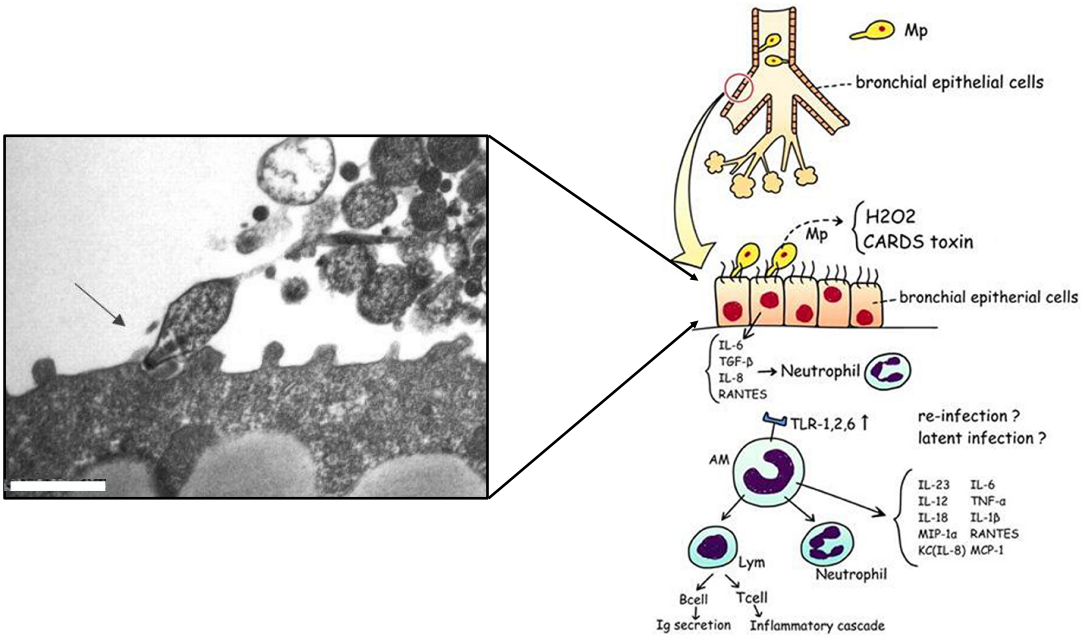


Fig. 6. Proceso de infección de *M. pneumoniae*. En el panel de la izquierda se muestra una imagen de microscopía electrónica de transmisión de tejido traqueal de hámster infectado por *M. pneumoniae* (editado de J.L. Jordam y D.C. Krause) y en el panel de la derecha se muestra un esquema postulado para la patogénesis de la neumonía provocada por *M. pneumoniae* (T.Saraya et al, Frontiers in Microbiology, 2014). Escala 500nm.

I) Fusión de la membrana: En 1998 se observó por primera vez, mediante microscopía electrónica, la fusión de la membrana del micoplasma con la membrana de las células huésped, esto lo permite la falta de pared celular del micoplasma [78]. Esta fusión va a desencadenar varios sucesos como: cambios en los sitios de reconocimiento a receptores (por la introducción de componentes de la membrana del micoplasma en la membrana de la célula huésped), introducción de moléculas como nucleasas, peróxido de hidrógeno (H_2O_2)... y la internalización al interior de la célula del micoplasma (**Figura 7**) [79].

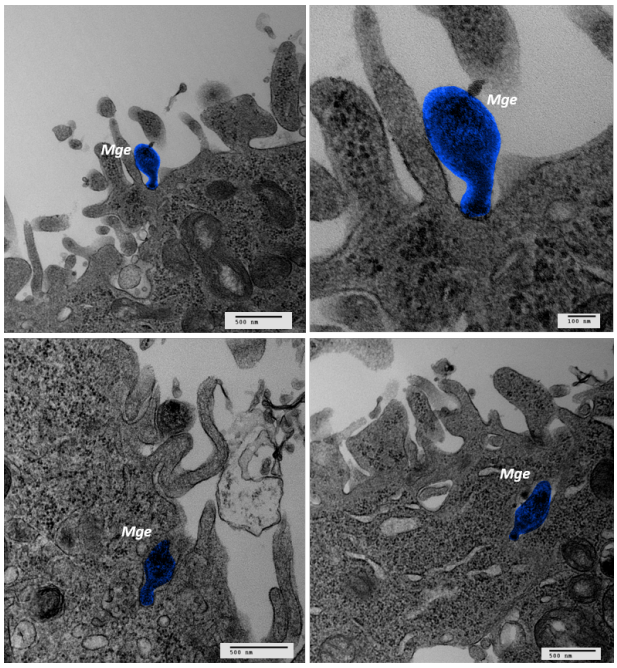


Fig. 7. Internalización de *M. genitalium* en células humanas. Imágenes tomadas por microscopía de transmisión electrónica en la cual se puede observar a *M. genitalium* adherido a la superficie de las células (paneles superiores) y penetrando al interior de las células del huésped (paneles inferiores). Imagen realizada por el grupo del Dr. J.Piñol de la Universidad Autónoma de Barcelona.

II) Nutrición: En apartados anteriores se comentó que las rutas de síntesis de algunas moléculas importantes para la supervivencia del micoplasma son limitadas. Una vez adherido el micoplasma, en su membrana hay una serie de transportadores capaces de obtener sustancias como oxígeno, glucosa, aminoácidos…. del interior de la célula del huésped [80,81].

III) Invasión de la célula huésped: Como se comentó en el apartado III) Evasión del sistema inmune, el micoplasma se puede internalizar al interior de la célula pudiendo aumentar su supervivencia. Además, se va a nutrir de los metabolitos de la célula huésped, como los precursores para la síntesis de nucleótidos obtenidos a través de las nucleasas que secretan [13] y que están relacionadas con la inducción de apoptosis por la fragmentación internucleosomal de la cromatina [82].

IV) Radicales libres de oxígeno: La generación de la especie H_2O_2 es un posible subproducto del metabolismo del glicerol a través de la enzima L-alfa-glicerofosfato oxidasa del propio micoplasma [83]. El H_2O_2 liberado por el patógeno o por el propio sistema inmune, no puede ser eliminado de las células del huésped debido a la inhibición de la enzima catalasa por el micoplasma, lo que originará que la célula del huésped sea más susceptible al estrés oxidativo producido a lo largo del proceso de infección [84]. Ensayos realizados en células y ratones demuestran que el aumento de H_2O_2 provoca una disminución tanto de la actividad en proteínas de repuesta a estrés oxidativo como el aclaramiento del líquido pulmonar provocado por una alteración del transporte de iones en los alveolos [85]. También, se ha comprobado que existe una relación entre el incremento de H_2O_2 con el aumento en la presión de oxígeno en el interior de las células del huésped, lo que desencadenará finalmente la muerte celular [83]. La presencia de H_2O_2 producirá la lisis celular de los eritrocitos por varias vías, como por ejemplo, la desnaturalización de la hemoglobina [86].

V) Community-acquired respiratory distress syndrome toxin (CARDS toxin): Es una toxina que se encuentra exclusivamente en *M. pneumoniae* y es secretada desde el interior del micoplasma al citoplasma de las células alveolares por endocitosis mediada por clatrina [87,88]. La toxina CARDS desencadena en la célula huésped la vacuolización celular, la exfoliación de las células de la mucosa, la ciliostasis y también está involucrada en procesos proinflamatorios [89-91].

VI) Inflamación: El proceso de inflamación por el micoplasma empieza con la activación de receptores tipo *Toll*-1, *Toll-2* y *Toll*-6 por las lipoproteínas de la membrana del micoplasma [92]. Los receptores tipo *Toll*, a su vez activan el factor nuclear κβ (NF-κβ)

que sintetiza citoquinas proinflamatorias como la interleucina-1β (IL-1β) y la interleucina-8 (IL-8) y el factor de necrosis tumoral alfa (TNF-α) [93]. En humanos y animales el efecto de las citoquinas es el de actuar como reguladores de los eventos que desempeñan las células del sistema inmunitario y minimizar la enfermedad, aunque la síntesis de interleucinas de tipo proinflamatorio puede exacerbarla a través del desarrollo de lesiones inmunológicas. Por lo que, cuanto más se estimulen dichas citoquinas más se agrava la enfermedad clínica y las lesiones pulmonares. Otra de las vías para la producción de citoquinas proinflamatorias es por la toxina CARDS a través de la activación de los inflamosomas. El inflamosoma es un complejo multiproteico presente en el citoplasma y que está constituido por un receptor NOD (NLR), un receptor AIM2, la proteína tipo punto asociada a apoptosis (ASC) y la caspasa-1. La toxina CARDS activa la caspasa-1, quien escindirá la pro-IL-1β para generar la IL-1β [94]. Otra de las vías para la inducción de respuesta inflamatoria es por autofagia, mediante la activación de los receptores tipo *Toll*-4 [95].

VII) Sistema inmune humoral: Como se citó anteriormente, algunos micoplasmas poseen IgA proteasas y nucleasas que degradan inmunoglobulinas de tipo A y neutrófilos, respectivamente [73,74]. En pacientes de *M. pneumoniae* se ha observado una estrecha correlación entre la presencia de P1 y un aumento de las IgE (inmunoglobulina involucrada en procesos alérgicos). Uno de los efectos de las reacciones alérgicas provocadas por IgE, es la inflamación de las vías respiratorias promoviendo la aparición de asma. También, se ha observado que las infecciones por *M. pneumoniae* ocasionan un aumento de Th2 y células inmunes, así como cambios fisiológicos que son factores proactivos para la aparición o agravamiento del asma [96-98]. Aunque en la infección por micoplasma se hayan detectado moléculas que estén presentes en pacientes con asma, no hay evidencias directas del desarrollo de esta enfermedad por infecciones de *M. pneumoniae*. Sin embargo, en los diferentes estudios realizados se detectó la presencia de este patógeno en un elevado porcentaje de los pacientes con asma (10-50%) [99-101].

VIII) Daño inmune celular: En las infecciones por micoplasma se produce una reducción en la proporción del número de células T CD4+:CD8+ por la muerte celular de

los T CD4+ o por la expansión de los T CD8+. Esta ratio es globalmente utilizada como marcador del estado del sistema inmune de una persona, cuando el valor es inferior a 1 está relacionada con enfermedades que afectan al sistema inmune, por ejemplo, infecciones por VIH o tuberculosis [102]. Otra relación que se ve afectada es la de los linfocitos Th1:Th2 (linfocitos diferenciados a partir de los T que cooperan mediante la producción de citoquinas), aunque esta parece ser más controvertida ya que en ensayos con ratas se observó que en algunas de ellas dominaba Th1 que producía una agregación de linfocitos peribronquiales. Otros análisis presentaban un valor de Th2 más elevado que exhibían hiperplasia de células mesenquimatosas de los alveolos. Como comentamos en el apartado anterior, el aumento de producción de estas moléculas Th2 puede estar relacionado con la aparición del asma [103].

3.5 División celular

La división celular en micoplasmas del *cluster* de *Pneumoniae* sigue siendo una incógnita debido a que se desconoce el mecanismo por el cual se rige. Aparentemente, la fisión binaria debería ser el mecanismo más probable, ya que es el esencial en el mundo bacteriano. Sin embargo, un estudio demostró que sin FtsZ (proteína fundamental para llevar a cabo la división celular bacteriana) se podía llevar a cabo la división en *M. genitalium* usando la fuerza originada por la OT [104]. Así mismo, el gen que codifica la FtsZ no es esencial en *M. genitalium*, ya que su eliminación no provoca alteraciones en la división celular, en los componentes del citoesqueleto ni en los parámetros del crecimiento celular. Los resultados obtenidos en ensayos indican que la OT desarrolla un papel fundamental en este proceso [104]. Un estudio *in vivo* con proteínas de la OT de *M. pneumoniae* marcadas con fluorescencia mostró que la duplicación de la OT se produce inmediatamente antes que la citoquinesis, es decir, antes que se de la división de las células, y que los procesos de adhesión y *gliding* son fundamentales en la división celular ya que ayudan a separar las dos OTs de las células hijas [105]. Además, se probó que la duplicación del contenido genómico ocurre antes que la citoquinesis [106]. Otras proteínas que participan en la división del micoplasma son la topoisomerasa IV y la MraZ, la primera se encarga de relajar la tensión del ADN superenrrollado [107], mientras que la segunda (aún sin ser una proteína esencial) participa de forma directa o indirecta

en este proceso [105]. Con respecto a la duplicación de la OT existen dos teorías plausibles propuestas, una semi-conservativa y otra *de novo*. En la primera de ellas se propone que las proteínas de la zona electrodensa de la OT se separan y sirven como molde para una nueva OT [108], mientras que en la segunda teoría las OTs hijas se originarían *de novo* sin usar un molde [105] (**Figura 8**).

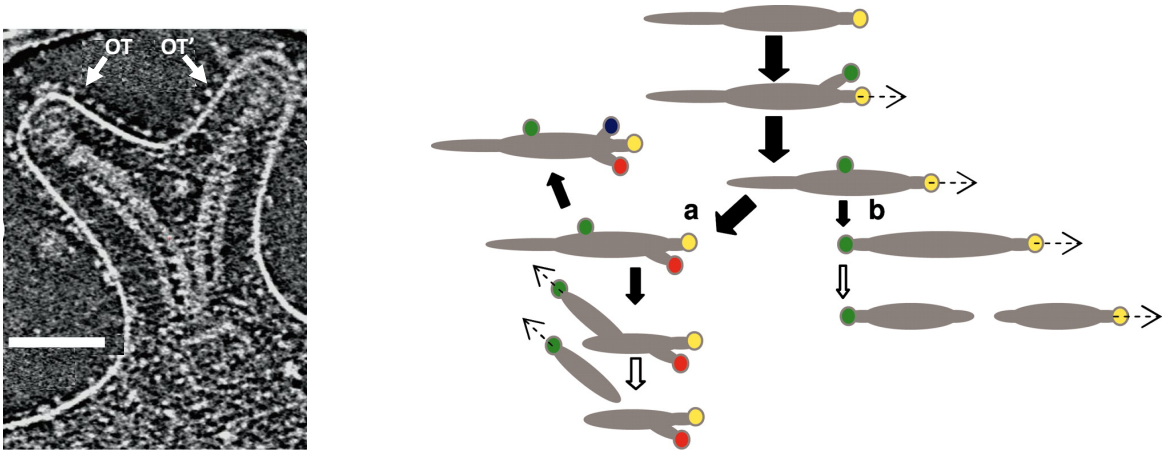


Fig. 8. Proceso de división del *M. pneumoniae*. a) Tomograma de la OT en el proceso de división. b) Modelo de la duplicación de la OT y el ciclo de crecimiento de *M. pneumoniae*. El círculo amarillo corresponde a la OT inicial y el resto de los círculos a las que aparecen posteriormente. Las flechas discontinuas indican el movimiento del orgánulo y las flechas blancas la citoquinesis. En la mayoría de los casos se producen varias multiplicaciones (1), aunque rara vez puede aparecer solamente una (2). Imágenes adaptadas de A.Kawamoto, American society for microbiology, 2016 y M.Benjamin et al, PNAS, 2006.

4 Epidemiología y tratamiento

4.1 Epidemiología

El pequeño tamaño que presenta *M. pneumoniae* le facilita su propagación por vía oral y es capaz de causar infecciones respiratorias superiores e inferiores creando epidemias y situaciones endémicas. Se estima que esta bacteria es responsable de hasta el 40% de las neumonías adquiridas en personas de todas las edades [109]. Además, *M. pneumoniae* puede inducir manifestaciones clínicamente significativas en sitios extrapulmonares y/o efectos inmunológicos en hasta un 25% de las infecciones [2]. Una vez que el micoplasma se introduce en el huésped, el periodo de incubación puede ser de hasta tres semanas. Tras este periodo, se desarrolla la denominada neumonía atípica

(tos seca, fiebre baja…) y pueden aparecer otros síntomas como otitis, congestión, faringitis… (**Figura 9**). Los síntomas suelen ser leves, pero existe un 18% de casos que necesitan hospitalización en los cuales el oxígeno en sangre es bajo o tienen dificultades para respirar [110]. Sin embargo, las ratios de mortalidad son bajos y oscilan dependiendo de la población, como también ocurre actualmente con la epidemia del virus SARS Covid-19, entre un 3-5%. Pero en casos de graves sintomatologías este porcentaje sube hasta alrededor del 30% [111,112]. También, se ha demostrado que *M. pneumoniae* está presente en coinfecciones con otro tipo de bacterias que afectan a las vías respiratorias como puede ser *S. pneumoniae* y también con virus, como el recientemente nombrado SARS Covid-19, exacerbando los síntomas y aumentando la tasa de mortalidad [113].

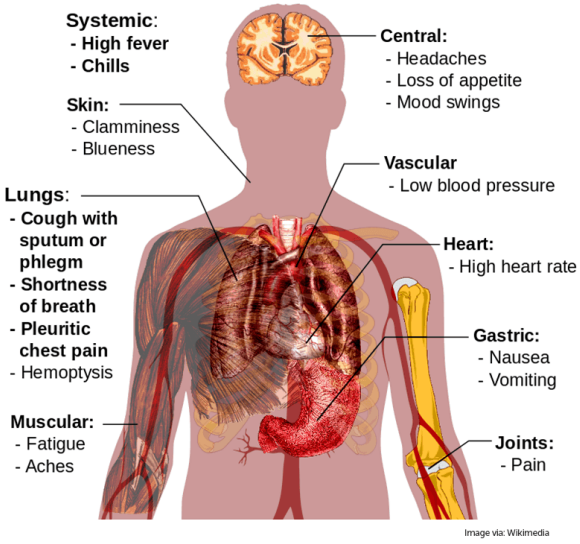


Fig. 9. Esquema de las manifestaciones clínicas más habituales por infección de *M. pneumoniae* (S.Aryal, Microbe Notes, 2017)

M. pneumoniae se encuentra distribuida por todo el mundo, aunque los mayores focos de infección son Japón, China y varios puntos de Europa y América (**Figura 10**). Al parecer, hay ciclos de los diferentes tipos y subtipos de *M. pneumoniae*, ya que en estudios reportados desde Japón demostraron que las infecciones entre los años 1995-2001 fueron por P1-tipo2, mientras que, desde 2002-2005 fue por P1-tipo1 [114]. Aunque también pueden coexistir diferentes tipos en una misma población [114]. La tasa de *M. pneumoniae* es otro aspecto que también sufre variaciones, se ha observado que la epidemia ocurre de forma cíclica cada 3-5 años. Un estudio reportó que esta epidemia

cíclica estuviera relacionada con cambios climáticos, ya que en América de Norte el brote ocurre a finales de verano u otoño, mientras que en Japón la incidencia mensual de casos de *M. pneumoniae* parece corresponderse con la oscilación del Sur-El Niño y el Dipolo del Océano Índico [115].

La epidemiología asociada a *M. genitalium* es completamente diferente de *M. pneumoniae. M. genitalium* es un patógeno de las vías urogenitales de transmisión sexual, provoca sobre todo cervicitis en mujeres y uretritis en hombres, pudiendo llegar en casos extremos a la infertilidad [116]. *M. genitalium* también posee una clasificación en tipos, pero no hay estudios sobre variables climáticas o estacionales que puedan afectar a la tasa de infección, aunque sí se ha observado la variación según el género, siendo mayor su prevalencia en hombres que en mujeres [117]. La incidencia es del 6%, pero la ratio puede variar según la población, ya que en personas de alto riesgo de transmisión sexual sube entre un 9-50% [118]. A diferencia de *M. pneumoniae*, las infecciones *por M. genitalium* no causan ningún tipo de sintomatología que pueda desencadenar en la muerte. También existe coinfección con otros microorganismos como se han reportado con el virus VIH [119].

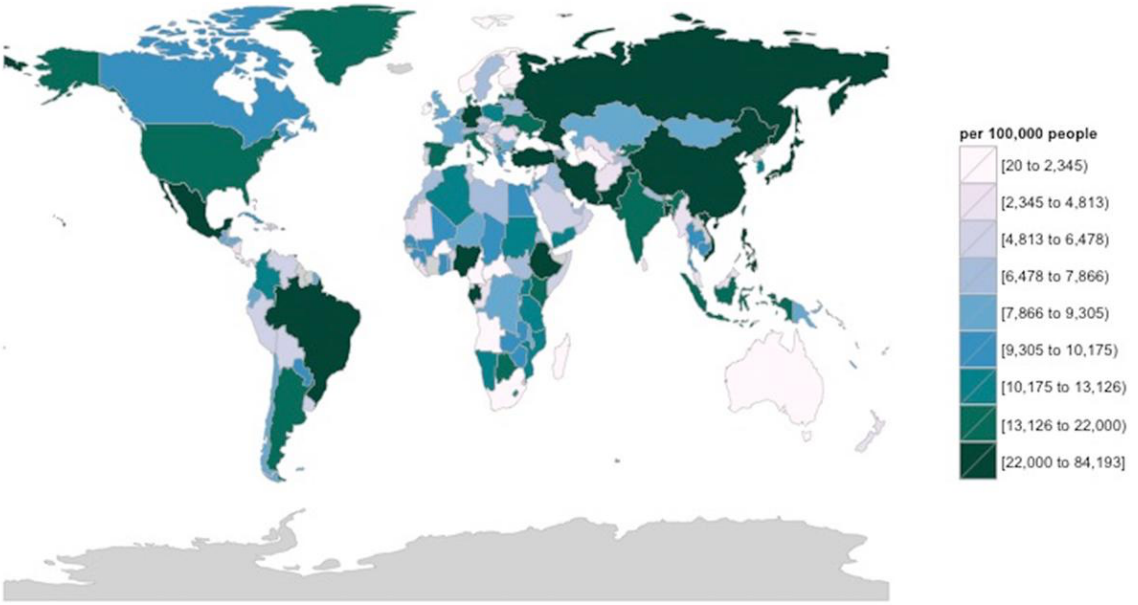


Fig. 10. Mapa de distribución de neumonía provocada por *M. pneumoniae* interpolada entre los años 2000-2012, según literatura (G.K. Parrott et al, Frontiers in Microbiology, 2016).

4.2 Diagnóstico y tratamiento

M. pneumoniae está siendo considerado en cualquier infección de las vías respiratorias, especialmente en pacientes que no respondan ante un tratamiento de antibióticos del tipo β-lactámico (no tiene efecto al no tener pared bacteriana). En Japón han desarrollado un diagnóstico rápido de neumonía típica o atípica mediante un sistema de puntuación según varios parámetros. Dependiendo de los resultados se suministra macrólidos o tetraciclinas, aunque estudios recientes han observado mutaciones en la subunidad 23s del ribosoma del *M. pneumoniae* que confiere resistencia a esta familia de antibióticos [120]. Las pruebas más fiables para el diagnóstico de neumonía por *M. pneumoniae* son las serológicas y PCRs. Las serológicas son principalmente los inmunoensayos, concretamente los ELISAs que consisten en la detección de la unión antígeno-anticuerpo. El fragmento del dominio C de P1 se utiliza como antígeno para este método de ensayo (DiaSorin). En estos ELISAs, se pueden dar falsos positivos, ya que se ha demostrado que existe una reactividad cruzada entre P1 de *M. pneumoniae* y P140 de *M. genitalium* [121]. Sin embargo, la diferencia entre las sintomatologías producidas por cada patógeno debería ser suficiente para aceptar o rechazar un resultado. Los ensayos de PCRs se realizan con oligos de secuencias específicas de *M. pneumoniae*. Otras técnicas que se utilizan menos frecuentemente debido a su baja eficiencia y/o rapidez son los cultivos de bacterias, las radiografías y las pruebas de aglutinación. Las pruebas de aglutinación reconocen anticuerpos aglutinantes (anticuerpos inespecíficos que aglutinan eritrocitos humanos a bajas temperaturas) que se forman tras la infección por *M. pneumoniae* [122]. El principal medio para el tratamiento son los antibióticos (no β-lactámicos). Si persiste la enfermedad se puede suministrar corticosteroides (para modular la inflamación) o terapia inmunomoduladora.

Con respecto a *M. genitalium*, los métodos de diagnóstico y tratamiento utilizados son muy similares a los de *M. pneumoniae*, se basan en la recogida de una muestra de hisopos vaginales o primera orina de la mañana en el caso de mujeres y hombres respectivamente, a continuación, se realiza una PCR con oligos específicos. En cuanto a los tratamientos, el más común es la azitromicina, que pertenece al grupo de

los macrólidos (siempre que no se detecte mutación en la subunidad 23s del ribosoma) o en otros casos de persistente enfermedad se suministran antibióticos no macrólidos o β-lactámicos, como pueden ser la doxiciclina o la moxifloxacina.

A pesar de que la localización y la sintomatología de *M. pneumoniae* y *M. genitalium* son distintas, presentan una alta homología en varías proteínas, lo que hace que tanto su morfología como la capacidad de adherirse, moverse, dividirse e infectar a células sean muy similares. Según diversas investigaciones, estas capacidades se pueden desarrollar gracias al papel que juega la OT, convirtiéndola en una de las principales dianas de estudio en micoplasmas. Aunque la sintomatología de estos microorganismos difiere, el diagnóstico y el tratamiento son parecidos, con *M. pneumoniae* presentando cuadros clínicos más graves que pueden llegar incluso a producir la muerte.

Objetivos

OBJETIVOS

La presente tesis está enfocada principalmente a la caracterización biológica y estructural de las proteínas de *M. pneumoniae* que componen el Nap, una de las subestructuras de la *Organela Terminal*, y que son esenciales para la patogenicidad del microorganismo. También se ha estudiado el dominio C de P140, proteína ortóloga a P1 en *M. genitalium*, y participado en la estructura completa de la misma. El objetivo general de la investigación se enmarca en los procesos de motilidad y adhesión en micoplasmas del *cluster* de *Pneumoniae*.

Los objetivos específicos se presentan a continuación de acuerdo con la cronología con que se fueron desarrollando los distintos estudios:

1. Determinación estructural del dominio C de P140 de *M. genitalium*.

2. Colaboración en la resolución de la estructura de P140 y determinación de la organización del Nap en *M. genitalium*.

3. Determinación estructural, implicaciones funcionales, biológicas e inmunológicas de las proteínas que componen el Nap en *M. pneumoniae*.

Capítulo 1

Conformación alternativa del dominio C de la proteína P140 de *Mycoplasma genitalium*

Vizarraga, D.; Pérez-Luque, R.; Martín, J.; Fita, I.; Aparicio, D.

Resumen

El patógeno humano *Mycoplasma genitalium* es responsable de uretritis en hombres y de cervicitis y enfermedad inflamatoria pélvica en mujeres. La adherencia de *M. genitalium* a las células epiteliales humanas está mediada por un complejo de adherencia llamado Nap, que es esencial para la infectividad. El Nap es un dímero de heterodímeros formado por las proteínas transmembrana e inmunodominantes P110 y P140. El genoma de *M. genitalium* contiene múltiples copias de porciones que comparten homología con las regiones extracelulares de P140 y P110, codificadas por los genes *mg191* y *mg192*, respectivamente. La recombinación homóloga entre los genes y las copias permite la generación de una gran diversidad de variantes de P140 y P110 para superar la vigilancia del sistema inmunológico del huésped. Curiosamente, el dominio C de la región extracelular de P140 presenta un bajo grado de variabilidad de secuencia y parece ser esencial para la función de Nap. Este dominio C confiere alta flexibilidad en su unión con el dominio N y además ancla la proteína a la membrana del micoplasma. En el presente trabajo, se describen las estructuras tridimensionales obtenidas por rayos X de dos formas cristalinas de una construcción del dominio C de P140. En ambas formas cristalinas, se forma un octámero compacto con simetría de grupo puntual D4. La estructura del dominio C determinada en este trabajo presenta diferencias significativas con respecto a la estructura del dominio C encontrado recientemente en P140 entera. La plasticidad estructural del dominio C parece ser un posible mecanismo que puede ayudar en el funcionamiento del complejo de adhesión de los micoplasmas.

electronic reprint

STRUCTURAL BIOLOGY
COMMUNICATIONS

ISSN: 2053-230X
journals.iucr.org/f

Alternative conformation of the C-domain of the P140 protein from *Mycoplasma genitalium*

David Vizarraga, Rosa Pérez-Luque, Jesús Martín, Ignacio Fita and David Aparicio

Acta Cryst. (2020). F**76**, 508–516

STRUCTURAL BIOLOGY COMMUNICATIONS

ISSN 2053-230X

Alternative conformation of the C-domain of the P140 protein from *Mycoplasma genitalium*

David Vizarraga, Rosa Pérez-Luque, Jesús Martín, Ignacio Fita and David Aparicio*

Department of Structural Biology, Institut de Biologia Molecular de Barcelona (IBMB–CSIC), Parc Científic de Barcelona, Baldiri Reixac 10-12, 08028 Barcelona, Spain. *Correspondence e-mail: daacri@ibmb.csic.es

Received 8 July 2020
Accepted 7 September 2020

Edited by B. Kobe, University of Queensland, Australia

Keywords: *Mycoplasma genitalium*; adhesins; Nap; gliding motility; infection.

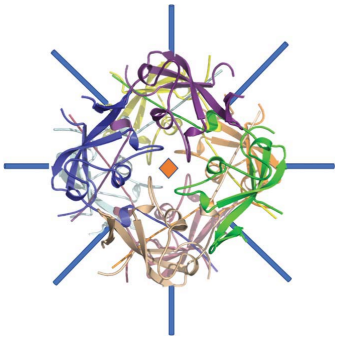
PDB references: P140 C-domain, space group *C*2, 6rcd; space group *I*422, 6rcc

Supporting information: this article has supporting information at journals.iucr.org/f

The human pathogen *Mycoplasma genitalium* is responsible for urethritis in men, and for cervicitis and pelvic inflammatory disease in women. The adherence of *M. genitalium* to host target epithelial cells is mediated through an adhesion complex called Nap, which is essential for infectivity. Nap is a transmembrane dimer of heterodimers of the immunodominant proteins P110 and P140. The *M. genitalium* genome contains multiple copies of portions that share homology with the extracellular regions of P140 and P110 encoded by the genes *mg191* and *mg192*, respectively. Homologous recombination between the genes and the copies allows the generation of a large diversity of P140 and P110 variants to overcome surveillance by the host immune system. Interestingly, the C-terminal domain (C-domain) of the extracellular region of P140, which is essential for the function of Nap by acting as a flexible stalk anchoring the protein to the mycoplasma membrane, presents a low degree of sequence variability. In the present work, the X-ray crystal structures of two crystal forms of a construct of the P140 C-domain are reported. In both crystal forms, the construct forms a compact octamer with *D*4 point-group symmetry. The structure of the C-domain determined in this work presents significant differences with respect to the structure of the C-domain found recently in intact P140. The structural plasticity of the C-domain appears to be a possible mechanism that may help in the functioning of the mycoplasma adhesion complex.

1. Introduction

Mycoplasmas are microorganisms that are characterized by their reduced genomes and their lack of a cell wall, presenting only a single phospholipid bilayer, the plasma membrane (Razin, 1992; Razin *et al.*, 1998). These features make mycoplasmas unique among bacteria, explaining why they are used as model organisms in systems and synthetic biology (Xavier *et al.*, 2014; Gibson *et al.*, 2008; Glass *et al.*, 2006). Many mycoplasmas live in close association with eukaryotic cells, either intracellularly or at the cell surface (Rottem, 2003). *Mycoplasma genitalium*, the causative agent of urethritis in men and cervicitis and pelvic inflammatory disease in women, presents arguably the smallest genome amongst organisms that are capable of self-replication and axenic growth (McGowin & Anderson-Smits, 2011; Glass *et al.*, 2006). The adherence of *M. genitalium* to host target epithelial cells is mediated by an ∼0.5 MDa transmembrane complex called Nap, which is composed of a dimer of the heterodimer formed by the P110 and P140 proteins (Aparicio *et al.*, 2018, 2020; Burgos *et al.*, 2006; Mernaugh *et al.*, 1993; Scheffer *et al.*, 2017). In *M. genitalium* and in other members of the pneumonia cluster of mycoplasmas, such as the important human pathogen *M. pneumoniae*, the Nap adhesion complex also participates in



a unique type of gliding cell motility that is essential for infectivity (Krause, 1996; Krause & Baseman, 1982; Radestock & Bredt, 1977; Burgos *et al.*, 2006; García-Morales *et al.*, 2016; Nakane *et al.*, 2011; Seto *et al.*, 2005). Besides their roles in adhesion and gliding motility, the P110 and P140 proteins are probably the most immunodominant proteins in *M. genitalium* (Svenstrup *et al.*, 2002; Morrison-Plummer *et al.*, 1987). The genome of *M. genitalium* contains several non-identical copies of portions of the extracellular regions of P110 and P140, named MgPar sequences (Ma *et al.*, 2010; Peterson *et al.*, 1995). By orthologous recombination with the corresponding genes, the copies allow the generation of a large diversity of variants of P110 and P140 in order to overcome surveillance by the host immune system (Iverson-Cabral *et al.*, 2007; Ma *et al.*, 2007). The only genetically invariable continuous stretches in the extracellular regions of P110 and P140 correspond to the C-terminal domain (C-domain) and the hinge connecting the N-domain and C-domain (Ma *et al.*, 2010). The C-domain of P140 is even highly conserved when compared with the corresponding domain of the orthologous protein P1 from *M. pneumoniae*, suggesting a critical functional role that, together with its low genetic variability, makes this domain a promising therapeutic target. A better understanding of the cytoadherence and gliding mechanisms could facilitate the implementation of more effective antimicrobial approaches based on anti-adhesion therapies.

This work presents X-ray crystal structures of a construct from the P140 C-domain from *M. genitalium* in two different crystal forms. The structure of the isolated domain determined here shows major differences with respect to the recently reported structure of the C-domain when forming part of the intact Nap. Possible roles for the structural plasticity of the C-domain in the functioning of Nap are discussed.

2. Experimental procedures

2.1. Cloning, expression and protein purification

The C-domain of P140 was amplified from a synthetic clone of *M. genitalium* gene *mg191* (GenScript) using the primers P140cF and P140cR (Supplementary Table S1). The PCR fragment was cloned to generate a construct comprising residues 1220–1351 with a His tag at the N-terminal end containing an upstream proteolytic sequence for protease 3C (Supplementary Fig. S1). Recombinant protein was obtained after 4 h of expression at 37°C in *Escherichia coli* B834 (DE3) cells (Merck) upon induction with 0.5 m*M* isopropyl *β*-D-1-thiogalactopyranoside. The cells were harvested and lysed by sonication in 1× PBS and 40 m*M* imidazole (binding buffer) followed by centrifugation at 20 000*g* at 4°C. The supernatant was charged onto a pre-equilibrated HisTrap 5 ml column (GE Healthcare) with binding buffer and was eluted with binding buffer containing 400 m*M* imidazole. Soluble aliquots were treated overnight with protease 3C in a 1:100 molar ratio (protease 3C:P140 C-domain) and subsequently passed through a reverse HisTrap column. Finally, the aliquots were pooled and loaded onto a Superdex 200 GL 10/300 column (GE Healthcare) pre-equilibrated with 20 m*M* Tris pH 7.4, 150 m*M* NaCl. To prepare a selenomethionine (SeMet) derivative of the C-domain, bacteria were grown in Seleno-Methionine Medium Complete (Molecular Dimensions) using the protocol provided with the kit. Purification of the SeMet-labeled C-domain was performed following the same steps as described for the native protein.

2.2. Crystallization, data collection and structure determination

Crystallization screening was performed using either a Cartesian or a Crystal Phoenix robot by mixing 150 nl crystallization condition with 150 nl protein solution at 6 mg ml^{-1} in 96-well plates. The best diffracting crystals belonged to the monoclinic space group *C*2 (unit-cell parameters a = 114.5, b = 83.0, c = 118.2 Å, β = 117.5°) and the tetragonal space group *I*422 (unit-cell parameters a = 90.1, b = 90.1, c = 58.0 Å). The monoclinic crystals were obtained in crystallization buffer consisting of 0.15 *M* DL-malic acid, 0.1 *M* bis-Tris propane pH 7.0 and the tetragonal crystals were obtained in crystallization buffer consisting of 20% PEG 3350, 0.02 *M* $MgCl_2{\cdot}6H_2O$, 0.1 *M* Tris pH 8.5. All crystals used for data collection were flash-cooled in liquid nitrogen with 15% glycerol as a cryo-protectant.

X-ray data sets for the monoclinic and tetragonal crystal forms were collected to resolutions of 1.98 and 1.43 Å, respectively. X-ray data collection was carried out on the XALOC beamline (Juanhuix *et al.*, 2014) at the ALBA Synchrotron, Spain. SAD data were collected from an SeMet-derivatized tetragonal crystal at the Se *K*-edge energy. Data were processed in *xia*2 (Winter, 2010) using *XDS* (Kabsch, 2010), and *AIMLESS* and *POINTLESS* (Evans, 2006) from the *CCP*4 suite of programs (Winn *et al.*, 2011). The structure of the C-terminal domain construct was solved by SAD from a tetragonal SeMet-derivatized crystal at 1.65 Å resolution using *SHELX*/*E* (Sheldrick, 2015). The structures of the native tetragonal and monoclinic crystals were then solved by molecular replacement using *Phaser*, and contained one and eight subunits in the asymmetric unit, respectively. The structures were refined by alternating interactive and automatic cycles with *Coot* (Emsley *et al.*, 2010) and *REFMAC* (Murshudov *et al.*, 2011), giving agreement factors R and $R_{\rm free}$ of 18.3% and 20.0%, respectively, for the *I*422 structure, and 17.6% and 20.6%, respectively, for the *C*2 structure (Table 1). The final refined structures have been deposited in the PDB as entry 6rcc for the *I*422 structure and 6rcd for the *C*2 structure.

2.3. SEC-MALS analysis

The molecular weight and oligomerization of the construct in solution were measured using a Superdex 75 10/300 GL column (GE Healthcare) in a Prominence liquid-chromatography system (Shimadzu) connected to a DAWN HELEOS II multi-angle light-scattering (MALS) detector and an Optilab T-rEX refractive-index (dRI) detector (Wyatt Technology). The *ASTRA* 7 software (Wyatt Technology) was used for data

Table 1
X-ray data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

	P140 C-domain, SeMet derivative	P140 C-domain	P140 C-domain
Data-collection statistics			
Space group	*I*422	*I*422	*C*2
a, *b*, *c* (Å)	90.21, 90.21, 58.17	90.10, 90.10, 57.98	114.51, 83.02, 118.16
α, *β*, *γ* (°)	90.00, 90.00, 90.00	90.00, 90.00, 90.00	90.00, 117.50, 90.00
Unique reflections	14807 (2121)	20651 (3175)	67417 (9766)
Resolution (Å)	48.89–1.65 (1.74–1.65)	45.04–1.43 (1.51–1.43)	52.41–1.98 (2.09–1.98)
Wavelength (Å)	0.97911	1.07216	1.07216
R_{merge}† (%)	0.104 (1.565)	0.09 (1.73)	0.08 (0.90)
$\langle I/\sigma(I)\rangle$	16.3 (2.0)	11.50 (1.5)	10.50 (1.9)
$CC_{1/2}$	0.997 (0.644)	0.998 (0.71)	0.998 (0.64)
CC_{ano}	0.289 (0.011)		
Completness (%)	100.00 (99.90)	92.41 (99.8)	98.30 (99.5)
Multiplicity	6.4 (4.1)	6.3 (6.3)	3.40 (3.4)
Refinement statistics			
Resolution		45.04–1.43 (1.51–1.43)	82.41–1.98 (1.99–1.98)
No. of reflections		20650 (3012)	67416 (1284)
R_{cryst}‡ (%)		18.30 (23.84)	17.80 (22.00)
R_{free}§ (%)		19.90 (26.87)	20.60 (27.24)
No. of residues (in unit cell)		99	782
No. of ligands		2	0
Solvent content (%)		51.80	54.43
Average *B* factor (Å^2)		29.94	46.22
Coordinate error¶ (Å)		0.20	0.23
R.m.s. deviation, bond lengths (Å)		0.01	0.01

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of an observation and $\langle I(hkl)\rangle$ is the mean value of observations for an unique reflection. ‡ $R_{cryst} = \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure-factor amplitudes, respectively. § R_{free} was calculated with 5% of the data, which were excluded from refinement. ¶ Based on maximum likelihood.

processing and analysis of the results. A d*n*/d*c* value of 0.185 ml g^{-1} (typical of proteins) was assumed for calculations.

3. Results

3.1. Structure determination of the C-domain from P140

Tetragonal and monoclinic crystals were obtained from a construct of the C-domain of P140 from *M. genitalium* spanning residues Lys1220–Asp1351 (Supplementary Fig. S1) with a two-residue extension (Gly-Pro) at the N-terminal end (remaining after the removal of the 6×His tag). The sequence contained two methionines, which allowed the preparation of an SeMet derivative in order to collect an anomalous data set (at 1.65 Å resolution) from the tetragonal crystals that was solved by SAD using *SHELXD* and *SHELXE* (Sheldrick, 2015; Table 1). These tetragonal crystals, which belonged to space group *I*422, contained one subunit in the crystal asymmetric unit, with a solvent content of 52% by volume. A complete model was built in the experimental map (Fig. 1*a*), which was then refined with *REFMAC* (Murshudov *et al.*, 2011) using a diffraction data set at 1.4 Å resolution obtained from a native protein crystal, giving final agreement factors *R* and R_{free} of 18.3% and 20.0%, respectively (Table 1). The monoclinic crystals, which belonged to space group *C*2, were then solved by molecular replacement with *Phaser* (McCoy, 2007) using the structure determined from the tetragonal crystals as a search model. The monoclinic crystals contained eight subunits in the asymmetric unit, with a solvent-content volume of 54%, and the structure was refined using *REFMAC* to a resolution of 1.98 Å using noncrystallographic symmetry restraints, giving final agreement factors *R* and R_{free} of 17.6% and 20.6%, respectively (Fig. 1*b* and Table 1).

3.2. Overall structure of the isolated C-domain

The C-domain construct presents a structure that is essentially identical in the tetragonal and monoclinic crystal forms. In both crystals, the closest neighboring subunits are organized as compact octamers with *D*4 point-group symmetry (Fig. 1*c*). In the tetragonal crystal form, with only one subunit in the asymmetric unit, the fourfold and the twofold axes of the octamer are coincident with crystallographic symmetry axes, implying that all of the subunits in the octamer are identical. In the monoclinic crystal, a whole octamer is found in the asymmetric unit, with the eight subunits related by accurate fourfold and twofold noncrystallographic symmetries. The packing of the octamers is very similar in both crystal forms, despite the differing space groups (Fig. 1*c*).

The structure determined contains a few disordered residues at the N-terminal end, which were not included in the model. These are followed by an extended stretch of about 20 residues that protrudes away from the subunit (Fig. 1*a*) and interacts extensively with other subunits in the octamer (Figs. 2*a* and 2*b*). There is then a zeta-like motif, approximately from Pro1268 to Ala1298, with two loosely antiparallel helices and an irregular stretch that reaches the first *β*-strand of a three-stranded antiparallel *β*-sheet. The third *β*-strand, which is the last strand in the *β*-sheet, is interrupted by a ten-residue *β*-bulge (Pro1330–Pro1339) that is exposed to the solvent and contains a few disordered residues in its central part, with the coordinates corresponding to residues Asn1333–

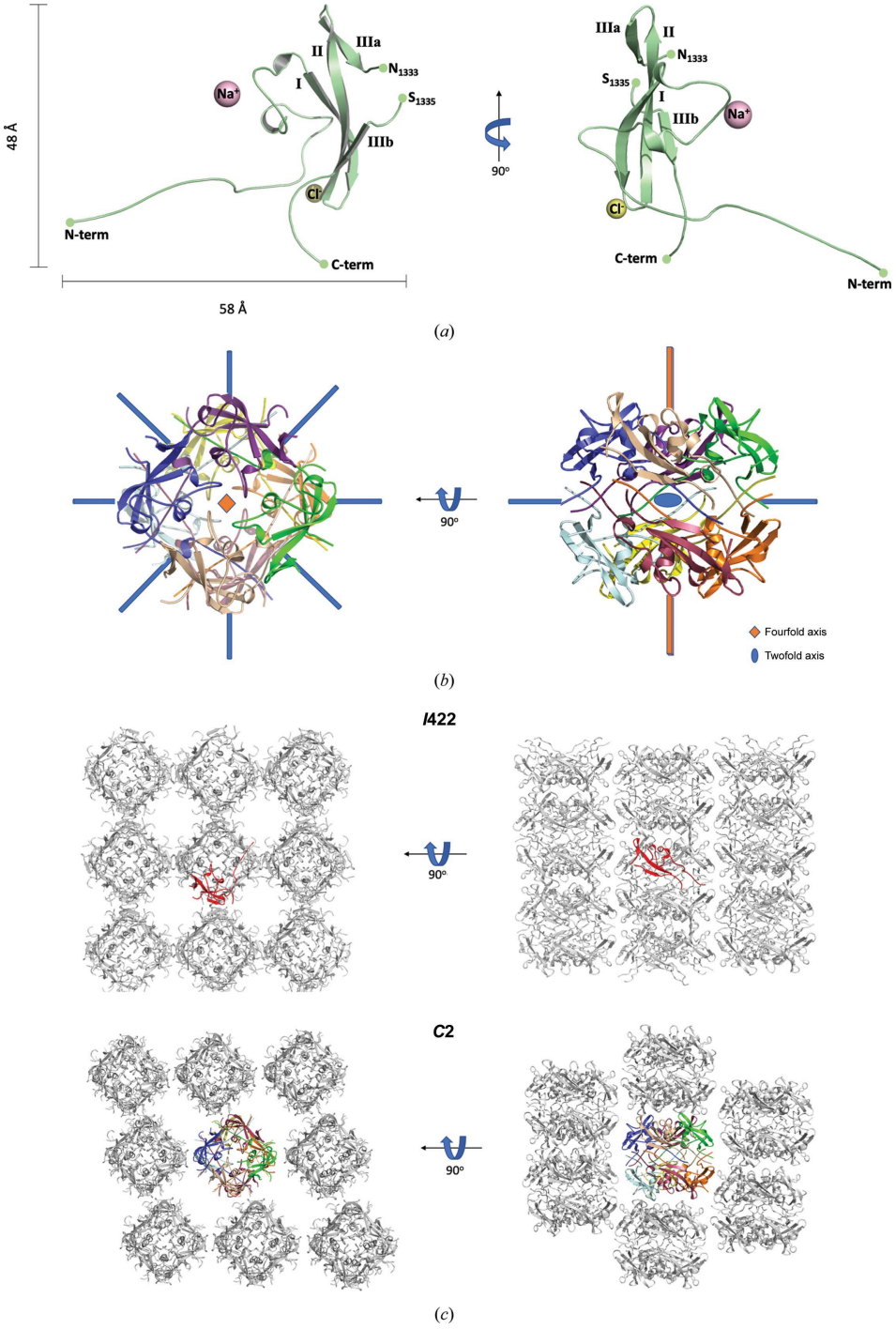


Figure 1
Structure of the isolated C-domain from P140. (*a*) Two views (90° apart from each other) of the structure of one subunit of the C-domain from P140 as found in the asymmetric unit of the tetragonal crystals. The three strands in the β-sheet are indicated. The third strand is interrupted by an insertion of ten residues that is partially disordered. The positions of an Na^+ ion and a Cl^- ion are also depicted. (*b*) Two views, 90° apart, down and perpendicular to the fourfold axis, of the octamer with *D*4 symmetry found in the asymmetric unit of the monoclinic crystals. (*c*) In spite of the differences in the unit-cell content, the structures of the subunits and the packing interactions are similar in both the tetragonal (*I*422, upper panels, with one subunit in the asymmetric unit shown in red) and the monoclinic (*C*2, lower panels, with a whole octamer shown in color) crystals.

Ser1336 missing in most of the subunits. The interaction between the zeta-like motif and the β-sheet defines a hydrophobic core of the domain, in which not even a single solvent (water) molecule has been found despite the large number of solvent molecules that were modeled in the tetragonal crystal, with a ratio of 1.3 waters per residue, owing to the quality and the resolution available.

In the octamer, interactions between subunits cover a large surface across the fourfold axis, with an averaged contact area between each neighboring pair of subunits of about 1200 Å^2 and an estimated free energy of −14 kcal mol^{-1} according to *PISA* (Krissinel & Henrick, 2007; Figs. 2*a* and 2*b*). Interactions between pairs of subunits cover a smaller surface across the twofold axis, but each subunit interacts with three other subunits through the twofold symmetries (with contact areas of about 500, 250 and 30 Å^2, respectively), which add up in the octamer (Figs. 2*a* and 2*c*). Therefore, although the tendency to form dimers between subunits might be weak, dimerization of the tetramers is expected to be favorable, with a contacting surface between tetramers of about 2750 Å^2 and

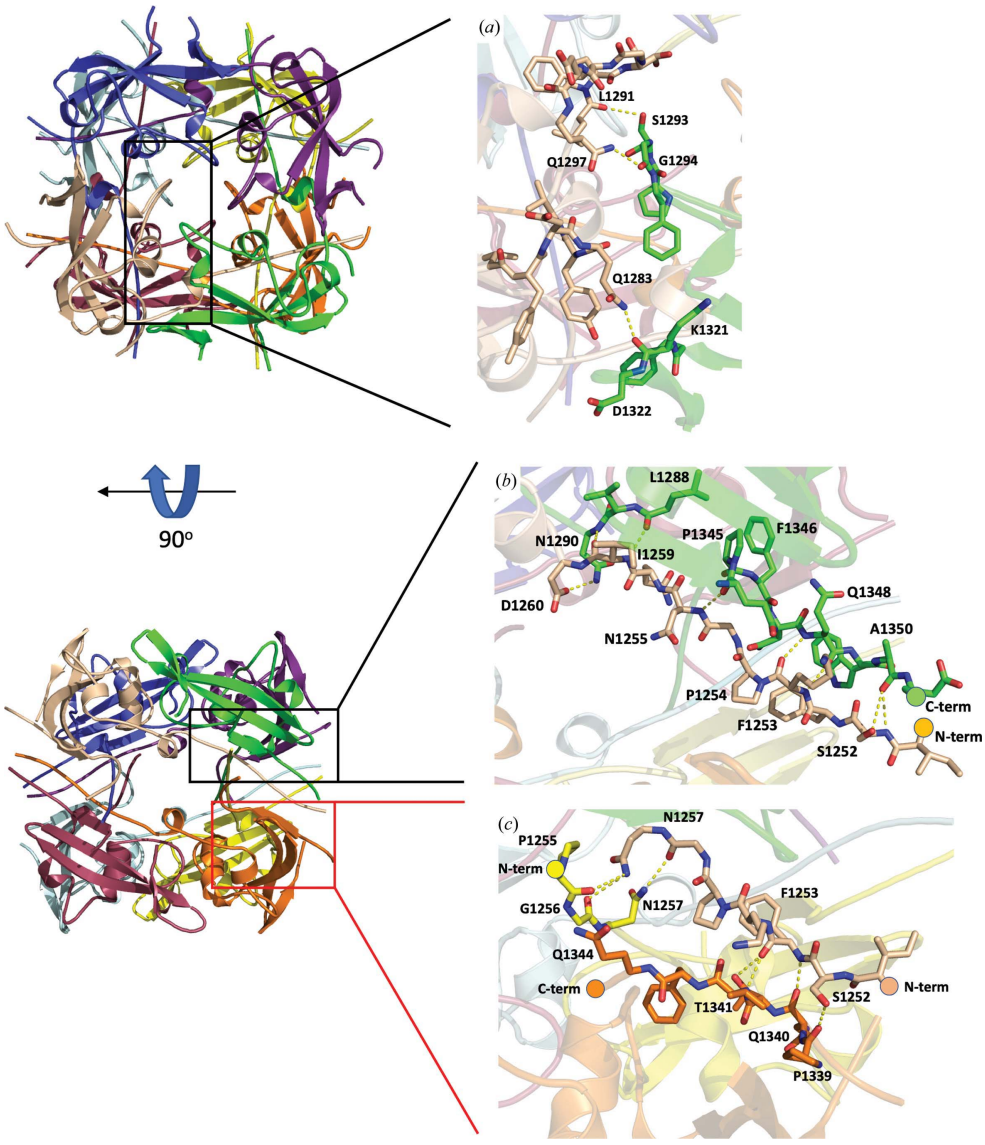


Figure 2
Interactions between subunits in the octamer. Two views, 90° apart from each other, showing the intersubunit interactions within the octamer. The insets show details of the residues involved in the interactions, with hydrogen bonds depicted as dashed lines. (*a*) View down the fourfold axis, showing interactions between subunits related by fourfold symmetry. (*b*) View down a twofold axis perpendicular to that in (*a*). (*c*) View down a twofold axis, showing interactions between the N- and C-terminal tails of the subunits related by this symmetry.

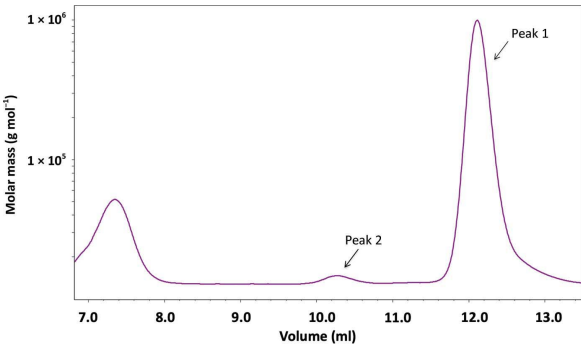
with the observed *D*4 octamers as the most stable oligomer. However, SEC-MALS analysis indicates that at the concentrations used (see Section 2) the C-domain construct remains monomeric in solution (Fig. 3). The interface between the two tetramers exclusively involves the N- and C-terminal ends of the eight subunits that are sandwiched at both sides by the four domain cores and contains a large number of solvent (water) molecules, reflecting that it is a hydrophilic interface. The entropic cost of fixing the many solvent molecules and all of the subunit ends together with the hydrophilic character of the interface might explain the presence of monomers in solution at the concentrations tested despite the large inter-subunit contacting surface in the octamers.

3.3. Comparison of the C-domain structures

Superposition of the structures of the C-domain construct solved in this work with the C-domain in the recently reported structure of P140 gives an r.m.s.d. of 1.0 Å for 55 equivalent residues (Fig. 4*a*). Thus, only about half (53%) of the residues are structurally equivalent between the two C-domain structures, while the other half of the residues show large structural differences, although the sequence is the same in both structures. The hydrophobic core, which includes the zeta-like motif and the three antiparallel strands of the β-sheet described above, is well conserved. In contrast, the N-terminal end region and the β-bulge (Pro1330–Pro1339), which emerges from the middle of the third β-strand in the three-stranded β-sheet, adopt different structures (Fig. 4*a*). There are also some structural differences for the residues at the very C-terminal end.

The contacting surface between the N- and C-domains in P140 is only about 650 Å^2 (according to *PISA*; Krissinel & Henrick, 2007; Fig. 4*b*). In the C-domain these interactions only involve the β-bulge and five other residues from the first and second β-strands of the β-sheet (Lys1303, Ile1305, Ser1306 and Val1316-Arg1317). As the conformations of these residues and of the corresponding strands remain essentially unchanged in the C-domain structures, only the absence of the interactions between the N-domain and the β-bulge appears to be responsible for the changes that appear in the isolated C-domain. The absence of these interactions can directly explain the conformational change of the β-bulge, which is extended in the structure solved in this work and bent in P140 (Fig. 4*b*). In turn, the extended β-bulge in the isolated C-domain lacks interactions with the N-terminal end region of the domain in P140, which changes the conformation of this N-terminal end region (Figs. 4*c* and 4*d*). Therefore, it appears that the β-bulge acts like a trigger that can be activated, one way or another, by just a few interactions with the N-domain of P140. Finally, it is worth mentioning that there are no interactions similar to those found between the subunits in the octamer that forms the isolated C-domain either in P140 or in the whole Nap (the structure of which has also recently been reported at ∼15 Å resolution from cryo-electron tomography and subtomogram averaging studies; Aparicio *et al.*, 2020).

Superposition of the C-domains from the Nap proteins P140 and P110 (Aparicio *et al.*, 2018), which present a unique topology (according to *DALI*; Holm & Rosenström, 2010), gives an r.m.s.d. of 2.2 Å for 70 equivalent residues (Fig. 5*a*). A β-bulge is also present in the corresponding β-strand from P110 and interacts with the N-domain (Fig. 5*b*). However, in P110 the β-bulge is extended and oriented in a direction that is not like that in P140, but is close to that found in the structure of the isolated C-domain. In P110 the N-end region of the C-domain appears to be stabilized by an extra hairpin with respect to P140 (Fig. 5*a*), which suggests that the C-domain of P110 does not have a triggering mechanism like that in P140.



	Mn (kDa)	Mw (kDa)	Polydispersity (Mw/Mn)	Calculated mass (µg)	Mass fractior (%)
Peak 1	70±0.07	16.71±0.07	1.00±0.01	141.54	98.86
Peak 2	42.49±0.67	42.49±0.68	1.00±0.02	1.63	1.14

Figure 3
SEC-MALS analysis of the C-domain of P140. SEC-MALS analysis of purified samples of the C-domain from P140 detected two species. The most abundant (peak 1, 98.86%), with a molecular weight of 16.71 kDa, corresponds to the monomeric species. A second species is observed (peak 2), with a molecular weight of 42.49 kDa, that could correspond to either unstable oligomers or to some contamination.

4. Discussion

The extracellular region of P140, which is probably the most immunodominant protein from *M. genitalium*, was predicted to contain a large N-domain and a small C-domain inmediately followed by a transmembrane helix. Here, we present the structure of the isolated C-domain. Initially, the reasons for undertaking structure determination of the isolated C-domain were mainly methodological as a possible means of contributing to the structure determination of the whole extracellular region of P140, for which crystals were available but were difficult to solve. However, the structure of the isolated C-domain, determined in this work by SAD using an SeMet derivative and refined at high resolution, did not provide any match on molecular replacement with the P140 crystals, indicating important differences between the structures of the isolated domain and of the domain within P140. Recently, the structure of P140 has become available (Aparicio *et al.*, 2020), confirming the important structural differences between the two C-domain structures, with only about half of the residues retaining structurally equivalent positions.

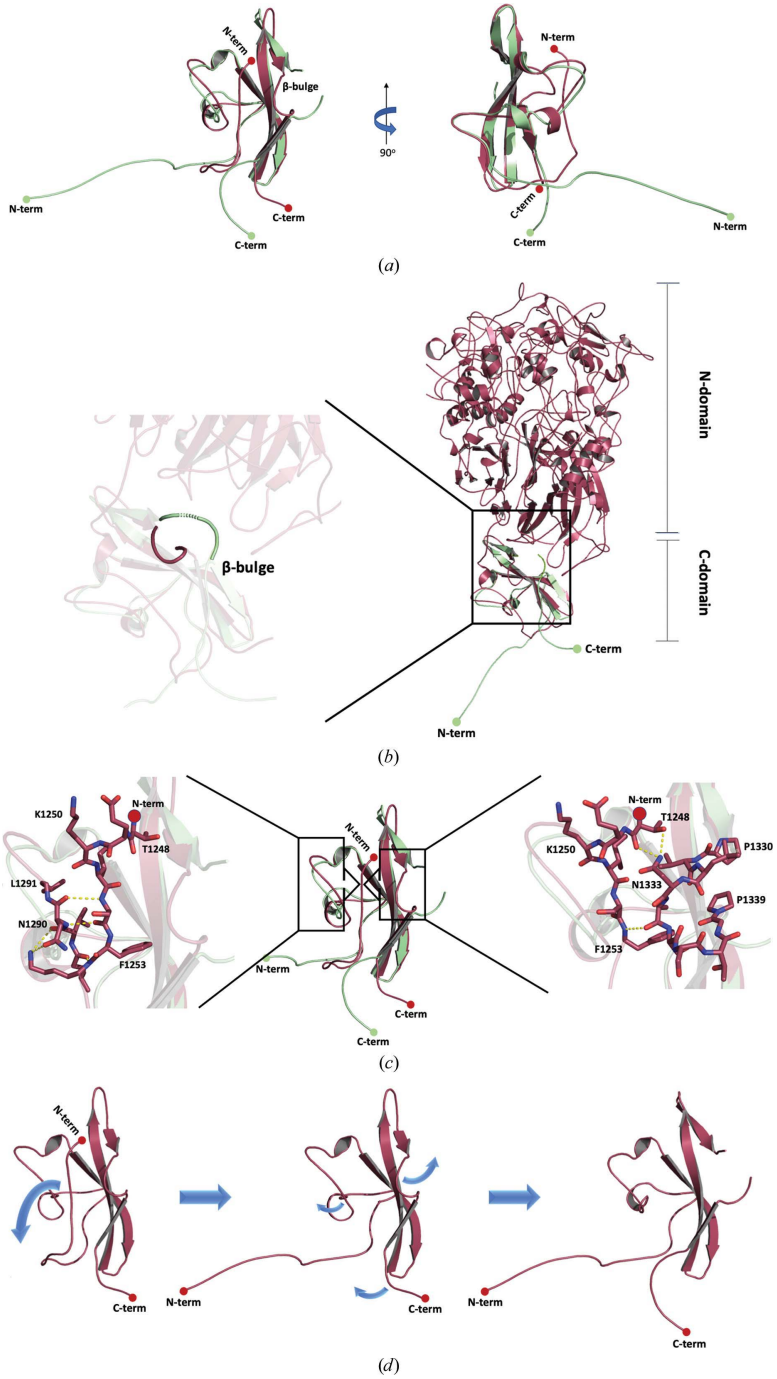
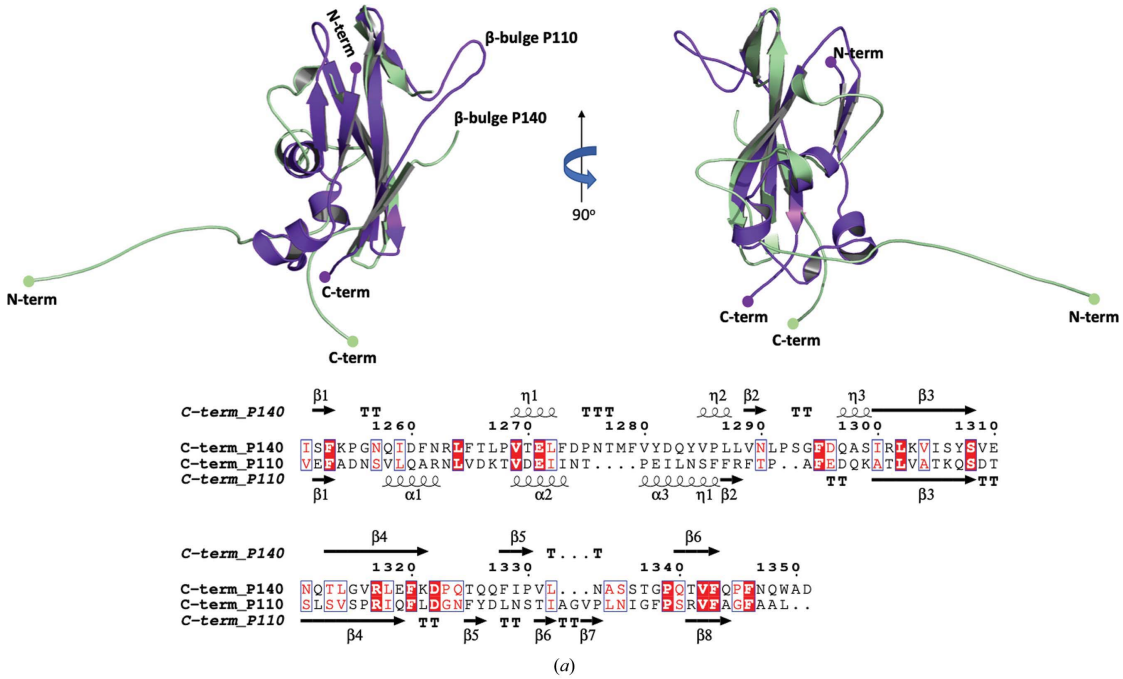


Figure 4
Comparison of the P140 C-domain structures. (*a*) Two views (90° apart from each other) of the superposed C-domain structures as found in this work (green) and in the structure of the whole P140 ectodomain (PDB entry 6s3u; red; Aparicio *et al.*, 2020). The largest differences are found in the N- and C-terminal tails and in the β-bulge. (*b*) The presence of the P140 N-domain would clash with the C-domain β-bulge as seen in this work, suggesting that movement of the β-bulge can be associated with movement of the N- and C-domains with respect to one another. (*c*) The insets show details of the interactions found in the whole P140 ectodomain structure that occur in the N-terminal end with either residues Leu1290 and Asn1291 (left) or with the β-bulge (right). (*d*) Scheme of the possible transition between the C-domain structures found in this work and in the whole P140 ectodomain. Disturbing the interactions of the N-terminus from the whole P140 ectodomain structure would destabilize this region, which in turn would eliminate the interactions with the β-bulge.

In the structure of P140, the interface between the N- and C-domains is small, involving only a few residues from the C-domain, all of which belong to a three-stranded β-sheet or to a β-bulge (Pro1330–Pro1339) that emerges from the middle of the third β-strand and is where the strongest interactions occur. The β-strands remain essentially unchanged between the two C-domain structures, while the β-bulge undergoes a complete rearrangement. The absence of interactions between the N-domain and the β-bulge might be the main trigger for the changes that are observed between the C-domain



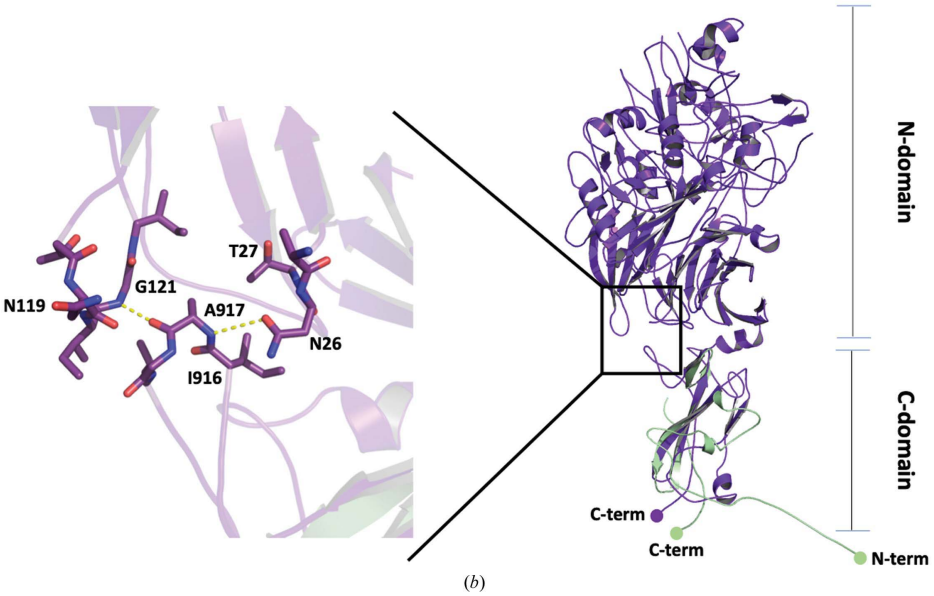


Figure 5
Comparison of P140 and P110 C-domain structures. (*a*) Two views (90° apart from each other) of the superposed C-domain structures as found in this work (green) and in the structure of the whole P110 ectodomain (PDB entry 6r3t; purple; Aparicio *et al.*, 2018). Both structures present a β-bulge, with a similar extended conformation, inserted at the same position in the structurally equivalent β-strand. (*b*) The relative position of the N- and C-domains differs between the P140 and P110 structures determined. The inset shows details of the interactions between the extended β-bulge and the P110 N-domain. Similar interactions might take place in P140 when the β-bulge adopts the extended conformation.

structures. It is interesting that these few interactions could be critical for the structural stability of about half of the residues in the C-domain. It is also surprising to find that in the second half of the domain the residues remain essentially unchanged.

The C-domain of P140 appears to be designed to contain a malleable region that is very sensible to interactions with the N-domain, and a second region where the transmembrane helix is connected that is not affected by interactions with the N-domain. From the experimental information on P140, it is clear that the N- and C-domains can move with respect to each another (Aparicio *et al.*, 2020). These movements may be required by the conformational transitions that Nap experiences during the cycle of attachment to and release from the sialic acid oligosaccharide cell host receptors. Therefore, even subtle signals from the N-domain, possibly initiated by the binding/release of Nap to the cell receptors, that alter the interaction with the C-domain β-bulge are expected to trigger structural changes in the C-domain that allow a global conformational transition of Nap. The residues involved in this mechanism, in particular the presence of the β-bulge in the C-domain, are well preserved in P1, the orthologue of P140 from *M. pneumoniae*, suggesting that this might be a general characteristic of the adhesion/release process in mycoplasmas from the pneumoniae cluster.

5. Related literature

The following reference is cited in the supporting information for this article: McGuffin *et al.* (2000).

Acknowledgements

Many thanks are given to the Crystallization Platform at the PCB and particularly to Dr Roman Bonet. The authors also acknowledge the beamtime granted through proposal 2017062252 at BL13-XALOC, ALBA Synchrotron, Spain.

Funding information

This work was supported by grant BFU2018-101265-B-100 from MINECO (Spain).

References

Aparicio, D., Scheffer, M. P., Marcos-Silva, M., Vizarraga, D., Sprankel, L., Ratera, M., Weber, M. S., Seybert, A., Torres-Puig, S., Gonzalez-Gonzalez, L., Reitz, J., Querol, E., Piñol, J., Pich, O. Q., Fita, I. & Frangakis, A. S. (2020). *Nat. Commun.* **11**, 2877.
Aparicio, D., Torres-Puig, S., Ratera, M., Querol, E., Piñol, J., Pich, O. Q. & Fita, I. (2018). *Nat. Commun.* **9**, 4471.
Burgos, R., Pich, O. Q., Ferrer-Navarro, M., Baseman, J. B., Querol, E. & Piñol, J. (2006). *J. Bacteriol.* **188**, 8627–8637.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.
García-Morales, L., González-González, L., Querol, E. & Piñol, J. (2016). *Mol. Microbiol.* **100**, 125–138.
Gibson, D. G., Benders, G. A., Andrews-Pfannkoch, C., Denisova, E. A., Baden-Tillson, H., Zaveri, J., Stockwell, T. B., Brownley, A., Thomas, D. W., Algire, M. A., Merryman, C., Young, L., Noskov, V. N., Glass, J. I., Venter, J. C., Hutchison, C. A. & Smith, H. O. (2008). *Science*, **319**, 1215–1220.
Glass, J. I., Assad-Garcia, N., Alperovich, N., Yooseph, S., Lewis, M. R., Maruf, M., Hutchison, C. A., Smith, H. O. & Venter, J. C. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 425–430.
Holm, L. & Rosenström, P. (2010). *Nucleic Acids Res.* **38**, W545–W549.
Iverson-Cabral, S. L., Astete, S. G., Cohen, C. R. & Totten, P. A. (2007). *Mol. Microbiol.* **66**, 55–73.
Juanhuix, J., Gil-Ortiz, F., Cuní, G., Colldelram, C., Nicolás, J., Lidón, J., Boter, E., Ruget, C., Ferrer, S. & Benach, J. (2014). *J. Synchrotron Rad.* **21**, 679–689.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Krause, D. C. (1996). *Mol. Microbiol.* **20**, 247–253.
Krause, D. C. & Baseman, J. B. (1982). *Infect. Immun.* **37**, 382–386.
Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
Ma, L., Jensen, J. S., Mancuso, M., Hamasuna, R., Jia, Q., McGowin, C. L. & Martin, D. H. (2010). *PLoS One*, **5**, e15660.
Ma, L., Jensen, J. S., Myers, L., Burnett, J., Welch, M., Jia, Q. & Martin, D. H. (2007). *Mol. Microbiol.* **66**, 220–236.
McCoy, A. J. (2007). *Acta Cryst.* D**63**, 32–41.
McGowin, C. L. & Anderson-Smits, C. (2011). *PLoS Pathog.* **7**, e1001324.
McGuffin, L. J., Bryson, K. & Jones, D. T. (2000). *Bioinformatics*, **16**, 404–405.
Mernaugh, G. R., Dallo, S. F., Holt, S. C. & Baseman, J. B. (1993). *Clin. Infect. Dis.* **17**, S69–S78.
Morrison-Plummer, J., Jones, D. H., Daly, K., Tully, J. G., Taylor-Robinson, D. & Baseman, J. B. (1987). *Isr. J. Med. Sci.* **23**, 453–457.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Nakane, D., Adan-Kubo, J., Kenri, T. & Miyata, M. (2011). *J. Bacteriol.* **193**, 715–722.
Peterson, S. N., Bailey, C. C., Jensen, J. S., Borre, M. B., King, E. S., Bott, K. F. & Hutchison, C. A. (1995). *Proc. Natl Acad. Sci. USA*, **92**, 11829–11833.
Radestock, U. & Bredt, W. (1977). *J. Bacteriol.* **129**, 1495–1501.
Razin, S. (1992). *FEMS Microbiol. Lett.* **100**, 423–431.
Razin, S., Yogev, D. & Naot, Y. (1998). *Microbiol. Mol. Biol. Rev.* **62**, 1094–1156.
Rottem, S. (2003). *Physiol. Rev.* **83**, 417–432.
Scheffer, M. P., Gonzalez-Gonzalez, L., Seybert, A., Ratera, M., Kunz, M., Valpuesta, J. M., Fita, I., Querol, E., Piñol, J., Martín-Benito, J. & Frangakis, A. S. (2017). *Mol. Microbiol.* **105**, 869–879.
Seto, S., Kenri, T., Tomiyama, T. & Miyata, M. (2005). *J. Bacteriol.* **187**, 1875–1877.
Sheldrick, G. M. (2015). *Acta Cryst.* C**71**, 3–8.
Svenstrup, H. F., Nielsen, P. K., Drasbek, M., Birkelund, S. & Christiansen, G. (2002). *J. Med. Microbiol.* **51**, 361–373.
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. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A. & Wilson, K. S. (2011). *Acta Cryst.* D**67**, 235–242.
Winter, G. (2010). *J. Appl. Cryst.* **43**, 186–190.
Xavier, J. C., Patil, K. R. & Rocha, I. (2014). *Microbiol. Mol. Biol. Rev.* **78**, 487–509.

Acta Cryst. (2020). F**76**, doi:10.1107/S2053230X20012297 Supporting information

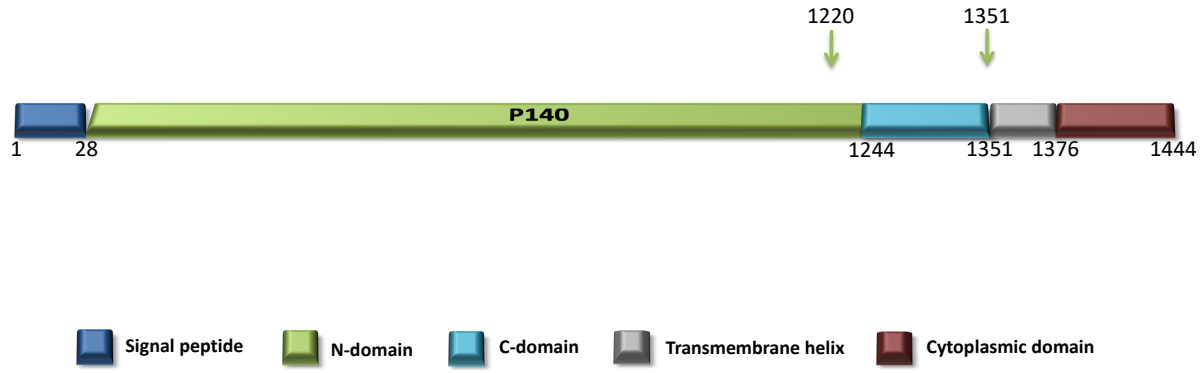
STRUCTURAL BIOLOGY
COMMUNICATIONS

Volume 76 (2020)

Supporting information for article:

Alternative conformation of the C-domain of the P140 protein from *Mycoplasma genitalium*

David Vizarraga, Rosa Pérez-Luque, Jesús Martín, Ignacio Fita and David Aparicio



Supplementary Figure S1. P140 regions and C-Domain construct.

Schematic representation of the P140 regions predicted with the Psipred server[1]: The signal peptide (blue), the extracellular region (orange), the transmembrane helix (gray) and the cytoplasmic region (brown). Green arrows indicate the C-domain construct used for crystallographic and biological studies.

Primers	Sequence
P140cF	AAGTTCTGTTTCAGGGCCCGAAAATGGATGGTAAAC
P140cR	ATGGTCTAGAAAGCTTTAATCTGCCCACTGATTAAAC

Supplementary Table S1.

Primers used in this study

References for the Supplementary information

1. McGuffin, L.J., Bryson, K. & Jones, D.T. The PSIPRED protein structure prediction server. Bioinformatics 16, 404-5 (2000).

Capítulo 2

Estructura y mecanismo del complejo de adhesión Nap del patógeno humano *Mycoplasma genitalium*

David Aparicio [1,4], Margot P. Scheffer [2,4], Marina Marcos-Silva [3,4], David Vizarraga [1], Lasse Sprankel [2], Mercè Ratera [1], Miriam S. Weber [2], Anja Seybert [2], Sergi Torres-Puig [3], Luis Gonzalez-Gonzalez [3], Julian Reitz [2], Enrique Querol [3], Jaume Piñol [3], Oscar Q. Pich [3], Ignacio Fita [1] y Achilleas S. Frangakis [2].

Resumen

Mycoplasma genitalium es un patógeno humano que se adhiere a células epiteliales humanas y causa uretritis, cervicitis y enfermedad inflamatoria pélvica. El complejo de adhesión transmembrana llamado Nap (comprende las proteínas P110 y P140) resulta ser esencial para la infectividad. Aquí presentamos la estructura cristalina de P140 tanto sola como en complejo con el dominio N de P110. Por crio-microscopía electrónica (crio-EM) y tomografía (crio-ET) encontramos conformaciones del Nap cerradas y abiertas, determinadas a 9,8 y 15 Å, respectivamente. Tanto las estructuras cristalinas como la estructura crio-EM se encuentran en una conformación cerrada, donde se ocluye el sitio de unión del ácido siálico en P110. Por el contrario, la estructura crio-ET muestra una conformación abierta, donde el sitio de unión es accesible. La información estructural, en combinación con los estudios funcionales, sugiere un mecanismo de *M. genitalium* para la unión y liberación al receptor de la célula huésped en el que las conformaciones del Nap se alternan para mantener la motilidad y garantizar la infectividad.

nature COMMUNICATIONS

ARTICLE

https://doi.org/10.1038/s41467-020-16511-2 OPEN

Structure and mechanism of the Nap adhesion complex from the human pathogen *Mycoplasma genitalium*

David Aparicio[1,4], Margot P. Scheffer[2,4], Marina Marcos-Silva[3,4], David Vizarraga[1], Lasse Sprankel[2], Mercè Ratera[1], Miriam S. Weber[2], Anja Seybert[2], Sergi Torres-Puig[3], Luis Gonzalez-Gonzalez[3], Julian Reitz[2], Enrique Querol[3], Jaume Piñol[3], Oscar Q. Pich[3✉], Ignacio Fita[1✉] & Achilleas S. Frangakis[2✉]

Mycoplasma genitalium is a human pathogen adhering to host target epithelial cells and causing urethritis, cervicitis and pelvic inflammatory disease. Essential for infectivity is a transmembrane adhesion complex called Nap comprising proteins P110 and P140. Here we report the crystal structure of P140 both alone and in complex with the N-terminal domain of P110. By cryo-electron microscopy (cryo-EM) and tomography (cryo-ET) we find closed and open Nap conformations, determined at 9.8 and 15 Å, respectively. Both crystal structures and the cryo-EM structure are found in a closed conformation, where the sialic acid binding site in P110 is occluded. By contrast, the cryo-ET structure shows an open conformation, where the binding site is accessible. Structural information, in combination with functional studies, suggests a mechanism for attachment and release of *M. genitalium* to and from the host cell receptor, in which Nap conformations alternate to sustain motility and guarantee infectivity.

[1] Instituto de Biología Molecular de Barcelona (IBMB-CSIC) and Maria de Maeztu Unit of Excellence, Parc Científic de Barcelona, Baldiri Reixac 10, 08028 Barcelona, Spain. [2] Buchmann Institute for Molecular Life Sciences, Max-von-Laue Str. 15, 60438 Frankfurt, Germany. [3] Institut de Biotecnologia i Biomedicina and Departament de Bioquímica i Biologia Molecular, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain. [4] These authors contributed equally: David Aparicio, Margot P. Scheffer, Marina Marcos-Silva. ✉email: oquijada@tauli.cat; ifrcri@ibmb.csic.es; achilleas.frangakis@biophysik.org

The human pathogen *Mycoplasma genitalium*, a member of the *pneumoniae* cluster of mycoplasmas, binds to eukaryotic cells by means of its adhesion complex, the Nap. This complex is formed by two heterodimers, each consisting of proteins P110 and P140[1–6]. In addition to their roles in cytadherence and motility, P110 and P140 are immunodominant proteins and constitute the main target of host antibodies during infection[7–9]. Antibiotic resistance to human pathogens from the *pneumoniae* cluster[10–13] is increasing at an alarming rate, making it necessary to explore novel therapeutic strategies. Anti-adherence molecules, aimed at preventing the establishment of infection, are attractive potential antimicrobial drugs[14,15]. A deep understanding of the Nap structure and adhesion mechanism will facilitate the development of anti-adherence therapies. Recently, we determined the crystal structure of the extracellular region of P110 and demonstrated its binding to sialic acid receptors[6]. Here, we address the structure and mechanism of the Nap adhesion complex and reveal an intricate interplay between P110 and P140.

Results

Crystal structure of P140 and in complex with P110N. Crystals were obtained from the extracellular region of P140 (residues 23–1351) (Fig. 1, Supplementary Figs. 1 and 2), both alone and in complex with the N-terminal domain of P110 (P110N: residues 23–827) (Fig. 2a, Supplementary Fig. 2). The structure of P140, for which there are no molecular models or experimental phases available, was determined by density modification techniques, starting with a mask derived from the sub-tomogram-averaged map of the whole Nap obtained by cryo-electron tomography (cryo-ET) (see Methods). With four heterodimers in the asymmetric unit, the P140–P110N crystals were refined at 2.65 Å resolution to a final model with agreement R and R_{free} factors of 18.7 and 22.4, respectively (Supplementary Table 1).

Structural similarities of P140 and P110. The structure of the extracellular region of P140, with a bulky N-terminal domain (residues 23–1243) and a small C-terminal domain (residues 1244–1351), has an overall shape resembling the capital letter P (Fig. 1a). The N-terminal domain consists of a seven-bladed (β-sheet) propeller and a "crown" formed by the clustering of the long polypeptide segments that emerge from the propeller (Supplementary Fig. 3). β-Sheets I to VI each have four strands, while β-sheet VII, the last β-sheet in the propeller that connects directly with the C-terminal domain, has only two strands (Fig. 1b). P140 and P110 share many features of domain organization and the topology of the secondary structural elements, suggesting a common ancestor, although the degree of conservation differs for the N-terminal and C-terminal domains (Supplementary Fig. 4). The N-terminal domains, with an RMSD of 3.5 Å between the Cα atoms of 359 structurally equivalent residues (~28%), are markedly different in the crown. In contrast, the C-terminal domains are closely related, with 74 equivalent residues (~71%) and an RMSD of 2.2 Å.

Binding site analysis of the Nap. The binding site for the sialylated oligosaccharides, identified in the structure of P110[6], is located at the interface between P140 and P110 in the crystal structure of the P140–P110N complex (Fig. 2a–d). Interaction of the two subunits changes the position of the sialic binding β-hairpin while the interfering loop (P140 residues 807–827) inserts into the binding pocket of P110N, sterically interfering with the binding of oligosaccharides to the complex (Fig. 2c–d). In agreement with this, surface plasmon resonance analysis shows that, in solution, sialylated compounds 3SL and 6SL (neuraminic acid forming an α2–3 or an α2–6 linkage to a lactose monosaccharide, respectively), which bind to P110 alone[6], do not bind either to P140 alone or to P140–P110 complexes (Supplementary Fig. 5). The structure of the P140–P110N complex suggests that P110 residues Gln460–Asp461, from the binding β-hairpin, and

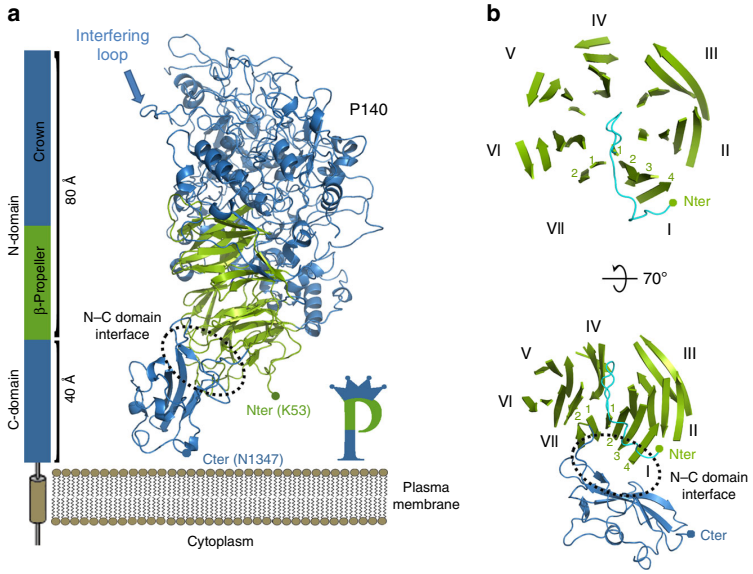
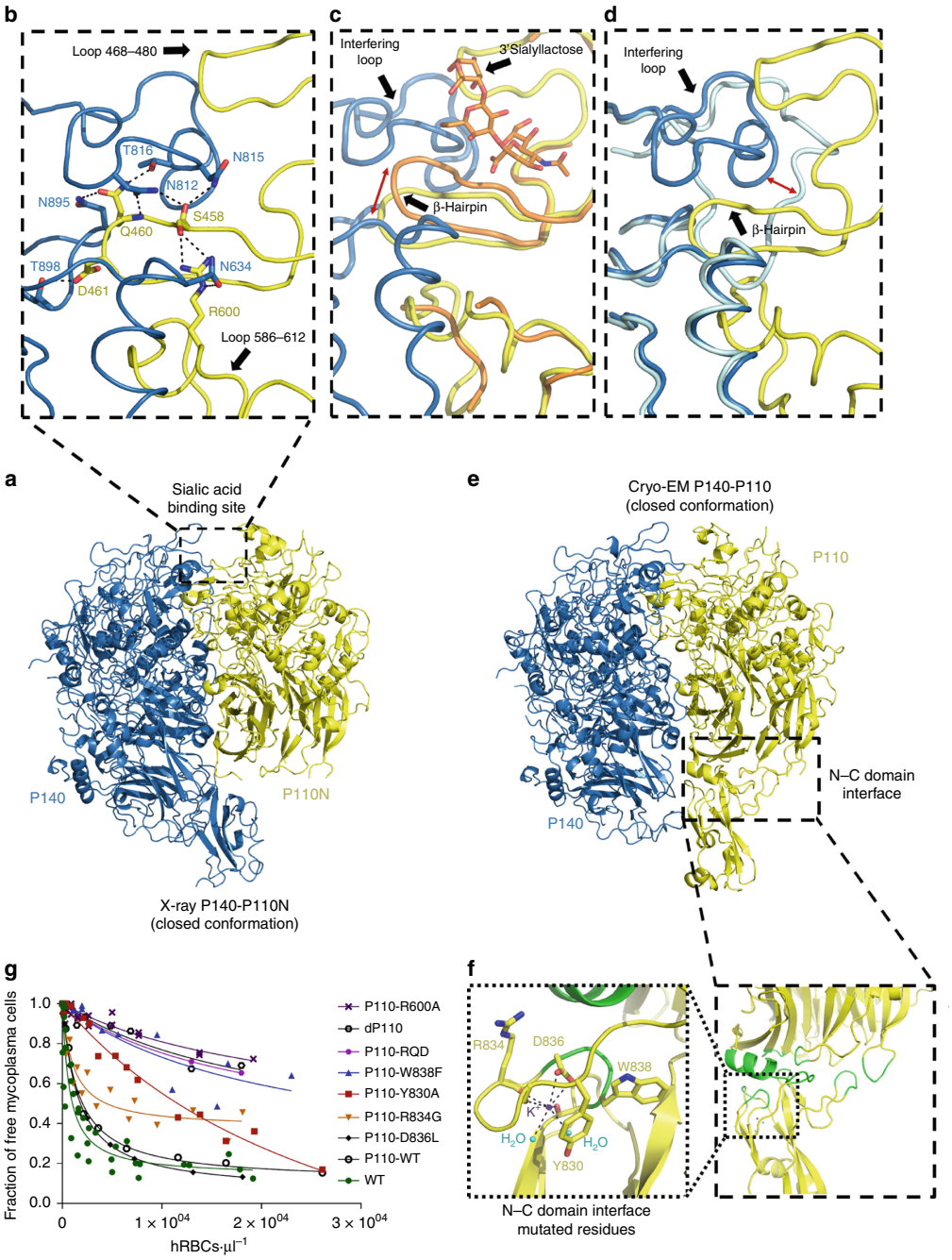


Fig. 1 Structure of P140. a P140 contains a large N-terminal domain and a smaller C-terminal domain, which is followed in sequence by a predicted transmembrane helix. The N-terminal domain has two distinct regions: the β-propeller (green) and the crown (blue). **b** Two views of the β-propeller, 70° apart. The first and last β-sheets of the propeller (I and VII, respectively) interact with the C-terminal domain and are structurally contiguous in the propeller ring. A large β-bulge (residues 69–77, in cyan) occludes the center of the propeller ring. The presence of only two strands in the last β-sheet (VII) constrains the interface between the N-terminal and C-terminal domains.



Arg600, close to the receptor-binding site, play an important role in the interaction with P140 (Fig. 2b, d). P110 variants carrying an R600A substitution or the triple substitution RQD (R600A, Q460A, and D461A) were introduced by transposon delivery into an *M. genitalium* P110 null mutant[1]. Strains expressing the P110-RQD variant protein, which was barely detectable, showed a null binding capacity phenotype (Fig. 2g, Supplementary Fig. 6b). The variant protein P110-R600A was well expressed, but the strain presented no capacity for adherence and characterization of cell motility was not feasible.

Single-particle cryo-EM of the P140–P110 extracellular region. Using a sample of P140–P110 complexes, with the complete extracellular region included for both subunits (P140 residues 23–1351 and P110 residues 23–938), we performed single-particle cryo-electron microscopy (cryo-EM). We obtained a map with an overall resolution of 4.1 Å, although non-isotropic (Fig. 2e, Supplementary Table 2, Supplementary Figs. 7 and 8). The P140–P110N X-ray structure could be fitted as a rigid-body without modifications into the P140–P110 cryo-EM map with UCSF Chimera[16] (Supplementary Table 3, Supplementary Fig. 7,

Fig. 2 Closed conformation of the P140–P110 heterodimers by X-ray and cryo-EM. a X-ray crystal structure of the P140–P110N heterodimer. **b** Network of hydrogen bonds in the vicinity of the oligosaccharide-binding site between P140 (blue) and P110N (yellow). Rearrangements experienced by P110N (**c**) and P140 (**d**) with respect to the structures of P110 alone (orange) and P140 alone (light blue), respectively. Rearrangements of the ß-hairpin and the interfering loop are indicated with red arrows. Oligosaccharide 3SL (**c**) is depicted as it was found in the structure of P110 alone[6], showing that it would have steric clashes in the closed conformation of the P140–P110N complex. **e** Cryo-EM structure of the P140–P110 heterodimer. The C-terminal domain of P140 is missing in the structure because it was not defined in the cryo-EM map (Supplementary Fig. 7). The C-terminal domain of P110 was visible in the cryo-EM map, although with a local resolution lower than that for the N-terminal domains. **f** Details of the interface (depicted in green) between the N-terminal and C-terminal domains, close to where a potassium ion was found in the structures of P110 alone. **g** Hemadsorption assays by fluorescence-activated cell sorting (FACS) analysis. Graphic represents inverse Langmuir plots (see also Supplementary Information) containing a fixed amount of mycoplasma cells and increasing amounts of human red blood cells (hRBCs). Plots represent the best-fitting curves to a series of hemadsorption measurements obtained from at least two biological repeats for each strain. We performed a double-gating strategy, using a preliminary FL3-H/FL2-H gate following an SSC-H/FL1-H gate. Mutations are either in the vicinity of the oligosaccharide-binding site at the interface between P140 and P110 (residues explicitly shown in panel **b**) or close to the interface between the N-terminal and C-terminal domains of P110 (residues shown in panel **f**). R^2 for each fitting curve: WT (0.8576), dP110 (0.9054), P110-WT (0.9921), P110-Y830A (0.9663), P110-W838F (0.8339), P110-R834G (0.8757), P110-D836L (0.9993), P110-R600A (0.8663), and P110-RQD (0.9999).

Supplementary Movie 1). Therefore, the structure of the P140–P110 complex found by cryo-EM corresponds to the conformation of the X-ray P140–P110N structure, where access to the sialylated oligosaccharides binding site is occluded (Supplementary Movie 1). In the P140–P110 cryo-EM map, there is no density for the C-terminal domain of P140, whereas the C-terminal domain of P110, which is absent in the P140–P110N complex, is visible, albeit with a weak density (Fig. 2e, Supplementary Fig. 7). This indicates a significant flexibility of the C-terminal domains with respect to the bulkier N-terminal domains. In the P140–P110 complex found by cryo-EM, the interface between subunits spans 2758 Å^2 with 30 hydrogen bonds and an estimated Gibbs free energy of −20 kcal/mol, resulting in a 100% probability of the formation of the complex (PISA server[17]).

Motility analysis of mutants. Five mutations were introduced close to the interface between the N-terminal and C-terminal domains, to check if either adhesion or motility was affected. The five mutated residues (Y830A, R834G, D836L, W838F, and G839F) were chosen in the vicinity of the potassium-binding site found in P110[6] (Fig. 2e, f, Supplementary Table 4). The P110-G839F strain showed no detectable levels of adhesins, while P110-D836L exhibited hemadsorption values similar to those of wild-type cells (Fig. 2g). The P110-Y830A and P110-R834G variants showed intermediate hemadsorption phenotypes, characterized by unusually high B_{max} values (see Methods). Hemadsorption of the P110-W838F mutant was severely impaired, with high K_d and B_{max} values. Motility was examined by time-lapse microcinematography to monitor the movements of individual cells for 120 s (Supplementary Fig. 6, Supplementary Table 5). In the P110-W838F mutant, cells were completely non-motile and, in agreement with this extreme phenotype, phase contrast images revealed the presence of large aggregates resulting from the inability of cells to spread. The Y830A, R834G, and D836L variants also exhibited altered gliding velocities, indicating that structural integrity at the interface between the N-terminal and C-terminal domains, away from the cell receptor-binding site, is critical for motility in *M. genitalium*.

Single-particle analysis of the Nap. Next, we performed single-particle cryo-EM using a purified sample of Nap complexes, obtained as previously described[3] (see Methods). The purified Nap complexes contain full-length P140 and P110 proteins, including the transmembrane helices and cytoplasmic regions, which are required for formation of tetramers. Classification indicated that only ~15–20% of the images corresponded to complete Nap particles where the extracellular region was well defined. The best map, with an overall resolution of 9.8 Å, allowed accurate rigid-body fitting with Chimera[16] of the P140 structure alone and the P110 structure alone, confirming the presence of a dimer of P140–P110 heterodimers in each Nap (Fig. 3a–c, Supplementary Fig. 9, Supplementary Table 2). The arrangement of P140–P110 heterodimers is essentially identical to the X-ray and cryo-EM P140–P110 structures, with the interactions between both subunits preventing access to the sialylated oligosaccharide-binding site. This is the "closed" conformation of the Nap (Supplementary Movie 2). C-terminal domains act like stalks connecting the large N-terminal domain of each subunit with the outer surface of the cell membrane. The two P110 subunits are almost parallel to the dimer axis and face away from each other, while the two P140 subunits of the Nap adopt a "V-shaped" arrangement with the C-terminal domains very close to each other (Fig. 3c). The distance between heterodimers is large (Fig. 3b, c), suggesting that their interaction is weak. This is in agreement with MALS measurements where only heterodimers are detected when mixing equimolecular amounts of constructs from the extracellular regions of P110 and P140 (Supplementary Fig. 10). Therefore, there are two interfaces in the Nap between the extracellular regions of P140 and P110, which we name the "tight" and "loose" interfaces (Fig. 3b, e). The cryo-EM map of the Nap shows density corresponding to the Nap transmembrane region and the positions of the transmembrane helices (their N-terminal ends) can be seen (Fig. 3a, c). The close proximity of the P140 C-terminal domains also brings the N-terminal ends of the corresponding transmembrane helices closer together. For P140 and P110, sequence analysis indicates the presence of one transmembrane helix containing one and two Engelman motifs (GXXXG, with X any residue, in general hydrophobic), respectively, which are characteristic of high-affinity interactions in membrane helices[18,19].

Cryo-electron tomography of the in situ Nap. The structure of the in situ Nap complex was also determined by cryo-ET from mildly lysed *M. genitalium* cells[3]. Classification of Nap volumes and subsequent sub-tomogram averaging provided an improved map for the most abundant class (~85%) at 15 Å resolution, in which the four subunits, two from P110 and two from P140, are clearly distinguished and present a nearly perfect twofold symmetry (Fig. 3d–f, Supplementary Fig. 11). The rigid-body fitting of the structures of P110 alone and P140 alone into the cryo-ET density with Chimera[16] reveals major differences from the cryo-EM structure of the Nap (Fig. 3, Fig. 4a, Supplementary Table 3). In the cryo-ET structure, the longest axis of the four subunits runs parallel to the axis of twofold symmetry of the Nap. This

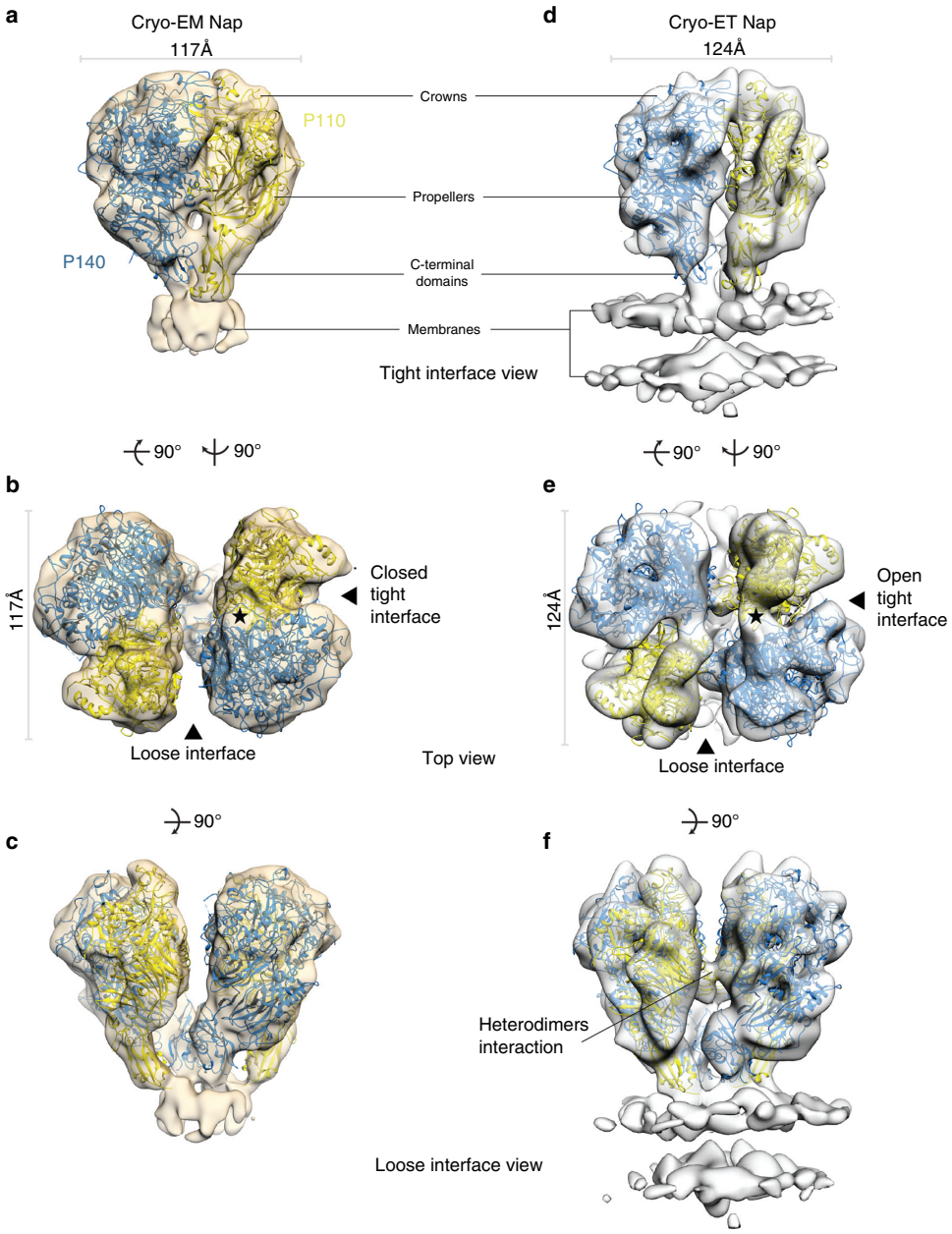


Fig. 3 The structure of Nap by single-particle cryo-EM and in situ cryo-ET. a–c Three different views of the single-particle cryo-EM from the Nap (beige surface) with the structures of P140 (blue) and P110 (yellow) fitted into the density. In **a** the crown, propeller, and C-terminal domain are indicated. **b** Top view of the Nap, depicting the loose and tight interfaces. An approximate twofold axis perpendicular to the cell membrane relates the pair of P110–P140 heterodimers of a Nap. Fitting of the P140 and P110 subunits is unambiguous. The interface between the P110 and P140 subunits, defined as tight in the cryo-EM structure of the Nap, is the same found in the cryo-EM structure of the P140–P110 heterodimer. The interfering loop and the binding site are indicated with a black star. **c** The lateral view along the loose interface shows the V-like shape adopted by the two P140 subunits. **d–f** Three views of the cryo-ET map of the Nap, with the structures of P140 and P110 accurately fitted. The membrane bilayer of the mycoplasma cell is clearly defined in the lower part of the lateral views (**d**, **f**). **e** The top view shows that the tight interface is wider in the cryo-ET "open" conformation than in the cryo-EM "closed" conformation of the Nap. **f** The lateral view along the loose interface shows an interaction between heterodimers.

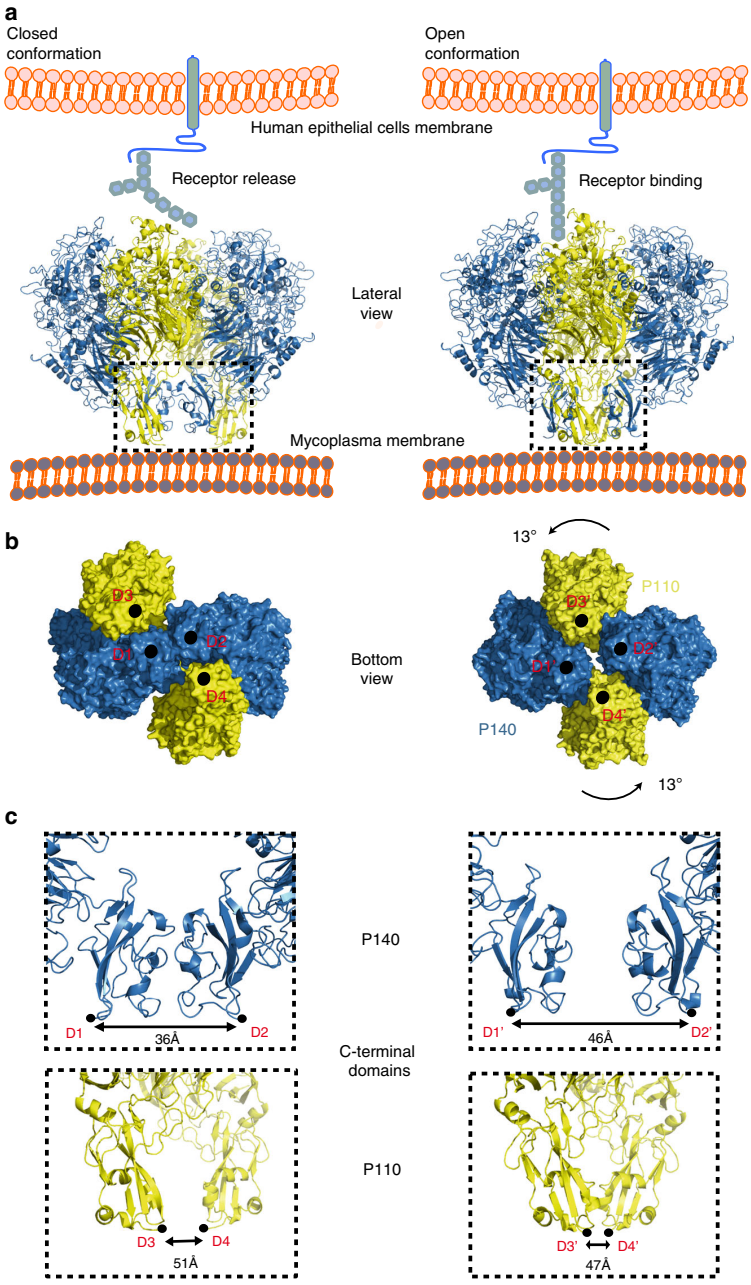


Fig. 4 Structural rearrangements between "open" and "closed" conformations. a "Closed" (left panel) and "open" (right panel) conformations of the Nap complex. The "closed" conformation associated with the release state of the sialylated oligosaccharides cell receptors. The "open" conformation associates with the bound and ready-to-bind state. **b** Bottom view along the Nap twofold axis, depicting points of contact (black circles) between subunits and the mycoplasma cell outer membrane. These points also mark the start (the N-terminal ends) of the transmembrane helices. The N-terminal end positions are changing from a squared to rhomboid shape, when transitioning from the closed to the open conformation. **c** Changes to the relative positions of the C-terminal domains from P140 (upper panels) and P110 (lower panels).

implies that between the cryo-ET and cryo-EM structures there is a rotation of P110 relative to P140 in the heterodimer and an increase of 10 Å in the closest distance between the C-terminal domains of P140 subunits (Fig. 4b, c, Supplementary Fig. 12, Supplementary Movies 3 and 4). The loose interface between P140 and P110 subunits, which is very wide in the cryo-EM structure of the Nap, is narrower in the cryo-ET structure, with a few interactions between P140 and P110 (in particular, P140 residues 1175–1179 interacting with the P110 loop 750–755) that are absent in the cryo-EM Nap (Fig. 3c, f). The tight interface

between P140 and P110, which is in the "closed" conformation in the cryo-EM structure of the Nap, undergoes major rearrangements in the cryo-ET structure, becoming on average ~7 Å wider (Fig. 3a, b, d, e). This widening "opens" the sialylated oligosaccharide-binding site, making it accessible for binding in the cryo-ET structure of the Nap. Therefore, the cryo-ET and cryo-EM structures of the Nap correspond to conformations that, with respect to the sialylated host cell receptors, are respectively "open", with the binding site accessible and ready for binding, and "closed", with the binding site occluded and not accessible for binding. In the cryo-ET map, the cell membrane surrounding the Nap is visible as a double layer that is perpendicular to the Nap twofold axis. The positions of the C-terminal domains are largely different between the cryo-ET and cryo-EM Nap structures, changing from a squared to rhomboid shape, respectively (Fig. 4b). These differences also affect the positions of the transmembrane helices that follow in sequence immediately after the C-terminal domains, which provide a mechanism for communication between the Nap extracellular and intracellular regions. On the intracellular side, the cryo-ET map shows a compact density, similar to the one reported previously[3], with a volume suited to about 400 residues and an approximate twofold symmetry, where the four subunits appear to interact closely as they cannot be individualized at the resolution available.

Discussion

Integration of the information obtained from the structural approaches shows that the Nap is composed of two P140–P110 heterodimers with a loose interface between them in the extracellular region, suggesting that subunits in the Nap are held together mainly by interactions in the intracellular and transmembrane regions. Each heterodimer of the Nap appears to undergo large structural rearrangements between the "open" and "closed" conformations, which are associated with attachment and release to cell receptors, respectively (Fig. 4, Supplementary Fig. 12, Supplementary Movie 4). The "open" conformation corresponds to a state in which the P140 and P110 extracellular regions from the four subunits of the Nap interact only weakly with each other, allowing P110 to remain in a conformation that is ready to bind or is bound to sialylated cell receptors. The "closed" conformation corresponds to a state in which the tight interaction of the extracellular regions of P140 and P110 occlude the cell receptor-binding site. The "closed" conformation could occur rapidly to release the Nap from the sialylated cell receptor. To avoid been trapped in the overall most stable state, the cycling between "open" and "closed" conformations would require a net input of energy in each Nap complex. The structural rearrangements in the Nap extracellular region can be transmitted to the cell interior through repositioning of the transmembrane helices. In vivo variants of Nap complexes indicate that structural rearrangements can play a critical role in both cell adhesion and motility. Results from this work, together with previous data, could form the basis for developing therapeutic strategies targeting the adhesion machinery of mycoplasmas from the *pneumoniae* cluster.

Methods

Cloning, expression, and purification of P110N, P140, and P110. The region corresponding to the MG_192 gene from *M. genitalium* (strain G37, residues 23–827, P110N) was amplified from a synthetic clone (Supplementary Table 6) using primers P110N-F and P110N-R as Forward and Reverse, respectively (Supplementary Table 7). The PCR fragment was cloned into the expression vector pOPINE[20] (gift from Ray Owens plasmid #26043, Addgene, Watertown, USA) to generate a C-end His_6-tagged protein. The recombinant protein was obtained after expression in B834 (DE3) cells (Merck, Darmstadt, Germany) at 20 °C o/n upon induction with 1 mM IPTG at 0.6 OD_{600}. Cells were harvested and lysed in 1× phosphate-buffered saline (PBS) buffer by sonication. Subsequently, the cell extract was centrifuged at 25,000 × *g* at 4 °C and the supernatant applied to a 5 ml Histrap column (GE Healthcare, Pittsburgh, USA) equilibrated with 1× PBS as binding buffer and 1× PBS with 500 mM imidazole as elution buffer. Soluble aliquots of His_6-tagged P110N were pooled and loaded onto a HiLoad Superdex 200 16/60 column (GE Healthcare) in buffer consisting of 50 mM Tris pH 7.4 and 150 mM NaCl. The region corresponding to the MG_192 gene from *M. genitalium* (strain G37, residues 23–938, P110) was amplified from a synthetic clone (Supplementary Table 6) using primers P110-F and P110-R as Forward and Reverse, respectively (Supplementary Table 4). The region corresponding to the MG_191 gene from *M. genitalium* (strain G37, residues 23–1351) was amplified from a synthetic clone (Supplementary Table 6) using primers P140F and P140R as Forward and Reverse, respectively (Supplementary Table 4). Protein production and purification for P140 and P110 followed the same protocol as for P110N[6].

Preparation of P140–P110 and P140–P110N complexes and SEC-MALS analysis. Central fractions of the Superdex 200 16/60 column (GE Healthcare) from P140 and either P110N or P110 were mixed in a 1:1 ratio to a final concentration of ~11 mg/ml (measured with an absorption coefficient of 1.32). Each mixture was incubated for 30 min at room temperature to obtain the heterodimeric complex. For protein oligomerization experiments, prior to sample injections, the multi-angle light scattering (MALS, WYATT Technologies & Corporation) detector was normalized with 25 µl of BSA at 5 mg/ml. Then 150 µl of each purified P140, P110, and P140–P110 sample at 0.6, 2, and 0.5 mg/ml, respectively, were injected at 0.5 ml/min using a Superdex 200 increase 10/300 GL column (GE Healthcare) equilibrated with a running buffer consisting of 50 mM Tris pH 7.4 and 150 mM NaCl. All samples were purified by filtering through 0.22 µm filters.

Preparation of purified Nap complexes. A P110His strain was generated for purification of the Nap complexes from *M. genitalium* G37 cells (ATCC 33530). Produced by genetic engineering[3], the strain carries a 6xHistag insertion in the MG192 gene. Four liters of the P110His strain grown in SP4 medium in suspension at 37 °C at 150 r.p.m. was harvested by centrifugation (16,000 × *g*, 4 °C, 30 min). The pellet was washed twice with PBS with calcium and magnesium, followed by cell disruption by sonication in PBS in the presence of 1 mM EDTA, 5 mM β-mercaptoethanol, 0.1 mM PMSF, and EDTA-free cocktail of protease inhibitors (Roche Diagnostics, Mannheim, Germany). The pellet generated after centrifugation (70,000 × *g*, 4 °C) was resuspended in 75 mM Tris pH 7.4, 400 mM NaCl, 5% glycerol, and 2% n-octyl-β-D-glucopyranoside detergent by homogenization in a glass homogenizer. Solubilization of membranes was done overnight at 4 °C in an orbital shaker. Solubilized membranes were centrifuged at 50,000 × *g*, 30 min (4 °C) and the supernatant was purified by Ni^{2+}-affinity chromatography in 75 mM Tris pH 7.4, 400 mM NaCl, 5% glycerol, and 0.5% octylglucoside detergent. The purified Nap complex was obtained by Superose 6 size-exclusion chromatography equilibrated with the same buffer.

Crystallization of P140 alone and the P140–P110N complex. Screening for initial crystallization conditions was performed with 150 nl droplets in 96-well plates on a Cartesian robot (Cartesian (TM) Dispensing Systems) for both P140 alone and the P140–P110N complex. Optimized crystals from P140 were prepared by mixing 1 µl P140 at 6.5 mg/ml and 1 µl reservoir solution containing 20% PEG3350, 0.1 M BisTris Propane pH 7.5, and 0.2 M sodium sulfate hydrate at 20 °C in hanging drop in 24-well plates. Crystals from the P140–P110N complex were prepared by mixing 1 µl of P140–P110N complex at 10.95 mg/ml and 1 µl reservoir solution containing 23% PEG 500MME, 4% PGA, and 0.1 M sodium cacodylate pH 6.5 at 20 °C in hanging drop in 24-well plates. All crystals were flash-cooled in liquid nitrogen with 20% glycerol as a cryo-protectant.

X-ray data collection and structure determination. X-ray diffraction experiments were performed at the Xaloc Beamline (ALBA, Spain). Data were processed with Xia2[21] using XDS[22], Aimless and Pointless[23] from the CCP4i suite of programs[24]. P140 crystals belong to the space group C2 with six subunits in the asymmetric unit, while P140–P110N crystals belong to the space group $P2_1$ and contain four heterodimers in the crystal asymmetric unit. For the P140–P110N crystals, a partial molecular replacement solution was obtained with Phaser software[25] using the N-terminal domain of the available structures of P110 (PDB code 6RT3) as a search model. In these P140–P110N crystals, an initial mask was tentatively defined for P140 using the sub-tomogram averaged map of a whole Nap complex obtained from cryo-ET images of *M. genitalium* ghost cells. Density modification in the crystal of the P140–P110N complex, by iterative non-crystallographic symmetry averaging and solvent flattening followed by phase extension with DM[26], allowed us to obtain a rough density for P140, in which a few secondary structural elements were recognizable. This density was sufficient to provide an initial molecular replacement solution for the P140 crystals (with six subunits in the asymmetric unit) and for crystals of the orthologous protein P1 from *M. pneumoniae* (now deposited in the PDB with code 6RC9). Averaging within and between crystals, while updating and refining the masks, provided electron density maps from which it was possible to progressively build in parallel the structures of P140 and P1. The weak selenium methionine anomalous data, available from both the P140 and the P140–P110N crystals, were used to confirm

the sequence assignment to the P140 structure. Final models were traced with Coot[27] and refined with Refmac5[28] (Supplementary Table 1).

Cryo-ET and sub-tomogram averaging. Ghost cells were prepared in a similar manner as reported previously[3]. Briefly, cells of the adherently growing *M. genitalium* G37 strain (ATCC 33530) were cultivated on 300-mesh lacey carbon EM grid (PLANO GmbH, Wet zlar, No. S166-3) in 3 ml SP4 medium, rinsed for 1 min with PBS before incubation with 20 mM TEA pH 7.5, 0.5 M KCl, and 1% (v/v) Triton X-100 for 1–5 min in the Vitrobot Mark IV (Thermo Fisher Scientific, Waltham, USA) humidity chamber at 95% humidity and 15 °C. Before plunge freezing, 5 nm gold fiducial markers (Protein A-conjugated to gold beads, Cell Biology Department, UMC Utrecht) were added. Tilt-series were recorded using SerialEM 3.5.8[29] on a 300 kV Titan Krios (Thermo Fisher Scientific) in energy-filtered transmission electron microscopy (EFTEM) mode using a GIF Quantum® SE post-column energy filter (Gatan, Inc., California, USA) set to 20 eV slit width. A K2 Summit detector (Gatan Inc.) was operated in counting mode with a dose rate of ~1.0e^-/pixel/s, using a frame rate of 0.25 s and a total dose in the range of 40–50e^-/Å^2. A nominal magnification of ×64,000 was used with the K2 in super-resolution mode, resulting in a pixel size of 1.1 Å/pixel. The tilt-series generally covered an angular range from −60° to +60° with an angular increment of 3° and defocus set between −2 and −4 µm using either dose symmetric or bidirectional acquisition. Tomograms were reconstructed by super-sampling SART[30] with 3D contrast transfer function (CTF) correction. For volume visualization and isosurface rendering, the EMpackage in Amira was used (Thermo Fisher Scientific)[31]. Sub-tomogram averaging was performed in a similar manner as reported previously[3]. In short, Nap complexes were manually selected and pre-aligned according to their orientation on the membrane, which provided a strong constraint lowering the degrees of freedom for sub-tomogram averaging. This first pre-aligned average was used as the starting reference template for the iterative sub-tomogram averaging procedure, thus completely avoiding any initial model bias. The "open" structure for the cryo-ET data is maintained when using the "closed" structure as a starting reference (Supplementary Fig. 13). We then averaged 11,000 Nap sub-tomograms. Post-processing involved refinement by projection matching and dose weighting of the wedge components. After classification, 8800 Nap sub-tomograms were included in the final average. All processing was performed using sub-tomograms with a voxel size of 0.44 nm. The 0.5 criterion was used for resolution estimation. For classification of sub-tomograms, the averaged particles were normalized, and single projection slices in *X*, *Y*, and *Z* directions were extracted using custom MATLAB scripts, which are available on GitHub (https://github.com/uermel/Artiatomi). Subsequently, the two-dimensional (2D) projection slices were imported into Relion[32] and 2D classification without image alignment was performed. Constituent sub-tomograms from the respective 2D classes were summed together to create 3D classes. Molecular graphics rendering and analyses were performed with either UCSF Chimera[16] or UCSF ChimeraX[33] (Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, USA). In all cases, rigid-body transformations were used to fit X-ray structures into the cryo-EM density maps using the fit-in-map function in UCSF Chimera.

Single-particle cryo-EM. For single-particle cryo-EM, a 3.5 µl aliquot of 20 µg/ml purified P140–P110 heterodimer, or 25 µg/ml purified Nap complex, was applied to a (30 s) glow-discharged R1.2/1.3 Quantifoil grid (Quantifoil, Großlöbichau, Germany), and plunge-frozen in liquid ethane (Vitrobot Mark IV, Thermo Fisher Scientific) at 100% humidity, 4 °C, blot force −3, wait time 60 s, with a blotting time of 15 s for the heterodimer and 20 s for the Nap complex. Dose-fractionated movies were collected at a nominal magnification of ×130,000 (1.05 Å per pixel) in nanoprobe EFTEM mode at 300 kV with a Titan Krios (Thermo Fisher Scientific) electron microscope using a GIF Quantum S.E. post-column energy filter in zero loss peak mode and a K2 Summit detector (Gatan Inc.). For the heterodimer, 32 frames of 0.2 s each were collected and for the Nap complex 34 frames per micrograph were collected. The camera was operated in dose-fractionation counting mode with a dose rate of ~7.7 electrons per Å^2 s^{-1} for the heterodimer and ~7.4 electrons per Å^2 s^{-1} for the Nap complex, resulting in both cases in a total dose of ~50 electrons per Å^2. Defocus values ranged from −1 to −4 µm with marginal (<0.1 µm) astigmatism.

Single-particle cryo-EM image processing. Relion 3.0 was used for the whole-image processing workflow[32] unless stated otherwise. Beam-induced motion correction was performed using UCSF MotionCor2[34]. The CTF of each micrograph was estimated using GCTF[35]. As reference for autopicking the heterodimer, the crystal structure filtered to 20 Å was used. As a reference for the Nap complex, the crystal structure of the dimer was fitted into the cryo-ET data and subsequently lowpass filtered to 20 Å. Cross-checking experiments with Laplacian-of-Gaussian autopicking were performed for the Nap to verify that the structure was not reference-biased. For the heterodimer, 1,024,402 particles were extracted from 2751 micrographs, and for the Nap complex, 160,203 particles were extracted from 2225 micrographs, using a 270 pixel box in both cases. All images were normalized to make the average intensity of the background equal to zero during pre-processing. False-positive particles were removed manually or by unsupervised 2D classification. The CTF-refinement function from Relion 3.0 was used to perform per-particle defocus estimation using a search range from 300 nm around the estimated values from the whole micrographs. To correct for local motion and for radiation damage, we used the Bayesian polishing function of Relion 3.0, in which the resolution-dependent decay caused by radiation damage is taken into account[36]. The polished particles were subjected to another round of 2D classification to remove remaining junk particles. The remaining 404,064 particles for the heterodimer and 28,000 particles for the Nap complex were used for 3D classification using the initially generated density map filtered to 60 Å as reference. For the heterodimer, six classes were generated using 3D classification, of which one class (37,009 particles) showed electron density in the stalk region for P110 and was refined using 3D auto-refine. For the Nap complex, six classes were generated using 3D classification, of which two classes (11,737 particles) were tetrameric. The one tetrameric class containing 5612 particles was refined using 3D auto-refine. Gold-standard Fourier shell correlations (FSCs) were calculated during the 3D refinement between two independently refined halves of the data. The resulting map was post-processed to exclude solvent regions from the FSC calculation and to perform sharpening with a temperature factor of −78 Å^2 for the heterodimer and −218 Å^2 for the Nap complex resulting in a global resolution of 4.1 Å for the heterodimer and 9.8 Å for the Nap complex, using the 0.143 criterion. Local resolution was estimated using Relion 3.0, showing high resolution in the core regions of the dimer and low resolution in the surface and stalk regions (Supplementary Figs. 8 and 9).

Surface plasmon resonance. For binding assay experiments, a Biacore 3000 biosensor platform (GE Biosystems) equipped with a research-grade streptavidin-coated biosensor chip SA was used. The chip was docked into the instrument and preconditioned with three 1-min injections of 1 M NaCl in 50 mM NaOH. Both 3SL-PAA-biotin and 6SL-PAA-biotin (Carbosynth) oligosaccharides were injected over the second and third flow cell, respectively, at 10 µg/ml diluted in HBS-P (10 mM Hepes, pH 7.4, 0.15 M NaCl, and 0.005% P20). The first cell was left blank to serve as a reference. The running buffer consisted of HBS-P at a flow rate of 30 µl/min and the immobilization levels acquired were ~160 and ~180 response units for 3SL-PAA-biotin and 6SL-PAA-biotin, respectively. A series of diluted purified extracellular P140 and P140–P110 samples in HBS-P (1.25, 2.5, 5, 10, and 20 µM) were injected over the flow cell surface at 30 µl/min. Interaction analysis were performed at 25 °C and the protein was allowed to associate and dissociate for 60 and 90 s, respectively, followed by a 30 s regeneration injection step of 0.05% SDS at 30 µl/min.

Strains, culture conditions, and primers. *M. genitalium* was grown in SP4 medium (Hardy Diagnostics) at 37 °C in tissue culture flasks. Mutants were isolated on SP4 agar plates supplemented with puromycin (3 µg/ml). All *M. genitalium* strains used in this work are listed in Supplementary Table 4. *Escherichia coli* XL1-Blue strain (Agilent, Santa Clara, USA) was used for cloning and plasmid amplification purposes. It was grown in LB broth or on LB-agar plates containing 100 µg/ml ampicillin. Primers used are listed in Supplementary Table 7.

DNA manipulation and mutant construction. Plasmid DNA was purified using GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific). PCR products and DNA fragments were recovered from agarose gels using NucleoSpin Gel and PCR Clean-up Kit (Macherey-Nagel, Düren, Germany), and digested using the corresponding restriction enzymes (Thermo Fisher Scientific) when necessary. For transformation of *M. genitalium*, plasmids were purified using the GenElute HP Midiprep Kit (Sigma-Aldrich, St. Louis, USA) following the manufacturer's instructions.

P110 variants were generated in a two-step PCR procedure using DNA from plasmid pTnPacP110-WT as a template[6]. For each mutant, the first PCR round was performed using primer COMmg192-F and the specific mg192 reverse primer, or primer COMmg192-R and the specific mg192 forward primer (Supplementary Table 4). To reconstitute the mutated full-length MG_192 alleles, we conducted splicing by overhang extension (SOE) PCR using the specific amplicon pair obtained for each mutant as a template (R600A, Y830A, R834G, D836L, W838F, and G839F) and primers COMmg192-F and COMmg192-R. Then, the mutated P110 alleles were digested with *Apa*I and *Xho*I and ligated into a similarly digested pMTnPac plasmid backbone[6] to generate the corresponding pMTnPacP110 plasmid series. The P110 variant carrying the triple substitution RQD was obtained using primers Q460AD461Amg192-F and -R to introduce the double substitution Q460A and D461A into the pMTnPacP110-R600A plasmid, which was used as the DNA template for the PCR reaction. Sequencing analysis of the different TnPacP110 minitransposons, using primers Tnp3, RTPCR192-F, RTPCR192-R, and PacUp, ruled out the presence of additional mutations in the MG_192 sequence. These plasmids were transformed into a *M. genitalium* MG_192 null mutant to create the different P110 variant strains. Identification of the minitransposon insertion site in the individual clones was done by sequencing using the PacDown primer and chromosomal DNA as a template.

Transformation and screening. *M. genitalium* MG_192 null mutants were transformed by electroporation using 5 µg of plasmid DNA of the different minitransposons, as previously described[37]. Puromycin-resistant colonies were

picked, propagated, and stored at −80 °C. For screening purposes, strains were further propagated in 25 cm^2 tissue culture flasks with puromycin and lysed using 0.1 M Tris-HCl pH 8.5, 0.05% Tween-20, and 250 µg/ml Proteinase K for 1 h at 37 °C. Then, Proteinase K was inactivated at 95 °C for 10 min. *M. genitalium* lysates were screened by sequencing using the PacDown primer. By contrast, the MG_192 alleles were fully re-sequenced to rule out the presence of undesired mutations.

Sequencing reactions. Sequencing reactions were performed with the BigDye® v3.1 Cycle Sequencing kit using 2.5 µl of genomic DNA, following the manufacturer's instructions (Thermo Fisher Scientific). All sequencing reactions were analyzed using an ABI PRISM 3130xl Genetic Analyser at the Servei de Genòmica i Bioinfòrmatica (UAB).

SDS-PAGE. Whole-cell lysates were obtained from mid-log phase cultures grown in 75 cm^2 flasks. Protein concentration was determined with the Pierce™ BCA Protein Assay Kit (Thermo Fisher Scientific), and similar amounts of total protein were separated by SDS-PAGE following the standard procedures.

Quantitative hemadsorption assay. We used 10^9 mycoplasma cells during the hemadsorption assay. Fluorescence-activated cell sorting (FACS) data were acquired using a FACSCalibur (Becton Dickinson, Franklin Lakes, USA) equipped with an air-cooled 488 nm argon laser and a 633 nm red diode laser and analyzed with the CellQuest-Pro and FACSDiva software (Becton Dickinson). Hemadsorption was quantified using flow cytometry as previously described[38] with few modifications. Binding of mycoplasma cells to red blood cells can be modeled in an inverse Langmuir isothermal kinetic function:

$$M_f = 1 - \frac{B_{max}[RBC]}{K_d + [RBC]}. \quad (1)$$

The plots represent the best-fitting curves to a series of hemadsorption measurements obtained from at least two biological repeats for each strain. We performed a double-gating strategy, using a preliminary FL3-H/FL2-H gate following an SSC-H/FL1-H gate, as described previsouly[38].

Time-lapse microcinematography. Gliding properties of the wild type and mutant strains were analyzed by time-lapse microcinematography as previously described[5]. Samples from mid-log phase cell cultures were grown overnight on eight-well ibiTreat µ-slides (Ibidi, Gräfelfing, Germany) filled with 200 µl of SP4 medium. Prior to the observation, medium was replaced with fresh SP4 pre-warmed at 37 °C and cell movement was examined at 37 °C and 5% CO_2 using a Nikon Eclipse TE 2000-E inverted microscope equipped with a CO_2 Microscope Cage Incubation System (Okolab, San Bruno, USA). Images were captured at 2 s intervals for a total of 2 min for all different strains. The frequency of motile cells was determined by examining approximately 250 isolated cells of each strain. Given that cells of the P110-W838F mutant were mainly aggregated, only 64 individual cells could be analyzed. Mean velocities were obtained from the analysis of at least 25 motile, isolated cells, by measuring the traveled distance and dividing this value by the time spent in microcinematography. Analysis of the gliding properties was performed using the ImageJ software with the MTrack2, MTrackJ, wrMTrack, and Stack Deflicker plugins (https://imagej.nih.gov/ij/).

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Atomic coordinates and structure factors for the reported crystal structures of P140 and the P140–P110N complex have been deposited in the Protein Data Bank under accession codes 6RUT and 6S3U, respectively. The cryo-electron microscopy density for the 4.1 Å density map of the heterodimer was deposited in the Electron Microscopy Data Bank (EMDB) under the accession code EMD-10890 and the fitted P140–P110N complex under the accession code 6YRK. The 9.8 Å density map of the single-particle cryo-EM density map of the Nap was deposited in the EMDB under the accession codes EMD-10260. The 15 Å in situ density map of the in situ cryo-ET Nap was deposited in the EMDB under the accession code EMD-10259. The source data underlying Fig. 2g are provided as a Source Data file. Other data are available from the corresponding authors upon reasonable request.

Code availability

The reconstruction, sub-tomogram averaging, and classification code referenced in the methods is available:

(i) Artiatomi: http://github.com/uermel/Artiatomi

(ii) Classification: https://github.com/lsprankel/Aparicio_2020.

Received: 22 November 2019; Accepted: 6 May 2020;
Published online: 08 June 2020

References

1. Burgos, R. et al. *Mycoplasma genitalium* P140 and P110 cytadhesins are reciprocally stabilized and required for cell adhesion and terminal-organelle development. *J. Bacteriol.* **188**, 8627–37 (2006).
2. Mernaugh, G. R., Dallo, S. F., Holt, S. C. & Baseman, J. B. Properties of adhering and nonadhering populations of *Mycoplasma genitalium*. *Clin. Infect. Dis.* **17**(Suppl 1), S69–78 (1993).
3. Scheffer, M. P. et al. Structural characterization of the NAP; the major adhesion complex of the human pathogen *Mycoplasma genitalium*. *Mol. Microbiol.* **105**, 869–879 (2017).
4. Nakane, D., Adan-Kubo, J., Kenri, T. & Miyata, M. Isolation and characterization of P1 adhesin, a leg protein of the gliding bacterium *Mycoplasma pneumoniae*. *J. Bacteriol.* **193**, 715–22 (2011).
5. Garcia-Morales, L., Gonzalez-Gonzalez, L., Querol, E. & Pinol, J. A minimized motile machinery for *Mycoplasma genitalium*. *Mol. Microbiol* **100**, 125–38 (2016).
6. Aparicio, D. et al. *Mycoplasma genitalium* adhesin P110 binds sialic-acid human receptors. *Nat. Commun.* **9**, 4471 (2018).
7. Ma, L., Jensen, J. S., Mancuso, M., Myers, L. & Martin, D. H. Kinetics of genetic variation of the *Mycoplasma genitalium* MG192 Gene in experimentally infected chimpanzees. *Infect. Immun.* **84**, 747–53 (2015).
8. Wood, G. E. et al. Persistence, immune response, and antigenic variation of *Mycoplasma genitalium* in an experimentally infected pig-tailed macaque (Macaca nemestrina). *Infect. Immun.* **81**, 2938–51 (2013).
9. Wood, G. E., Patton, D. L., Cummings, P. K., Iverson-Cabral, S. L. & Totten, P. A. Experimental infection of pig-tailed macaques (Macaca nemestrina) with *Mycoplasma genitalium*. *Infect. Immun.* **85**, e00738-16.(2017).
10. Cao, B. et al. High prevalence of macrolide resistance in *Mycoplasma pneumoniae* isolates from adult and adolescent patients with respiratory tract infection in China. *Clin. Infect. Dis.* **51**, 189–194 (2010).
11. Parrott, G. L., Kinjo, T. & Fujita, J. A compendium for *Mycoplasma pneumoniae*. *Front. Microbiol.* **7**, 513 (2016).
12. Bradshaw, C. S. et al. Azithromycin failure in *Mycoplasma genitalium* urethritis. *Emerg. Infect. Dis.* **12**, 1149–52 (2006).
13. Couldwell, D. L. & Lewis, D. A. *Mycoplasma genitalium* infection: current treatment options, therapeutic failure, and resistance-associated mutations. *Infect. Drug Resist.* **8**, 147–61 (2015).
14. Ofek, I., Hasty, D. L. & Sharon, N. Anti-adhesion therapy of bacterial diseases: prospects and problems. *FEMS Immunol. Med. Microbiol.* **38**, 181–91 (2003).
15. Sharon, N. Carbohydrates as future anti-adhesion drugs for infectious diseases. *Biochim. Biophys. Acta* **1760**, 527–537 (2006).
16. Pettersen, E. F. et al. UCSF chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–12 (2004).
17. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–97 (2007).
18. Russ, W. P. & Engelman, D. M. The GxxxG motif: a framework for transmembrane helix-helix association. *J. Mol. Biol.* **296**, 911–9 (2000).
19. Seto, S., Kenri, T., Tomiyama, T. & Miyata, M. Involvement of P1 adhesin in gliding motility of *Mycoplasma pneumoniae* as revealed by the inhibitory effects of antibody under optimized gliding conditions. *J. Bacteriol.* **185**, 1875–7 (2005).
20. Berrow, N. S. et al. A versatile ligation-independent cloning method suitable for high-throughput expression screening applications. *Nucleic Acids Res.* **35**, e45 (2007).
21. Winter, G. xia2: an expert system for macromolecular crystallography data reduction. *J. Appl. Crystallogr.* **43**, 186–190 (2010).
22. Kabsch, W. Xds. *Acta Crystallogr D Biol. Crystallogr.* **66**, 125–32 (2010).
23. Evans, P. Scaling and assessment of data quality. *Acta Crystallogr. Sect. D* **62**, 72–82 (2006).
24. CCP4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D. Biol. Crystallogr* **50**, 760–763 (1994).
25. McCoy, A. J. Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **63**, 32–41 (2007).
26. Cowtan, K. D. & Main, P. Phase combination and cross validation in iterated density-modification calculations. *Acta Crystallogr. Sect. D* **52**, 43–48 (1996).
27. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol. Crystallogr* **60**, 2126–32 (2004).
28. Murshudov, G. N. et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 355–367 (2011).
29. Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36–51 (2005).
30. Kunz, M. & Frangakis, A. S. Three-dimensional CTF correction improves the resolution of electron tomograms. *J. Struct. Biol.* **197**, 114–122 (2017).
31. Pruggnaller, S., Mayr, M. & Frangakis, A. S. A visualization and segmentation toolbox for electron microscopy. *J. Struct. Biol.* **164**, 161–5 (2008).
32. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).

33. Goddard, T. D. et al. UCSF ChimeraX: meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).
34. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
35. Zhang, K. Gctf: real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016).
36. Scheres, S. H. A Bayesian view on cryo-EM structure determination. *J. Mol. Biol.* **415**, 406–418 (2012).
37. Pich, O. Q., Burgos, R., Planell, R., Querol, E. & Pinol, J. Comparative analysis of antibiotic resistance gene markers in *Mycoplasma genitalium*: application to studies of the minimal gene complement. *Microbiology* **152**, 519–527 (2006).
38. Garcia-Morales, L., Gonzalez-Gonzalez, L., Costa, M., Querol, E. & Pinol, J. Quantitative assessment of mycoplasma hemadsorption activity by flow cytometry. *PLoS ONE* **9**, e87500 (2014).

Acknowledgements

We thank Utz Ermel for providing custom Matlab scripts. This work was supported by grants BFU2018-101265-B-100 and BIO2017-84166-R from MINECO (Spain). A.S.F. acknowledges the Deutsche Forschungsgemeinschaft Grant FR1653/6-1 and SFB902, as well as the Loewe Dynamem funding for the project. D.A. acknowledges María de Maeztu Unit of Excellence grant MDM-2014-0435. Many thanks are given to the XALOC beamline team at ALBA for all the support during data collection and, in particular, to Dr. Barbara M. Calisto for her many contributions to the mycoplasma structural studies. Many thanks are also given to the Crystallography Platform at the Barcelona Science Park (PCB).

Author contributions

Conceived and designed the experiments: D.A., M.P.S., J.P., O.Q.P., I.F., and A.S.F. Performed the experiments: D.A., M.P.S., M.M.-S., D.V., M.R., M.S.W., A.S., J.R., L.S., S.T.-P., L.G.-G., and O.Q.P. Analyzed the data: D.A., M.P.S., M.M.-S., D.V., M.S.W., J.R., L.S., J.P., O.Q.P., I.F., and A.S.F. Contributed reagents/materials/analysis tools: E.Q., J.P., I.F., and A.S.F. Wrote the paper: D.A., M.S., O.Q.P., I.F., and A.S.F.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41467-020-16511-2.

Correspondence and requests for materials should be addressed to O.Q.P., I.F. or A.S.F.

Peer review information *Nature Communications* thanks Jeremy Derrick, Makoto Miyata, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer review reports are available.

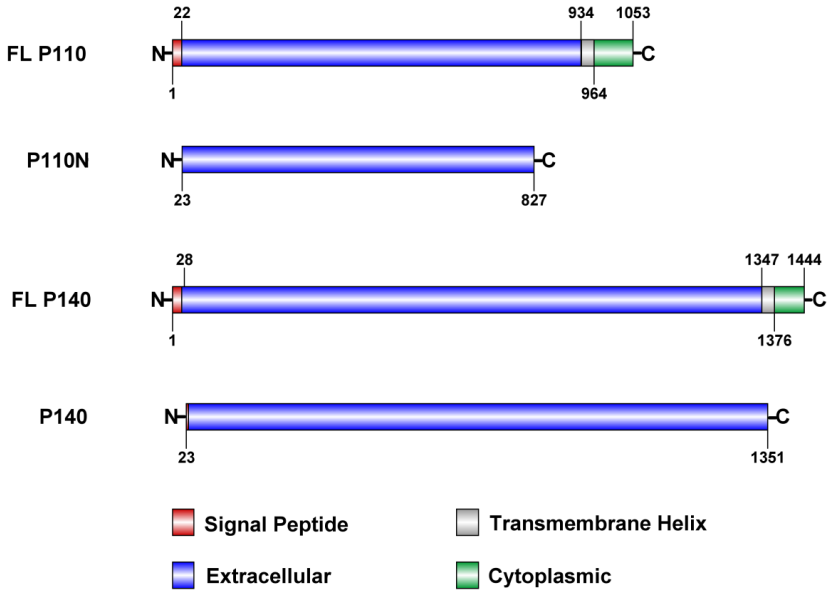
Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary Information

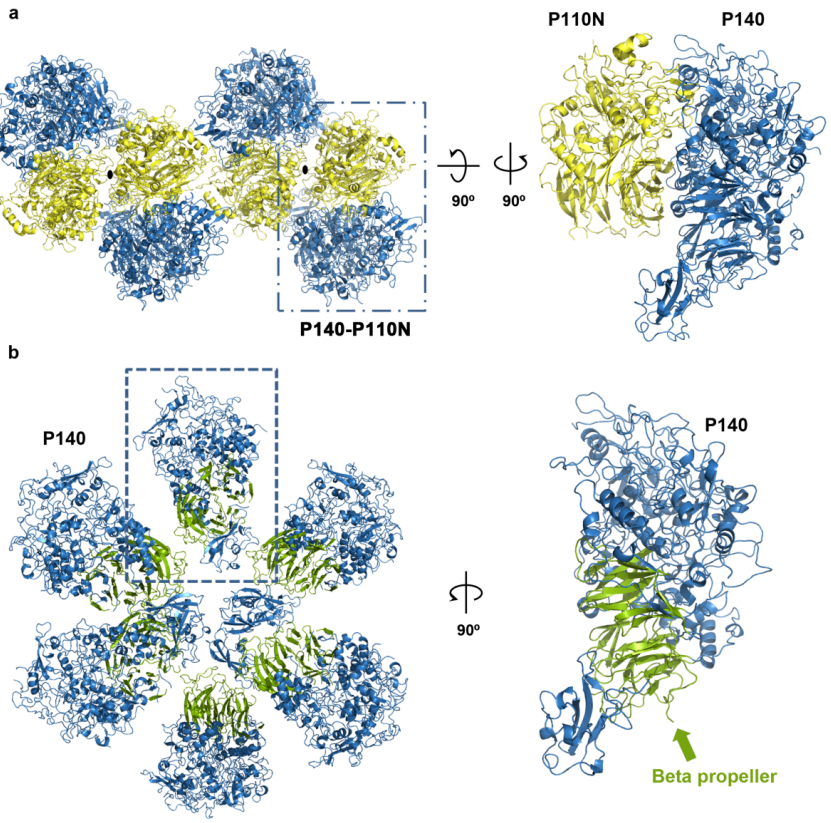
Structure and mechanism of the Nap adhesion complex from the human pathogen *Mycoplasma genitalium*

David Aparicio, Margot P. Scheffer, Marina Marcos-Silva et al.



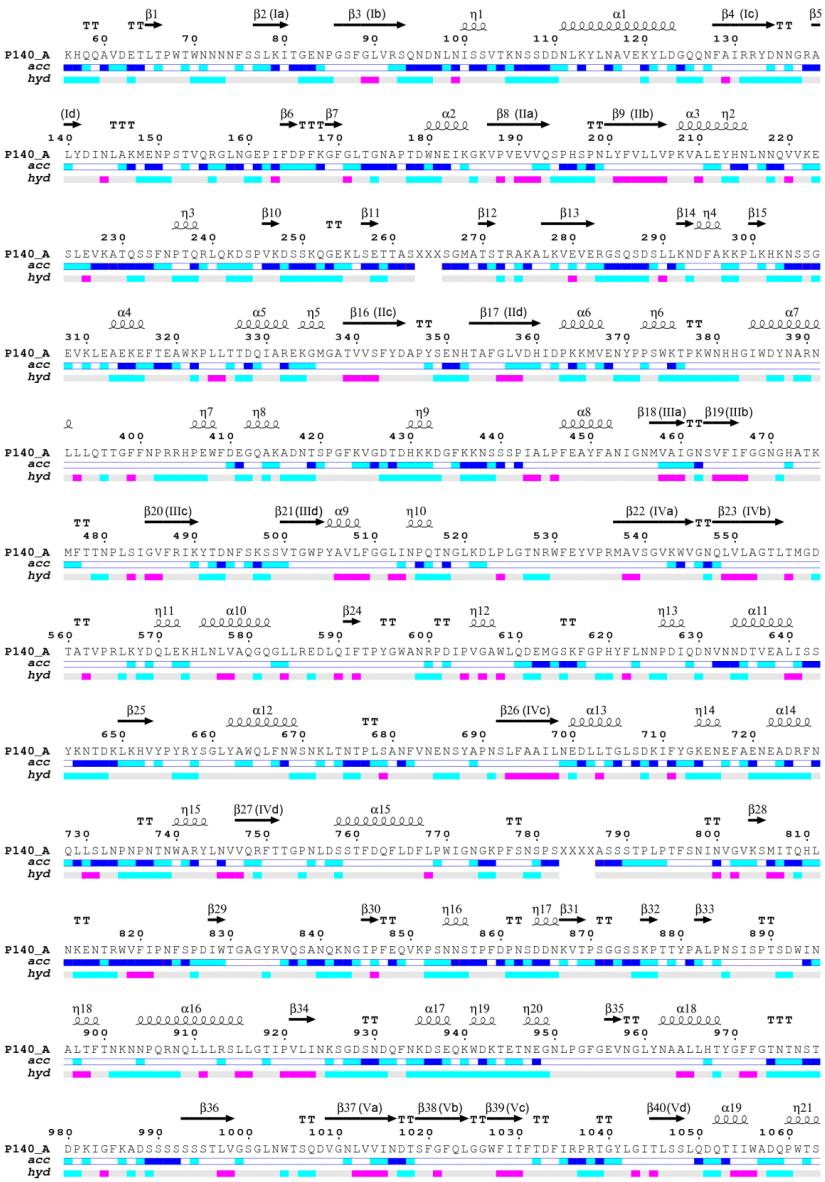
Supplementary Figure 1. **Proteins and constructs used in this study**

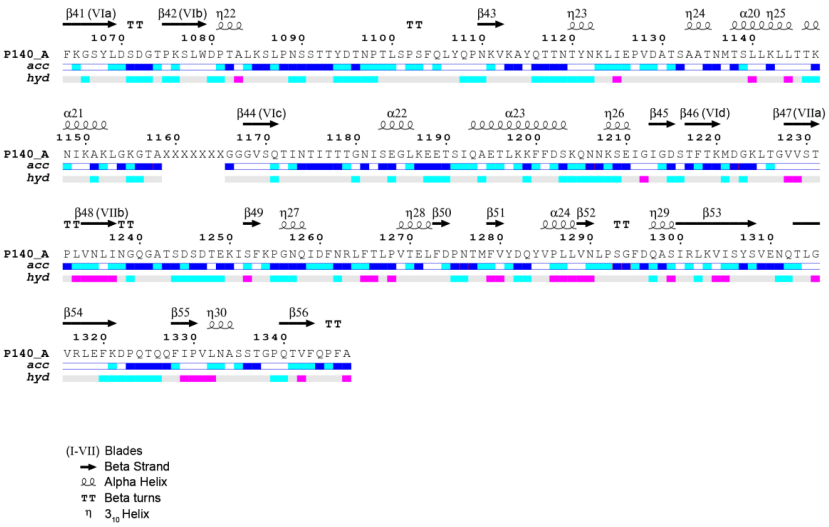
The Psi-Pred server[1] was used to predict the location of signal peptides (red), extracellular regions (lilac), transmembrane helices (grey) and cytoplasmic regions (green) in the full length (FL) P140 and P110 proteins. Constructs of the extracellular regions of P140 and P110 are also shown. In P110N the C-domain was excluded. Numbers indicate the residues at the boundaries.



Supplementary Figure 2. **Asymmetric units content of P140-alone and P140-P110N crystals**

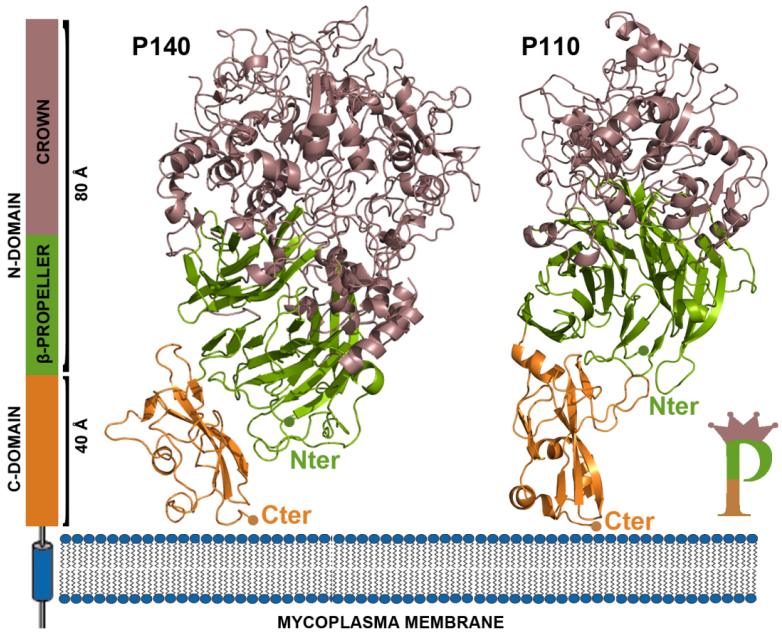
a) The P140-P110N crystals contain four P140-P110N heterodimers in the asymmetric unit that are arranged as two dimers with a local, non-crystallographic, two-fold axis (black oblong dots). The P140 construct (blue) includes the N- and C-domains, while the P110N construct (yellow) corresponds only to the N-domain of P110. b) The P140-alone crystals contain six P140 subunits in the asymmetric unit. β-propellers are explicitly indicated (green).


P140_A
acc
hyd



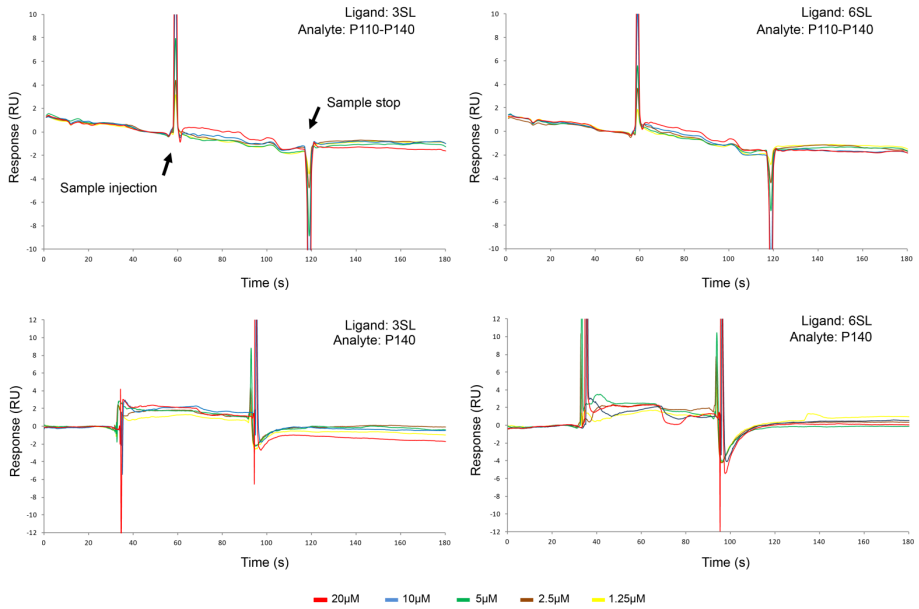
Supplementary Figure 3. **P140 sequence and secondary structural elements**

Amino acid sequence of the extracellular region of P140, with the corresponding secondary structural elements (helices with squiggles, strands with arrows and turns with TT letters) (Endscript2 [2]). The seven blades (or β-sheets) in the propeller are represented by roman numerals (I-VII). Bars below the sequence show solvent accessibility (blue is accessible, cyan is intermediate, white is buried) and hydropathy (pink is hydrophobic, cyan is hydrophilic).

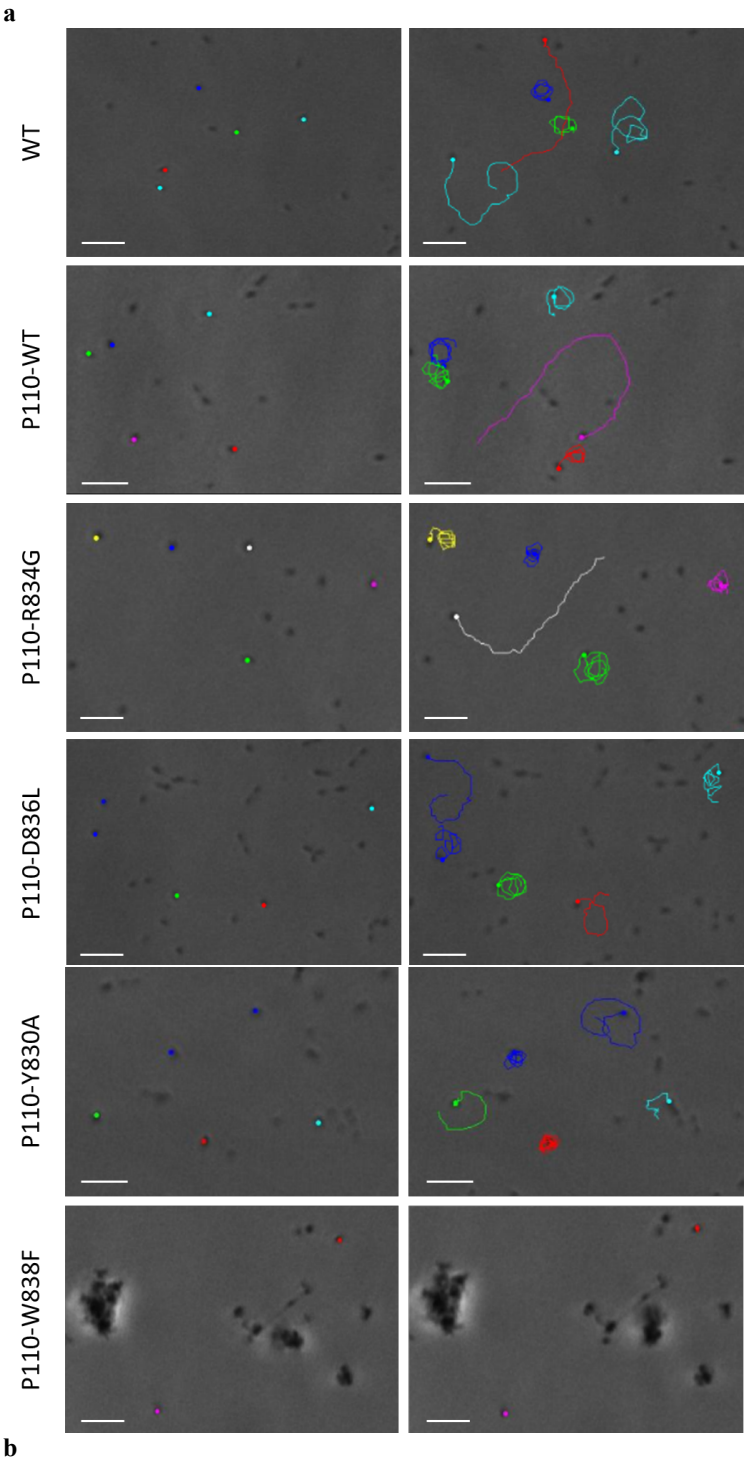


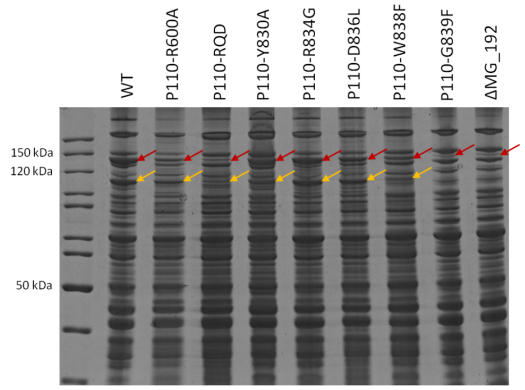
Supplementary Figure 4. **Structural comparison of P140 and P110**

P140 and P110 present a similar domain organization with small C-domains (orange) and large N-domains in the extracellular regions of the proteins. The N-terminal domains contain two distinct regions, a seven blade β-propeller (green) and a crown (brown), with the β-propellers situated between the C-terminal domains and the crowns. Also in both proteins, the C-terminal domains are followed in sequence by a (predicted) transmembrane helix and a cytosolic region with about a hundred residues.



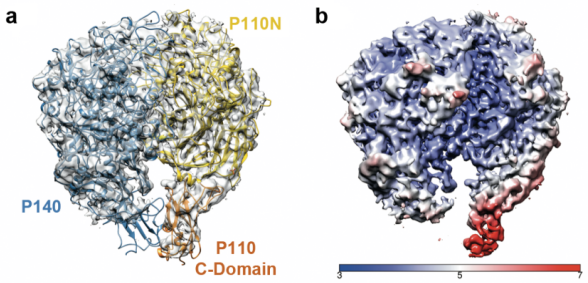
Supplementary Figure 5. Surface Plasmon Resonance (SPR)
Surface Plasmon Resonance sensorgrams showing no binding for the extracellular P140-P110 heterodimer (top panels) and for P140 (bottom panels) against increasing amounts of 3SL (left panels) and 6SL (right panels).


a
WT
P110-WT
P110-R834G
P110-D836L
P110-Y830A
P110-W838F
b



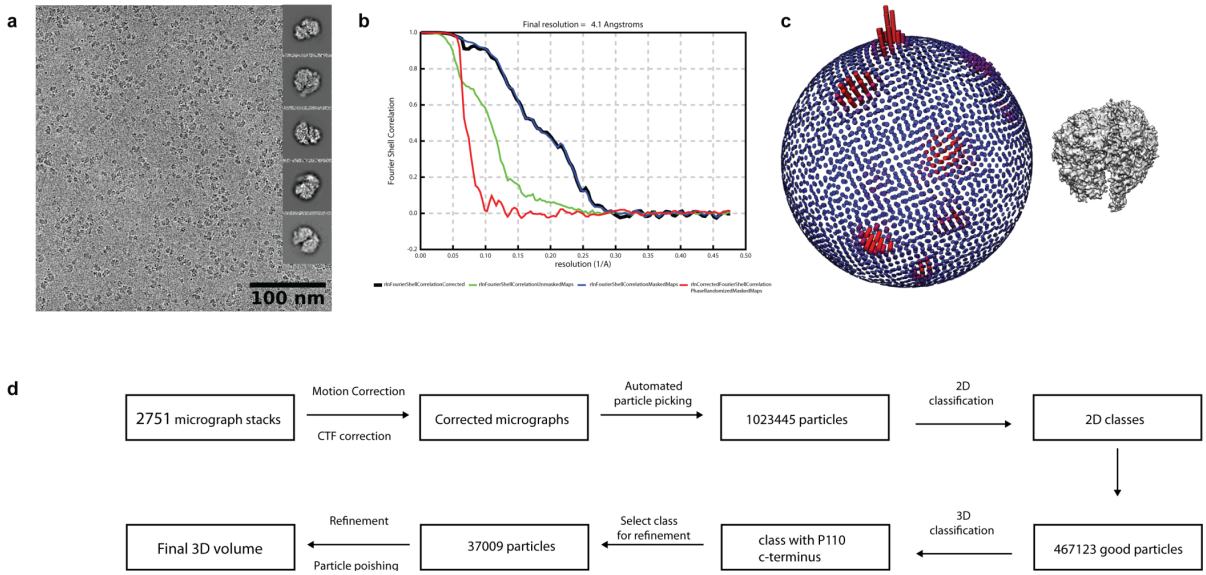
Supplementary Figure 6. Time-lapse microcinematography and expression analysis for WT and selected strain variants

a) Representative tracks of *M. genitalium* G37 Wild Type (WT) and P110-WT cells, and P110-R834G, P110-D836L, P110-Y830A and P110-W838F mutant cells. Although cells from P110-W838F were found to have severe hemadsorption deficiencies by flow cytometry, these cells attached properly to the plastic surface of Ibidi slides. This behaviour is somewhat common and has been previously described [3]. Bar is 5 µm. b) SDS-PAGE of whole cell lysates from the WT and the different adhesin mutants. Red and orange arrows indicate the presence of P140 and P110, respectively. Between two and four independent clones from each strain were analysed by SDS-PAGE and representative clones from each strain are shown here (original SDS-PAGE is provided independently).

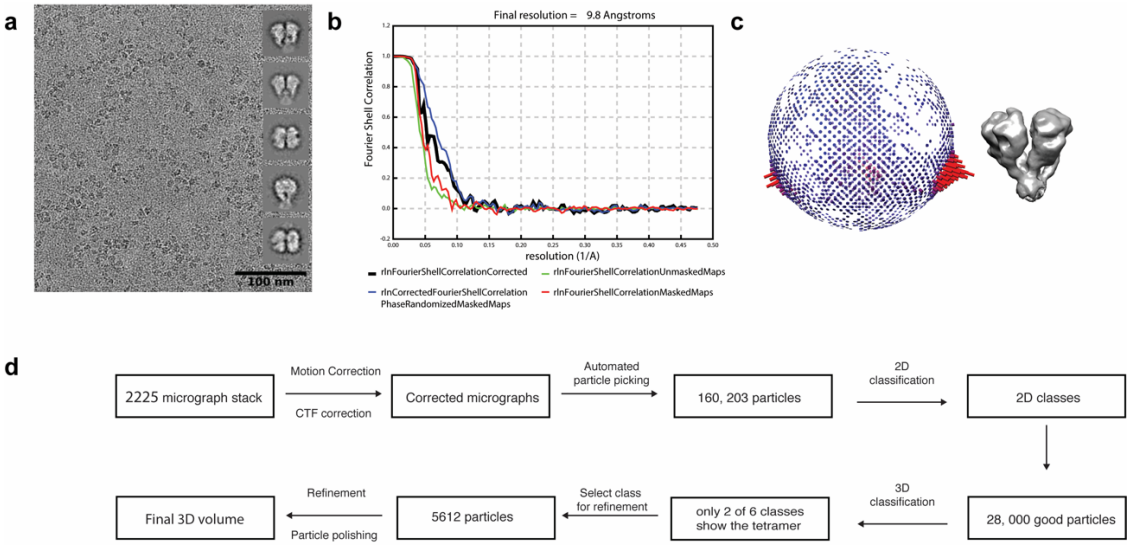


Supplementary Figure 7. Cryo-EM of the P140-P110 extracellular region

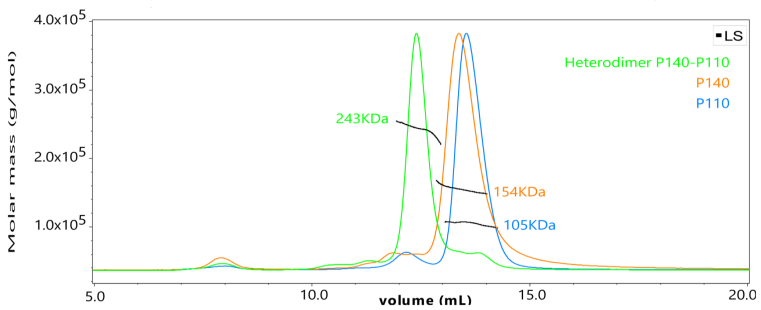
a) Cryo-EM map of the extracellular region of the P140-P110 heterodimer. The structure of P140-P110N, fitted into the map density, is also shown with P140 in blue and P110 in yellow. There is no density for the C-domain of P140, while density is visible, although weak, for the C-domain of P110 that was fitted independently into the cryo-EM map (orange). b) Local resolution of the cryo-EM map. The lowest resolution corresponds to the C-terminal domains, which indicates a significant flexibility of these domains, in particular of the C-terminal domain from P140, with respect to the bulkier N-domains.



Supplementary Figure 8. **Processing of Cryo-EM data from the extracellular regions of P140-P110** a) Representative cryo-electron micrograph of P140-P110 heterodimers with a subset of 2D classes as insets (2751 micrographs were recorded in total). b) Fourier shell correlation showing a resolution of 4.1 Å using the 0.143 criteria. c) Angular distribution plot of images corresponding to the final 3D volume that displays anisotropy. d) Flowchart showing the processing pipeline from the micrograph stacks to the final 3D volume.

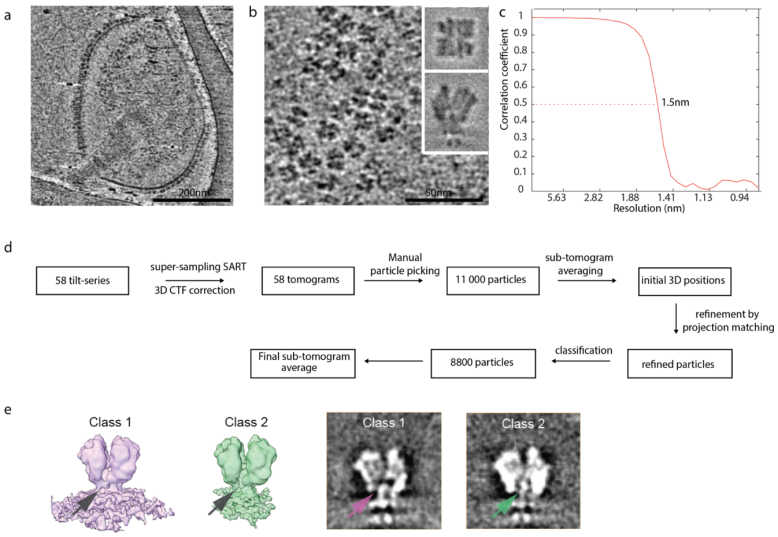


Supplementary Figure 9. **Image and data cryo-EM processing of the Nap complex** a) Representative cryo-electron micrograph of Nap complexes with a subset of 2D classes as insets (2225 micrographs were recorded in total). b) Fourier shell correlation of the cryo-EM map showing a resolution of 9.8 Å. c) Angular distribution plot of the final 3D volume. d) Flowchart showing the processing pipeline from the micrograph stacks to the final 3D volume.

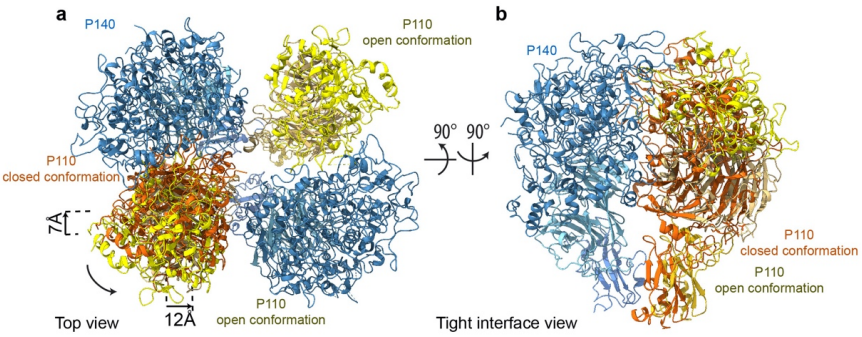


Supplementary Figure 10. SEC-MALS measurements

MALS measurements indicate that essentially only heterodimers are detected when mixing equimolecular amounts of P110 and P140.

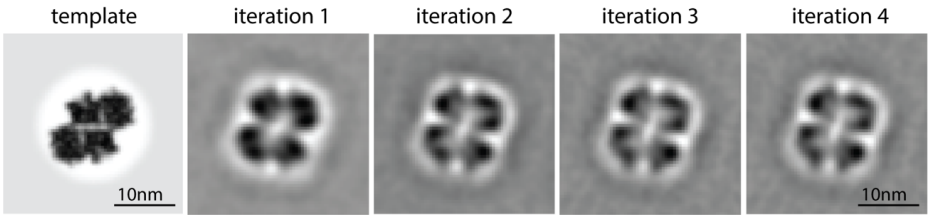


Supplementary Figure 11. **Cryo-ET of *M. genitalium* cells**. a) A 0.88 nm thick slice through a representative tomogram (58 tomograms were used in total) of a detergent-permeabilized *M. genitalium* cell. b) A zoomed in top view of the Nap complexes found on the cell in (a). Insets show slices through the top view (upper inset) and side view (lower-inset) of the sub-tomogram average. c) Fourier shell correlation of the sub-tomogram average showing a resolution of 15 Å at the 0.5 criterion. d) Flowchart showing the processing pipeline from the raw tilt-series to the final sub-tomogram average. e) Different classes observed during subtomogram averaging, class 1 (in purple) shows well resolved stalks (arrows), whereas class 2 (in green) shows less well-resolved stalks. Particles from class 1 were used in the final average.



Supplementary Figure 12. Transition between the open and closed conformations through the movement of P110.

(a) Top view and (b) a tight interface view of a dimer of P140-P110 heterodimers showing the relative movement of P110 from the closed conformation (orange) as it retracts 7Å from its position at the tight interface and moves 12 Å closer to P140 across the loose interface to form the open conformation (yellow). P140 (blue) is kept unchanged as a reference.



Supplementary Figure 13. Sub-tomogram averaging experiment using the crystal structure tetramer as a starting reference

When using the ‘closed' tetramer structure (crystal structure packing unit) as a starting reference, the sub-tomogram averaging procedure still converges to the open structure after 5 iterations. Scale bar is 10 nm.

Supplementary Tables

Data collection and Refinement Statistics [a]		
	P140	**P140-P110N**
Data collection		
Space group	C2	$P2_1$
Cell dimensions		
a,b,c (Å)	459.19, 116.65, 285.64	150.99, 157.28, 192.36
α,β,γ (°)	90.00, 124.20, 90.00	90.00, 93.81, 90.00
Unique reflections	123183 (6159)	189359 (784)
Resolution (Å)	189.88-3.26 (3.58-3.26)	157.28-2.65 (2.70-2.65)
Wavelength (Å)	0.9791	0.9789
Rmeas (%)[b]	0.15 (0.67)	0.58 (2.11)
I/σI	5.8 (1.6)	6.6 (1.5)
CC(1/2)	0.98 (0.74)	0.87 (0.343)
Completeness (Spherical)	63.5 (13.2)	73.3 (6.1)
Completeness (Ellipsoidal)	86.1 (57.9)	90.8 (44.4)
Redundancy	3.1 (2.6)	6.8 (6.8)
Refinement Statistics		
Resolution	37.36-3.26	121.66-2.65
Num. of reflections	123103	189358
R_{Work} (%)[c]	18.4	18.70
R_{free} (%)[d]	20.5	22.4
No. residues	1292	2134
No. ligands	0	0
Solvent content (%)	67.36	46
Av. B-factor (Å^2)	80.85	40.36
Coor. Error (Å)[e]	0.42	0.36
Rms dev.bonds (Å)	0.01	0.01
Rms dev.angles (°)	1.27	1.25

[a] **Values in parentheses correspond to the highest resolution shell.**
[b] **$R_{sym} = \Sigma_{hkl}\Sigma_i|I_i(hkl) - \langle I(hkl)\rangle| / \Sigma_{hkl}\Sigma_i I_i (hkl)$, where $I_i(hkl)$ is the intensity of an observation and $\langle I(hkl)\rangle$ is the mean value of observations for a unique reflection.**
[c] **$R_{cryst} = \Sigma_h|Fo(h) - Fc(h)| / \Sigma_h|Fo(h)$, where Fo and Fc are the observed and calculated structure-factor amplitudes, respectively.**
[d] **R_{free} was calculated with 5% of data, which was excluded from the refinement.**
[e] **Based on maximum likelihood.**

Supplementary Table 1. Data collection and refinement statistics

	P140-P110N	NAP complex
Data collection		
Microscope	FEI Titan Krios	FEI Titan Krios
Detector	Gatan K2 Summit	Gatan K2 Summit
Acquisition software	SerialEM 3.7	SerialEM 3.7
Magnification	130,000x	130,000x
Pixel size (Å)	1,05	1,05
Total electron dose (e- / Å²)	50	50
Dose rate (Å2 / s-1)	7,7	7,4
number of frames	32	34
Defocus range (μm)	-1 to -4	-1 to -4
Micrographs used	2.751	2.225
Processing		
Software	Relion 3.0 beta	Relion 3.0 beta
Motion correction	UCSF MotionCor2	UCSF MotionCor2
CTF estimation	GCTF v 1.06	GCTF v 1.06
Total extracted particles	1.024.402	160.203
After 2D Classification	467.123	28.000
Number of 3D classes	6	6
Number of refined particles	37.009	5.612
Symmetry	/	C2
Map sharpening B factor (Å²)	-78	-218
Resolution FSC 0.143 (Å)	4,1	9,8

Supplementary Table 2. Single particle analysis data collection

	Correlation#/clash-score^		
	Cryo-EM (P140-P110 heterodimer)	**Cryo-EM ("Closed" Nap conformation)**	**Cryo-ET ("Open" Nap conformation)**
P140	**0.74 / 13.55**	-	-
P110	**0.73 / 13.55**	-	-
P140	-	**0.95 / 12.13**	**0.94 / 10.39**
P110	-	**0.94 / 12.42**	**0.94 / 10.39**

Real-space correlation of the fitted models into the corresponding electron density maps. Fitting was based on the local optimisation algorithm of Chimera[4]
^The "clash-score" is defined as the number of serious overlaps (> 0.4 Å) per 1000 atoms (Molprobity [5]).

Supplementary Table 3. Correlation and clash-score values from rigid-body fitted crystal structures

Strain Name	Genotype	Reference
G37	Wild-type	ATCC 33530
ΔMG_192	Deletion of the MG_192 gene by allelic exchange (G37 ΔMG_192::*tet*M438)	Burgos et al. 2006
P110-WT	Re-introduction of a MG_192 wild-type allele in a ΔMG_192 mutant	Aparicio et al. 2018
P110-R600A	Introduction of a MG_192 allele bearing a R600A substitution in a ΔMG_192 mutant	This work
P110-RQD	Introduction of a MG_192 allele bearing a R600A, Q460A and D461A substitutions in a ΔMG_192 mutant	This work
P110-Y830A	Introduction of a MG_192 allele bearing a Y830A substitution in a ΔMG_192 mutant	This work
P110-R834G	Introduction of a MG_192 allele bearing a R834G substitution in a ΔMG_192 mutant	This work
P110-D836L	Introduction of a MG_192 allele bearing a D836L substitution in a ΔMG_192 mutant	This work
P110-W838F	Introduction of a MG_192 allele bearing a W838F substitution in a ΔMG_192 mutant	This work
P110-G389F	Introduction of a MG_192 allele bearing a G839F substitution in a ΔMG_192 mutant	This work

Supplementary Table 4. Strains used in this study

Strain	n	Motile cells (%)	Velocity (µm/s)[a]	Motility loss (%)
G37 (WT)	~250	88	0.126±0.002	-
P110-WT	~250	77	0.119±0.014	5.6
P110-Y830A	~250	78	0.100±0.011*	20.6
P110-R834G	~250	86	0.114±0.007	9.5
P110-D836L	~250	86	0.102±0.009*	19.0
P110-W838F	~64	0	0	100

[a]Velocities are shown as mean values +/- standard error. n=25 biologically independent cells, from different fields, different preparations and different microcinematographies.
* Statistically significant values (T(25)=2.06; p<0.05). Statistical significance was assessed with the Paired Student's two-sided T test.

Strain	Difference of means	Confidence interval (+/-)	Effect size	Degrees of freedom	p-value
WT	-	-	-	-	-
P110-WT	0.007	0.01663	0.16286	24	0.57
P110-Y830A	0.026	0.01522	0.69427	24	0.02
P110-R834G	0.012	0.00988	0.48446	24	0.10
P110-D836L	0.024	0.01230	0.81200	24	0.01
P110-W838F	-	-	-	-	-

Supplementary Table 5. Gliding motility parameters of *M. genitalium* G37 wild-type cells and all different mutant strains.

Synthetic clone for P140

CGAATTGAAGGAAGGCCGTCAAGGCCGCATTTAATTAAACAAGTTTGTACAAAAAAGCAGGCTCTGGAAGTTCTGTTTCAGGGTCCGAGCGGTGCAATGCATC
AGCCGAAAAAACGTCTGGCAAAAAAAAGCTGGGCATTTCTGACCGCAGCACTGACCCTGGGTGTTATTACCGGTGTTGGTGGTTATTTTCTGTTTAACCAGAAT
AAACAGCGTAGCAGCGTTAGCAATTTTGCATATCAGCCGAAACAGCTGAGCGTTAAACATCAGCAGGCAGTTGATGAAACCCTGACCCCGTGGACCTGGAATA
ATAATAATTTTTCAAGCCTGAAAATTACCGGTGAAAATCCGGGTAGCTTTGGTCTGGTTCGTAGCCAGAATGATAATCTGAATATTAGCAGCGTGACCAAAAAT
AGCAGCGACGATAATCTGAAATATCTGAATGCCGTCGAGAAATATCTGGATGGTCAGCAGAATTTTGCCATTCGTCGCTATGATAATAATGGTCGTGCCCTGTA
TGATATTAACCTGGCAAAAATGGAAAATCCGAGCACCGTTCAGCGTGGTCTGAATGGTGAACCGATTTTTGATCCGTTTAAAGGTTTTGGTCTGACCGGTAATG
CACCGACCGATTGGAATGAAATTAAAGGTAAAGTTCCGGTTGAGGTTGTTCAGAGTCCGCATAGCCCGAATCTGTATTTTGTTCTGCTGGTTCCGAAAGTTGCC
CTGGAATACCATAATCTGAACAATCAGGTGGTGAAAGAAAGCCTGGAAGTTAAAGCAACCCAGAGCAGCTTTAATCCGACCCAGCGTCTGCAGAAAGATAGT
CCGGTTAAAGATAGCAGCAAACAGGGTGAAAAACTGAGCGAAACCACCGCAAGCAGCATGAGCAGCGGTATGGCAACCAGCACCCGTGCAAAAGCACTGAA
AGTTGAAGTTGAACGTGGTAGCCAGAGCGATAGCCTGCTGAAAAATGATTTTGCAAAAAAACCGCTGAAGCATAAAAACAGCAGCGGTGAAGTTAAACTGGA
AGCCGAAAAAGAATTTACCGAAGCATGGAAACCGCTGCTGACCACCGATCAGATTGCACGTGAAAAAGGTATGGGTGCAACCGTTGTTAGCTTTTATGATGCA
CCGTATAGCGAAAATCATACCGCATTTGGTCTGGTGGATCATATTGATCCGAAAAAAATGGTCGAGAATTATCCGCCTAGCTGGAAAACCCCGAAATGGAATC
ATCATGGCATCTGGGATTATAATGCACGTAATCTGCTGCTGCAGACCACCGGTTTTTTTAATCCGCGTCGTCATCCGGAATGGTTTGATGAAGGTCAGGCAAAA
GCAGATAATACCAGTCCGGGTTTTAAAGTTGGTGATACCGATCATAAAAAAGATGGCTTTAAAAAAAATAGCTCCAGCCCGATTGCACTGCCGTTTGAAGCAT
ATTTTGCCAATATTGGTAATATGGTGGCCATTGGTAACAGCGTGTTTATTTTTGGTGGTAATGGTCATGCCACCAAAATGTTTACCACCAATCCGCTGAGCATTG
GTGTGTTTCGTATCAAATATACCGATAATTTTAGCAAAAGCAGCGTTACCGGTTGGCCGTATGCAGTTCTGTTTGGTGGTCTGATTAATCCGCAAACCAATGGT
CTGAAAGATCTGCCGCTGGGCACCAATCGTTGGTTTGAATATGTTCCGCGTATGGCAGTTAGCGGTGTTAAATGGGTTGGTAATCAGCTGGTTCTGGCAGGCAC
CCTGACCATGGGTGATACCGCAACCGTTCCGCGTCTGAAATATGATCAGCTGGAAAAACATCTGAATCTGGTTGCACAGGGTCAGGGTCTGCTGCGTGAAGAT
CTGCAGATTTTTACCCCGTATGGTTGGGCAAATCGTCCGGATATTCCGGTTGGTGCATGGCTGCAGGATGAAATGGGTAGCAAATTTGGTCCGCACTATTTTCT
GAACAACCCGGATATTCAGGACAATGTGAATAATGATACCGTGGAAGCACTGATCAGCAGCTATAAAAATACCGACAAACTGAAACACGTGTATCCGTATCGT
TATAGCGGTCTGTATGCATGGCAGCTGTTTAATTGGAGCAATAAACTGACCAATACACCGCTGAGCGCCAATTTTGTTAACGAAAATAGCTATGCACCGAATA
GCCTGTTTGCAGCAATTCTGAATGAAGATCTGCTGACCGGTCTGAGCGATAAAATTTTTTATGGCAAAGAGAACGAATTTGCCGAAAATGAAGCCGATCGTTTT
AATCAGCTGCTGAGCCTGAATCCGAATCCGAATACCAATTGGGCACGTTATCTGAATGTTGTTCAGCGTTTTACCACCGGTCCGAATCTGGATAGCAGCACCTT
TGATCAGTTTCTGGATTTTCTGCCGTGGATTGGTAATGGTAAACCGTTTAGCAATAGCCCGAGCCCGAGCACCAGCGCCAGCAGTAGTACACCGCTGCCGACCT
TTAGCAATATTAATGTTGGTGTGAAAAGCATGATCACCCAGCATCTGAATAAAGAAAATACCCGCTGGGTGTTTATCCCGAATTTTAGTCCGGATATTTGGACC
GGTGCAGGTTATCGTGTTCAGAGCGCAAATCAGAAAAATGGTATTCCGTTCGAACAGGTTAAACCGAGCAATAATAGCACCCCGTTTGATCCGAATAGCGACG
ATAACAAAGTTACCCCGAGCGGTGGTAGCAGCAAACCGACCACCTATCCGGCACTGCCGAATAGCATTAGCCCGACCAGCGATTGGATTAATGCACTGACCTT
TACCAACAAAAATAATCCGCAGCGTAACCAACTGCTGCTGCGTAGCCTGCTGGGCACCATTCCGGTTCTGATTAATAAAAGCGGTGATAGCAACGATCAGTTC
AACAAAGATAGCGAACAGAAATGGGACAAAACCGAAACCAATGAAGGTAATCTGCCTGGTTTTGGTGAAGTTAATGGTCTGTATAATGCAGCTCTGCTGCACA
CCTATGGTTTTTTTGGCACCAATACCAATAGCACCGATCCGAAAATTGGCTTTAAAGCAGATAGCAGCAGCTCAAGTAGCAGCACCCTGGTTGGTAGCGGTCTG
AATTGGACCAGCCAGGATGTTGGTAATCTGGTTGTTATTAACGATACCAGCTTTGGTTTTCAGCTGGGTGGTTGGTTTATTACCTTTACCGATTTCATTCGTCCG
CGTACCGGTTATCTGGGTATTACCCTGAGCAGCCTGCAGGATCAGACCATTATTTGGGCAGATCAGCCGTGGACCAGCTTTAAAGGTAGCTATCTGGATAGTGA
TGGCACCCCGAAAAGCCTGTGGGATCCGACCGCACTGAAAAGCCTGCCGAATAGTAGCACCACCTATGATACCAATCCGACCCTGAGCCCGAGCTTTCAGCTG
TATCAGCCGAATAAAGTTAAAGCATATCAGACCACCAATACCTATAATAAACTGATTGAACCGGTTGATGCAACCAGCGCAGCAACCAATATGACCAGTCTGC
TGAAACTGCTGACCACGAAAAACATTAAAGCCAAACTGGGTAAAGGCACCGCCAGCAGCCAGGGTAATAATAATGGTGGTGGTGTTAGCCAGACCATTAATA
CCATTACCACCACCGGTAATATTAGCGAAGGTCTGAAAGAAGAAACCAGCATTCAGGCAGAAACCCTGAAAAAATTCTTCGATAGCAAACAGAATAATAAAA
GCGAAATCGGCATTGGCGATAGCACCTTTACCAAAATGGATGGTAAACTGACCGGTGTTGTTAGCACACCGCTGGTTAATCTGATTAATGGTCAGGGTGCAAC
CAGCGATAGCGATACCGAAAAAATTAGCTTTAAACCGGGTAACCAGATCGACTTTAATCGTCTGTTTACCCTGCCGGTTACCGAACTGTTTGATCCTAATACCA
TGTTCGTGTATGATCAGTATGTTCCGCTGCTGGTGAATCTGCCGAGCGGTTTTGATCAGGCAAGCATTCGTCTGAAAGTGATTAGCTATAGCGTGGAAAATCAG
ACCCTGGGCGTTCGTCTGGAATTTAAAGATCCGCAGACCCAGCAGTTTATTCCGGTGCTGAATGCAAGCAGCACCGGTCCGCAGACCGTTTTTCAGCCGTTTAA
TCAGTGGGCAGATTATGTTCTGCCGCTGATTGTTACCGTTCCGATTGTTGTGATTATTCTGAGCGTTACCCTGGGTCTGACCATTGGTATTCCGATGCATCGTAA
TAAAAAAGCACTGCAGGCAGGTTTTGACCTGAGCAATAAAAAAGTTGACGTTCTGACCAAAGCCGTTGGTAGCGTTTTAAAGAAATTATTAACCGCACCGGC
ATTAGCAATGCACCGAAAAAACTGAAACAGGCAACCCCGACCAAACCGACCCCGAAAACCCCTCCGAAACCGCCTGTTAAACATAAGTAAACCCAGCTTTCTT
GTACAAAGTGGTGGCGCGCCCTGGGCCTCATGGGCCTTCCTTTCACTGCC

Synthetic clone for P110

CGAATTGAAGGAAGGCCGTCAAGGCCGCATTTAATTAAACAAGTTTGTACAAAAAAGCAGGCTCTGGAAGTTCTGTTTCAGGGTCCGAGCGGTGCA
ATGAAAACCATGCGTAAACAAATTTATAAAAAAGCCTATTGGCTGCTGCTGCCGTTTCTGCCGCTGGCACTGGCAAATACCTTTCTGGTTAAAGAGGATAGCA
AAAACGTTACCGCATATACCCCGTTTGCAACCCCGATTACCGATAGCAAAAGCGATCTGGTTAGCCTGGCACAGCTGGATAGCAGCTATCAGATTGCAGATCA
GACCATTCATAACACCAACCTGTTTGTGCTGTTCAAAAGCCGTGATGTGAAAGTGAAATATGAAAGCAGCGGCAGCAACAACATTAGCTTTGATAGCACCAGC
CAGGGTGAAAAACCGAGCTATGTTGTTGAATTTACCAACAGCACCAATATCGGCATTAAATGGACCATGGTGAAGAAATATCAGCTGGATGTTCCGAATGTGA
GCAGCGATATGAATCAGGTTCTGAAAAACCTGATTCTGGAACAGCCGCTGACCAAATATACCCTGAATAGCAGCCTGGCAAAAGAAAAAGGTAAAACCCAGC
GTGAAGTTCATCTGGGTAGCGGTCAGGCAAATCAGTGGACCAGCCAGCGTAATCAGCATGATCTGAATAATAATCCGAGCCCGAATGCAAGCACCGGTTTTAA
ACTGACCACCGGTAATGCATATCGTAAACTGAGCGAAAGCTGGCCGATTTATGAACCGATTGATGGCACCAAACAGGGTAAAGGTAAAGATAGCAGCGGTTG
GAGCAGCACCGAAGAAAATGAAGCAAAAAATGATGCACCGAGCGTTAGCGGTGGTGGTAGCAGCAGCGGCACCTTTAATAAATATCTGAATACCAAACAGGC
CCTGGAAAGCATTGGTATTCTGTTTGATGATCAGACACCGCGTAATGTTATTACCCAGCTGTATTATGCAAGCACCAGCAAACTGGCAGTGACCAATAATCATA
TTGTGGTTATGGGCAATAGCTTTCTGCCGAGCATGTGGTATTGGGTTGTTGAACGTAGCGCACAAGAAAATGCCAGCAATAAACCGACCTGGTTTGCAAATAC
CAATCTGGATTGGGGTGAGGATAAACAGAAACAGTTTGTTGAAAACCAGCTGGGCTATAAAGAAACCACCAGTACCAATAGCCACAACTTTCATAGCAAAAG
CTTTACCCAGCCTGCATATCTGATTAGCGGTATTGATAGCGTGAACGATCAGATTATCTTCAGCGGTTTTAAAGCAGGTAGCGTTGGTTATGATAGCAGCAGCA
GCAGTAGCAGCTCAAGCAGCAGCTCAACCAAAGATCAGGCACTGGCATGGTCAACCACCACCAGTCTGGATAGCAAAACCGGTTATAAAGATCTGGTGACCA
ATGATACCGGTCTGAATGGTCCGATTAATGGTAGCTTTAGCATCCAGGATACCTTTAGCTTTGTTGTTCCGTATAGCGGCAATCATACCAATAATGGCACCACC
GGTCCGATTAAAACCGCATATCCGGTTAAAAAAGATCAGAAAAGCACCGTGAAAATTAATAGCCTGATTAATGCCACACCGCTGAATAGCTATGGTGATGAAG
GTATTGGTGTTTTTGATGCACTGGGCCTGAACTATAACTTCAAAAGCAATCAAGAACGTCTGCCGAGCCGTACCGATCAGATTTTTGTTTATGGTATTGTGAGC
CCGAATGAACTGCGTAGCGCAAAAAGCAGCGCAGATAGCACCGGTAGCGATACCAAAGTTAATTGGAGCAATACCCAGAGCCGTTATCTGCCGGTTCCGTATA
ATTATAGCGAAGGTATTATTGATGCCGACGGTTTTAAACGTCCGGAAAATCGTGGTGCAAGCGTTACCACCTTTAGCGGTCTGAAAAGCATTGCACCGGATGG
TTTTGCAAATAGCATTGCAAATTTTAGCGTGGGTCTGAAAGCAGGTATTGATCCGAATCCGGTTATGAGCGGTAAAAAGCAAATTATGGTGCCGTTGTTCTGA
CCCGTGGTGGTGTTGTTCGTCTGAATTTTAATCCGGGTAATGATAGCCTGCTGAGCACCACCGATAATAATATTGCACCGATCAGCTTTAGCTTTACCCCGTTTA
CCGCAGCAGAAAGCGCAGTTGATCTGACCACCTTTAAAGAAGTGACCTATAACCAAGAAAGCGGTCTGTGGTCCTATATTTTTGATAGCAGCCTGAAACCGAG
CCATGATGGTAAACAGACACCGGTTACCGACAATATGGGTTTTAGCGTTATTACCGTTAGCCGTACCGGTATTGAACTGAATCAGGATCAGGCAACCACCACC
CTGGATGTTGCACCGAGCGCACTGGCAGTTCAGAGCGGTATTCAGAGCACCACCCAGACCCTGACCGGAGTTCTGCCGCTGAGCGAAGAATTTAGCGCAGTTA
TTGCAAAAGATAGCGATCAGAATAAAATCGACATTTACAAAAATAATAATGGCCTGTTTGAAATTGATACCCAGCTGAGCAATAGCGTTGCAACCAATAATGG
TGGTCTGGCACCGAGCTATACCGAAAATCGTGTTGATGCATGGGGCAAAGTTGAATTTGCAGATAATTCAGTTCTGCAGGCACGTAATCTGGTTGATAAAACC
GTGGATGAGATTATTAACACACCGGAAATCCTGAACAGCTTTTTCGTTTTACCCCTGCATTCGAAGATCAGAAAGCAACCCTGGTGGCAACCAAACAGAGCG
ATACCAGCCTGAGCGTTAGTCCGCGTATTCAGTTTCTGGATGGCAATTTTTATGACCTGAATAGCACCATTGCCGGTGTTCCGCTGAATATTGGTTTTCCGAGCC
GTGTTTTTGCAGGTTTTGCAGCACTGCCTGCATGGGTTATTCCGGTTAGCGTTGGTAGCAGCGTTGGCATTCTGTTTATTCTGCTGGTTCTGGGTCTGGGTATTG
GTATTCCGATGTATCGTGTTCGTAAACTGCAGGATGCCAGCTTTGTGAACGTGTTTAAAAAAGTTGATACCCTGACCACCGCAGTTGGTAGCGTTATAAAAAG
ATTATCACCCAGACCGGTGTGGTTAAAAAAGCACCGAGTGCACTGAAAGCAGCAAATCCGAGCGTTAAAAAACCGGCAGCATTTCTGAAACCGCCTGTTCAGC
CTCCGAGCAAACCGGAAGGTGAACAGAAAGCAGTTGAAGTTAAAAGCGAGGAAACCAAAAGCTAAACCCAGCTTTCTTGTACAAAGTGGTGGCGCGCCCTGG
GCCTCATGGGCCTTCCTTTCACTGCC

Supplementary Table 6. Synthetic clones obtained from Geneart. Start and stop codons of the full-length MG191 and MG192 genes are highlighted.

Primer name	Sequence
COMmg192-F	AGTGGGCCCACTAACAAAAAACAAATTAGTGATGTTGTTAGTGATTGTGTGAAAAAAATTGTTTATAATTAAGTTTGTATGAAAACAATGAGAAAACAG
COMmg192-R	AGTCTCGAGCTAACTTTTGGTTTCTTCTG
R600Amg192-F	CCTGAAAACGCGGGTGCTAGT
R600Amg192-R	ACTAGCACCCGCGTTTTCAGG
Q460AD461Amg192-F	AGTTTTTCAATCGCTGCCACCTTCAGCTTTG
Q460AD461Amg192-R	AAAGCTGAAGGTGGCAGCGATTGAAAAACTCCC
Y830Amg192-F	GGTTAGCACCTAGTGCCACAGAAAACAGGG
Y830Amg192-R	CCCTGTTTTCTGTGGCACTAGGTGCTAACC
R834Gmg192-F	GTTACACAGAAAACGGGGTTGATGCATGGG
R834Gmg192-R	CCCATGCATCAACCCCGTTTTCTGTGTAAC
D836Lmg192-F	CAGAAAACAGGGTTTTGGCATGGGGTAAAG
D836Lmg192-R	CTTTACCCCATGCCAAAACCCTGTTTTCTG
W838Fmg192-F	GTTGATGCATTTGGTAAAGTTG
W838Fmg192-R	CAACTTTACCAAATGCATCAAC
G839Fmg192-F	GATGCATGGTTTAAAGTTGAG
G839Fmg192-R	CTCAACTTTAAACCATGCATC
PacUp	GTAGCTAATCTAACAGTAGG
PacDw	GTCCTAGAACTTGGTGTATG
RTPCRmg192-F	TCCCCTAATGAATTGCGAAG
RTPRCmg192-R	CAGGGGCAATTGATTTAAGC
Tnp3	GATTCATGATTATATCGATCAAC
P110-F	AGGAGATATACCATGGCACTGGCAAATACCTTTC
P110-R	GTGATGGTGATGTTTAGGCAGTGCTGCAAAC
P110N-F	AGGAGATATACCATGGCACTGGCAAATACCTTTC
P110N-R	GTGATGGTGATGTTTTGCCAGACCACCATTATTGG
P140-F	AGGAGATATACCATGGGTGTTATTACCGGTGTTG
P140-R	GTGATGGTGATGTTTATCTGCCCACTGATTAAACG

Supplementary Table 7. Primers used in this study

References

1. Jones, D.T. Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol* **292**, 195-202 (1999).
2. Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic acids research* **42**, W320-W324 (2014).
3. Aparicio, D. et al. Mycoplasma genitalium adhesin P110 binds sialic-acid human receptors. *Nature Communications* **9**, 4471 (2018).
4. Pettersen, E.F. et al. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-12 (2004).
5. Davis, I.W., Murray, L.W., Richardson, J.S. & Richardson, D.C. MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Res* **32**, W615-9 (2004).

Capítulo 3

Proteínas inmunodominantes P1 y P40/P90 del patógeno humano *Mycoplasma pneumoniae*

David Vizarraga [1], Akihiro Kawamoto [2,3], U. Matsumoto [4], Ramiro Illanes [1], Rosa Pérez-Luque [1], Jesús Martín [1], Rocco Mazzolini [5], Paula Bierge [6], Oscar Q. Pich [6,7], Mateu Espasa [8], Isabel Sanfeliu [8], Juliana Esperalba [9], Miguel Fernández-Huerta [9],
Margot P. Scheffer [10], Jaume Pinyol [7,] Achilleas S. Frangakis [10], Maria Lluch-Senar [5], Shigetarou Mori [11], Keigo Shibayama [11], Tsuyoshi Kenri [11], Takayuki Kato [2,3] Keiichi Namba [2,12,13], Ignacio Fita [1], Makoto Miyata [4,14] y David Aparicio [1].

Resumen

Mycoplasma pneumoniae es una bacteria patógena humana que causa neumonía atípica primaria. La motilidad e infectividad de *M. pneumoniae* están mediadas por las proteínas inmunodominantes P1 y P40/P90, que forman un complejo transmembrana de adhesión. Aquí presentamos la estructura de P1, determinada por cristalografía de rayos X y crio-microscopía electrónica (crio-EM), y la estructura de P40/P90 por rayos X. Al contrario de lo que se había sugerido, el sitio de unión para el ácido siálico se encontró en P40/P90 y no en P1. La variabilidad genética y clínica se concentra en las superficies del dominio N de P1 y P40/P90. Los anticuerpos policlonales generados contra el dominio C (mayormente conservado de P1) inhibieron la adhesión de *M. pneumoniae* y los ensayos de serología con sueros de pacientes infectados fueron positivos cuando se probaron contra este dominio C. P40/P90 también mostró una fuerte reactividad contra sueros infectados de humanos. Los elementos arquitectónicos determinados para P1 y P40/P90 abren nuevas posibilidades en el desarrollo de vacunas contra las infecciones por *M. pneumoniae*.

nature COMMUNICATIONS

ARTICLE

https://doi.org/10.1038/s41467-020-18777-y OPEN

Immunodominant proteins P1 and P40/P90 from human pathogen *Mycoplasma pneumoniae*

David Vizarraga[1], Akihiro Kawamoto[2,3], U. Matsumoto[4], Ramiro Illanes[1], Rosa Pérez-Luque[1], Jesús Martín[1], Rocco Mazzolini[5], Paula Bierge[6], Oscar Q. Pich[6,7], Mateu Espasa[8], Isabel Sanfeliu[8], Juliana Esperalba[9], Miguel Fernández-Huerta[9], Margot P. Scheffer[10], Jaume Pinyol[7], Achilleas S. Frangakis[10], Maria Lluch-Senar[5], Shigetarou Mori[11], Keigo Shibayama[11], Tsuyoshi Kenri[11], Takayuki Kato[2,3], Keiichi Namba[2,12,13], Ignacio Fita[1], Makoto Miyata[4,14]✉ & David Aparicio[1]✉

Mycoplasma pneumoniae is a bacterial human pathogen that causes primary atypical pneumonia. *M. pneumoniae* motility and infectivity are mediated by the immunodominant proteins P1 and P40/P90, which form a transmembrane adhesion complex. Here we report the structure of P1, determined by X-ray crystallography and cryo-electron microscopy, and the X-ray structure of P40/P90. Contrary to what had been suggested, the binding site for sialic acid was found in P40/P90 and not in P1. Genetic and clinical variability concentrates on the N-terminal domain surfaces of P1 and P40/P90. Polyclonal antibodies generated against the mostly conserved C-terminal domain of P1 inhibited adhesion of *M. pneumoniae*, and serology assays with sera from infected patients were positive when tested against this C-terminal domain. P40/P90 also showed strong reactivity against human infected sera. The architectural elements determined for P1 and P40/P90 open new possibilities in vaccine development against *M. pneumoniae* infections.

[1] Instituto de Biología Molecular de Barcelona (IBMB-CSIC), Parc Científic de Barcelona, Baldiri Reixac 10, 08028 Barcelona, Spain. [2] Graduate School of Frontier Biosciences, Osaka University, 1-3 Yamadaoka, Suita, Osaka 565-0871, Japan. [3] Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565-0871, Japan. [4] Graduate School of Science, Osaka City University, Osaka 558-8585, Japan. [5] EMBL/CRG Systems Biology Research Unit, Centre for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology, Dr Aiguader 88, 08003 Barcelona, Spain. [6] Laboratori de Recerca en Microbiologia i Malalties Infeccioses, Institut d'Investigació i Innovació Parc Taulí (I3PT), Hospital Universitari Parc Taulí, Universitat Autònoma de Barcelona, 08208 Sabadell, Spain. [7] Departament de Bioquímica i Biologia Molecular, Institut de Biotecnologia i Biomedicina, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain. [8] Departament de Microbiologia, Hospital Universitari Parc Taulí, Universitat Autònoma de Barcelona, 08208 Sabadell, Spain. [9] Departament de Microbiologia, Hospital Universitari Vall d´Hebron, Universitat Autònoma de Barcelona, 08035 Barcelona, Spain. [10] Buchmann Institute for Molecular Life Sciences, Max-von-Laue Str. 15, 60438 Frankfurt, Germany. [11] Department of Bacteriology II, National Institute of Infectious Diseases, Tokyo, Japan. [12] RIKEN Center for Biosystems Dynamics Research and SPring-8 Center, 1-3 Yamadaoka, Suita, Osaka 565-0871, Japan. [13] JEOL YOKOGUSHI Research Alliance Laboratories, Osaka University, 1-3 Yamadaoka, Suita, Osaka 565-0871, Japan. [14] The OCU Advanced Research Institute for Natural Science and Technology (OCARINA), Osaka City University, Osaka 558-8585, Japan. ✉email: miyata@sci.osaka-cu.ac.jp; daacri@ibmb.csic.es

Mycoplasma pneumoniae is a human pathogen responsible for upper and lower respiratory tract infections[1]. It is estimated that this bacterium is responsible for up to 40% of community-acquired pneumonias in persons of all ages[2]. In addition to being a severe respiratory pathogen, *M. pneumoniae* may induce clinically significant manifestations in extrapulmonary sites and/or immunologic effects in as many as 25% of the infections[3]. Unlike for other important respiratory pathogens, such as *Streptococcus pneumoniae* and *Haemophilus influenzae*, a vaccine for *M. pneumoniae* is not yet available despite the considerable efforts[4].

M. pneumoniae binds to host target cells by means of a polar structure known as the attachment organelle, and glides in the direction of this differentiated structure at a speed of ~1 µm/s (Fig. 1a)[5]. Gliding motility in *M. pneumoniae* and related species such as *Mycoplasma genitalium* is caused by a unique mechanism completely different from the motility of *Mycoplasma mobile*, which has been studied in further detail[6–8]. Adherence of *M.*

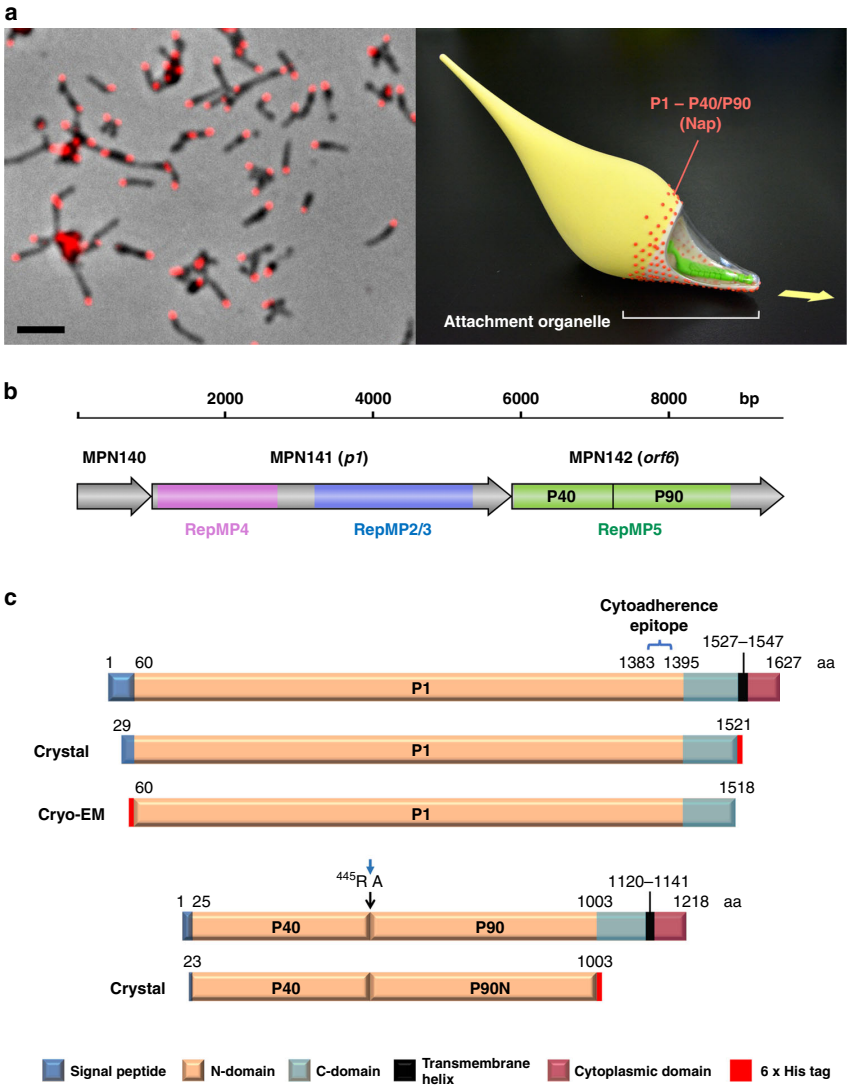


Fig. 1 The Nap complex in *M. pneumoniae*. a A phase-contrast immunofluorescence microscopy image of *M. pneumoniae* cells (left). Locations of the Nap (surface of the attachment organelle) were visualized by using anti-P1 monoclonal antibody and Alexa Fluor® 546 secondary antibody (red). Bar = 2 µm. Three-dimensionally printed model of *M. pneumoniae* cell (right). The positions of the attachment organelle and the Nap structure are indicated. The cell surface is partly modeled as a transparent canopy to show the internal core structure of the organelle. The arrow indicates gliding direction. **b** Scheme of the *p1* operon that contains three genes: MPN140, MPN141(*p1*), and MPN142(*orf6*). The locations of repetitive elements (RepMPs) in MPN141 and MPN142 are shown by colors. The regions corresponding to P40 and P90 are also shown in MPN142. **c** Scheme of native P1 and P40/P90 proteins and the recombinant constructs used for crystallization and cryo-EM analysis. The major cytoadherence epitope of P1 reported previously is indicated (1383–1395). The boundary between P40 and P90N proteins identified in this study (445 R/A) is indicated by arrows (Supplementary Fig. 5). Positions of 6 × His-tag used for purification of recombinants are shown by red boxes.

pneumoniae to cells of the respiratory tract is mediated by a network of adhesins and cytoadherence accessory proteins[9,10]. Within this network the 170 kDa protein P1 was identified as a major determinant for cytoadherence and gliding motility with antibodies against P1 preventing both adhesion and motility[9–12]. For about 40 years, P1 has been assumed to be responsible for binding to sialic acid oligosaccharide receptors from the host cells[13]. Because all these relevant properties, P1 has been attracting attention since the late 1970s, although it was soon recognized that accessory proteins were also required for its functioning[14]. P1, together with the P40/P90 polypeptides, forms a transmembrane complex called the "Nap"[15,16]. The structure of the Nap complex of *M. genitalium*, a human urogenital pathogen closely related to *M. pneumoniae*, was recently revealed by single-particle cryo-electron microscopy (cryo-EM) at 9 Å and cryo-electron tomography (cryo-ET) at ~15 Å resolution[17]. The *M. genitalium* Nap structure consists of a dimer of a P140–P110 complex protruding outward from the mycoplasma membrane and forming a large knob with a diameter of ~15 nm. P140 and P110 are the *M. genitalium* orthologues of P1 and P40/P90, respectively. Therefore, it was most unexpected to find that in *M. genitalium* the binding site for sialic acid oligosaccharides is in P110 and not in P140[17,18]. P1 is considered to be one of the most immunodominant proteins in *M. pneumoniae* cells, playing a major role in the immune response of infected patients and accordingly in diagnosis and epidemiologic studies[3,19,20]. The *M. pneumoniae* genome contains repeated regions, denominated RepMPs. The majority (75%) of RepMPs has homology with MPN141 (P1) and MPN142 (P40/P90) (Fig. 1b). Homologous recombination between RepMPs and either MPN141 or MPN142 generate variability within antigenic regions of P1 and P40/P90, respectively, providing an essential strategy to evade the immune host system[21–23]. P40/P90 consists of two polypeptides from the proteolytic cleavage of a 130 kDa translate encoded by MPN142[24] (Fig. 1b). Historically little attention has been dedicated to P40/P90 and the contribution of P40/P90 to cytoadherence is much less studied than for P1. Even the presence of the P40 polypeptide in the Nap has been questioned and the reason for and mechanism of the MPN142 cleavage remain unknown[25].

In the present work, we report the structures of *M. pneumoniae* proteins P1 and P40/P90 and of complexes of P40/P90 bound to sialylated oligosaccharides. The results elucidate the antigenic organization of these immunodominant proteins allowing the mapping of the genetic and clinical variability. Moreover, serological tests on different fragments of P1 reveal the recognition preference of IgG antibodies for conserved over variables regions, suggesting an immune system strategy of selecting these regions for long-term immunity. The information is considered pivotal for clarifying the motility, infectivity, and epidemics of *M. pneumoniae*.

Results

Structure determination of P1. Crystals were obtained of the *M. pneumoniae* Nap protein P1 extracellular region (residues Thr29-Asp1521). This P1 construct, empirically chosen after many crystallization attempts, excludes the predicted N-end of the signal secretion peptide (spanning residue Met1 to Leu59 according to the structural properties prediction program Psipred[26]). In this construct are also excluded the (predicted) transmembrane and intracellular regions at the C-end of P1 (residues Tyr1522 to Ala1627) (Fig. 1c). These P1 crystals were solved simultaneously with the determination of the crystal structure from the *M. genitalium* orthologous protein P140, by averaging between crystals from both proteins (with 41% sequence identity) despite neither molecular models nor experimental phases were available (see "Methods") (Supplementary Fig. 1). The final refined model of P1, at 1.9 Å resolution, has agreement R and R_{free} factors of 18.7 and 22.9, respectively (Fig. 2a, b and Supplementary Table 1).

The structure of the ectodomain from P1 was also studied by single-particle analysis using cryo-EM (Figs. 1c and 2b and Supplementary Fig. 2). Although P1 particles were clearly observed in the cryo-EM micrographs, the quality of the map was too poor to refine the atomic model owing to severe orientation bias. To alleviate this problem, we collected data by using grids covered with a thin film of graphene oxide (GO), which drastically improved the data quality[27] (Supplementary Fig. 2a). A total of 68,014 selected particles yielded a 3D EM map at an overall resolution of ~2.9 Å, according to the gold-standard Fourier shell correlation (FSC) 0.143 criterion (see "Methods") (Supplementary Table 2).

The structure of P1 consists of a large N-terminal domain, with Asn60 as the first residue visible in the X-ray and the cryo-EM maps, and a smaller C-terminal domain (residues Gly1395–Asp1521) (Fig. 2). The N-terminal domain of P1 has a β-sheet propeller topology of seven consecutive blades or β-sheets, each with four antiparallel strands, except for β-sheet VII that contains only two strands. Interestingly, these two strands correspond exactly to a peptide with important immune properties, which had been defined as the cytoadherence epitope[28] (Fig. 2c). The lengths of connections between the β-sheets of the propeller are very diverse, ranging from just seven residues between sheets VI and VII, to 315 residues between sheets IV and V. Links between adjacent β-strands in the β-sheets of the propeller, also vary widely in length, ranging from just 11 to 219 residues. The longest intra and inter β-sheet connections cluster together creating a crown-like structure on one side of the β-propeller opposite to the position where the C-terminal domain is located. In the crown there are 13 disordered loops, spanning 117 residues, which could not be built (Fig. 2c). These disordered regions, essentially the same in the X-ray and cryo-EM structures, are spread over the N-terminal domain surface (Supplementary Figs. 1 and 3a). Therefore, the N-terminal domain of P1 can be described as a structured core surrounded by a fuzzy surface. The C-terminal domain contains only one disordered loop (residues Thr1483–Asn1495) and has an elongated fold with the same unique topology (according to DALI[29]) as the C-terminal domain of adhesin P110 from *M. genitalium*.

The cryo-EM and X-ray maps of P1 are well defined for most of the N-terminal domain with the residues clearly identifiable, in general, in both maps (Fig. 2b). However, from residue Pro1388 to the C-end, a region that includes the hinge between domains and the whole C-terminal domain, the density in the cryo-EM map becomes weak, indicating that there are hinge-like movements between domains. In the crystal structure, the C-terminal domain is anchored by crystal contacts and the density is clear and crisp, while in the top part of the crown, with few crystal contacts, temperature factors are high (Supplementary Fig. 3b) and chain tracing was difficult. Together the cryo-EM and X-ray maps provide accurate structural information about the whole P1 ectodomain, proving the mobility of domains with respect to one another and confirming the characteristic presence of a large number of disordered loops.

Crystal structure of P40/P90N. Diffracting crystals could not be obtained with constructs of the complete extracellular region of P40/P90, which is predicted by Psipred[26] to span from Ala25 to Pro1113 (with Met1 to Leu24 corresponding to a signal N-peptide) (Fig. 1c). After numerous attempts, high-resolution diffracting crystals were grown from a construct, referred henceforth as P40/P90N, spanning residues Ser23-Val1003 that, by sequence alignment, corresponds to the N-terminal domain of the *M.*

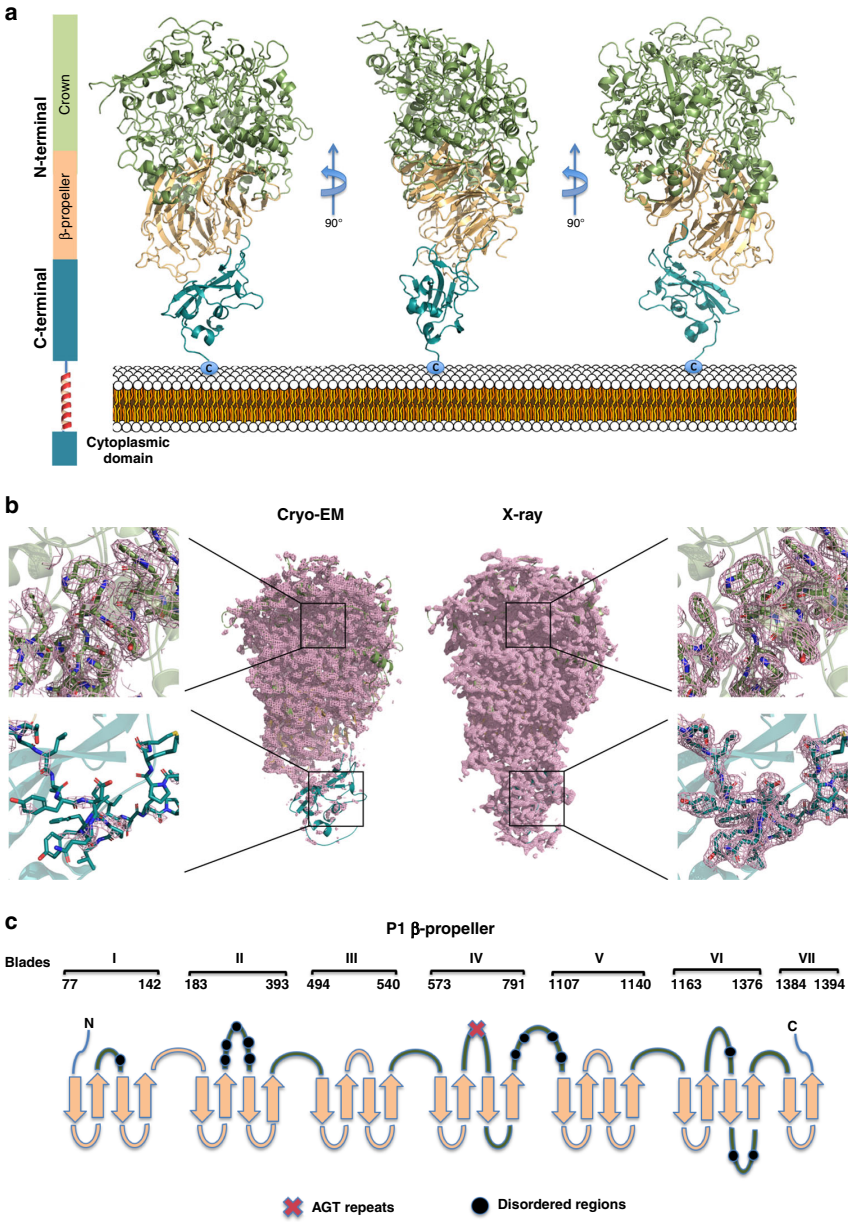


Fig. 2 X-ray and cryo-EM structures of P1. a Ribbon representation, with three 90° apart views, of the complete extracellular region of P1. The large N-terminal domain is organized around a seven-blade β-propeller (light orange) from where the crown (green) emerges. The C-terminal domain (cyan) is immediately followed by a predicted transmembrane helix, which requires the proximity of this domain to the cell membrane. **b** P1 map densities from cryo-EM (left panels) and X-ray crystallography (right panels). Insets show representative regions from the N-terminal domain (top) and the C-terminal domain (bottom). The quality of the cryo-EM density provides the clear identification of residues in the N-terminal domain (at an estimated resolution of 2.9 Å). On the contrary, the C-terminal domain is poorly defined in the cryo-EM map, indicating high flexibility with respect to the N-terminal domain. The core of the N-terminal domain and, especially, the C-terminal domain are well defined in the X-ray map, while the very top of the N-terminal domain presents high-temperature factors and was difficult to trace. **c** Topology of the β-propeller in P1. Loops contributing to the crown are colored green. Disordered loops are indicated with black dots and the stretch of serines (the AGT repeats) with a red cross. Strands and loops are not in scale.

genitalium orthologous protein P110 (Supplementary Fig. 4). The P40/P90N crystals, with two subunits in the asymmetric unit, were solved by molecular replacement using as a searching model the structure of P110 (PDB entry 6R3T), with a sequence identity between the extracellular regions of 44%. The refined P40/P90N structure, at 2.65 Å resolution, has agreement R and R_{free} factors of 21.4 and 23.4, respectively (Supplementary Table 1).

The structure of P40/P90N reveals a large globular domain that is structurally closely related to the N-terminal domain of P110, with a RMSD of 1.2 Å for 742 equivalent residues[18] (Fig. 3a and

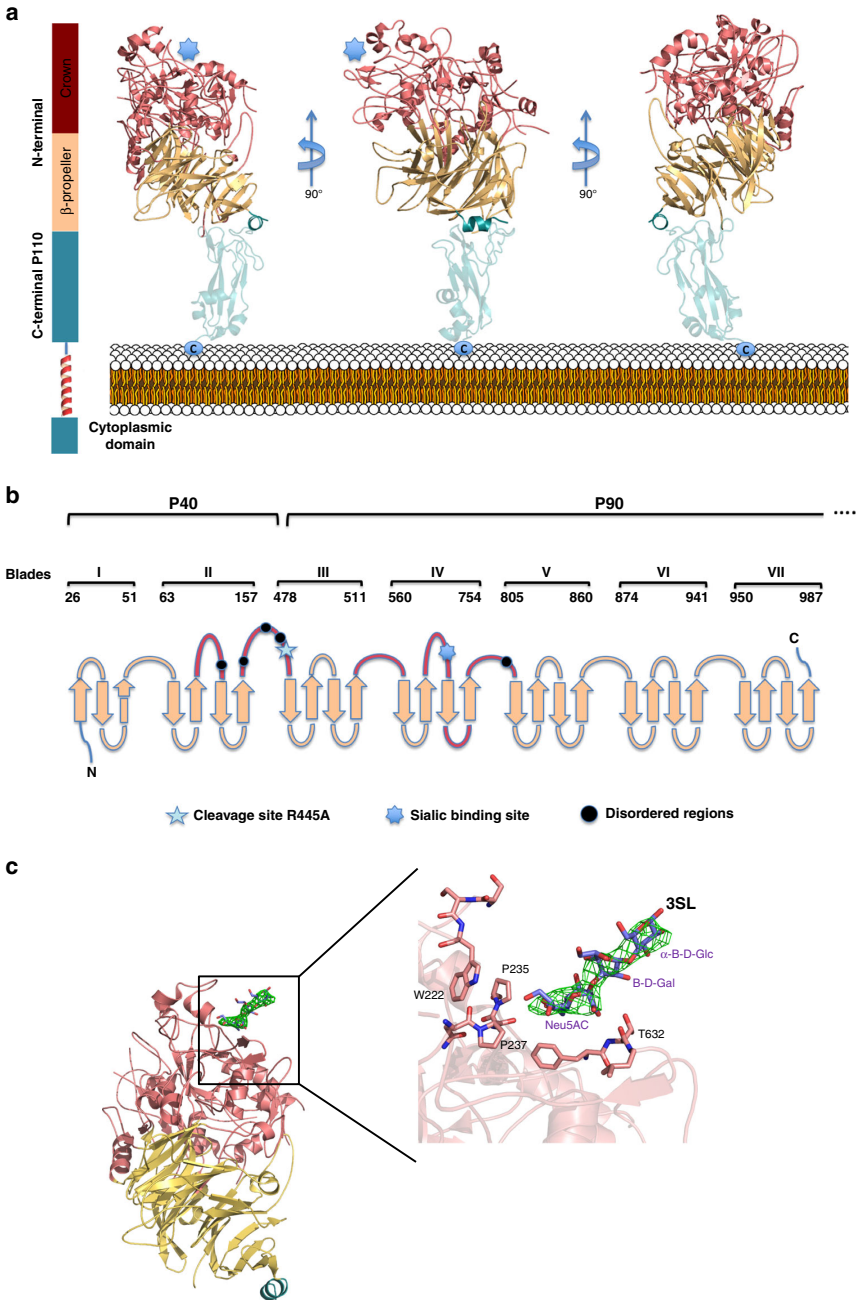


Fig. 3 X-ray structure of P40/P90 and of its complex with oligosaccharide 3SL. a Ribbon representation, with three 90° apart views, of the crystal structure from the P40/P90 N-terminal domain that is organized, similarly to P1, around a seven-blade β-propeller (light orange) from where it emerges the crown (pink). The sialic binding site found in P40/P90 is explicitly indicated. The predicted C-terminal domain of P40/P90 is also shown (cyan), modeled from the C-terminal domain of the orthologous adhesin P110 from *M. genitalium*. **b** Topology of the β-propeller from P40/P90. Loops contributing to the crown are colored pink. The first blade β-sheet contains only three strands in P40/P90 placing the N-end at the propeller face opposite to the one in P1. Localizations of the disordered loops and sialic binding site are indicated with black dots (black) and a star (blue), respectively. The position of the cleavage site, which results in the corresponding P40 and P90 separate polypeptides, is depicted with a pentagram (pale blue). Strands and loops are not in scale. **c** Binding of 3SL to the N-terminal domain of P40/P90. Ribbon representation of the P40/P90N structure in complex with oligosaccharide 3SL. The inset shows a detail of the residues shaping the binding pocket. The electron density omit map (at 0.9 sigma) corresponding to 3SL is also shown.

Supplementary Fig. 4). The sequence of P40/P90N presents two long insertions with respect to P110 of more than 60 residues each (Gly297–Thr368 and Leu395–Thr462 that are referred as S1 and S2, respectively) (Fig. 3b and Supplementary Fig. 4). Both insertions, completely disordered in the crystal structures determined, are found between β-propeller blades II and III. P40/P90, like P1, has a seven-blade β-propeller topology with many of the connections between the β-strands clustering together and forming the crown (Fig. 3b).

P40 and P90 had often been considered as two independent proteins resulting from the posttranslational processing of the MPN142 gene product[25,30]. In our hands, the purified P40/P90 samples present a characteristic SDS-PAGE gel pattern indicating a slow progression of cleavage in a few specific locations (Supplementary Fig. 5a). The primary scission occurs between Arg445 and Ala446, located toward the C-end of insertion S2 (Fig. 3b and Supplementary Figs. 4 and 5a). The two polypeptides produced by this primary cleavage, the shorter spanning from the N-end to Arg445 and the longer from Ala446 to the C-end, can correspond to the subunits P40 and P90 that are found in mycoplasma cells[30]. When the primary scission site was mutated to avoid cleavage, the P40/P90 variant protein (Ser445–Ser446) presented, unexpectedly, a similar proteolysis pattern with cleavage taking place between Arg455 and Ala456, still in the S2 insertion (Supplementary Fig. 5a). Studies with mycoplasma cells had reported two cleavage sites occurring between Arg445–Ala446[30] and (probably) Leu454–Arg455[25,31], in good agreement with the two cleavage Arg–Ala sites (RA motifs) now found in vitro. Regardless of the extend of proteolysis, all P40/P90N samples behave as a single and sharp peak by SEC–MALS (Supplementary Fig. 5b) indicating that cleavage takes place only when folding is completed and that a unique three-dimensional structure of P40/P90N is maintained. Structural stability of the cleaved protein probably relies on the large interface between the P40 and P90 polypeptides, with 75 hydrogen bonds, 6 salt bridges, and a buried surface area of 10,480 Å^2, which gives an estimated free energy of −55 kcal with an expected 100% probability of formation according to the Complexation Significance Score algorithm in PISA[32] (Supplementary Fig. 6a). Addition of EDTA to the P40/P90 samples halts proteolysis completely, which suggests a key role for divalent metals. No metals were found in the crystal structures of P40/P90, but EDTA was required for the crystallization conditions. The presence of significant amounts of Zn^{2+} ions in a purified P40/P90 sample was detected by inductive coupled plasma–mass spectrometry (ICP–MS), although the concentration of protein was higher than the one found for the ion (see "Methods"). The possible presence in the P40/P90 samples of an undetected protease, which could be coming from the over-expression in *Escherichia coli*, was analyzed by SEC–MALS. Different constructs of P40/P90N, with and without the C-terminal domain, always produced a sharp peak at the molecular weight corresponding to the construct, without any indication of other species. Altogether these results suggest that cleavage of P40/P90 is a self-proteolytic process. However, we have not yet been able to find the residues that could be responsible for this enzymatic activity. Besides insertions S1 and S2, the structure of P40/P90 presents only three short disordered loops (residues Ser118–Gly127, Ser168–Gly175, and Ala774–Arg777) and the N-terminal domain from P40/P90 can be seen as a well-defined globular structure with two long floppy pendants (Supplementary Fig. 3a). The complete extracellular region of P40/P90 was generated by combining the structure of the N-terminal domain with the structure of the Pro1004–Pro1113 region of P40/P90 modeled according to the P110 C-terminal domain because the sequence identity is high (68%) and the expected RMSD low (~0.6 Å)[33] (Fig. 3a and Supplementary Fig. 4). In solution, P1 and P40/P90 form an heterodimer showing a single peak in a gel filtration profile (Supplementary Fig. 5c).

Complexes of P40/P90N with sialylated oligosaccharides. The presence of a sialic binding site in P40/P90 was confirmed by the crystal structures of complexes of P40/P90N with oligosaccharides 3SL (neuraminic acid forming an α2–3 linkage to a lactose monosaccharide) and 6SL (with an α2–6 linkage) (Supplementary Table 1). This result is consistent with the facts that sialylated oligosaccharides act as cell receptors for *M. pneumoniae* and *M. genitalium*[34–36], and that adhesin P110 from *M. genitalium*, the orthologue of P40/P90 contains a binding site specific for neuraminic acid[18]. The sialic binding site of P40/P90 is in a location structurally equivalent to the one in P110, in the upper part of the crown (Fig. 3a, c and Supplementary Fig. 7). Only the neuraminic acid moiety interacts directly with the P40/P90 β-hairpin that contains the X-Tyr/Phe–Ser/Thr motif (Leu630–Phe631–Thr632) characteristic for the binding to sialic compounds[18,34]. P40 functions as the external rim of the sialic binding site pocket, which would become exposed without P40 (Supplementary Fig. 6b). Surface plasmon resonance (SPR) indicates that sialylated oligosaccharides bind to P40/P90 with dissociation constants (K_d) in the micromolar range for 3SL (~8.0) and 6SL (~2.9), similar to the values reported for P110[18] (Supplementary Fig. 8) although, according to in vivo studies, 6SL binds at a much lower affinity than 3SL[37,38]. For P1, SPR shows no binding to sialylated oligosaccharides, in agreement with what is observed for *M. genitalium*[17].

Structural relationships of P1 and P40/P90. The structural results reveal that P1 (1627 residues) and P40/P90 (1218 residues) present a similar overall organization and topology in the extracellular regions. The similarity is maintained between the transmembrane helices and the cytoplasmic regions where the sequence identity of P1 and P40/P90 is high (43%). Despite the different size, the close relationship of both proteins suggests that P1 and P40/P90 have a recent common phylogenetic ancestor. Superposition of N-terminal domains from P1 and P40/P90, with a sequence identity of only 12%, gives a RMSD of 3.5 Å for 361 equivalent residues, which correspond mainly to residues from the β-sheets of the propellers. However, β-sheet I has only three strands in P40/P90, but four in P1 and this "extra" first strand in the first β-sheet of P1 places the N-end in the propeller face that is the opposite to the one in P40/P90 (Fig. 4). The "extra" first strand of β-sheet I from P1 starts with a large β-bulge (residues 69–77) occluding the center of the propeller, contrary to P40/P90 where the center is empty. Linkage between the N- and C-terminal domains of P1 changes with respect to P40/P90, due to the absence of the last two strands in β-sheet VII. Therefore, β-sheets I and VII, structurally contiguous within the propeller, present the largest differences between P1 and P40/P90. The hypothesis of a recent common origin is further supported by the close structural relationships between the C-terminal domains from the non-orthologous proteins P1 and P110, with a sequence identity of 19% and a RMSD of 2.0 Å for 73 equivalent residues (Supplementary Fig. 9a), The hypothesis is also reinforced by the presence of the characteristic "AGT repeats" motif in the same loop of the propellers, between the second and third β-strands of the fourth blade, of P1 and P110 (Supplementary Fig. 9b).

Mapping of epitopes and genetic variability onto P1 and P40/P90. Epitopes reported for six monoclonal antibodies (mAbs) against P1 with inhibiting cytoadherence activities for *M. pneumoniae* (aN, P1.62, P1.26, M51, M58, and 6E7)[28,39] (Supplementary Fig. 10) are all exposed on the surface of P1, as expected

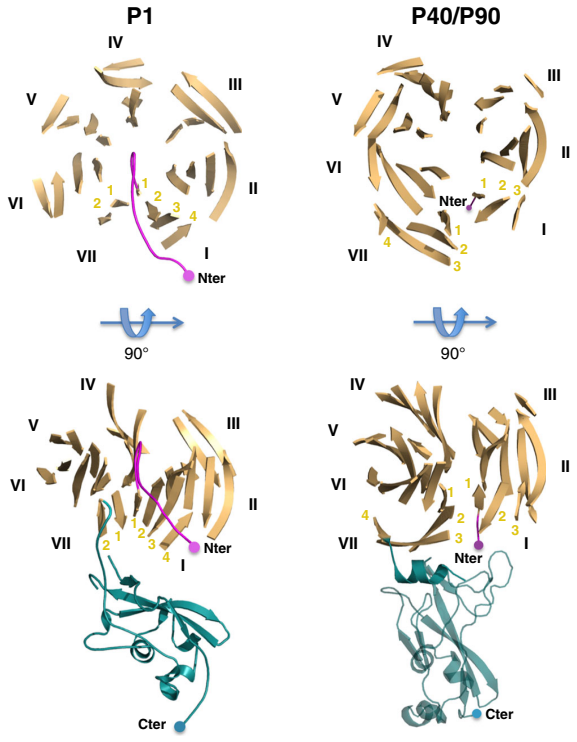


Fig. 4 Comparison of P1 and P40/P90. Comparison of the β-propeller structures and of their connections with the corresponding C-terminal domains in P1 (left panels) and P40/P90 (right panels). The C-terminal domain is tangential to the propeller ring in P1 and protrudes radially in P40/P90, where the C-terminal domain was modeled from the orthologous adhesin P110 from *M. genitalium*. The N-ends are situated in opposite faces of the propeller in P1 and P40/P90. The center of the propeller is occluded in P1 but empty in P40/P90.

to be accessible by the antibodies (Fig. 5a). Epitopes from mAbs aN, P1.62, P1.26, and 6E7 are located in the N-terminal domain, whereas epitopes from mAbs M51 and M58 are in the C-terminal domain. The epitope from mAb 6E7 corresponds to the strands from the last β-sheet of the β-propeller, close to the C-terminal domain. The mAbs, P1.62 and P1.26 are reported to bind to dual epitopes (Supplementary Fig. 10), separated on the structure of P1 (Fig. 5a). Although dual epitopes of P1.26 are both exposed in the structure of P1, one of them positioned on Asn980–Lys987 might become hidden in the Nap complex.

Two other reported epitopes (Trp810–Tyr817 and Phe1124–Arg1131), recognized by sera from many of the *M. pneumoniae* infected patients[20], are located in the N-terminal domain (Fig. 5a). The first of these exist inside the P1 molecule, while the second is exposed in the P1 surface, although its accessibility might also be limited in the Nap complex. These two epitopes might be immunodominant for antibody production; however, antibodies against these epitopes may not have cytoadherence-inhibitory activities.

Sequences of P1 and P40/90 exhibit sequence variations depending on the *M. pneumoniae* strain. The two major variation types, known as 1 and 2, are harbored in strains M129 and FH, respectively[22,31,40]. For P1 the main source of variability is RepMP4 and RepMP2/3, while for P40/P90 is RepMP5[41]. It has been reported that mAb P1.62 only binds to type 2 P1 protein[42], consistent with the fact that the epitope of this mAb is located at a variable site. The structures of P1 and P40/P90 served as a topographic map to determine positions of known variable regions. RepMPs from P1 and P40/P90 are exclusively located at the N-terminal domains of these proteins, with the C-terminal domains remaining genetically conserved (Fig. 5b, c).

Polyclonal antibodies against constructs of the C-terminal domain. Adherence-inhibiting antibodies recognizing the constant C-terminal domain of P1 could be more effective in eliciting immunoprotection than antibodies to variable regions within the N-terminal domain. Previous studies had shown that gliding and binding of *M. pneumoniae* were affected by addition of a monoclonal antibody generated against a large recombinant fragment of P1 (residues Ala1160–Gln1518) that contains the C-terminal domain[12]. To consolidate these results and to better delineate the binding site of antibodies during an immune response, polyclonal antibodies were generated against two constructs of P1 spanning residues Lys1376–Asp1521 and Ala1400–Asp1521 (see "Methods"). The shortest construct corresponds accurately to the C-terminal domain, while the longest construct also includes the hinge between domains and the two strands from blade VII (Fig. 6a). Quantitative PCR (qPCR) was used to measure the amount of genomic DNA (gDNA) from the *M. pneumoniae* cells in solution, the fraction not adhered to the glass (see "Methods"). The *M. pneumoniae* cells had been previously incubated with different percentages of sera from rats challenged with the P1 constructs. The analysis determines how much the *M. pneumoniae* cells are affected by the presence of the polyclonal antibodies. Without exposure to sera, 81% of the *M. pneumoniae* cells attached to the plastic surface upon incubation of these (scraped) cells for 2 h in 24-well plates. Then, to evaluate the possible toxicity by other components of the sera, such as the complement that enhances the ability of antibodies and phagocytic cells to clear microbes, different percentages (1, 3, 6, and 10%) of serum obtained from pre-immune bleed (PPI) of challenged animals were tested. The total number of recovered mycoplasma after incubation only decreased with the highest concentrations (6 and 10%) of PPI when compared with samples not treated with serum (Supplementary Fig. 11a). Thus the low concentrations (1 and 3%) were selected to test the inhibitory effects on the adhesion capability of *M. pneumoniae* cells (Fig. 6b). Toxicity effects can be subtracted, at least in part, by normalization with the corresponding values of PPI serum and the measures for concentrations of 10% are also reported (Supplementary Fig. 11b). As a final checking, the specificity of antibodies for the C-terminal constructs was also assessed. Serum from challenged rats presented no significant differences with their PPI when incubated with the corresponding constructs before the quantitative adhesion assay.

Incubation of serum from rats challenged with the longest C-terminal domain construct resulted in the number of the *M. pneumoniae* cells attached being decreased significantly with reduction ratios with respect to PPI samples of 0.31 and 0.56 for added serum concentrations of 1% and 3%, respectively (p values of 9×10^{-6} and 1×10^{-6} by applying t-test one-tailed distribution; two samples equal variance). Reduction ratios for the short construct were 0.22 and 0.33, for added serum concentrations of 1% and 3%, respectively (p values of 0.016 and 3×10^{-4}). Together these results reveal that *M. pneumoniae* adhesion capabilities are significantly reduced by polyclonal antibodies generated against C-terminal domain constructs of P1, with inhibition being highest for the longest construct that includes the cytoadherence peptide.

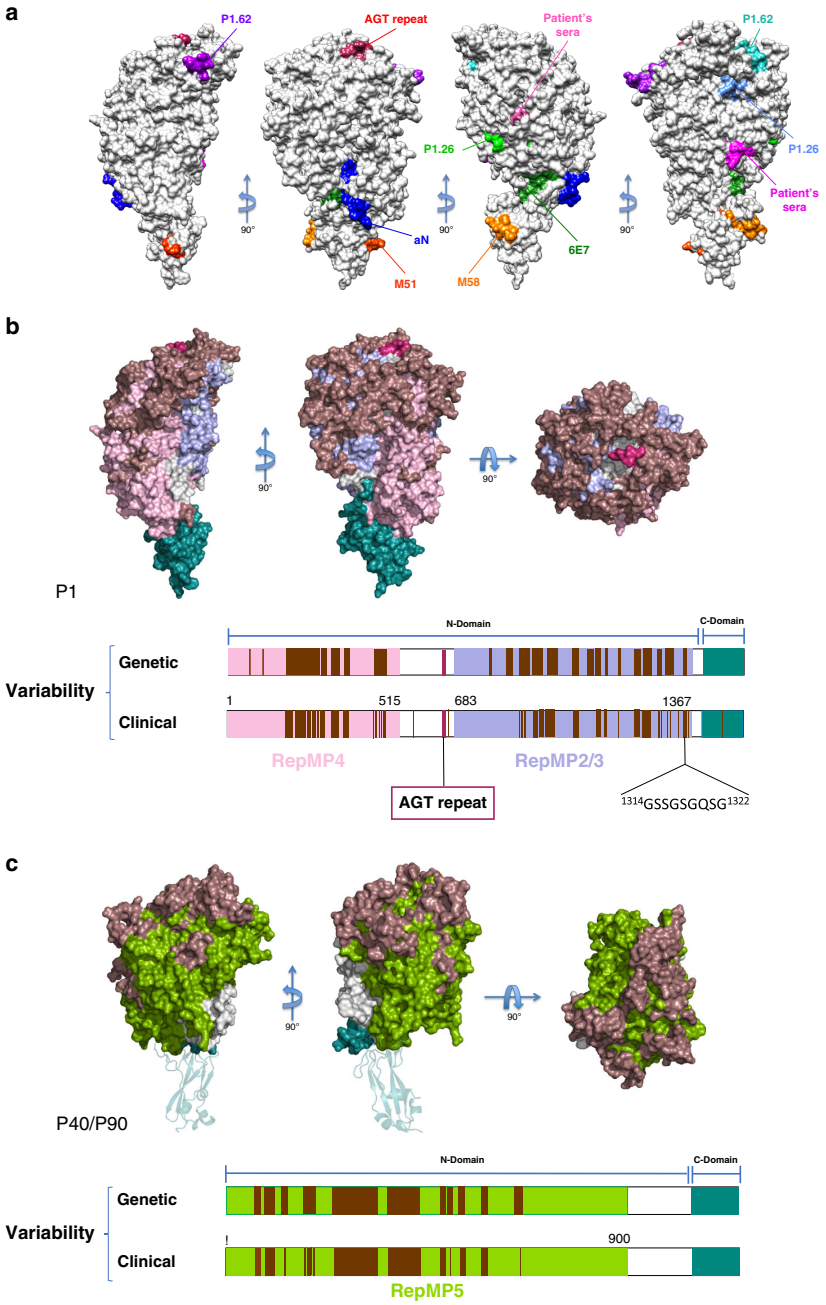


Fig. 5 Mapping of P1 epitopes. Variability in P1 and P40/P90. a Mapping into the P1 structure of all known epitopes (Supplementary Fig. 10). Most epitopes are located on the protein surface, although accessibility by antibodies appears to be difficult for a few of them. **b** Three views of the P1 surface, 90° apart from each other, with the genetically variable regions (brown), RepMP4 (light pink), RepMP2/3 (light blue), and the C-terminal domain (cyan). **c** The corresponding three views of the P40/P90 surface, with the genetically variable regions (brown), RepMP5 (green), and the C-terminal domain (cyan). The C-terminal domain, modeled from the structure of P110, is depicted as ribbons. Sites of genetic and clinical variability are also represented for P1 and P40/P90 in the lower part of the corresponding panels. Clinical variability refers to the variation exhibit by clinical isolates characterized so far. In turn, "genetic variability" indicates regions that are potentially variable by DNA recombination between RepMP elements in *M. pneumoniae* genome. Presently, genetic variability regions are wider than clinical variability.

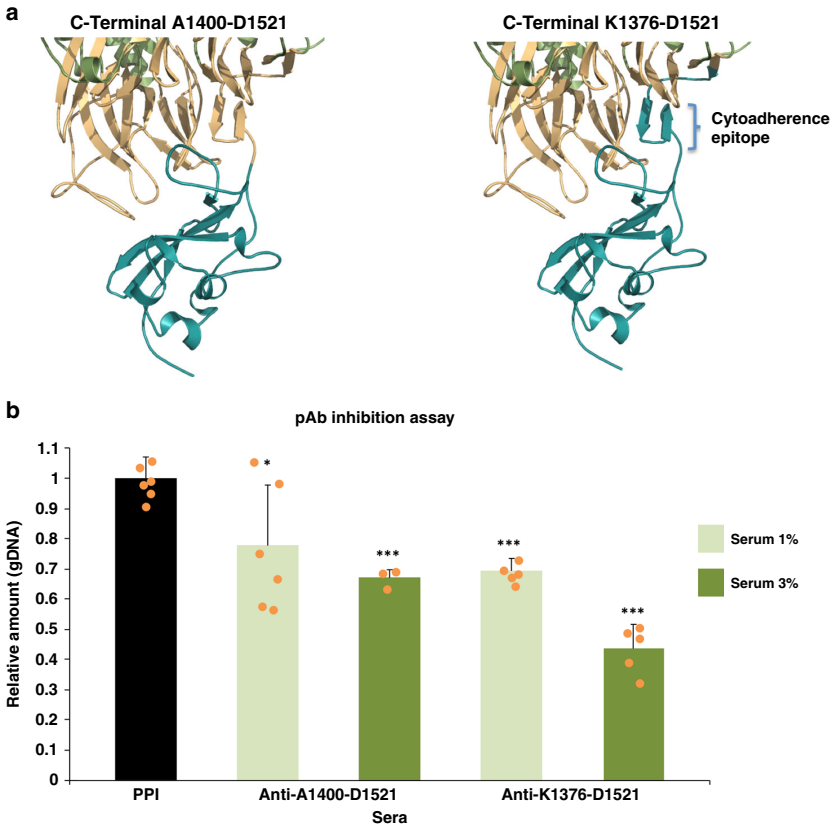


Fig. 6 Inhibition adhesion assays of polyclonal antibodies generated against constructs of the P1 C-terminal domain. a Residues corresponding to the P1 constructs used as antigens to generate polyclonal antibodies are colored in cyan in the structure of P1. The short construct (left, residues Ala1400–Asp1521) includes only the C-terminal domain. The long construct (right, residues Lys1376–Asp1521) includes also the hinge, which connects the N-terminal domain with the C-terminal domain, and the two strands from the last propeller β-sheet, also known as the cytoadherence epitope. **b** Sera raised against any of the two constructs inhibit adhesion of *M. pneumoniae* cells even at low concentrations, with 3% showing a significantly higher inhibition than 1%. Error bars represent the mean ± standard deviation. $*p < 0.05$ and $***p < 0.001$. Source data are provided as a Source Data file.

Recognition of P1 and P40/P90 by sera from *M. pneumoniae* patients. Determination of the structural organization from the P1 and P40/P90 ectodomains encouraged us to assess the immunogenic activity of different constructs from these proteins using total sera from 19 patients with positive diagnostic for *M. pneumoniae* based on serological tests (Table 1). Clinically relevant information, including the IgG and IgM values, was available for all sera. In addition, sera from five patients that tested negative for *M. pneumoniae* were also included in the study for control purposes. First, the immunodominant nature of the adhesins was confirmed, revealing that 17 out of the 19 sera (90%) tested positive for P1. Similarly, 14 sera (74%) tested positive for P40/P90; all positive sera for P40/P90 were positive for P1. Of note, eight sera (42%) were also positive for P140. This result is not unexpected because P1 and P140 show extensive antibody cross-reactivity[43]. Alternatively, some of the P140 positive patients might have been infected with *M. genitalium* in addition to *M. pneumoniae*. Next, the sera response was examined with the same two C-terminal domain constructs of P1 (Lys1376–Asp1521 and Ala1400–Asp1521) described before in the generation of polyclonal antibodies. The number of positive sera, 5 out of 19 (26%), decreased substantially with respect to the full length P1, although importantly all these positive sera correspond to the ones with highest IgG levels.

For P40/P90, an additional evaluation of the sera reactivity with recombinant proteins lacking either insertion 1 (S1), or insertions 1 and 2 (S1S2) was performed to assess the contribution of the insertions and the loss of the proteolytic site. Construct S1 was detected by 14 out of 19 sera (74%), the same ratio to that of the entire P40/P90. In contrast, remarkably, there is an important decrease to just seven positive sera (37%) when the P40/P90 S1S2 construct was evaluated. Furthermore, even for most of these seven positive sera the response was substantially reduced. Therefore, results indicate that P40/P90 is a strong immunogenic protein, with insertion 2 (S2) acting as an immunodominant region, although recognized by sera with low IgG levels.

Discussion

Adherence of *M. pneumoniae* to sialylated receptors on the surface of respiratory cells was elucidated years ago[34,35,37,44]. Since the identification of the critical role of P1 in *M. pneumoniae* cytoadherence, it has been accepted that this protein was responsible for binding to sialic acid oligosaccharides[11–14,45]. In addition to P1, other cytadherence-accessory proteins had been reported as required for binding to sialic acids[14,46]. The present

Table 1 ELISA immunoassays for antigen detection by sera from patients infected with *M. pneumoniae*.

Antigen	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
P1	++	++	++	++	++	++	++	++	++	++	+++	++	+	++	++	−	+++	−	+++
C-terminal 1376	−	−	++	−	−	++	−	−	−	−	+++	−	−	−	−	−	+++	−	+++
C-terminal 1400	−	−	+	−	−	−	−	−	−	−	+++	−	−	+	−	−	+++	−	+++
P40/P90	++	++	−	++	++	+	++	++	++	−	+++	−	++	++	++	−	+++	−	+++
P40/P90s1	++	++	−	++	++	++	++	++	++	−	+++	−	++	++	++	−	+++	−	+++
P40/P90s1s2	−	−	−	−	−	−	++	−	++	−	+++	−	−	++	++	−	+++	−	+++
P140	−	++	++	+	−	−	−	−	−	−	++	−	−	−	++	++	++	−	++
IgG–IgM levels[a]																			
IgG (UA/ml)	42.8	35.7	62.6	4.65	0.44	0.72	34.1	0.25	2.15	14.3	>200	15.9	33.6	64.4	49.50	25.2	>200	44.5	72.4
IgM (UA/ml)	>27	2.8	17	14	15	>27	>27	12	>27	19	2.4	>27	>27	>27	12.00	2.80	6.20	1.80	7.40

−, negative (value < 0.080); +, weak positive (0.080 ≤ value ≤ 0.100); ++, strong positive (value > 0.100).
[a]IgG and IgM levels against a P1 C-terminal fragment of *M. pneumoniae* determined by the commercial Liaison *M. pneumoniae* IgG, IgM kit (DiaSorin).

study demonstrates that the binding in vitro to sialic acid receptors relies in fact on P40/P90 (Fig. 3c and Supplementary Fig. 7), however, affinities and specificities could be tuned through other proteins or due to the density and distribution of the oligosaccharides in the surface[37,38]. This important result is consistent with the recent finding that binding to sialic acid oligosaccharides is mediated by P110 in *M. genitalium*[18]. Cleavage of the P40/P90 polypeptide chain occurs in vitro at two exposed Arg–Ala motives (Arg445–Ala446 and Arg455–Ala456) located in the central part of the N-terminal domain, producing two polypeptide chains that can correspond to the P40 and P90 subunits found in vivo (Supplementary Figs. 4 and 5a). Cleavage appears to be a divalent metal ion dependent autoproteolysis taking place only after the folding is completed. Nevertheless, P40/P90 continues to behave as a single globular protein structurally similar to P110. This is in agreement with the extensive interface between the P40 and P90 polypeptides indicating that they form a very stable ensemble (according to PISA[32]) (Supplementary Fig. 6). However, P40 is located at the outer surface part of the Nap and due to the absence of covalent bonds might become separated from P90 and from the Nap complex under stringent or denaturing conditions. The loss of P40 only in certain conditions could explain the discrepancies in the past about whether or not P40 was part of the adhesion complex. The sialic acid oligosaccharides binding site is located in a loop of the P90 polypeptide with P40 contributing only to the entrance of the pocket (Fig. 3c and Supplementary Fig. 6).

There are clear structural relationships between P1 and P40/P90 suggesting a common protein ancestor, a hypothesis that is reinforced by the structural similarities found between P1 with the non-orthologous protein P110 from *M. genitalium* (Fig. 4 and Supplementary Fig. 9). This common ancestor could have been encoded in a mobile element, thus facilitating transmission within different species of the pneumoniae cluster. The presence of multiple copies of the adhesin genes in the genomes of several mycoplasmas supports this hypothesis. In each species, the adhesin genes may have evolved to mediate adherence to different hosts and tissues. Molecular systems where one of the catalytic sites, at the interface of two paralogs, is lost during evolution can be related with the situation in F-type ATPase/synthase, widely distributed among most living organisms[47], raising the possibility of a similar evolutionary process for the Nap. Interestingly, the sequence identity between P40/P90 and P1 is low in the N-terminal domains (12%), increases in the C-terminal domains (17%), and becomes high for the transmembrane and cytoplasmic regions (43%). The increased variation of the sequences from the most exposed regions of the Nap, the N-terminal domains, can reflect specific attachment and motility requirements but also a mycoplasma survival strategy related with enhanced antigenic variability.

To date, 13 antigenic variants have been identified in clinical isolates for P1 and 5 for P40/P90[48] (Fig. 5b, c). Regions of P1 and P40/P90 presenting genetic and clinical variability cluster at the surface of the upper part of the N-terminal domain, where antibodies could easily block access to the receptor binding site. In fact, mapping the known genetic variability of *M. pneumoniae* onto the Nap structure, modeled according to the information available from *M. genitalium* with the structures of P1 and P40/P90, reveals that the sialic binding site pocket is completely surrounded by these variable regions (Fig. 7). P1 and P40/P90 present many sequence insertions, which amount to more than 200 extra residues in each protein, with respect to their orthologues P140 and P110 in *M. genitalium* (Supplementary Figs. 1 and 4). Interestingly, most of the insertions correspond to disordered regions in the structures of the *M. pneumoniae* proteins. In P1 there are 14 disordered loops, while in P40/P90 there are 4,

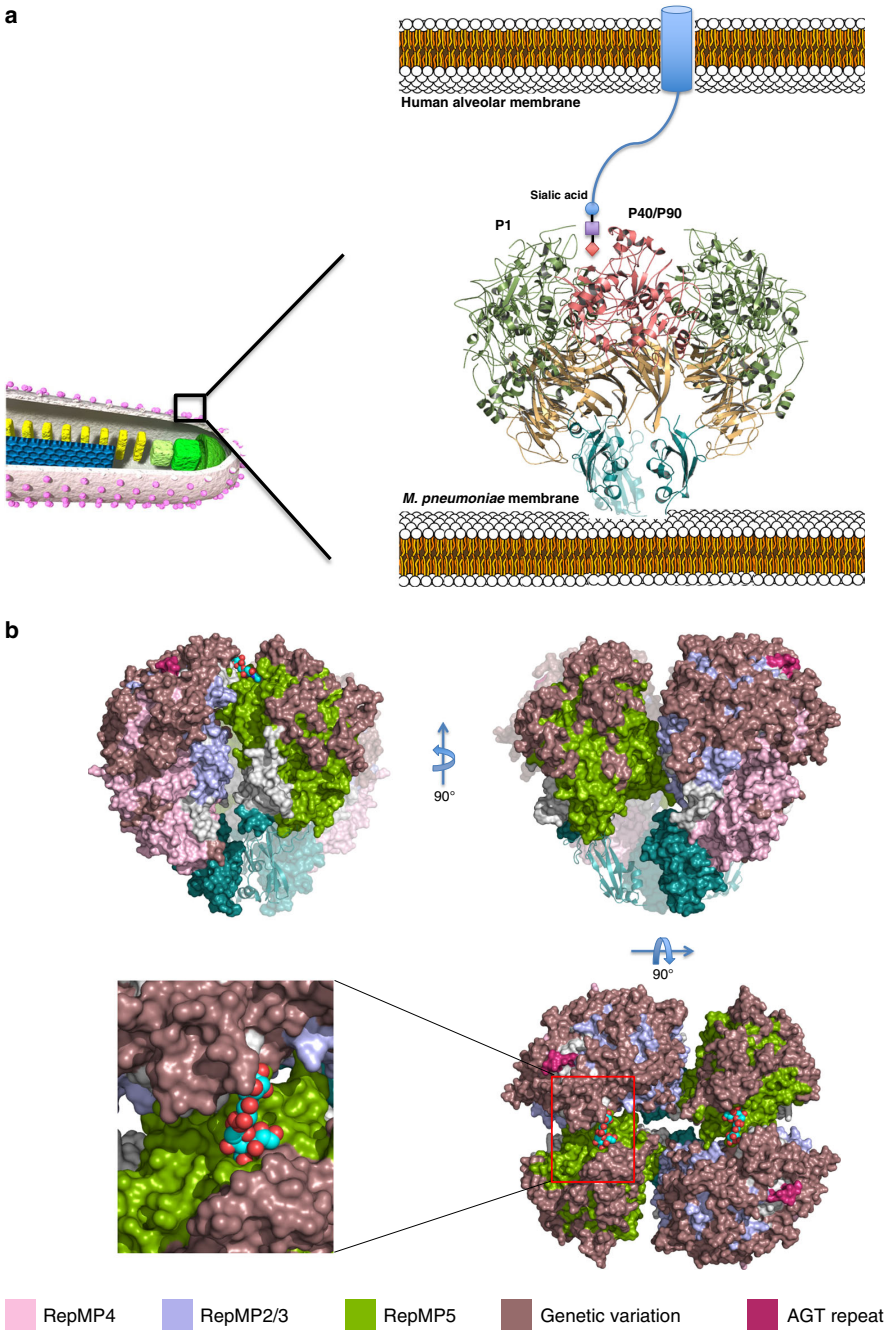


Fig. 7 Structure of the *M. pneumoniae* Nap. a Representation of the whole *M. pneumoniae* Nap ectodomain as a dimer of a P1-P40/P90 heterodimer, modeled from the information recently available for *M. genitalium*. Binding to a receptor from the host cell, a human alveolar epithelial cell, is indicated. The sialic acid binding site is located in a pocket at the top of the Nap that is not far from the twofold symmetry axis of the Nap. The crowns from P1 (green) and P40/P90 (pink) are separated by the β-propellers (brown) from the C-terminal domains (cyan), which anchor the Nap to the mycoplasma membrane. **b** Three views, 90° apart from each other, of the Nap surface depicting the sialic acid compound 3SL (spheres) bound into the binding site of the two P40/P90 subunits of a Nap. Regions with genetic variability of P1 and P40/P90 (brown) cover most of the surface from the upper part of the Nap. RepMP5 from P40P90 is depicted in green, while RepMP4 and RepMP2/3 from P1 are depicted in pink and blue, respectively. C-terminal domains from both P1 and P40/P90 are depicted in cyan. The inset shows the sialic binding site pocket, with a 3SL molecule bound, surrounded by genetically variable regions.

but 2 of them more than 60 residues long (insertions S1 and S2). The C-terminal domain from P1 contains only one disordered loop and no genetic variability, pointing to the interest of immunogenic studies with this domain.

Polyclonal antibodies against two constructs from the C-terminal domain of P1 (Lys1376–Asp1521 and Ala1400–Asp1521) inhibit the cytoadherence of *M. pneumoniae* cells (Fig. 6b). In addition, most of the sera from *M. pneumoniae* infected patients presenting a high IgG content also recognize these C-terminal domain constructs (Table 1). On the contrary, these C-terminal domain constructs were poorly recognized by sera from the early stages of infection where a high IgM content is present. The full length extracellular region of P1 was strongly recognized by all sera, independently of the IgG or IgM ratios, confirming the immunodominance of P1[19]. These results indicate that the C-terminal domain might be a relatively weak or poorly accessible antigenic region, in agreement with what the structural information suggests, but effective to generate more specific IgGs antibodies capable of inhibiting mycoplasma cytoadherence. Large P1 constructs, such as the whole extracellular region[49] or a long commercial peptide (residues 1160–1521, LIAISON *M. pneumoniae* IgM and IgG, Biotrin International Ltd), maybe better suited for diagnosis than small C-terminal domain constructs, which might, however, provide a significant protective response. P40/P90, often overlooked in immunological studies in spite of pioneering studies showing P90 was an important immunogen[50], is recognized by 74% of the sera from infected patients. Interestingly, the serologic response decreased sharply for P40/P90 constructs when the second insertion (S2 residues Leu395–Thr462) was eliminated. Most of the sera that recognize S2 have a high IgM content, dominant in the initial stages of the immune response. S2 presents the highest genomic (64%) and clinical (~80%) variability and is highly exposed and mobile on the surface of P40/P90. Exposure of S2 is likely increased by the cleavage at its C-end that generates the two polypeptides P40 and P90. Altogether these results suggest that S2, located within P40, might be acting as a decoy to divert the immune response (Table 1). The high immune response against P40/P90, with unique structural peculiarities, invites a reconsideration of its therapeutic possibilities.

Gliding of mycoplasmas from the pneumoniae cluster, including *M. pneumoniae* and *M. genitalium*, is thought to be caused by a mechanism where the force, generated by ATP energy[51], is transmitted to the Nap that appears to work as a leg/foot on the surface of the mycoplasma cell[7,8,12,52–56]. Gliding of *M. pneumoniae* cells is always in the direction of the attachment organelle, which indicates that the gliding machinery, including the Naps at the cell surface, should provide a well-defined directionality. To achieve a directional gliding, the release of the sialylated receptors is as essential as the binding because mycoplasma cells cannot move forward if they remain anchored to the surface receptors where they have first attached. In the Nap, the conformational changes required by the foot/leg activity during gliding have to be synchronized with the binding/release to the cell receptors, which are thought to be randomly distributed on the host surface. The binding site in P40/P90 could be regulated by the interaction with P1 as found in *M. genitalium*[17]. This may be the reason why P1 was expected to contain the binding site. In the cryo-EM map of P1, the weak density of the C-terminal domain indicates flexibility with respect to the N-terminal domain proving that P1 can experience significant conformational changes consistent with the observations that antibodies that bind to the C-terminal domain can interfere with the attachment and gliding of mycoplasma cells. Therefore, although P40/P90 contains the binding site to the sialic acid cell receptors, P1 has still to be seen as a major player in the Nap functioning during gliding and also in the adhesion to host cells. Moreover, proteins others than the Nap components, including P30, P65, HMW1, HMW2, or PrpC/PrkC, are known to be essential for binding and gliding[57–60]. These proteins might support functioning and conformation of the Naps to play its roles.

Integration of previous knowledge with the structural data obtained in this work provides important clues for a better understanding of the functioning of Naps in *M. pneumoniae*, although many questions remain open about the adhesion and gliding motility mechanisms. The structural framework obtained in this work also allows the mapping of epitopes and of the genetic and clinical variability of P1 and P40/P90 explaining many of the immunogenic properties of these immunodominant proteins in *M. pneumoniae*. These results can help in the development of vaccines and therapeutics for inhibiting the infectivity of this respiratory pathogen.

Methods

Cloning, expression, and purification of P1 and P40/P90 constructs. Regions corresponding to the MPN141 and MPN142 genes from *M. pneumoniae* were amplified from synthetic clones (Supplementary Tables 3 and 4), using primers P1F and P1R and P40/P90F and P40/P90R, respectively (Supplementary Table 5). The PCR fragments were cloned into the expression vector pOPINE[61] (gift from Ray Owens plasmid #26043, Addgene, Watertown, USA) to generate constructs, with a C-terminal His-tag, comprising residues 29–1521 for P1 and 23–1003 and 23–1114 for a short and a long construct of P40/P90, respectively. Recombinant proteins were obtained after expression at 22 °C in B834 (DE3) cells (Merck), upon induction with 0.8 mM IPTG at 0.6 OD_{600}. Cells were harvested and lysed by sonication in 1xPBS, 40 mM imidazole and centrifuged at 49,000 × *g* at 4 °C. Supernatant was loaded into a HisTrap 5 ml column (GE Healthcare) pre-equilibrated in 1xPBS with 40 mM imidazole as binding buffer and 1xPBS with 400 mM imidazole as elution buffer. Soluble aliquots were pooled and loaded onto a Superdex 200 GL 10/300 column (GE Healthcare) in buffer consisting of TRIS 20 mM pH 7.4 and 150 mM NaCl. Buffers prepared for P40/P90 purification were supplemented with 1 mM EDTA to prevent proteolysis. Mutants of P40/P90 were produced following the same protocol as for the wild type, but using the corresponding primers showed in Supplementary Table 5. Cleavage sites in P40/P90 were determined with N-terminal Edman sequencing of the bands obtained by SDS-PAGE of purified samples of the P40/P90 constructs (Supplementary Fig. 5a). Mixtures of P1 with either the short or the long P40/P90 constructs were used for the preparation of complexes in a 1:1 ratio. The presence of stable P1-P40/P90 heterodimers, verified by gel filtration, was the only complex detected in the mixtures (Supplementary Fig. 5c).

For cryo-EM data collection, P1 of *M. pneumoniae* M129 strain (MPN141) was expressed and purified as described previously[62]. The *p1* gene corresponding to the amino acid residues 60–1518 of P1 adhesin was codon-optimized, synthesized, and inserted between *Nde*I and *Xba*I sites in pCold I vector (TAKARA BIO, Shiga, JAPAN). The resulting plasmid, pP1-1, wherein *p1* was fused to an N-terminal 6 × His-tag and a Factor Xa site. BL21 or BL21 (DE3) pLysS cells carrying pP1-1 were cultured at 37 °C until the culture density reached an OD_{600} value of 0.3. The cells were subjected to cold shock at 15 °C for 30 min, followed by the addition of IPTG to a final concentration of 1.0 mM, and cultured overnight at 15 °C. The cells were harvested by centrifugation (4000 × *g*, 10 min, 4 °C) and washed twice with a TN buffer comprising 20 mM Tris-HCl and 500 mM NaCl. The suspension was frozen at −80 °C. The cell pellet was thawed, suspended in a binding buffer (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 20 mM imidazole) containing 1 mM PMSF. Cells were then disrupted with a Sonication (NIHONSEIKI, Tokyo, Japan). Unbroken cells were removed by centrifugation (60,000–140,000 × *g*, 60 min, 4 °C). The supernatant was loaded onto a HisTrap HP column (GE Healthcare, Milwaukee, WI) equilibrated with the binding buffer. The bound proteins were eluted with a linear gradient of 10–200 mM imidazole in TN buffer. Fractions containing P1 adhesin were dialyzed against a TNG buffer (20 mM Tris-HCl (pH 8.0), 150 mM NaCl) and loaded onto a HiLoad 16/60 Superdex 200 prep grade column (GE Healthcare) equilibrated with the TNG buffer at a flow speed of 1 mL/min. The collected P1 adhesin fractions were combined, dialyzed against 20 mM Tris-HCl (pH 8.0), and applied to an anion exchanger, Q Sepharose Fast Flow equilibrated with the buffer, and eluted with a linear gradient of NaCl from 0 to 0.4 M. The purified protein was concentrated using an Amicon Ultra 30k spin filter (Millipore, Darmstadt, Germany) if necessary.

Crystallization and X-ray data collection of P1 and P40/P90N. Ninety-six-well plates crystallization screenings were performed on both Cartesian (Cartesian (TM) Dispensing Systems) and Phenix (Art Robbins Instruments) robots mixing 150 nL of reservoir condition with 150 nL of protein solution at concentrations of 6 mg/mL for each of the P1 and P40/P90 proteins. The optimized crystals of P1 were

obtained after mixing 1 µL of P1 at 6 mg/mL with 1 µL 20% PEGMME 2000 and PCB (sodium propionate, sodium cacodylate, and BIS-TRIS propane pH 4) as a reservoir condition in a ratio of 2:1, respectively. P40/P90N optimized crystals were obtained mixing 1 µL of P40/P90N at 6 mg/mL with 1 µL 20% PEG 3350, 0.3% MetOH in a ratio 1:1, respectively. To obtain crystals of P40/P90N in complex with 3SL or 6SL, protein samples were previously incubated with 10 mM of each oligosaccharide 1 h at 20 °C. All attempts for P40/P90 crystallization were not successful. Crystals used for data collection were flash frozen in liquid nitrogen with 15% glycerol as cryo-protectant.

X-ray data collection was carried out at Xaloc beamline (ALBA Synchrotron, Spain). Data were processed with Xia2[63] using XDS[64], Aimless and Pointless[65] from the CCP4 suit of programs[66].

Determination of the P1 crystal structure. The crystals from P1 belong to space group C2 and could contain, by packing considerations, only one subunit in the asymmetric unit. We tried to solve these P1 crystals by heavy atoms and SeMet derivatives, but finally the resolution was achieved by density modification techniques applying solvent flattening and averaging with crystals from the *M. genitalium* orthologous protein P140, whose structure was also unknown at that time[17]. With this procedure the structures of P1 and of P140 (with a sequence identity of 41% between their ectodomains) were solved simultaneously. The P140 containing crystals used corresponded to crystals from the complex of P140 with the N-domain of P110 (P140–P110N) and to crystals from P140 alone. Steps followed for the structure determination of P1 and P140 were the following:

(1) A partial molecular replacement solution was obtained for the P140–P110N crystals using as searching model the N-domain from the P110 structure[17]. This partial solution of the P140–P110N crystals contained four P110N subunits in the asymmetric unit. The four P110N subunits were organized as two pairs and within each pair the two subunits were related by accurate local (non-crystallographic) twofold axis. In this partial solution, no density was visible that could correspond to P140.
(2) From this partial solution of the P140–P110N crystals, a difference (Fo–Fc) map was generated at about 3 Å resolution. Positive densities in this difference map, expected to correspond mainly to P140, were tentatively reinforced by averaging (without applying any mask) using one of the local twofold symmetries. This unmasked averaging also weakens density from other subunits in the crystal not related by the local twofold symmetry. In this averaged density it was possible to differentiate continuous regions of strong and of weak density, but it was not yet possible to identify any consisting molecular feature from P140.
(3) The cryo-electron tomography (CET) map of a whole *M. genitalium* Nap[17] was then used to define an initial possible mask for P140. The CET map, at ~17 Å resolution, corresponded to a dimer of P140–P110 complexes, where the fitting of the P110 structures was unambiguous. The CET map was placed on the (Fo–Fc) averaged map by superposing the corresponding fitted P110 structures. Then, a putative "P140 initial map" was generated as the (Fo–Fc) averaged density inside the CET envelop corresponding to P140. After some polishing (avoiding overlapping with neighbor subunits or retaining regions with continuous densities) the assumed "P140 initial map" was used to perform a molecular replacement search (using program PHASER) obtaining a reasonable possible solution with four P110N structures and four "P140 initial maps." Attempts to use the density from the CET maps directly for phasing, and not just as a mask, were unsuccessful in our hands.
(4) Iterative cycles of density averaging (and solvent flattening), alternating with manual readjustments of the P140 mask, were then performed with the P140–P110N crystals. The procedure converged giving clearer and more continuous density, although no secondary structures that could correspond to P140 were visible. Likely, the almost parallel orientation of the two non-crystallographic twofold axes weakened the phasing power of this averaging. In spite of the limitations, the new density (within the updated mask) allowed obtaining a molecular replacement solution for the P1 crystals. Surprisingly, the new density did not provide a solution for the P140 alone crystals.
(5) Density modification with averaging within the P140–P110N crystal and now also with the P1 crystals (using program DMMULTI[67]) improved quickly the maps. The iterative cycles of DMMULTI were alternated with cycles of phase extension for the P1 crystals (using the "autobuild" protocol in Phenix at 1.94 Å resolution) and with manual readjustments of the averaging masks (that were updated about forty times). Model building was carried forward for P1 and P140 in parallel. With about 50% of the structures traced, it was possible to obtain a molecular replacement solution for the P140 alone crystals that, with six subunits in the crystal asymmetric unit, facilitated the completion of the P1 and P140 structures.

P1 was then refined, giving agreement factors R and R_{free} of 18.7% and 22.9%, respectively (Supplementary Table 1). The difference between these two factors was used traditionally to assess how well the available molecular model explains the corresponding experimental diffraction data. The difference between both factors is related with the bias introduced in the model during refinement. Despite the quality of the final map (Fig. 2b), the P1 structure presented a significant amount of disordered residues located at the N-end (missing residues 29–59) and in 14 loops (102–105, 228–230, 259–268, 278–282, 298–300, 337–348, 831–847, 870–888, 923–928, 941–944, 1226–1232, 1308–1324, 1341–1349, and 1482–1495). The final refined structure has been deposited in the PDB with code 6RC9.

Determination of the P40/P90 structures. The P40/P90 crystals belong to space group $P2_12_12_1$ and contain two subunits in the asymmetric unit. The structure was solved by molecular replacement with program Phaser[68] using as searching model the structure from the orthologous protein P110 from *M. genitalium* (PDB code 6R3T). The structure of P40/P90 was refined at 2.65 Å resolution, giving agreement factors R and R_{free} of 21.4% and 23.4%, respectively (Supplementary Table 1). Crystals of complexes from P40/P90 with 3SL and 6SL belong to space C2, although with pseudo-orthorhombic unit cell dimensions, and contain one subunit in the asymmetric unit. The 3SL crystals showed a well-defined oligosaccharide with high occupancy, at 3.2 Å. The 6SL crystals, at 2.8 Å resolution, presented some disorder and an overall weak density for the oligosaccharide that was interpreted as partial occupancy.

Tracing and refinement of all the structures was performed alternating interactive and automatic cycles with programs Coot[69] and Buster[70], respectively. The final refined structures have been deposited in the PDB with codes: 6RJ1 for P40/P90, 6TLZ for the 3SL complex, and 6TM0 for the 6SL complex.

Single-particle cryo-EM. The purified P1 protein construct was applied onto a Quantifoil holey carbon grid (R1.2/1.3, Cu, 200 mesh) covered with a thin film of GO flakes (SIGMA-ALDRICH), blotted for 4.5 s at 4 °C in 100% humidity, and plunge frozen in liquid ethane by using a Vitrobot Mark IV (Themo Fisher Scientific). Cryo-EM imaging was obtained using a Titan Krios (Themo Fisher Scientific) operating at 300 kV acceleration voltage and equipped with a Cs corrector (CEOS, GmbH) and a Falcon III direct electron detector (Themo Fisher Scientific). A total of 5279 movies were obtained in the electron counting mode with a physical pixel size of 0.69 Å/pixel and total dose of 60 e/Å^2 with 36.33 s exposure time. The data were automatically collected using EPU software with a defocus range of −0.75 to −2.75 µm and were fractionated into 78 movie frames.

Single-particle cryo-EM image processing. The movie frames were subsequently aligned to correct for beam-induced movement and drift using MotionCor2[71], and contrast transfer function was estimated using Gctf[72]. A total of 4,574,424 particles images were automatically picked using Gautomatch (http://www.mrc-lmb.cam.ac.uk/kzhang/) and several rounds of 2D classification and 3D classification were performed using RELION-2.1[73]. A total of 295,195 particle images from good 2D classes were selected for making the initial model of P1 using cryoSPARC[74]. First and second 3D classification were calculated with angular sampling interval at 7.5 and 3.7 degrees, respectively. Finally, 68,014 particles from the best 3D class were 3D refined, producing a reconstruction with a resolution of 2.88 Å and a B-factor of 89 Å^2 with the gold-standard FSC criteria (FSC ≥ 0.143). The local resolution was estimated using RELION-2.1. The processing strategy is described in Supplementary Fig. 2a–d and the model refinement statistics in Supplementary Table 2.

SEC–MALS analysis. The molecular weight of the P40/P90 short construct was measured using a Superose 6 10/300 GL (GE HEalthcare) column in a Prominence liquid chromatography system (Shimadzu) connected to a DAWN HELEOS II multi-angle light scattering (MALS) detector and an Optilab T-REX refractive index (dRI) detector (Wyatt Technology). ASTRA 7 software (Wyatt Technology) was used for data processing and result analysis. A dn/dc value of 0.185 ml/g (typical of proteins) was assumed for calculations.

Antibody production and validation. Polyclonal antibodies against adhesin P1 were produced by Eurogentec. Briefly, rats were injected with peptides Lys1376–Asp1521 or Ala1400–Asp1521 and bleeds were collected at day 28. Pre-immunization bleeds were also collected as negative controls. Specificity of the polyclonal antibodies was tested by western blot against the recombinant peptides and protein extract of Mycoplasma cells. The dilution used for each antibody was 1:500.

The production of polyclonal antibodies has complied with ethical regulations for animal testing and research. It complies with the following association's requirements: Federation of European Laboratory animal Science Associations and UK Home Office Animals Scientific Procedures Act.

The regulations followed by Eurogentec are: 2010/63/EU and 01/2005/EU.

Quantitative adhesion assay. *M. pneumoniae* (ATCC 29343) was grown in a T75 flask in Hayflick medium during 24 h, harvested by scraping, and resuspended in Hayflick medium to a final concentration of 10^7 cells/ml. Incubation, during 1 h at room temperature on a rotating wheel, of 1 ml of a cell suspension with sera (at 1, 3, 6, and 10% concentrations) containing polyclonal antibodies raised against peptides Lys1376–Asp1521 or Ala1400–Asp1521, in presence or absence of 2 µg from the corresponding peptide. A preimmunization serum was used as a negative control. After incubation, cell suspensions were seeded in 24-well plates and

incubated at 37 °C to allow cells adhere to the plastic. After 2 h, supernatants (containing the fraction of the cells that did not adhere to the plastic) and attached cells were collected and processed for gDNA extraction using MasterPure DNA Purification Kit (EpiCentre #MCD85201). Samples were obtained in triplicate, and the relative amount of cells was quantified by qPCR with SYBR Green PCR Master Reagent (Thermo #4367659), using a LightCycler 480 Real-Time PCR machine (Roche) with the following conditions: denaturation at 95 °C for 10 s; 40 cycles of amplification of 95 °C for 15 s and 60 °C for 1 min; melting curve at 95° for 15 s, 60 °C for 15 s, 95° continuous. The following oligonucleotides were used: qPCR Myco FW 5′-ACGATGATTACAGGCGGTTC-3′ and qPCR Myco RV 5′- GTTG GTGGCCTCTTGTTGAT-3′.

Immunoassays. Diagnostic of *M. pneumoniae* was conducted using the Liaison *M. pneumoniae* IgG, IgM kit (DiaSorin) using a 1/100 dilution of the patient sera. Indirect ELISA assays were performed on 96-well plates Immulon 4 HBX 96-well plates (ThermoFisher) incubating 1 µg of each antigen at 4 °C overnight. 1/100 dilutions of each patient serum were added to the plate and detected using an anti-human IgG antibody conjugated with HRP (Thermofisher Scientific). Upon incubation for 30 min with 100 µl of substrate (Thermofisher Scientific), 100 µl of sulfuric acid 25% was added to stop the reaction and absorbance was read at 450 nm on a Triturus ELISA instrument (Grifols) device. Reference filter was set at 620 nm.

Ethical approval. Ethical approval for the study was obtained from the Ethics Committees for Research with Drugs from the Parc Taulí (Ref 2019/664) and Vall d'Hebron University Hospitals (PR(AG)24/2020).

Inductive coupled plasma–mass spectrometry (ICP–MS). Samples were digested with 1 ml of HNO_3 plus 0.5 ml of H_2O_2 in the closed teflon reactor during 24 h at 90 °C. The Zn measure was realized with ICP–MS in an instrument Agilent 7500ce with respect to a calibration line prepared by dilution from a ZN standard solution (Inorganic Ventures).

Surface plasmon resonance. Binding kinetics were determined by SPR using a Biacore 3000 biosensor platform (GE Biosystems) equipped with a research-grade streptavidin-coated biosensor chip SA. The chip was preconditioned with three 1-min injections of 1 M NaCl and 50 mM NaOH. Subsequently, the second and third flow cells of the chip were loaded, respectively, with oligosaccharides 6SL-PAA-biotin and 3SL-PAA-biotin (Carbosynth) at 10 µg/ml diluted in 20 mM TRIS pH 7.4, 150 mM NaCl, and 0.05 % tween20. It is noteworthy to mention that the ligands used contains a long polyacrylamide chain connecting the biotinylated chemically linked tag with the oligosaccharides, which should avoid most steric clashes between biotin and the P1 and P40/P90 samples used. The immobilization levels acquired were at ~200 response units and the first cell was left blank to serve as a reference. A series of diluted concentrations were injected using a flow rate of 30 µl/min at 25 °C (0.9, 1.9, 3.7, and 7.5 µM of P40/P90 for 6SL titration and 1.0, 2.0, 4.0, 16.0, and 32.0 µM of both P40/P90 and P1 for 3SL titration). The proteins were allowed to associate and dissociate for 60 and 120 s, respectively, for P1 and 120 and 240 s, respectively, for P40/P90. Dissociations were followed by a regeneration step of 30 s with 0.05% SDS at 30 µl/min. The data were fitted assuming a Langmuir 1:1 binding model using BiaEvaluation 3.1 software to determine the equilibrium dissociation constant K_D.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Atomic coordinates and structure factors for the reported crystal structures of P1, P40/P90, and the P40/P90-3SL and P40/P90-6SL complexes have been deposited into the Protein Data Bank (PDB) under accession codes 6RC9, 6RJ1, 6TLZ, and 6TM0, respectively. Cryo-EM density has been deposited in the EM Data Base under the accession code EMD-30233. Atomic coordinates of the Cryo-EM structure have been deposited in the PDB under the accession code 7BWM. Source data are provided with this paper. Other data are available from the corresponding authors upon reasonable request.

Received: 24 April 2020; Accepted: 8 September 2020;
Published online: 14 October 2020

References

1. Narita, M. Pathogenesis of extrapulmonary manifestations of *Mycoplasma pneumoniae* infection with special reference to pneumonia. *J. Infect. Chemother.* **16**, 162–169 (2010).
2. Ferwerda, A., Moll, H. A. & de Groot, R. Respiratory tract infections by *Mycoplasma pneumoniae* in children: a review of diagnostic and therapeutic measures. *Eur. J. Pediatr.* **160**, 483–491 (2001).
3. Atkinson, T. P., Balish, M. F. & Waites, K. B. Epidemiology, clinical manifestations, pathogenesis and laboratory detection of *Mycoplasma pneumoniae* infections. *FEMS Microbiol. Rev.* **32**, 956–973 (2008).
4. Parrott, G. L., Kinjo, T. & Fujita, J. A compendium for *Mycoplasma pneumoniae*. *Front. Microbiol.* **7**, 513 (2016).
5. Radestock, U. & Bredt, W. Motility of *Mycoplasma pneumoniae*. *J. Bacteriol.* **129**, 1495–1501 (1977).
6. Miyata, M. & Hamaguchi, T. Prospects for the gliding mechanism of *Mycoplasma mobile*. *Curr. Opin. Microbiol.* **29**, 15–21 (2016).
7. Miyata, M. & Hamaguchi, T. Integrated information and prospects for gliding mechanism of the pathogenic bacterium *Mycoplasma pneumoniae*. *Front. Microbiol.* **7**, 960 (2016).
8. Nishikawa, M. S. et al. Refined mechanism of *Mycoplasma mobile* gliding based on structure, ATPase activity, and sialic acid binding of machinery. *mBio* **10**, e02846-19 (2019).
9. Krause, D. C. *Mycoplasma pneumoniae* cytadherence: unravelling the tie that binds. *Mol. Microbiol.* **20**, 247–253 (1996).
10. Krause, D. C. & Baseman, J. B. *Mycoplasma pneumoniae* proteins that selectively bind to host cells. *Infect. Immun.* **37**, 382–386 (1982).
11. Feldner, J., Gobel, U. & Bredt, W. *Mycoplasma pneumoniae* adhesin localized to tip structure by monoclonal antibody. *Nature* **298**, 765–767 (1982).
12. Seto, S., Kenri, T., Tomiyama, T. & Miyata, M. Involvement of P1 adhesin in gliding motility of *Mycoplasma pneumoniae* as revealed by the inhibitory effects of antibody under optimized gliding conditions. *J. Bacteriol.* **187**, 1875–1877 (2005).
13. Hu, P. C. et al. *Mycoplasma pneumoniae* infection: role of a surface protein in the attachment organelle. *Science* **216**, 313–315 (1982).
14. Krause, D. C., Leith, D. K., Wilson, R. M. & Baseman, J. B. Identification of *Mycoplasma pneumoniae* proteins associated with hemadsorption and virulence. *Infect. Immun.* **35**, 809–817 (1982).
15. Burgos, R. et al. *Mycoplasma genitalium* P140 and P110 cytadhesins are reciprocally stabilized and required for cell adhesion and terminal-organelle development. *J. Bacteriol.* **188**, 8627–8637 (2006).
16. Scheffer, M. P. et al. Structural characterization of the NAP; the major adhesion complex of the human pathogen *Mycoplasma genitalium*. *Mol. Microbiol.* **105**, 869–879 (2017).
17. Aparicio, D. et al. Structure and mechanism of the Nap adhesion complex from the human pathogen *Mycoplasma genitalium*. *Nat. Commun.* **11**, 2877 (2020).
18. Aparicio, D. et al. *Mycoplasma genitalium* adhesin P110 binds sialic-acid human receptors. *Nat. Commun.* **9**, 4471 (2018).
19. Chourasia, B. K., Chaudhry, R. & Malhotra, P. Delineation of immunodominant and cytadherence segment(s) of *Mycoplasma pneumoniae* P1 gene. *BMC Microbiol.* **14**, 108 (2014).
20. Jacobs, E., Pilatschek, A., Gerstenecker, B., Oberle, K. & Bredt, W. Immunodominant epitopes of the adhesin of *Mycoplasma pneumoniae*. *J. Clin. Microbiol.* **28**, 1194–1197 (1990).
21. Dandekar, T. et al. Re-annotating the *Mycoplasma pneumoniae* genome sequence: adding value, function and reading frames. *Nucleic Acids Res.* **28**, 3278–3288 (2000).
22. Su, C. J., Chavoya, A. & Baseman, J. B. Regions of *Mycoplasma pneumoniae* cytadhesin P1 structural gene exist as multiple copies. *Infect. Immun.* **56**, 3157–3161 (1988).
23. Wenzel, R. & Herrmann, R. Repetitive DNA sequences in *Mycoplasma pneumoniae*. *Nucleic Acids Res.* **16**, 8337–8350 (1988).
24. Sperker, B., Hu, P. & Herrmann, R. Identification of gene products of the P1 operon of *Mycoplasma pneumoniae*. *Mol. Microbiol.* **5**, 299–306 (1991).
25. Layh-Schmitt, G. & Herrmann, R. Localization and biochemical characterization of the ORF6 gene product of the *Mycoplasma pneumoniae* P1 operon. *Infect. Immun.* **60**, 2906–2913 (1992).
26. McGuffin, L. J., Bryson, K. & Jones, D. T. The PSIPRED protein structure prediction server. *Bioinformatics* **16**, 404–405 (2000).
27. Pantelic, R. S., Meyer, J. C., Kaiser, U., Baumeister, W. & Plitzko, J. M. Graphene oxide: a substrate for optimizing preparations of frozen-hydrated samples. *J. Struct. Biol.* **170**, 152–156 (2010).
28. Dallo, S. F., Su, C. J., Horton, J. R. & Baseman, J. B. Identification of P1 gene domain containing epitope(s) mediating *Mycoplasma pneumoniae* cytoadherence. *J. Exp. Med.* **167**, 718–723 (1988).
29. Holm, L. & Rosenstrom, P. Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38**, W545–W549 (2010).
30. Widjaja, M., Berry, I. J., Pont, E. J., Padula, M. P. & Djordjevic, S. P. P40 and P90 from Mpn142 are targets of multiple processing events on the surface of *Mycoplasma pneumoniae*. *Proteomes* **3**, 512–537 (2015).
31. Ruland, K., Himmelreich, R. & Herrmann, R. Sequence divergence in the ORF6 gene of *Mycoplasma pneumonia*. *J. Bacteriol.* **176**, 5202–5209 (1994).

32. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).
33. Chothia, C. & Lesk, A. M. The relation between the divergence of sequence and structure in proteins. *EMBO J.* **5**, 823–826 (1986).
34. Bensing, B. A. et al. Structural basis for sialoglycan binding by the *Streptococcus sanguinis* SrpA adhesin. *J. Biol. Chem.* **291**, 7230–7240 (2016).
35. Manchee, R. J. & Taylor-Robinson, D. Utilization of neuraminic acid receptors by mycoplasmas. *J. Bacteriol.* **98**, 914–919 (1969).
36. Sobeslavsky, O., Prescott, B. & Chanock, R. M. Adsorption of *Mycoplasma pneumoniae* to neuraminic acid receptors of various cells and possible role in virulence. *J. Bacteriol.* **96**, 695–705 (1968).
37. Kasai, T. et al. Role of binding in *Mycoplasma mobile* and *Mycoplasma pneumoniae* gliding analyzed through inhibition by synthesized sialylated compounds. *J. Bacteriol.* **195**, 429–435 (2013).
38. Williams, C. R. et al. Sialylated receptor setting influences *Mycoplasma pneumoniae* attachment and gliding motility. *Mol. Microbiol.* **109**, 735–744 (2018).
39. Gerstenecker, B. & Jacobs, E. Topological mapping of the P1-adhesin of *Mycoplasma pneumoniae* with adherence-inhibiting monoclonal antibodies. *J. Gen. Microbiol.* **136**, 471–476 (1990).
40. Su, C. J., Chavoya, A., Dallo, S. F. & Baseman, J. B. Sequence divergency of the cytadhesin gene of *Mycoplasma pneumoniae*. *Infect. Immun.* **58**, 2669–2674 (1990).
41. Spuesens, E. B. M. et al. Sequence variations in RepMP2/3 and RepMP4 elements reveal intragenomic homologous DNA recombination events in *Mycoplasma pneumoniae*. *Microbiology* **155**, 2182–2196 (2009).
42. Jacobs, E., Vonski, M., Oberle, K., Opitz, O. & Pietsch, K. Are outbreaks and sporadic respiratory infections by *Mycoplasma pneumoniae* due to two distinct subtypes? *Eur. J. Clin. Microbiol. Infect. Dis.* **15**, 38–44 (1996).
43. Morrison-Plummer, J., Lazzell, A. & Baseman, J. B. Shared epitopes between *Mycoplasma pneumoniae* major adhesin protein P1 and a 140-kilodalton protein of Mycoplasma genitalium. *Infect. Immun.* **55**, 49–56 (1987).
44. Baseman, J. B., Cole, R. M., Krause, D. C. & Leith, D. K. Molecular basis for cytadsorption of *Mycoplasma pneumoniae*. *J. Bacteriol.* **151**, 1514–1522 (1982).
45. Razin, S. & Jacobs, E. *Mycoplasma* adhesion. *J. Gen. Microbiol.* **138**, 407–422 (1992).
46. Waldo, R. H. 3rd, Jordan, J. L. & Krause, D. C. Identification and complementation of a mutation associated with loss of *Mycoplasma pneumoniae* virulence-specific proteins B and C. *J. Bacteriol.* **187**, 747–751 (2005).
47. Guo, H., Suzuki, T. & Rubinstein, J. L. Structure of a bacterial ATP synthase. *Elife* **8**, e43128 (2019).
48. Kenri, T. et al. Periodic genotype shifts in clinically prevalent *Mycoplasma pneumoniae* strains in Japan. *Front. Cell. Infect. Microbiol.* **10**, 385 (2020).
49. Svenstrup, H. F., Nielsen, P. K., Drasbek, M., Birkelund, S. & Christiansen, G. Adhesion and inhibition assay of *Mycoplasma genitalium* and *M. pneumoniae* by immunofluorescence microscopy. *J. Med. Microbiol.* **51**, 361–373 (2002).
50. Leith, D. K., Trevino, L. B., Tully, J. G., Senterfit, L. B. & Baseman, J. B. Host discrimination of *Mycoplasma pneumoniae* proteinaceous immunogens. *J. Exp. Med.* **157**, 502–514 (1983).
51. Mizutani, M. & Miyata, M. Behaviors and energy source of *Mycoplasma gallisepticum* gliding. *J. Bacteriol.* **201**, e00397-19 (2019).
52. García-Morales, L., González-González, L. Querol, E. and Piñol, J. A minimized motile machinery for *Mycoplasma genitalium*. *Mol. Microbiol.* **100** 125–138 (2016).
53. Miyata, M. Centipede and inchworm models to explain *Mycoplasma* gliding. *Trends Microbiol.* **16**, 6–12 (2008).
54. Seybert, A. et al. Cryo-electron tomography analyses of terminal organelle mutants suggest the motility mechanism of *Mycoplasma genitalium*. *Mol. Microbiol.* **108**, 319–329 (2018).
55. Nakane, D., Kenri, T., Matsuo, L. & Miyata, M. Systematic structural analyses of attachment organelle in *Mycoplasma pneumoniae*. *PLoS Pathog.* **11**, e1005299 (2015).
56. Kawamoto, A. et al. Periodicity in attachment organelle revealed by electron cryotomography suggests conformational changes in gliding mechanism of *Mycoplasma pneumoniae*. *mBio* **7**, e00243-16 (2016).
57. Chang, H. Y., Jordan, J. L. & Krause, D. C. Domain analysis of protein P30 in *Mycoplasma pneumoniae* cytadherence and gliding motility. *J. Bacteriol.* **193**, 1726–1733 (2011).
58. Hasselbring, B. M., Sheppard, E. S. & Krause, D. C. P65 truncation impacts P30 dynamics during *Mycoplasma pneumoniae* gliding. *J. Bacteriol.* **194**, 3000–3007 (2012).
59. Page, C. A. & Krause, D. C. Protein kinase/phosphatase function correlates with gliding motility in *Mycoplasma pneumoniae*. *J. Bacteriol.* **195**, 1750–1757 (2013).
60. Romero-Arroyo, C. E. et al. *Mycoplasma pneumoniae* protein P30 is required for cytadherence and associated with proper cell development. *J. Bacteriol.* **181**, 1079–1087 (1999).
61. Berrow, N. S. et al. A versatile ligation-independent cloning method suitable for high-throughput expression screening applications. *Nucleic Acids Res.* **35**, e45 (2007).
62. Kenri, T. et al. Production and characterization of recombinant P1 adhesin essential for adhesion, gliding, and antigenic variation in the human pathogenic bacterium, *Mycoplasma pneumoniae*. *Biochem. Biophys. Res. Commun.* **508**, 1050–1055 (2019).
63. Winter, G. xia2: an expert system for macromolecular crystallography data reduction. *J. Appl. Crystallogr.* **43**, 186–190 (2010).
64. Kabsch, W. XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
65. Evans, P. Scaling and assessment of data quality. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 72–82 (2006).
66. Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760–763 (1994).
67. Cowtan, K. D., Zhang, K. Y. J. and Main, P. International Tables for Crystallography. Vol. F, ch. 15.3, pp. 407–412 (2012).
68. McCoy, A. J. Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr. D Biol. Crystallogr.* **63**, 32–41 (2007).
69. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D. Biol. Crystallogr.* **60**, 2126–2132 (2004).
70. Bricogne, G. et al. BUSTER 2.11.12. Cambridge, United Kingdom: Global Phasing Ltd. (2017).
71. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
72. Zhang, K. Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016).
73. Kimanius, D., Forsberg, B. O., Scheres, S. H. & Lindahl, E. Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. *Elife* **5**, e18722 (2016).
74. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).

Acknowledgements

This work was supported by grants BFU2018-101265-B-100 (MINECO) to I.F., BIO2017-84166-R (MINECO) to J.P., and BES-2015-076104 (MINECO) to M.L.S., JSPS KAKENHI Grant Number JP25000013 to K.N., and a FEDER project from Instituto de Salud Carlos III (ISCIII, Acción Estratégica en Salud 2016). This work has also been funded by the Platform Project for Supporting Drug Discovery and Life Science Research (BINDS) from AMED under Grant Number JP19am0101117 to K.N. (support number 1282), by the Cyclic Innovation for Clinical Empowerment (CiCLE) from AMED under Grant Number JP17pc0101020 to K.N., and by JEOL YOKOGUSHI Research Alliance Laboratories of Osaka University to K.N. This work was supported by a Grant-in-Aid for Scientific Research on the Innovative Area "Harmonized Supramolecular Motility Machinery and Its Diversity" (MEXT KAKENH, JP24117002 to M.M., JP25117530 and JP15H01337 to T.K.), Grants-in-Aid for Scientific Research (B) and (A) (MEXT KAKENHI, JP24390107, JP17H01544), JST CREST (JPMJCR19S5), Osaka City University (OCU) Strategic Research Grant 2018 for top priority researches to M.M. D.A. acknowledges a María de Maeztu Unit of Excellence grant MDM-2014-0435. Thanks are given to O. Conchillo and Prof. X. Daura for their enlightening suggestions and to Prof. P. Loewen for a careful reading of the manuscript. Many thanks are given to the XALOC beamline team at ALBA for their support during data collection and to the Crystallography Platform at the Barcelona Science Park (PCB), especially to Dr. Roman Bonet for his expert advice during SEC–MALS analysis. We also acknowledge to Dr. Sílvia Barceló from the Molecular Interactions Unit, IDIBELL and to Dr. Marta Taules from the Scientific and Technologic Department, UB for their technical support during SPR experiments. Finally, we are very grateful to Dr. Juan Delgado from the Immunology Department of the Hospital Universitari Parc Taulí for his helpful advices on performing ELISA analyses.

Author contributions

Conceived and designed the experiments: D.V., I.F., T.K., K.N., M.M., and D.A. Performed the experiments: D.V., A.K., U.M., R.I., R.P.L., J.M., R.M., P.B., O.Q.P., M.P.S., T.K., T.A.K., and D.A. Analyzed the data: D.V., A.K., U.M., R.I., P.B., O.Q.P., M.E., I.S., J.E., M.F.H., M.P.S., J.P., A.S.F., M.L.S., S.M., K.S., T.K., T.A.K., I.F., and D.A. Contributed reagents/materials/analysis tools: A.S.F., M.E., K.N., M.M., and I.F. Wrote the paper: D.V., A.S.F., O.Q.P., M.L.S., T.K., M.M., I.F., and D.A.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41467-020-18777-y.

Correspondence and requests for materials should be addressed to M.M. or D.A.

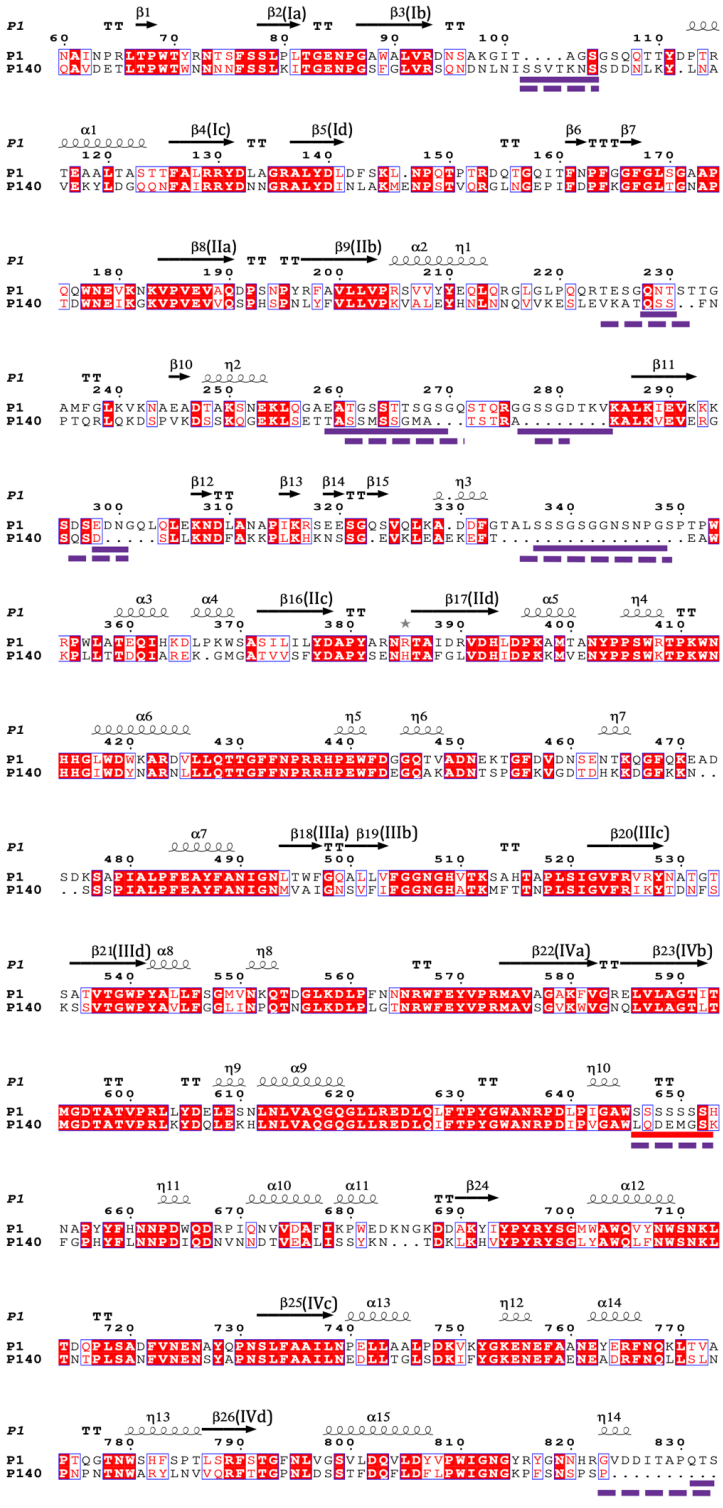
Peer review information *Nature Communications* thanks Jeremy Derrick and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

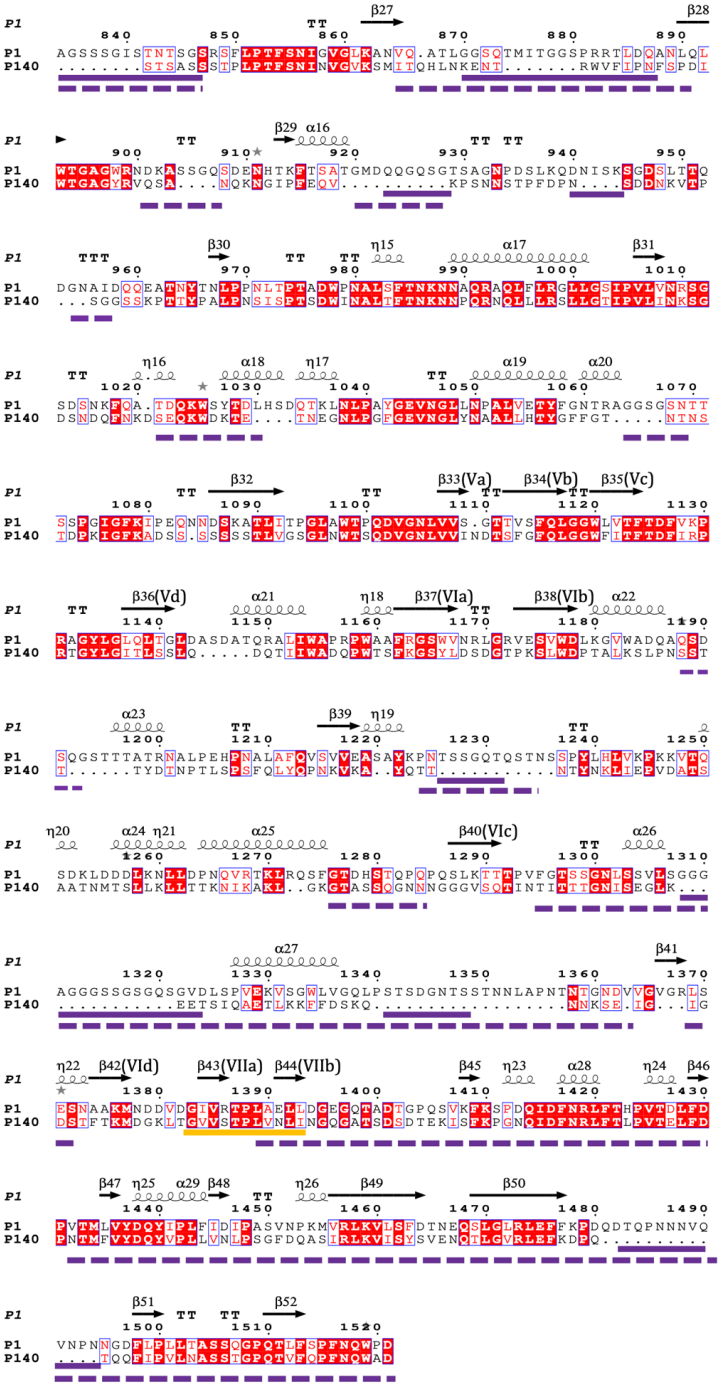
Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary Information

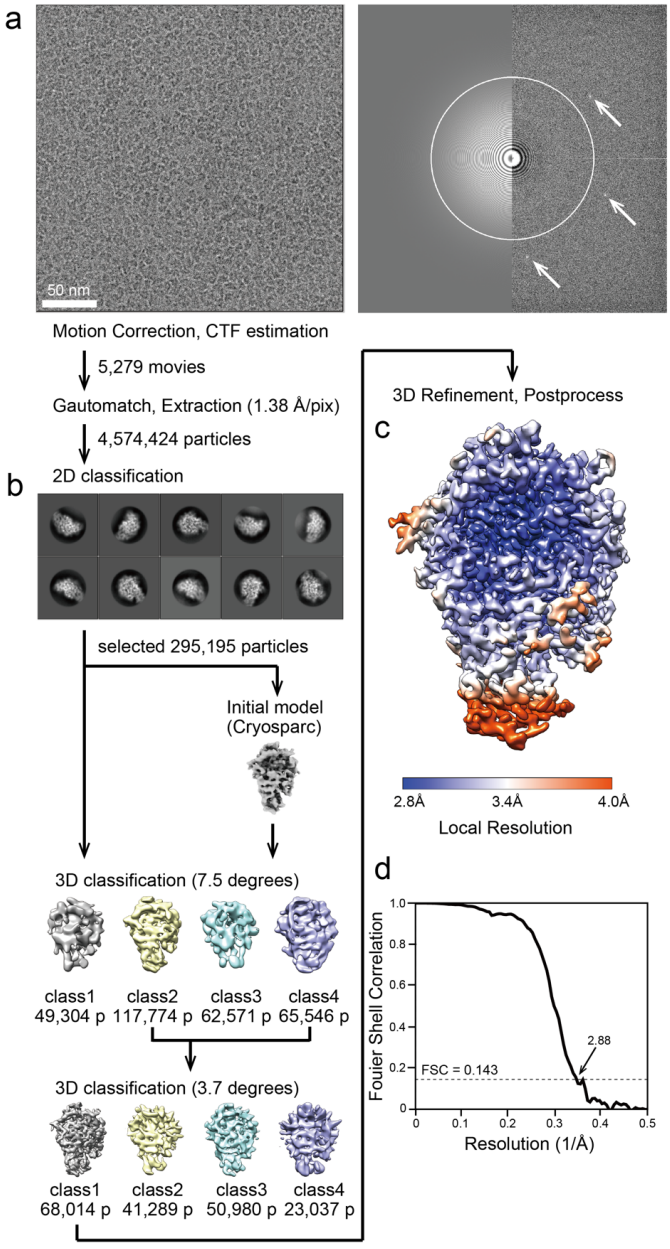
Immunodominant proteins P1 and P40/P90 from human pathogen *Mycoplasma pneumoniae*



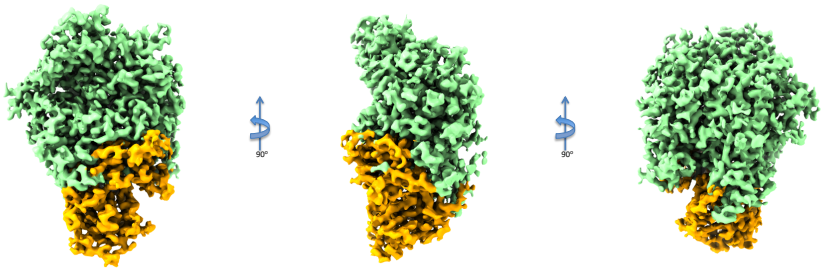


Supplementary Figure 1. **Structural sequence alignments of P1**

Structural sequence alignments obtained by the superposition between P1 and the orthologous protein P140 from *M. genitalium* and represented with program Espript [1]. The DSSP assignment of secondary structure elements for P1 are represented above the sequence, as squiggles (α-helices), arrows (β-strands) and TT (turns). Missing residues, corresponding to disordered regions, in the X-ray or in the Cryo-EM structures are indicated with a continuous or dashed purple underline, respectively. The AGT repeats are underlined red and the cytoadherence epitope in orange.

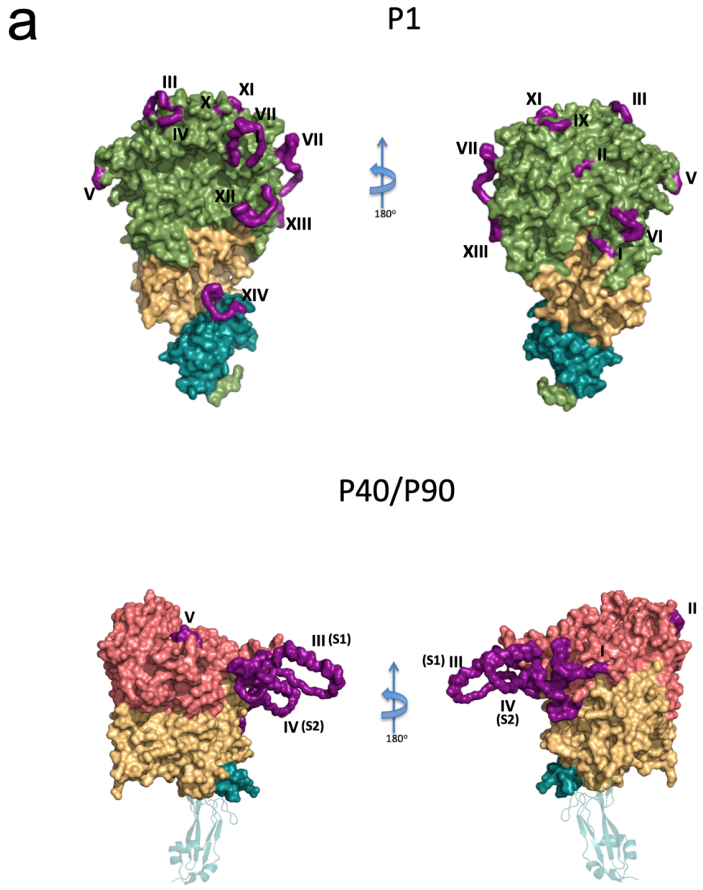

a
50 nm
Motion Correction, CTF estimation
5,279 movies
Gautomatch, Extraction (1.38 Å/pix)
4,574,424 particles
2D classification
b
selected 295,195 particles
Initial model
(Cryosparc)
3D classification (7.5 degrees)
class1
49,304 p
class2
117,774 p
class3
62,571 p
class4
65,546 p
3D classification (3.7 degrees)
class1
68,014 p
class2
41,289 p
class3
50,980 p
class4
23,037 p
3D Refinement, Postprocess
c
2.8Å
3.4Å
4.0Å
Local Resolution
d
Fourier Shell Correlation
FSC = 0.143
2.88
Resolution (1/Å)

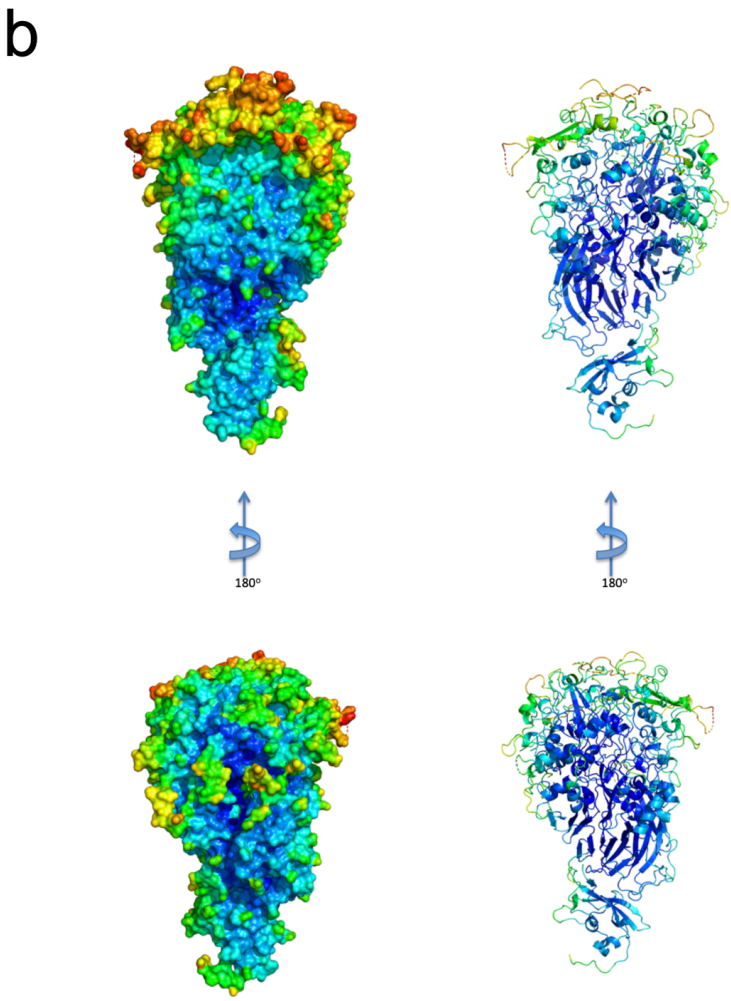
e



Supplementary Figure 2. Flow chart of cryo-EM image processing of P1

a) Representative microscope image and power spectra from the image. White arrow shows diffraction pattern derived from GO. **b)** 2D classification image of P1. **c)** Final 3D reconstruction map color-coded according to local resolution. **d)** Gold-standard Fourier Shell Correlation (FSC) curve of final map. **e)** Three 90° apart views of the cryo-EM map.

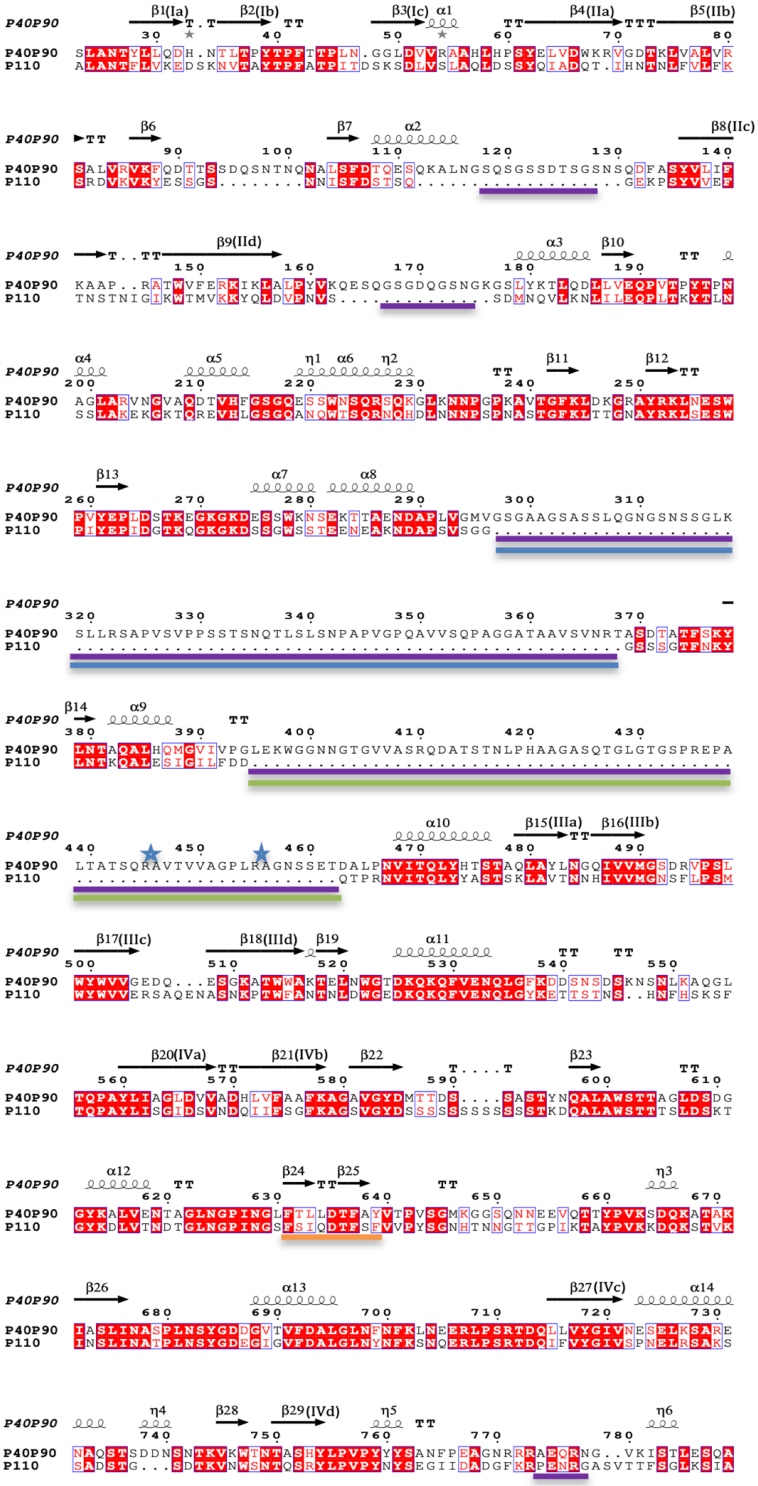

a
P1
III
XI
X
VII
IV
VII
V
XII
XIII
XIV
180°
XI
III
IX
VII
II
V
VI
XIII
I
P40/P90
V
III (S1)
IV (S2)
180°
II
(S1) III
I
IV
(S2)

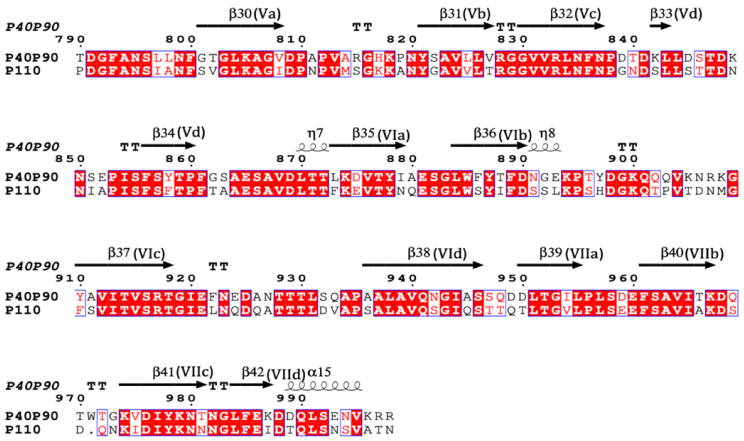


Supplementary Figure 3. Disordered regions in the structures of P1 and P40/P90

a) Protein surface representation indicating the disordered regions (purple) found in the crystal structures of P1 and P40/P90 (top and bottom panels, respectively). Most of these disordered regions correspond to insertions in the sequence with respect to the orthologue proteins P140 and P110 from *M. genitalium*. Only one disordered region, numbered XIV, is found in the C-terminal domain of P1 (cyan), while the crown of P1 contains thirteen

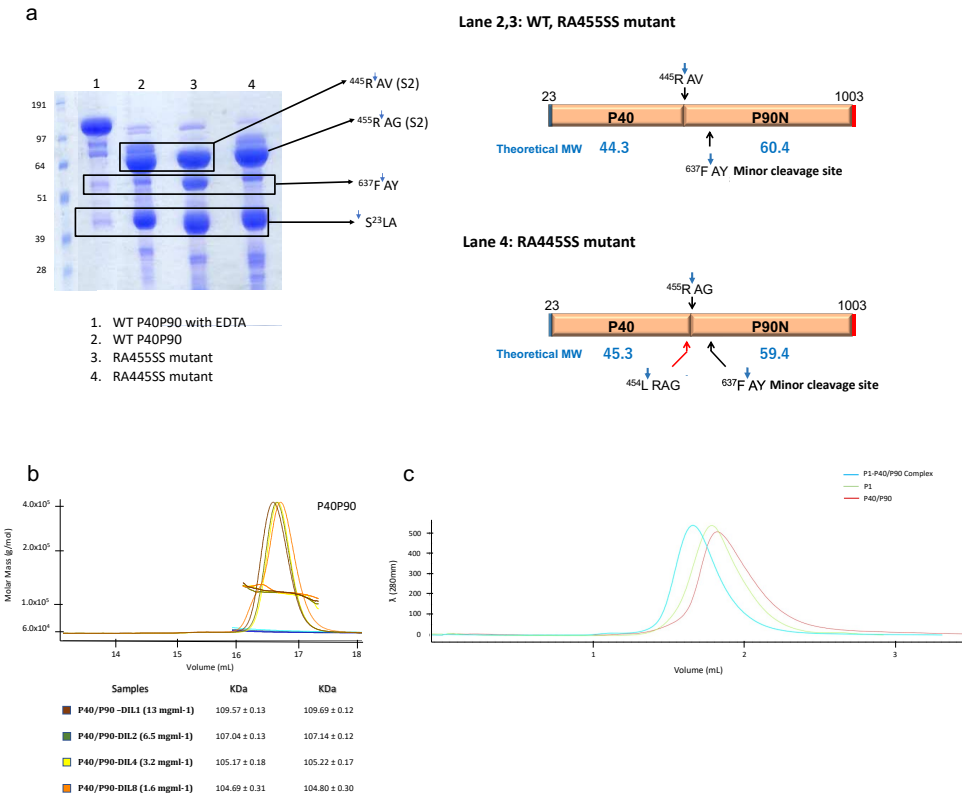
disordered regions, each with just a few residues. In P40/P90 there are only five disordered regions, but two of them, named as S1 and S2 (see in the text) with more than sixty residues each. There are no disordered regions in the β-propellers (brown). The disordered loops depicted in this figure, only for illustration purposes, have been built using the Sloop database [2]. **b)** Two 180° apart views of the temperature factors in the crystal structure of P1. The warmest places (red) are located in the upper part of the crown, while the coldest (dark blue) are in the β-propeller. The C-terminal domain is quite cold, indicating low mobility, contrary to what is found in the cryo-EM structure of P1.





Supplementary Figure 4. **Structural sequence alignments of P40/P90**

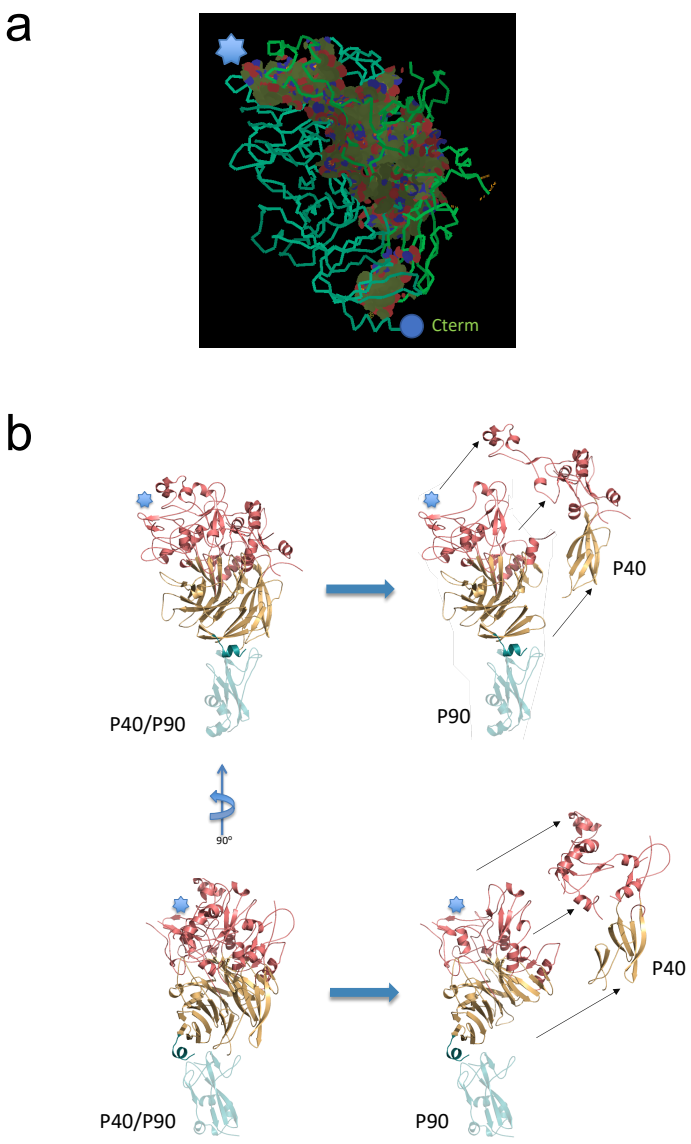
Structural sequence alignments obtained by the superposition between P40/P90 and the orthologous protein P110 from *M. genitalium* and represented with program Espript [1]. The DSSP assignment of secondary structure elements for P40/P90 are represented above the sequence, as squiggles (α-helices), arrows (β-strands) and TT (turns). Missing residues, corresponding to disordered regions of P40/P90 in the X-ray structure, are indicated with a purple line. The long insertions 1 and 2 (named as S1 and S2) are underlined in blue and green, respectively. An orange line indicates the sialic acid oligosaccharide binding loop and the two blue stars correspond to the Arg445-Ala446 and Arg455-Ala456 cleavage sites.



Supplementary Figure 5. Purification and cleavages of P40/P90

a) SDS-PAGE of purified samples from the P40/P90 ectodomain (residues from Ser23 to Pro1113) always showed several bands (left panel). For the construct with the wild type (WT) sequence and in the presence of EDTA (lane 1) the dominant band corresponds to the full construct of ~105 kDa. The pattern changed completely in the absence of EDTA (lane 2), with the band corresponding to the full construct almost disappearing and two new strong bands corresponding to polypeptides starting at Ser23 and Ala446 and with molecular weights of about 45 and 60 kDa, respectively. In the absence of EDTA the degradation pattern remains similar for the P40/P90 variant where the two residues from the cleavage motif (Arg445-Ala446) were replaced by serines (lane 4). However, now the band corresponding to the large polypeptide starts with Ala456. Finally, the pattern is

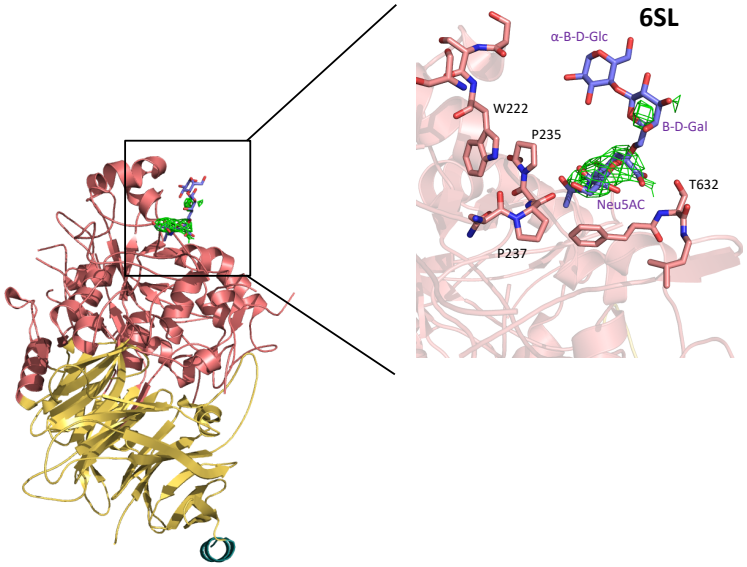
again very similar to the wild type for the P40/P90 variant where the two residues of the second RA cleavage motif Arg455-Ala456 were replaced by serines (lane 3). The relative size of the small and large polypeptides (right panel) could well correspond to the P40 and P90 subunits if the trans-membrane and cytoplasmic regions are added to the large polypeptide. Previously described cleavage site ^{454}LRAG is also shown [3] **b)** Analysis by MALS of purified samples of P40/P90N detected only one monomeric specie with a molecular weight of ~107 kDa**. c)** Chromatographic profiles (with a Superdex 200 5/150) of P1 (green), P40/P90 (red) and the P1-P40/P90 heterodimer (light blue).



Supplementary Figure 6. P40 and P90 polypeptides in P40/P90

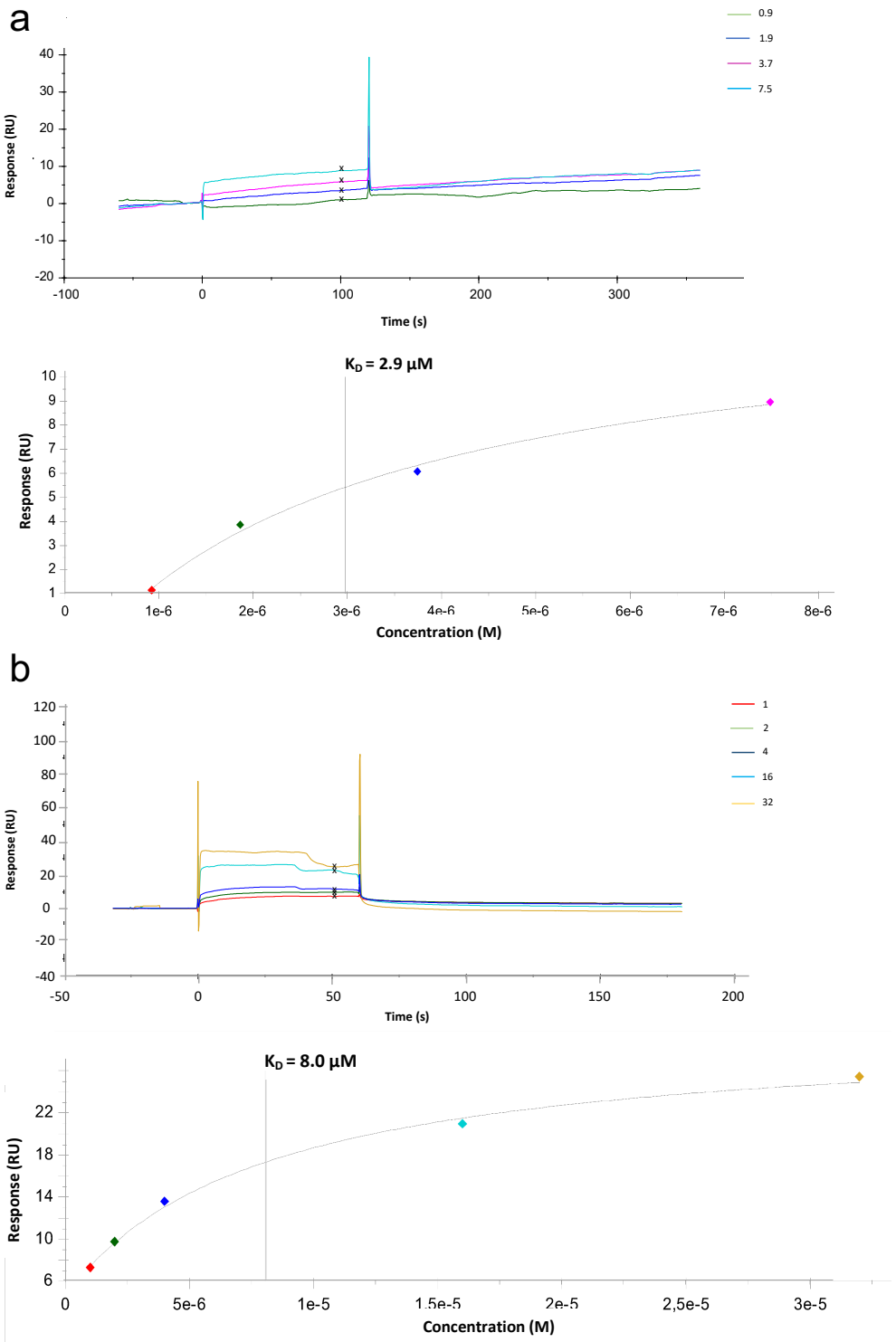
a) The P40 polypeptide (solid representation) sits on the surface of the P90 polypeptide, with a large interacting surface. The sialic acid binding site is indicated with a blue star. Colored blobs indicate interface contacts. **b)** Separation of the two polypeptides would

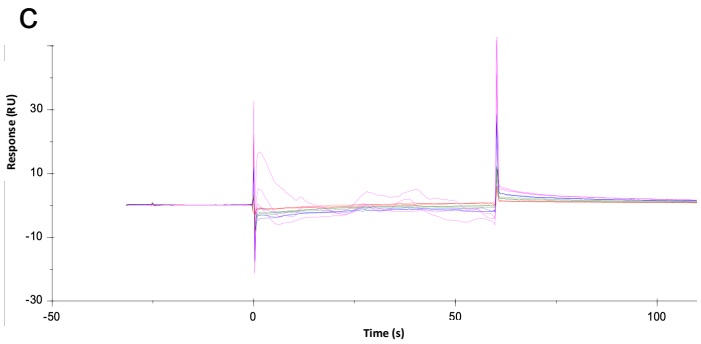
divide the β-propeller and alter the sialic binding site, although all the residues interacting directly with the neuraminic moiety belong to the P90 polypeptide.



Supplementary Figure 7. Binding of oligosaccharide 6SL to P40/P90

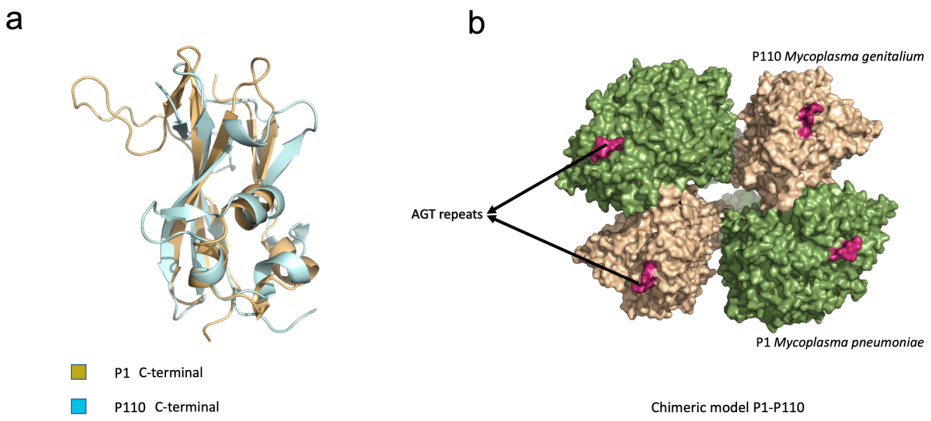
Ribbon representation of the P40/P90N structure in complex with oligosaccharide 6SL. The inset shows a detail of the residues shaping the binding pocket. The electron density omit map (at 0.9 sigma) corresponding to 6SL is also shown.


a
0.9
1.9
3.7
7.5
Response (RU)
Time (s)
K_D = 2.9 µM
Response (RU)
Concentration (M)
b
1
2
4
16
32
Response (RU)
Time (s)
K_D = 8.0 µM
Response (RU)
Concentration (M)



Supplementary Figure 8. Binding of oligosaccharides 3SL and 6SL to P40/P90

Sensorgrams (upper panels) and affinity plots (lower panels), obtained by Surface Plasmon Resonance (SPR), for the binding to protein P40/P90 of **a)** 6SL **b)** and 3SL. **c)** For protein P1 no binding was detected with 3SL. Source data are provided as a Source Data file.



Supplementary Figure 9. Structural relationships between P1 and P110

a) Superposition of the C-terminal domains from P1 and from the *M.genitalium* non-orthologous adhesin P110. **b)** Modelling of a chimeric Nap with P1 and P110, from the structure determined in *M. genitalium*. AGT repeats are colored in warm pink. Similarities between the two structures reinforces the idea of a common ancestor and the possibility of a four-fold symmetry organization in the Nap precursor.

Adherence inhibiting monoclonal antibodies

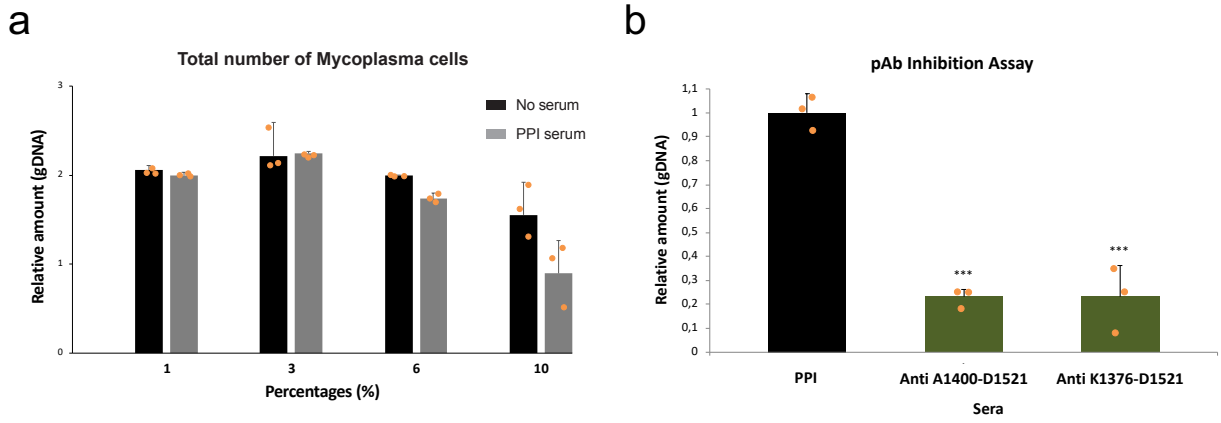
mAb	Epitopes	References
aN	NAINPRLTPWTYRN	
P1.26	NALSFTNK DVVGVGRL	
P1.62	EVKKKSDS ENHTKFTS	5
M58	SVNPKMVR	
M51	NEQSLGLR	
6E7	GIVRTPLAELLDG	4

Patient's sera

	Imunodominant epitopes	Reference
Patient's sera	WIGNGYRY FTDFVKPR	6

Supplementary Figure 10. Reported epitopes for P1

Mapping of epitopes was generated using data from references [4-6]



Supplementary Figure 11. Toxicity and adhesion inhibition of Mycoplasma cells

a) Relative amount of genomic DNA from *M. pneumoniae* with respect to the increasing percentages of PPI serum. No serum is the control value for non-toxicity. Sera concentrations of 6% show some toxicity that increases at 10%. **b)** Binding inhibition observed in cells when incubated with 10% serum generated against the P1 C-terminal domain short and long constructs Ala1400-Asp1521 and Lys1376-Asp1521, respectively. Results have been normalized with PPI value to minimize the effect of toxicity. Error bars represent the mean ± standard deviation. *** $p < 0{,}001$. Source data are provided as a Source Data file.

Supplementary Table 1. X-ray Data collection and Refinement statistics

Data collection[a]	P1	P40/P90	P40P90-3SL	P40P90-6SL
Space Group	C 1 2 1	P 21 21 21	C 1 2 1 *	C 1 2 1 *
Cell dimensions				
a,b,c (Å)	132.90, 115.64, 96.34	97.8,0 114.43, 165.05	116.12, 107.31, 162.03	116.33, 107.38, 160.45
α, β, γ (º)	90.00, 117.67, 90.00	90.00, 90.00, 90.00	90.00, 90.12, 90.00	90.00, 90.10, 90.00
Unique reflections	92405 (13355)	54524 (7870)	35571 (5211)	471848 (6819)
Resolution (Å)	72.20-1.94 (2.04-1.94)	97.8-2.65 (2.79-2.65)	81.01-3.10 (3.27-3.10)	80.22-2.80 (2.95-2.80)
Wavelength (Å)	0.97897	0.97927	0.97926	0.97926
Rmeas (%) [b]	0.143 (1.505)	0.145 (1.872)	0.149 (1.281)	0.121 (1.324)
I/Isigma	6.9 (1.1)	8.6 (1.0)	5.3 (1.0)	7.4 (1.0)
CC (½)	0.993 (0.658)	0.996 (0.621)	0.989 (0.315)	0.996 (0.252)
Completness (%)	96.5 (95.7)	99.8 (99.7)	97.4 (97.9)	97.0 (96.6)
Redundancy	3.5 (3.6)	6.5 (6.3)	3.1 (3.1)	2.9 (2.9)
Refinement statistcs[a]				
Resolution	72.20-1.94 (1.99-1.94)	82.53-2.65 (2.67-2.65)	81.01-3.10 (3.12-3.10)	80.22-2.80 (2.82-2.80)
Num. reflections	92369 (6370)	54428 (1044)	35571 (661)	47182 (913)
R_{Cryst} (%) [c]	18.7 (23.28)	21.4 (25.70)	24.40 (23.10)	20.90 (23.35)
R_{free} (%) [d]	22.9 (25.69)	23.4 (32.60)	27.90 (24.75)	26.50 (25.02)
Model Composition				
Nº Residues [e]	1339	1618	1613	1622
Nº Ligands	0	0	2	2
Solvent content (%)	40.0	44.4	57.4	56.6
R.m.s deviations				
Bond lengths (Å^2)	0.01	0.008	0.008	0.01
Bond angles (º)	1.15	1.08	1.03	1.19
B factor (Å²)				
Protein	60.7	89.5	147.2	96.1
Solvent	43.5	59.7	-	-
Ligand	-	-	245.0 **	116.5
Validation				
MoltProbity score	2.26	2.18	2.56	2.58
Clashscore	3.70	2.50	4.40	5.50
Poor rotamers (%)	6.23	9.50	10.20	13.00
Ramachandran plot				
Favored (%)	91.9	93.9	89.0	91.4
Allowed (%)	5.6	4.7	9.1	7.0
Disallowed (%)	2.5	1.4	1.9	1.6

[a] Values in parentheses correspond to the highest resolution Shell.

[b] $R_{sym} = \Sigma_{hkl} \Sigma_i I_i(hkl) - \langle I(hkl) \rangle / \Sigma_{hkl} \Sigma_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of an observation and $\langle I(hkl) \rangle$ is the mean value of observations for an unique reflection.

[c] $R_{cryst} = \Sigma_h F_0(h) - F_c(h) / \Sigma_h F_0(h)$, where F_o and F_C are the observed and calculated structure-factors amplitudes, respectively.

[d] R_{free} was calculated with 5% of data, which was excluded from refinement.

[e] Number of residues by unit cell.

[*] P40/P90 complexes presented a pseudorthorhombic packing within a monoclinic C2 space group.

[**] *B* factor obtained after refinement with ligand occupancy 1.0. The *B* factor decreased to 142.3 $Å^2$ for a ligand occupancy of 0.7.

Supplementary Table 2. Cryo-EM Data collection and Refinement statistics

Data collection and processing	P1
Magnification	96,000
Voltage (kV)	300
Electron exposure ($e^-/Å^2$)	60
Defocus range (μm)	-0.75 to -2.75
Pixel size (Å)	0,69
Symmetry imposed	C1
Initial particle images (nº)	4,574,424
Final particle images (nº)	68,014
Map Resolution	2.88
FSC threshold	0.143
Map resolution range (Å)	2.8 to 4.0
Refinement statistcs	
Initial model used	6RC9 (X-ray structure)
Model resolution (Å)	2.94
FSC threshold	0.5
Model resolution range (Å)	2.8 to 4.0
Map sharpening *B* factor ($Å^2$)	-89.4
Model composition	
Nonhydrogen atoms	9532
Protein residues	1219
N° Ligands	-
B factor ($Å^2$)	59.175
R.m.s deviations	
Bond lengths ($Å^2$)	0.005
Bond angles (°)	0.769
Validation	
MoltProbity score	2.45
Clashscore	7.37
Poor rotamer (%)	7.02
Ramachandran plot	
Favored (%)	93.96
Allowed (%)	6.04
Disallowed (%)	0.00

Supplementary Table 3. Synthetic DNA for P1

ATGCACCAGACCAAGAAAACCGCGCTGAGCAAGAGCACCTGGATCCTGATTCTGACCGCGACCGCGAGCCTGGCGACCGGTCTGACCGTGGTTGGTCACTTTACCAGCACCACCACCACC
CTGAAACGTCAGCAATTCAGCTATACCCGTCCGGACGAAGTTGCGCTGCGTCACACCAACGCGATCAACCCGCGTCTGACCCCGTGGACCTACCGTAACACCAGCTTCAGCAGCCTGCCG
CTGACCGGCGAGAACCCGGGTGCGTGGGCGCTGGTGCGTGACAACAGCGCGAAGGGTATTACCGCGGGTAGCGGCAGCCAGCAAACCACCTATGATCCGACCCGTACCGAAGCGGCGCT
GACCGCGAGCACCACCTTTGCGCTGCGTCGTTACGACCTGGCGGGTCGTGCGCTGTATGATCTGGACTTCAGCAAACTGAACCCGCAGACCCCGACCCGTGATCAAACCGGTCAGATCAC
CTTCAACCCGTTTGGTGGCTTCGGTCTGAGCGGTGCGGCGCCGCAGCAATGGAACGAGGTTAAAAACAAGGTGCCGGTTGAAGTGGCGCAAGACCCGAGCAACCCGTACCGTTTTGCGGT
TCTGCTGGTGCCGCGTAGCGTGGTTTACTATGAACAGCTGCAACGTGGTCTGGGCCTGCCGCAGCAACGTACCGAGAGCGGTCAGAACACCAGCACCACCGGTGCGATGTTCGGCCTGAA
AGTGAAGAACGCGGAAGCGGATACCGCGAAAAGCAACGAGAAGCTGCAGGGTGCGGAAGCGACCGGTAGCAGCACCACCAGCGGTAGCGGTCAAAGCACCCAGCGTGGTGGCAGCAGC
GGTGACACCAAAGTTAAGGCGCTGAAGATCGAGGTGAAGAAAAAGAGCGATAGCGAAGACAACGGCCAGCTGCAACTGGAGAAAAACGATCTGGCGAACGCGCCGATTAAGCGTAGCG
AGGAAAGCGGTCAAAGCGTTCAGCTGAAAGCGGACGATTTTGGTACCGCGCTGAGCAGCAGCGGTAGCGGTGGCAACAGCAACCCGGGTAGCCCGACCCCGTGGCGTCCGTGGCTGGCG
ACCGAGCAGATCCACAAAGACCTGCCGAAGTGGAGCGCGAGCATCCTGATTCTGTACGATGCGCCGTATGCGCGTAACCGTACCGCGATTGATCGTGTTGACCACCTGGATCCGAAGGCG
ATGACCGCGAACTACCCGCCGAGCTGGCGTACCCCGAAATGGAACCACCACGGTCTGTGGGACTGGAAGGCGCGTGATGTGCTGCTGCAGACCACCGGCTTCTTTAACCCGCGTCGTCAC
CCGGAATGGTTTGACGGTGGCCAAACCGTTGCGGATAACGAGAAGACCGGTTTTGATGTGGACAACAGCGAAAACACCAAACAAGGCTTCCAGAAAGAGGCGGATAGCGACAAAAGCG
CGCCGATTGCGCTGCCGTTTGAGGCGTATTTCGCGAACATTGGTAACCTGACCTGGTTTGGCCAGGCGCTGCTGGTTTTCGGTGGCAACGGTCATGTTACCAAGAGCGCGCACACCGCGCC
GCTGAGCATTGGTGTTTTTCGTGTGCGTTACAACGCGACCGGTACCAGCGCGACCGTTACCGGCTGGCCGTATGCGCTGCTGTTCAGCGGTATGGTGAACAAACAGACCGACGGCCTGAA
GGATCTGCCGTTTAACAACAACCGTTGGTTCGAATACGTTCCGCGTATGGCGGTGGCGGGTGCGAAATTTGTTGGCCGTGAGCTGGTGCTGGCGGGTACCATTACCATGGGTGACACCGCG
ACCGTTCCGCGTCTGCTGTATGATGAGCTGGAAAGCAACCTGAACCTGGTGGCGCAGGGTCAAGGCCTGCTGCGTGAAGACCTGCAGCTGTTCACCCCGTACGGTTGGGCGAACCGTCCG
GATCTGCCGATTGGCGCGTGGAGCTCTTCTAGCAGCAGCAGCCACAACGCGCCGTACTATTTTCACAACAACCCGGACTGGCAAGATCGTCCGATCCAGAACGTGGTTGACGCGTTCATTA
AACCGTGGGAGGATAAAAACGGCAAGGACGATGCGAAGTACATCTATCCGTACCGTTATAGCGGCATGTGGGCGTGGCAAGTTTATAACTGGAGCAACAAGCTGACCGACCAGCCGCTG
AGCGCGGATTTTGTGAACGAAAACGCGTACCAGCCGAACAGCCTGTTCGCGGCGATTCTGAACCCGGAGCTGCTGGCGGCGCTGCCGGACAAGGTTAAGTACGGCAAAGAGAACGAATT
TGCGGCGAACGAGTATGAACGTTTCAACCAAAAGCTGACCGTGGCGCCGACCCAGGGTACCAACTGGAGCCACTTCAGCCCGACCCTGAGCCGTTTTAGCACCGGTTTCAACCTGGTTGG
CAGCGTGCTGGACCAGGTTCTGGATTATGTGCCGTGGATCGGTAACGGCTACCGTTATGGTAACAACCACCGTGGCGTGGACGATATTACCGCGCCGCAAACCAGCGCGGGTAGCAGCAG
CGGCATTAGCACCAACACCAGCGGTAGCCGTAGCTTTCTGCCGACCTTCAGCAACATCGGTGTTGGCCTGAAAGCGAACGTGCAAGCGACCCTGGGTGGCAGCCAGACCATGATTACCGG
TGGCAGCCCGCGTCGTACCCTGGACCAAGCGAACCTGCAGCTGTGGACCGGTGCGGGCTGGCGTAACGACAAAGCGAGCAGCGGTCAGAGCGATGAAAACCACACCAAGTTTACCAGCG
CGACCGGTATGGACCAGCAGGGTCAGAGCGGTACCAGCGCGGGTAACCCGGACAGCCTGAAACAGGATAACATCAGCAAGAGCGGTGATAGCCTGACCACCCAAGACGGCAACGCGAT
TGATCAGCAAGAGGCGACCAACTACACCAACCTGCCGCCGAACCTGACCCCGACCGCGGACTGGCCGAACGCGCTGAGCTTTACCAACAAGAACAACGCGCAACGTGCGCAGCTGTTCCT
GCGTGGTCTGCTGGGCAGCATCCCGGTTCTGGTGAACCGTAGCGGTAGCGACAGCAACAAATTCCAAGCGACCGATCAAAAGTGGAGCTACACCGACCTGCACAGCGATCAGACCAAAC
TGAACCTGCCGGCGTATGGTGAAGTTAACGGCCTGCTGAACCCGGCGCTGGTTGAAACCTACTTTGGTAACACCCGTGCGGGTGGCAGCGGTAGCAACACCACCAGCAGCCCGGGTATCG
GCTTCAAAATTCCGGAACAGAATAACGACAGCAAAGCGACCCTGATTACCCCGGGCCTGGCGTGGACCCCGCAGGATGTTGGTAACCTGGTGGTTAGCGGCACCACCGTGAGCTTTCAAC
TGGGTGGCTGGCTGGTTACCTTTACCGACTTCGTGAAGCCGCGTGCGGGTTACCTGGGCCTGCAACTGACCGGTCTGGATGCGAGCGATGCGACCCAACGTGCGCTGATTTGGGCGCCGCG
TCCGTGGGCGGCGTTCCGTGGTAGCTGGGTTAACCGTCTGGGCCGTGTTGAGAGCGTGTGGGATCTGAAAGGTGTGTGGGCGGACCAGGCGCAAAGCGATAGCCAAGGTAGCACCACCAC
CGCGACCCGTAACGCGCTGCCGGAACACCCGAACGCGCTGGCGTTCCAAGTTAGCGTGGTTGAGGCGAGCGCGTATAAACCGAACACCAGCAGCGGTCAGACCCAAAGCACCAACAGCA
GCCCGTACCTGCACCTGGTTAAGCCGAAAAAGGTGACCCAGAGCGACAAACTGGACGATGACCTGAAGAACCTGCTGGATCCGAACCAAGTGCGTACCAAACTGCGTCAGAGCTTTGGT
ACCGACCACAGCACCCAGCCGCAACCGCAGAGCCTGAAGACCACCACCCCGGTTTTCGGTACCAGCAGCGGCAACCTGAGCAGCGTGCTGAGCGGTGGCGGTGCGGGCGGTGGCAGCAG
CGGTAGCGGCCAGAGCGGTGTTGATCTGAGCCCGGTTGAAAAGGTGAGCGGTTGGCTGGTTGGTCAACTGCCGAGCACCAGCGACGGTAACACCAGCAGCACCAACAACCTGGCGCCGA
ACACCAACACCGGCAACGATGTGGTTGGTGTTGGCCGTCTGAGCGAGAGCAACGCGGCGAAAATGAACGATGACGTTGACGGTATCGTTCGTACCCCGCTGGCGGAGCTGCTGGATGGTG
AGGGTCAGACCGCGGATACCGGTCCGCAGAGCGTTAAATTCAAGAGCCCGGACCAGATCGATTTTAACCGTCTGTTCACCCACCCGGTGACCGACCTGTTTGATCCGGTTACCATGCTGGT
GTACGACCAATATATTCCGCTGTTCATCGATATTCCGGCGAGCGTGAACCCGAAAATGGTTCGTCTGAAGGTGCTGAGCTTTGACACCAACGAACAGAGCCTGGGTCTGCGTCTGGAGTTC
TTTAAACCGGATCAAGACACCCAGCCGAACAACAACGTTCAAGTGAACCCGAACAACGGTGACTTCCTGCCGCTGCTGACCGCGAGCAGCCAAGGTCCGCAGACCCTGTTTAGCCCGTTC
AACCAGTGGCCGGATTACGTTCTGCCGCTGGCGATCACCGTGCCGATCGTGGTTATTGTTCTGAGCGTGACCCTGGGTCTGGCGATCGGCATTCCGATGCACAAAAACAAGCAAGCGCTGA
AAGCGGGTTTTGCGCTGAGCAACCAGAAAGTTGACGTGCTGACCAAGGCGGTTGGTAGCGTGTTCAAGGAGATCATTAACCGTACCGGCATTAGCCAAGCGCCGAAACGTCTGAAGCAG
ACCAGCGCGGCGAAACCGGGTGCGCCGCGTCCGCCGGTTCCGCCGAAACCGGGCGCGCCGAAGCCGCCGGTGCAGCCGCCGAAAAAGCCGGCGTAA

Supplementary Table 4. Synthetic DNA for P40/P90

ATGAAGAGCAAACTGAAGCTGAAACGTTATCTGCTGTTCCTGCCGCTGCTGCCGCTGGGTACCCTGAGCCTGGCGAACACCTATCTGCTGCAAGATCACAACACCCTGACCCCGTACACCC
CGTTTACCACCCCGCTGAACGGTGGCCTGGATGTGGTTCGTGCGGCGCACCTGCACCCGAGCTATGAGCTGGTTGACTGGAAACGTGTTGGTGATACCAAACTGGTGGCGCTGGTTCGTAG
CGCGCTGGTGCGTGTTAAGTTCCAGGACACCACCAGCAGCGATCAAAGCAACACCAACCAGAACGCGCTGAGCTTTGACACCCAAGAAAGCCAGAAAGCGCTGAACGGTAGCCAAAGCG
GCAGCAGCGACACCAGCGGTAGCAACAGCCAGGATTTCGCGAGCTATGTGCTGATTTTTAAAGCGGCGCCGCGTGCGACCTGGGTTTTCGAGCGTAAGATCAAACTGGCGCTGCCGTACG
TGAAACAAGAAAGCCAGGGTAGCGGCGATCAGGGCAGCAACGGTAAAGGCAGCCTGTATAAGACCCTGCAAGACCTGCTGGTTGAGCAACCGGTTACCCCGTACACCCCGAACGCGGGT
CTGGCGCGTGTGAACGGTGTTGCGCAAGATACCGTTCACTTCGGTAGCGGCCAGGAAAGCAGCTGGAACAGCCAACGTAGCCAGAAGGGTCTGAAAAACAACCCGGGCCCGAAAGCGGT
TACCGGTTTTAAGCTGGACAAAGGCCGTGCGTATCGTAAGCTGAACGAGAGCTGGCCGGTGTACGAACCGCTGGACAGCACCAAAGAGGGTAAAGGCAAAGATGAAAGCAGCTGGAAG
AACAGCGAGAAAACCACCGCGGAAAACGACGCGCCGCTGGTTGGTATGGTTGGTAGCGGTGCGGCGGGTAGCGCGAGCAGCCTGCAGGGTAACGGCAGCAACAGCAGCGGTCTGAAAA
GCCTGCTGCGTAGCGCGCCGGTGAGCGTTCCGCCGAGCAGCACCAGCAACCAGACCCTGAGCCTGAGCAACCCGGCGCCGGTGGGCCCGCAAGCGGTGGTTAGCCAACCGGCGGGTGGC
GCGACCGCGGCGGTGAGCGTTAACCGTACCGCGAGCGATACCGCGACCTTCAGCAAGTACCTGAACACCGCGCAAGCGCTGCACCAGATGGGTGTGATCGTTCCGGGCCTGGAGAAATG
GGGTGGCAACAACGGTACCGGCGTGGTTGCGAGCCGTCAAGACGCGACCAGCACCAACCTGCCGCATGCGGCGGGTGCGAGCCAGACCGGTCTGGGTACCGGTAGCCCGCGTGAGCCGG
CGCTGACCGCGACCAGCCAACGTGCGGTTACCGTTGTTGCGGGTCCGCTGCGTGCGGGTAACAGCAGCGAAACCGACGCGCTGCCGAACGTGATTACCCAGCTGTATCACACCAGCACCG
CGCAACTGGCGTACCTGAACGGTCAGATCGTGGTTATGGGCAGCGACCGTGTTCCGAGCCTGTGGTACTGGGTGGTTGGCGAGGATCAAGAAAGCGGCAAAGCGACCTGGTGGGCGAAG
ACCGAGCTGAACTGGGGTACCGATAAGCAAAAACAGTTCGTGGAAAACCAGCTGGGCTTTAAGGACGATAGCAACAGCGACAGCAAAAACAGCAACCTGAAGGCGCAAGGTCTGACCC
AGCCGGCGTATCTGATTGCGGGCCTGGACGTGGTTGCGGATCACCTGGTGTTCGCGGCGTTTAAAGCGGGTGCGGTTGGCTATGACATGACCACCGATAGCAGCGCGAGCACCTACAACC
AAGCGCTGGCGTGGAGCACCACCGCGGGTCTGGACAGCGATGGTGGCTACAAGGCGCTGGTTGAAAACACCGCGGGTCTGAACGGCCCGATCAACGGTCTGTTCACCCTGCTGGATACCT
TTGCGTATGTGACCCCGGTTAGCGGCATGAAAGGTGGCAGCCAAAACAACGAGGAAGTGCAGACCACCTACCCGGTTAAGAGCGACCAGAAAGCGACCGCGAAGATTGCGAGCCTGATC
AACGCGAGCCCGCTGAACAGCTATGGTGACGATGGCGTGACCGTTTTCGATGCGCTGGGTCTGAACTTCAACTTTAAGCTGAACGAGGAACGTCTGCCGAGCCGTACCGACCAACTGCTG
GTGTACGGCATTGTTAACGAGAGCGAACTGAAAAGCGCGCGTGAGAACGCGCAGAGCACCAGCGACGATAACAGCAACACCAAGGTTAAATGGACCAACACCGCGAGCCACTACCTGCC
GGTGCCGTACTATTACAGCGCGAACTTTCCGGAGGCGGGTAACCGTCGTCGTGCGGAACAACGTAACGGCGTGAAGATCAGCACCCTGGAAAGCCAGGCGACCGATGGTTTCGCGAACA
GCCTGCTGAACTTTGGCACCGGTCTGAAGGCGGGTGTGGACCCGGCGCCGGTTGCGCGTGGCCACAAGCCGAACTATAGCGCGGTGCTGCTGGTTCGTGGTGGCGTGGTTCGTCTGAACTT
TAACCCGGACACCGATAAACTGCTGGACAGCACCGATAAGAACAGCGAGCCGATTAGCTTCAGCTACACCCCGTTTGGTAGCGCGGAAAGCGCGGTGGACCTGACCACCCTGAAAGATG
TTACCTATATCGCGGAGAGCGGCCTGTGGTTCTACACCTTTGACAACGGTGAAAAACCGACCTATGATGGCAAGCAGCAACAGGTTAAGAACCGTAAAGGTTACGCGGTGATTACCGTTA
GCCGTACCGGCATCGAGTTCAACGAAGATGCGAACACCACCACCCTGAGCCAGGCGCCGGCGGCGCTGGCGGTGCAAAACGGTATTGCGAGCAGCCAGGACGATCTGACCGGCATCCTG
CCGCTGAGCGACGAGTTCAGCGCGGTGATTACCAAGGACCAAACCTGGACCGGTAAAGTTGATATCTATAAGAACACCAACGGCCTGTTTGAGAAAGACGATCAGCTGAGCGAAAACGT
GAAGCGTCGTGACAACGGTCTGGTTCCGATCTACAACGAAGGTATCGTGGACATTTGGGGCCGTGTTGATTTTGCGGCGAACAGCGTTCTGCAGGCGCGTAACCTGACCGACAAGACCGT
GGATGAGGTTATTAACAACCCGGACATCCTGCAAAGCTTCTTTAAGTTCACCCCGGCGTTTGATAACCAGCGTGCGATGCTGGTTGGTGAAAAAACCAGCGACACCACCCTGACCGTTAA
GCCGAAAATTGAGTATCTGGACGGTAACTTCTACGGCGAAGATAGCAAAATCGCGGGTATTCCGCTGAACATTGATTTTCCGAGCCGTATCTTTGCGGGTTTTGCGGCGCTGCCGAGCTGG
GTGATCCCGGTGAGCGTTGGTAGCAGCGTTGGCATCCTGCTGATTCTGCTGATCCTGGGCCTGGGTATCGGCATTCCGATGTATAAGGTTCGTAAACTGCAAGACAGCAGCTTCGTGGATG
TTTTTAAGAAAGTGGACACCCTGACCACCGCGGTGGGTAGCGTTTACAAGAAAATCATTACCCAGACCAGCGTGATCAAGAAAGCGCCGAGCGCGCTGAAAGCGGCGAACAACGCGGCG
CCGAAAGCGCCGGTTAAACCGGCGGCGCCGACCGCGCCGCGTCCGCCGGTGCAGCCGCCGAAGAAAGCGTAA,

Supplementary Table 5. Primers used in this study

Primers	Sequence
P1F	AGGAGATATACCATGACCGTGGTTGGTCACTTTACC
P1R	GTGATGGTGATGTTTATCCGGCCACTGGTTGAACGG
P40P90F	AGGAGATATACCATGAGCCTGGCGAACACCTATCTGCTG
P40P90GR	GTGATGGTGATGTTTAACCAGACCGTTGTCACGACGC
P40P90R	GTGATGTGTATGTTTGCTCGGCACGCGCCGCAAAACC
P40P90xINS1F	GGTATGGTTGGTAGCGCGAGCGATACCGCGACC
P40P90xINS1R	CGCGGTATCGCTCGCGCTACCAACCATACCAAC
P40P90xINS1&2F	GTTCCGGGCCTGGAGAGAGCGAAACCGACGCG
P40P90xINS1&2R	GTCGGTTTCGCTGCTCTCCAGGCCCGGAACGATC
P40P90RA445SSF	ACCGCGACCAGCCAAAGTTCGGTTACCGTTGTTGCG
P40P90RA445SSR	CGCAACAACGGTAACCGAACTTTGGCTGGTCGCGGT
P40P90RA455SSF	GTTGCGGGTCCGCTGAGTTCGGGTAACAGCAGCGAAACC
P40P90RA455SSR	GGTTTCGCTGCTGTTACCCGAACTCAGCGGACCCGCAAC
C-Domain K1376F	AGGAGATATACCATGAAAATGAACGATGACGTTG
C-Domain A1400F	AGGAGATATACCATGGCGGATACCGGTCCGCAG

Supplementary References

1. Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic acids research* **42**, W320-W324 (2014).
2. Burke, D.F., Deane, C.M. & Blundell, T.L. Browsing the SLoop database of structurally classified loops connecting elements of protein secondary structure. *Bioinformatics* **16**, 513-9 (2000).
3. Layh-Schmitt, G. & Herrmann, R. Localization and biochemical characterization of the ORF6 gene product of the Mycoplasma pneumoniae P1 operon. *Infect Immun* **60**, 2906-13 (1992).
4. Dallo, S.F., Su, C.J., Horton, J.R. & Baseman, J.B. Identification of P1 gene domain containing epitope(s) mediating Mycoplasma pneumoniae cytoadherence. *J Exp Med* **167**, 718-23 (1988).
5. Gerstenecker, B. & Jacobs, E. Topological mapping of the P1-adhesin of Mycoplasma pneumoniae with adherence-inhibiting monoclonal antibodies. *J Gen Microbiol* **136**, 471-6 (1990).
6. Jacobs, E., Pilatschek, A., Gerstenecker, B., Oberle, K. & Bredt, W. Immunodominant epitopes of the adhesin of Mycoplasma pneumoniae. *J Clin Microbiol* **28**, 1194-7 (1990).

Discusión general

M. pneumoniae y M. genitalium son dos importantes patógenos humanos que pueden producir complicaciones como neumonía atípica e infecciones urogenitales, respectivamente [2,3,4]. El mecanismo de infección para estos patógenos se inicia mediante la adhesión a las células del huésped a través de la OT [22-24]. Concretamente, a través de la macroestructura molecular denominada Nap, que es un complejo formado por un dímero de heterodímeros de las adhesinas P1 y P40/P90 en *M. pneumoniae* o P140 y P110 en *M. genitalium* [26]. Durante años, varios estudios han sugerido que el principal mecanismo para la adhesión de *M. pneumoniae* a los ácidos siálicos 3SL y 6SL se realizaba mediante la adhesina P1 [36,37]. Otras moléculas que también intervienen en este proceso y que son nombradas "complejos de adhesión accesorios" son P30, HMW1 y P40/P90. Pero no solamente la OT está relacionada con la adhesión celular, sino también con el movimiento de *gliding* del micoplasma [37]. A tal efecto, P1 también se ve involucrada en el *gliding* ya que existen varios reportes que mencionan que se encuentran anticuerpos contra esta proteína que paran el movimiento [37]. Además, el aumento de la proporción de 6SL con respecto a 3SL hace disminuir la frecuencia de movimiento de los micoplasmas, dando a entender que puede existir algún tipo de especificidad por el sustrato [33]. Para aumentar la subsistencia en el huésped, el micoplasma ha desarrollado varias estrategias para poder evadir el sistema inmune, una de ellas es la variación antigénica. Tanto P1 como P40/P90 son proteínas que presentan variaciones en sus secuencias sugiriendo que este mecanismo podría ayudar en la evasión del ataque de los anticuerpos [11,12]. En general, todos los resultados reportados demuestran la importancia del Nap y sus componentes accesorios en procesos de adhesión, *gliding* e inmunológicos de *M. pneumoniae* y *M. genitalium*.

El objetivo principal del presente trabajo es la obtención de las estructuras de las proteínas que conforman el Nap y dilucidar la implicación de este complejo en los procesos de adhesión, motilidad y inmunogenicidad en *M. pneumoniae* y *M. genitalium.*

En este trabajo se han determinado las estructuras tanto de las adhesinas que componen el Nap en *M. pneumoniae*, P1 y P40/P90, como también de la estructura ortóloga a P1, P140 y de su dominio C, en *M. genitalium*. El análisis estructural nos permite observar la alta similitud que existe entre las adhesinas de *M. pneumoniae* y *M. genitalium* para P1 con P140 y P40/P90 con P110 con unos RMSDs de 1.4 Å y 1.2 Å, respectivamente. La homología estructural se fundamenta principalmente en la similitud de la organización de los dominios N y C y en la alta homología secuencial que existen entre las adhesinas ortólogas siendo de alrededor del 40% para ambas. A pesar de ello, existen diferencias entre las adhesinas de ambos micoplasmas, la más importante es las inserciones que se encuentran a lo largo de *M. pneumoniae*. En P1 estas inserciones se componen de varios *loops* desordenados de entre 5 a 20 residuos, mientras que en P40/P90 existen esencialmente dos grandes inserciones desordenadas con respecto a P110, denominadas S1 y S2, de alrededor de 70 residuos cada una. La mayoría de estas inserciones, presentes tanto en P1 como en P40/P90, forman parte de la variabilidad genética de *M. pneumoniae* denominada RepMPs. Con la ayuda de las estructuras y de los análisis del genoma de micoplasma, se pudieron localizar cada uno de los RepMPs en la estructura. Se observó que la mayoría se encuentran en la zona de la corona de ambas adhesinas, accesibles al solvente para poder ser diana del ataque de los anticuerpos y que gran parte de las zonas comprendidas en los RepMPs se encontraban variadas en las cepas clínicas reportadas hasta el momento. Además, el modelo de la partícula del Nap mostró que el sitio de unión a los azúcares está rodeado por estas regiones variables, dificultando la unión a los anticuerpos que pudieran bloquear la adhesión a las células del huésped. Toda la información recogida en este estudio acerca de los RepMPs sugiere que se trata de un sistema muy sofisticado de evasión del sistema inmune, donde su función reside en captar la atención del sistema inmune para posteriormente variar y dificultar el reconocimiento de los anticuerpos creados para esa(s) región(es).

Las adhesinas que componen el Nap tanto en *M. genitalium* como en *M. pneumoniae* poseen la misma organización estructural: corona, β-propeller y un dominio C. Además, las proteínas no ortólogas P1 y P110 comparten una repetición de serinas (AGT r*epeats*). En conjunto, parecería indicar que las adhesinas de *M. genitalium*

y *M. pneumoniae* puedan tener un ancestro común. La adhesina ancestral formaría un tetrámero anclado a membrana con cuatro subunidades iguales como ocurre en las neuroaminidasas. Esta adhesina ancestral pudo haber evolucionado diferenciándose en dos proteínas independientes, según las funciones que pudieran desempeñar cada una de ellas. Este mecanismo ocurre, por ejemplo, en la ATPasa/sintasa del Tipo-F, donde se observó la pérdida de uno de los sitios catalíticos a lo largo de la evolución [123]. Para finalizar el análisis estructural entre las adhesinas del Nap, cabe mencionar que la identidad de secuencia es baja para el dominio N, del 12%, aumenta en el dominio C, 17%, y es muy alta para las regiones citoplasmáticas, 43%. Esta diferencia en la identidad puede estar relacionada con el posible ancestro común, ya que las zonas con baja homología serían aquellas que más hayan evolucionado, puesto que contienen el/los sitio(s) funcional(es) específico(s) o las variaciones antigénicas, mientras que aquellas zonas con alta homología no necesitan evolucionar al no desempeñar una función específica para una de las dos adhesinas.

Un estudio realizado en *M. pneumoniae,* concluyó que en la formación del complejo P1-P40/P90 la subunidad P40 no está presente tras realizarse la escisión entre P40 y P90, puesto que la primera es liberada tras la formación del complejo [45]. Según nuestros resultados, este corte entre las dos subunidades no ocurre solamente *in vivo* sino también *in vitro*, generando dos motivos R-A (R_{445}A y R_{455}A) ubicados al final de la S2, que como resultado se obtienen dos subunidades: P40 y P90. El corte se debe muy probablemente a una autoproteólisis mediada por metal que tiene lugar una vez la proteína ha completado su plegamiento y se ha dirigido hacia la membrana. Pero en contra del estudio comentado al principio del párrafo, tanto los datos estructurales obtenidos para P40/P90 como el gran número de puentes de hidrógeno que existen entre las dos subunidades, indican que las subunidades P40 y P90 permanecen unidas comportándose como una única proteína según los análisis realizados en el programa PISA [124]. Sin embargo, la falta de enlaces covalentes entre las dos subunidades puede hacer que en condiciones rigurosas o desnaturalizantes se pudieran llegar a separar, explicando el porqué en los complejos realizados hasta la fecha solamente apareciera la subunidad que se ancla a membrana, P90, liberando P40 del complejo.

El mecanismo de adhesión en *M. pneumoniae* y *M. genitalium* se ha estudiado extensivamente a lo largo de los años. Desde que se demostró que anticuerpos generados en ratón contra P1 bloqueaban la adhesión, se ha aceptado que esta proteína pudiera ser la principal adhesina en *M. pneumoniae* [36-38]. En contra de esto, nuestro trabajo demuestra con datos tanto cristalográficos, como los obtenidos por resonancia de plasmón superficial (SPR), que la adhesina encargada de la unión con los ácidos siálicos presentes en la membrana del huésped es P40/P90 y no P1. Estos resultados refuerzan el reciente descubrimiento en *M. genitalium,* donde P110 es la encargada de la unión a los ácidos siálicos [42]. Pero a pesar de que P40/P90 sea la adhesina principal, no se descarta que tanto P1 como otras proteínas próximas al Nap como la P30 puedan ayudar en el proceso de adhesión.

Relativo al proceso de adhesión, se muestran dos estados del heterodímero P110-P140 y del Nap, en los cuales se puede observar un estado abierto donde el sitio de unión a azúcares es accesible, mientras que en el cerrado no. Con estos dos estados del Nap, hemos confeccionado un posible modelo del ciclo de adhesión (**Figura 11**) que seguiría tanto *M. genitalium* como *M. pneumoniae*. El ciclo de adhesión empezaría con un estado abierto del Nap dispuesto a unirse al azúcar (etapa I), este estado abierto lo ocasionaría la estructuración de las zonas citoplasmáticas. Una vez que el micoplasma contacta con la célula huésped, P40/P90 o P110 se unen a los ácidos siálicos presentes en las membranas (etapa II). Cuando el azúcar se introduce en el interior del bolsillo de unión, se producen unos cambios conformacionales (señal 1) alrededor de este bolsillo que causarán que las proteínas del complejo se aproximen entre ellas (etapa II). El acercamiento de las adhesinas provocará una señalización interna, la cual activará una modificación postraducional (señal 2) que producirá que el Nap pase de un estado abierto a uno cerrado, liberando el azúcar del bolsillo de unión (etapa III). Existen diferentes posibilidades para esta modificación postraducional, puede tratarse de una fosforilación en la zona citoplasmática de P1 por el complejo PrkC/PrpC [56,57], una catálisis de algún sustrato o un proceso covalente. Por último, la finalización de la modificación postraducional anterior (señal 3) ocasionará que el Nap vuelva a un estado abierto dispuesto a volver a empezar el ciclo.

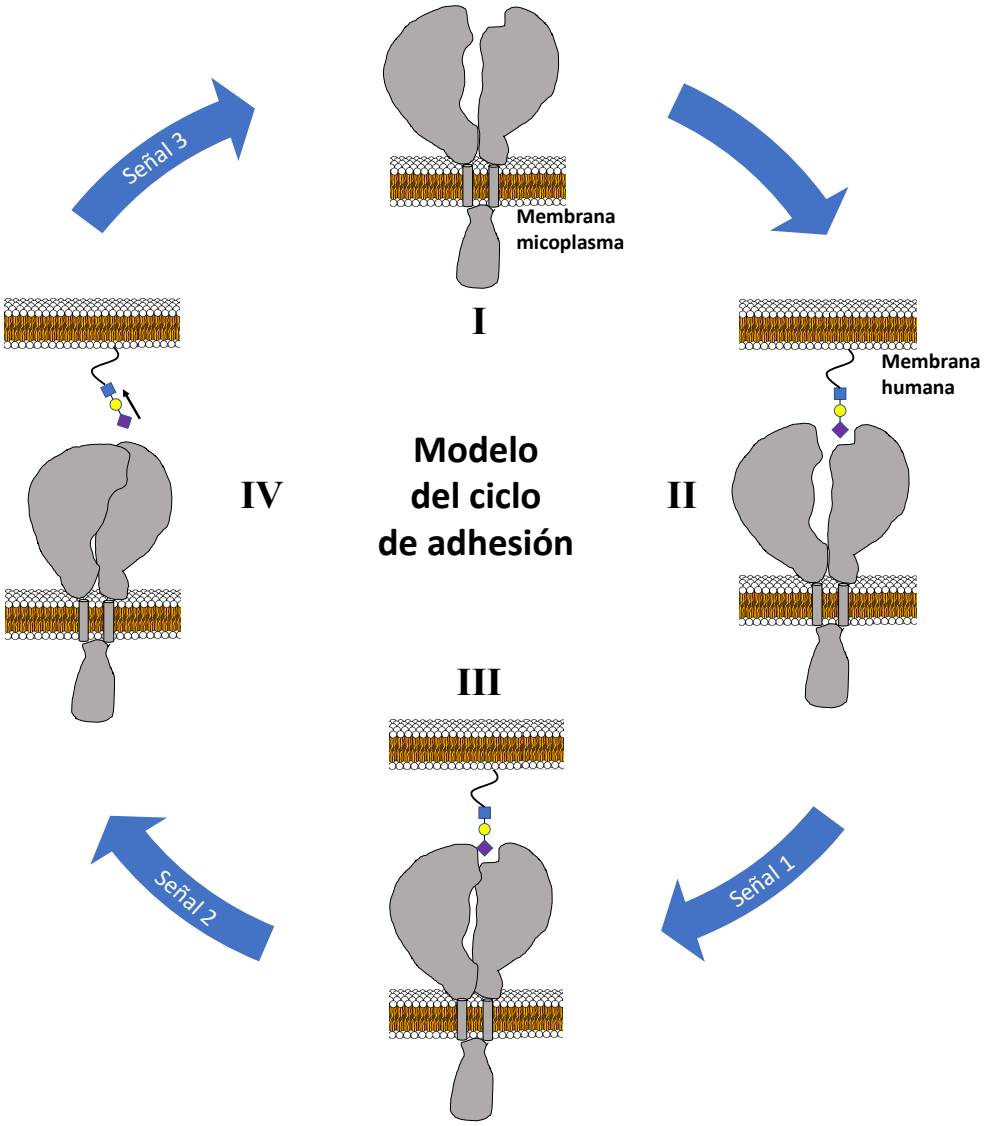


Fig. 11. Representación esquemática del ciclo de adhesión propuesto. Comienza con el Nap en un estado abierto (I) seguido de una unión a los azúcares sialilados a P40/P90 o P110 (II), lo que inicia algunos cambios conformacionales (señal 1) (III) que desencadenarían una modificación postraducional (señal 2) que lleva al Nap a un estado cerrado, donde se libera el azúcar (IV). Por último, la finalización de la modificación postraducional (señal 3) permite recuperar al Nap un estado abierto (I) completando el ciclo.

Estudios anteriores sugerían que el bloqueo de la adhesión y *gliding* se producía cuando los anticuerpos dirigidos hacia P1 tapaban el lugar de unión del receptor. Sin embargo, el ensayo de inhibición de adhesión en placa mostró que los anticuerpos de ratón se dirigen al dominio C de la proteína y a la hoja siete del β-propeller (denominada como péptido de citoadherencia), inhibiendo la adhesión de los micoplasmas a los ácidos siálicos. En contra de lo que se pensaba, estas zonas están bastante alejadas del

sitio de unión de azúcar, por lo que los anticuerpos inhiben la adhesión de una manera indirecta. Como se observa en la estructura de rayos X, este dominio C se conecta al dominio N mediante un *hinge* flexible que según los resultados de la crio-EM, le otorga cierta elasticidad entre los dos dominios. Además, la hoja siete del β-propeller no se encuentra expuesta en la partícula del Nap, por lo que en algún momento debe ser accesible para poder ser reconocido por los anticuerpos. Este hecho sugiere que durante el ciclo de adhesión y/o movimiento, el dominio C se puede desplazar con respecto al dominio N. Sin embargo, una vez el anticuerpo se une al dominio C, este no puede volver a su posición original, impidiendo la transición de un estado abierto a cerrado o viceversa, bloqueando una de las dos conformaciones. Por lo que, si el anticuerpo se une a un estado abierto y no deja transitar a un estado cerrado, el Nap se quedará continuamente en la forma abierta impidiendo soltarse de la célula huésped y también impidiendo su movimiento. Por otro lado, si se une a un estado cerrado impedirá la transición hacia el estado abierto del Nap y ocasionará que este no se pueda unir a la célula huésped y se desprenda. Otro posible mecanismo interesante a tener en cuenta analizando la estructura del dominio C, se encuentra en el β-bulge que interrumpe la tercera hoja. La ubicación entre los dominios N y C le permite actuar como una especie de pistón cuando el *hinge* sufre algún tipo de desplazamiento con respecto al dominio C, llegando a interaccionar con el dominio N, como ocurre en P110 de *M. genitalium*.

El ensayo immunológico mediante ELISA con suero de pacientes infectados con *M. pneumoniae* corroboró la alta inmunogenicidad presentada por la proteína P1 con el 90% de reconocimiento [39]. El estudio también confirmó el entrecruzamiento que existe entre P1 y P140 de *M. genitalium,* probablemente debido a que las similitudes en secuencia y estructura entre ambas proteínas dan lugar a epítopos parecidos [121]. Las dos construcciones del dominio C, que producen inhibición de la adhesión, solo fueron reconocidas en un 30%. Pero, los sueros que se reconocen son aquellos que tienen las IgGs más elevadas (supuestamente los correspondientes a pacientes con infecciones pasadas o en estadios tardíos). Este bajo reconocimiento en los sueros junto a unos valores de IgMs bajos (pacientes en estadios tempranos de la infección) puede indicar que los epítopos no estén muy expuestos/accesibles, o bien que haya un enmascaramiento con otras zonas de la proteína que sean más inmunológicas. Los

resultados serológicos junto con los de inhibición de la adhesión, indican que estos constructos del dominio C podrían ser utilizados para generar anticuerpos neutralizantes y detener la infección inhibiendo la adhesión del micoplasma. Estas propiedades convierten al dominio C de P1 en un buen candidato para ser utilizado como un posible antígeno en la producción de vacunas.

P40/P90 también se puede considerar una de las proteínas más inmunogénicas de *M. pneumoniae*, ya que a pesar de tener pocos datos, los resultados obtenidos en los ensayos ELISA muestran una reacción inmunogénica similar a la de P1. Además, los mutantes de P40/P90 testados mostraron que la inserción S2 es fundamental para la inmunogenicidad de la proteína, puesto que su deleción provoca una reducción a la mitad de la señal inmunogénica. S2, además de ser muy inmunogénica, es donde se produce la proteólisis entre las subunidades P40 y P90 y está comprendida en las regiones RepMPs, siendo la que más varía en las cepas clínicas. Los resultados sobre S2 sugieren que esta inserción actúa como una especie de "cebo" para el sistema inmunitario del huésped. El proceso comenzaría con la proteólisis, haciendo que la inserción S2 se sobreexponga y abarcando una mayor parte de la superficie de la proteína, implicando más accesibilidad para el ataque de los anticuerpos. Una vez captada la atención de los anticuerpos, la eliminación de esta inserción en otras cepas de micoplasmas hace que los anticuerpos ya sintetizados para los epítopos comprendidos en la S2 no surtan efecto, aumentando la efectividad en el proceso de infección. Estos resultados abren la vía para más estudios inmunológicos con P40/P90 y sugieren que la exposición de la inserción S2 es un posible mecanismo de evasión del sistema inmune.

Otro de los temas más estudiados en el grupo de los micoplasmas del *cluster Pneumoniae* es el *gliding* o movimiento por deslizamiento. Hasta la fecha, se han caracterizado varios sucesos importantes para este proceso, en particular: la fuente de energía es el ATP en *M. gallisepticum* [55] y en *M. mobile* [125], las fosforilaciones por el sistema PrkC/PrpC están involucradas en el movimiento [56,57] e incluso se han llegado a modelar dos posibles sistemas de movimiento descritos en el apartado de introducción.

Los resultados aportados no proporcionan una solución directa al *gliding*, pero sí que nos permiten aportar algunas ideas:

I) Los datos estructurales del Nap nos permiten descartar que el modelo de *gliding* sea parecido al de *M.mobile*, ya que el Nap es una estructura que tiene un eje de simetría binaria perpendicular a la membrana del micoplasma y aunque las partículas estuvieran alineadas sería imposible darle una direccionalidad al movimiento, por lo que descartaría que el Nap actuase como la "pata" de *M.mobile*, pues la direccionalidad tendría que venir del interior del micoplasma.

II) Los dos estados del Nap relacionados con el ciclo de adhesión también tienen gran importancia en el movimiento, ya que el ciclo de unión-desunión a los azúcares es crítico para que el micoplasma pueda desplazarse, puesto que debe existir una coordinación para que el movimiento se produzca una vez liberado el azúcar del Nap.

III) Aunque el Nap no sea el principal motor del micoplasma, por sí sola la transición entre una estructura abierta y cerrada va a proporcionar un movimiento rotatorio.

IV) Como hemos comentado anteriormente, el dominio C de P1 posee cierta flexibilidad que puede jugar un papel en el desarrollo del movimiento o del ciclo de adhesión. Los anticuerpos contra esta región también producen una inhibición del movimiento, bloqueando la plasticidad otorgada por este dominio C e impidiendo desarrollar correctamente el ciclo de adhesión y/o el movimiento.

Todas estas ideas junto con las aportadas en estudios previos permiten completar o aportar información extra al modelo de oruga para micoplamas del *cluster Pneumoniae*.

Los resultados obtenidos en este trabajo junto a los conocimientos previos permiten una mejor comprensión del funcionamiento de la partícula Nap en los procesos de adhesión y deslizamiento en *M. pneumoniae* y *M. genitalium*. También, el

mapeo de las variables genéticas y clínicas dentro de la estructura de P1 y P40/P90 explica uno de los principales mecanismos para la evasión del sistema inmune. Por último, los resultados obtenidos con el dominio C pueden ayudar a la confección de vacunas o terapias para combatir las infecciones de *M. pneumoniae* y *M. genitalium*.

Conclusiones

Conformación alternativa del dominio C de la proteína P140 de *Mycoplasma genitalium*

I) En solución el dominio C de P140 se encuentra en un estado monomérico, aunque para la formación de los cristales es necesario que el N-terminal del dominio se despliegue para poder interaccionar con otras moléculas, originando un octámero con simetría de grupo puntual D4 que permita generar la red cristalina.

II) La estructura del dominio C de P140 está compuesta por un N-terminal extendido seguido de dos hélices y tres hebras β antiparalelas, donde la tercera de ellas es interrumpida por un β-bulge.

III) El β-bulge puede participar en los cambios conformacionales experimentados en P140 durante el ciclo de adhesión o *gliding*.

Estructura y mecanismo del complejo de adhesión Nap del patógeno humano *Mycoplasma genitalium*

I) La estructura de P140 está compuesta por un gran dominio N (1243 residuos) y el dominio C (108 residuos). El dominio N tiene un *core* compuesto por un β-propeller de siete hojas β donde cada una de ellas está formada por cuatro hebras salvo la hoja siete, que tiene dos. En la zona más alejada a la membrana se encuentra la corona compuesta por varios *loops*, la mayoría desordenados.

II) En la comparación entre las dos estructuras de los dos dominios C de P140 se observa una plasticidad estructural que puede ser esencial para el funcionamiento del Nap. Este dominio podría actuar como una "pata" flexible anclada a la membrana, permitiendo la adherencia y *gliding*.

III) El Nap es un complejo compuesto por un dímero del heterodímero P140-P110. Los estados abierto y cerrado del Nap obtenidos a través de la tomografía de *M. genitalium* corresponderían al punto inicial y final de un ciclo de adhesión.

Proteínas inmunodominantes P1 y P40/P90 del patógeno humano *Mycoplasma pneumoniae*

I) La estructura de P1 presenta una alta homología estructural (1.4 Å) y secuencial (40%) con respecto a P140 de *M. genitalium*. La mayor diferencia entre ambas estructuras se encuentra en varias inserciones generalmente desordenadas repartidas a lo largo de P1.

II) La estructura de P40/P90N presenta una gran homología estructural (1.2 Å) y secuencial (44%) con respecto a P110 de *M. genitalium.* Como ocurre en P140 y P1, P40/P90N está compuesta por un dominio N, formado por una corona y un β-propeller de siete hojas β constituidas cada una por cuatro hebras salvo la hoja uno, que tiene tres.

III) P40/P90 es la proteína del Nap que se encarga directamente de la adhesión al ácido siálico. El sitio de unión está compuesto por una horquilla β con un motivo característico de unión a compuestos sialidados X-Y/F-T/S. Las afinidades obtenidas por SPR son semejantes entre los dos compuestos testados: 3SL y 6SL.

IV) La mayoría de las regiones variables RepMPs corresponden a *loops* expuestos distribuidos en la zona superior del Nap, principalmente alrededor del sitio de unión al receptor celular, proporcionando protección frente al sistema inmune del huésped.

V) La escisión de P40/P90 se trata de un proceso muy probablemente autoproteolítico dependiente de metal. En el WT corta en R^{445} | A^{446} y en el mutante $RA^{445}SS$ en R^{455} | A^{456}. Este corte otorgaría a la S2 una mayor exposición. A pesar del corte, ambas subunidades (P40 y P90) permanecen unidas a través de una interfaz muy extensa con 25 puentes de hidrógeno y 2 puentes salinos.

VI) Las similitudes estructurales de las adhesinas que conforman el Nap, tanto en *M. genitalium* como *M. pneumoniae,* plantea que ambas puedan provenir de un ancestro común y que a lo largo del tiempo se hayan diferenciado según la función que desempeñan dentro del micoplasma.

VII) Pruebas en sueros de pacientes dieron como resultado una alta inmunogenicidad para P1 y una actividad cruzada con P140, confirmando la similitud de los epítopos entre P1 y P140. El dominio C pierde inmunogenicidad con respecto a P1, pero es solamente reconocido por sueros con IgGs altas, sugiriendo que los anticuerpos pueden ser neutralizantes.

VIII) Pruebas en sueros de pacientes dieron como resultado una respuesta inmunológica alta para P40/P90, considerándola como una de las proteínas más inmunológicas de *M. pneumoniae*. La eliminación de la inserción S2 reduce en un 50% la respuesta inmunológica, desapareciendo principalmente la señal en aquellos sueros con IgMs altas.

IX) La alta inmunogenicidad junto a la variabilidad genética y la proteólisis que presenta la inserción S2, sugiere que pueda actuar como una especie de "cebo", exponiéndose hacia el exterior y atrayendo la atención de los anticuerpos, convirtiendo la S2 en un mecanismo para evadir el sistema inmune.

X) Los anticuerpos generados contra constructos que incluyen el dominio C y la hoja β siete de P1, inhiben la adhesión de *M. pneumoniae*. Estos resultados, junto con la conservación del dominio y la presencia de anticuerpos en sueros con IgGs elevadas, convierten a este dominio en un buen candidato para el desarrollo de vacunas.

XI) La inhibición de la adhesión y motilidad por anticuerpos contra el dominio C de P1 junto a la flexibilidad mostrada por este dominio, indica que los movimientos que se producen entre los dominios C y N son críticos para llevar a cabo la adhesión y el movimiento de *M. pneumoniae*.

Bibliografía

1. Wolf M, Müller T, Dandekar T, Pollack JD. Phylogeny of Firmicutes with special reference to Mycoplasma (Mollicutes) as inferred from phosphoglycerate kinase amino acid sequence data. *Int J Syst Evol Microbiol*. 2004;54(Pt 3):871-875.

2. Thomas Prescott Atkinson, Mitchell F. Balish, Ken B. Waites, Epidemiology, clinical manifestations, pathogenesis and laboratory detection of *Mycoplasma pneumoniae* infections, *FEMS Microbiology Reviews*, Volume 32, Issue 6, November 2008, Pages 956–973.

3. Sethi S, Singh G, Samanta P, et al. *Mycoplasma genitalium*: an emerging sexually transmitted pathogen. *Indian J Med Res* 2012; 136:942–55.

4. Cassell, G.H., Waites, K.B., Crouse, D.T., 1991. Perinatal mycoplasmal infections. *Clin. Perinatol*. 18, 241–262.

5. Ryan KJ, Ray CG (editors) (2004). *Sherris Medical Microbiology* (4th ed. edición). McGraw Hill. pp. 409-12.

6. Razin S, Argaman M, Avigan J. Chemical composition of mycoplasma cells and membranes. *J Gen Microbiol*. 1963;33:477-487.

7. Baseman JB, Tully JG. Mycoplasmas: sophisticated, reemerging, and burdened by their notoriety. *Emerg Infect Dis*. 1997;3(1):21-32.

8. Kahane I, Horowitz S. Adherence of mycoplasma to cell surfaces. *Subcell Biochem*. 1993;20:225-241.

9. Sun G, Xu X, Wang Y, Shen X, Chen Z, Yang J. *Mycoplasma pneumoniae* infection induces reactive oxygen species and DNA damage in A549 human lung carcinoma cells. *Infect Immun*. 2008;76(10):4405-4413.

10. H.Waites KB. New concepts of *Mycoplasma pneumoniae* infections in children. *Pediatr Pul-monol.* 2003;36 (4): 267-78.

11. R. Himmelreich, H. Hilbert, H. Plagens, E. Pirkl, B.-C. Li, R. Herrmann Complete sequence analysis of the genome of the bacterium *Mycoplasma pneumoniae Nucl. Acids Res*., 24 (1996), pp. 4420-4449.

12. Fraser CM, Gocayne JD, White O, et al. The minimal gene complement of Mycoplasma genitalium. *Science*. 1995;270(5235):397-403.

13. Pollack JD. Mycoplasma genes: a case for reflective annotation. *Trends Microbiol*. 1997 Oct;5(10):413-9.

14. Razin S, Yogev D, Naot Y. 1998. Molecular biology and pathogenicity of mycoplasmas. Microbiol. *Mol. Biol*. Rev. 62:1094–1156.

15. Großhennig S, Schmidl SR, Schmeisky G, Busse J, Stülke J. Implication of glycerol and phospholipid transporters in *Mycoplasma pneumoniae* growth and virulence. *Infect Immun*. 2013;81(3):896-904.

16. Chandonia JM, Kim SH, Brenner SE. Target selection and deselection at the Berkeley Structural Genomics Center. *Proteins*. 2006;62(2):356-370.

17. Dajani AS, Clynde WA Jr, Denny FW. Experimental infection with *Mycoplasma pneumoniae* (Eaton's agent). *J Exp Med*. 1965;121(6):1071-1086.

18. Eaton MD, Meiklejohn G, van Herick W, Corey M. Studies on the etiology of primary atypical pneumonia: II. Properties of the virus isolated and propagated in chick embryos. *J Exp Med*. 1945;82(5):317-328.

19. Chanock RM. *Mycoplasma pneumoniae*: proposed nomenclature for atypical pneumonia organism (Eaton agent). *Science* 1963; 140: 662.

20. Hutchison CA, Peterson SN, Gill SR, et al. Global transposon mutagenesis and a minimal Mycoplasma genome. *Science*. 1999;286(5447):2165-2169.

21. Tully J G, Taylor-Robinson D, Cole R M, Rose D L. A newly discovered mycoplasma in the human urogenital tract. *Lancet* 1981; 1: 1288-91.

22. Krause, D.C. and Balish, M.F. (2001), Structure, function, and assembly of the terminal organelle of *Mycoplasma pneumoniae*. *FEMS Microbiology Letters*, 198: 1-7.

23. Regula JT, Boguth G, Görg A, et al. Defining the mycoplasma 'cytoskeleton': the protein composition of the Triton X-100 insoluble fraction of the bacterium *Mycoplasma pneumoniae* determined by 2-D gel electrophoresis and mass spectrometry. *Microbiology (Reading)*. 2001;147(Pt 4):1045-1057.

24. Henderson GP, Jensen GJ. Three-dimensional structure of *Mycoplasma pneumoniae's* attachment organelle and a model for its role in gliding motility. *Mol Microbiol*. 2006;60(2):376-385.

25. Nakane D, Kenri T, Matsuo L, Miyata M. Systematic Structural Analyses of Attachment Organelle in Mycoplasma pneumoniae. *PLoS Pathog*. 2015;11(12):e1005299. Published 2015 Dec 3.

26. Burgos, R, Pich OQ, Ferrer-Navarro M, Baseman JB, Querol E, Piñol J. *Mycoplasma genitalium* P140 and P110 cytadhesins are reciprocally stabilized and required for cell adhesion and terminal-organelle development. *J Bacteriol* 188, 8627-37 (2006).

27. Layh-Schmitt, G. & Herrmann, R. Spatial arrangement of gene products of the P1 operon in the membrane of *Mycoplasma pneumoniae*. *Infect Immun*62, 974–979 (1994).

28. Baseman JB, Cole RM, Krause DC, Leith DK. Molecular basis for cytadsorption of *Mycoplasma pneumoniae*.*J Bacteriol*. 1982 Sep; 151(3):1514-22.

29. Manchee RJ, Taylor-Robinson D. Utilization of neuraminic acid receptors by mycoplasmas. *J Bacteriol*. 1969;98(3):914-919.

30. Sobeslavsky O, Prescott B, Chanock RM. Adsorption of *Mycoplasma pneumoniae* to neuraminic acid receptors of various cells and possible role in virulence. *J Bacteriol*. 1968 Sep; 96(3):695-705.

31. Varki A, Schauer R. Sialic Acids. In: Varki A, Cummings RD, Esko JD, et al., editors. Essentials of Glycobiology. 2nd edition. Cold Spring Harbor (NY): *Cold Spring Harbor Laboratory Press*; 2009. Chapter 14.

32. Wang B, Brand-Miller J. The role and potential of sialic acid in human nutrition. *Eur J Clin Nutr*. 2003 Nov;57(11):1351-69.

33. Kasai T, Nakane D, Ishida H, Ando H, Kiso M, Miyata M. Role of binding in Mycoplasma mobile and *Mycoplasma pneumoniae* gliding analyzed through inhibition by synthesized sialylated compounds. *J Bacteriol*. 2013 Feb; 195(3):429-35.

34. Lodish H, Berk A, Zipursky SL, et al. Molecular Cell Biology. 4th edition. *New York: W. H. Freeman*; 2000. Section 17.4, Translocation of Secretory Proteins across the ER Membrane.

35. Buchan DWA, Jones DT (2019). The PSIPRED Protein Analysis Workbench: 20 years on. Nucleic Acids Research.

36. Dallo SF, Su CJ, Horton JR, Baseman JB. Identification of P1 gene domain containing epitope(s) mediating *Mycoplasma pneumoniae* cytoadherence. *J Exp Med*. 1988;167(2):718-723.

37. S. Seto, T. Kenri, T. Tomiyama, M. Miyata. Involvement of P1 adhesin in gliding motility of *Mycoplasma pneumoniae* as revealed by the inhibitory effects of antibody under optimized gliding conditions *J. Bacteriol*., 187 (2005), pp. 1875-1877.

38. Svenstrup HF, Nielsen PK, Drasbek M, Birkelund S, Christiansen G. Adhesion and inhibition assay of *Mycoplasma genitalium* and *Mycoplasma pneumoniae* by immunofluorescence microscopy. *J Med Microbiol*. 2002;51(5):361-373.

39. Drasbek M, Nielsen PK, Persson K, Birkelund S, Christiansen G. Immune response to *Mycoplasma pneumoniae* P1 and P116 in patients with atypical pneumonia analyzed by ELISA. *BMC Microbiol*. 2004;4:7.

40. Su CJ, Chavoya A, Baseman JB. Regions of *Mycoplasma pneumoniae* cytadhesin P1 structural gene exist as multiple copies. *Infect Immun*. 1988 Dec;56(12):3157-61.

41. Ye Q, Mao JH, Shu Q, Shang SQ. *Mycoplasma pneumoniae* induces allergy by producing P1-specific immunoglobulin E. *Ann Allergy Asthma Immunol*. 2018;121(1):90-97.

42. Aparicio D, Torres-Puig S, Ratera M, Querol E, Piñol J, Pich OQ, Fita I. *Mycoplasma genitalium* adhesin P110 binds sialic-acid human receptors. *Nat Commun*. 2018 Oct 26;9(1):4471.

43. Layh-Schmitt, G. & Herrmann, R. Localization and biochemical characterization of the ORF6 gene product of the *Mycoplasma pneumoniae* P1 operon. *Infect Immun* 60, 2906-13 (1992).

44. Widjaja, M., Berry, I.J., Pont, E.J., Padula, M.P. & Djordjevic, S.P. P40 and P90 from Mpn142 are Targets of Multiple Processing Events on the Surface of *Mycoplasma pneumoniae*. *Proteomes* 3, 512-537 (2015).

45. Nakane D, Adan-Kubo J, Kenri T, Miyata M. Isolation and characterization of P1 adhesin, a leg protein of the gliding bacterium Mycoplasma pneumoniae. *J Bacteriol*. 2011;193(3):715-722.

46. Dallo SF, Lazzell AL, Chavoya A, Reddy SP, Baseman JB. Biofunctional domains of the *Mycoplasma pneumoniae* P30 adhesin. *Infect Immun*. 1996 Jul;64(7):2595-601.

47. Baseman JB, Morrison-Plummer J, Drouillard D, Puleo-Scheppke B, Tryon VV, Holt SC. Identification of a 32-kilodalton protein of *Mycoplasma pneumoniae* associated with hemadsorption. *Isr J Med Sci*. 1987 May;23(5):474-9.

48. Romero-Arroyo CE, Jordan J, Peacock SJ, Willby MJ, Farmer MA, Krause DC. *Mycoplasma pneumoniae* protein P30 is required for cytadherence and associated with proper cell development. *J Bacteriol*. 1999 Feb;181(4):1079-87.

49. Layh-Schmitt G., H. Hilbert, and E. Pirkl: A spontaneous hemadsorption-negative mutant of *Mycoplasma pneumoniae* exhibits a truncated adhesin-related 30-kilodalton protein and lacks the cytadherence-accessory protein HMW1. *J Bacterio*l 177, 843-846 (1995).

50. Burgos R, Pich OQ, Querol E, Piñol J. Functional analysis of the *Mycoplasma genitalium* MG312 protein reveals a specific requirement of the MG312 N-terminal domain for gliding motility. *J Bacteriol*. 2007;189(19):7014-7023.

51. Hasselbring BM, Jordan JL, Krause DC. Mutant analysis reveals a specific requirement for protein P30 in *Mycoplasma pneumoniae* gliding motility. *J Bacteriol*. 2005 Sep;187(18):6281-9.

52. McBride MJ. Bacterial gliding motility: multiple mechanisms for cell movement over surfaces. *Annu Rev Microbiol*. 2001;55:49-75.

53. Miyata M. Centipede and inchworm models to explain Mycoplasma gliding. *Trends Microbiol*. 2008 Jan;16(1):6-12.

54. Miyata M, Nakane D. 2013. Gliding mechanism of *Mycoplasma pneumoniae* subgroup implication from Mycoplasma mobile, p 237–252. *In* Browning G, Citti C (ed), Molecular and cell biology of Mollicutes. *Horizon Press*, Norfolk, United Kingdom.

55. Mizutani M, Miyata M. Behaviors and Energy Source of *Mycoplasma gallisepticum* Gliding. *J Bacteriol*. 2019;201(19):e00397-19. Published 2019 Sep 6.

56. Page CA, Krause DC. Protein kinase/phosphatase function correlates with gliding motility in *Mycoplasma pneumoniae*. *J Bacteriol*. 2013;195(8):1750-1757.

57. Schmidl SR, Gronau K, Hames C, Busse J, Becher D, Hecker M, Stülke J. The stability of cytadherence proteins in *Mycoplasma pneumoniae* requires activity of the protein kinase PrkC. *Infect Immun*. 2010 Jan;78(1):184-92.

58. Miyata M, Hamaguchi T. Prospects for the gliding mechanism of *Mycoplasma mobile*. *Curr Opin Microbiol*. 2016;29:15-21.

59. Nakane D, Miyata M. Cytoskeletal "jellyfish" structure of *Mycoplasma mobile*. *Proc Natl Acad Sci U S A.* 2007;104(49):19518-19523.

60. Adan-Kubo J, Uenoyama A, Arata T, Miyata M. Morphology of isolated Gli349, a leg protein responsible for *Mycoplasma mobile* gliding via glass binding, revealed by rotary shadowing electron microscopy. *J Bacteriol*. 2006;188(8):2821-2828.

61. Nonaka T, Adan-Kubo J, Miyata M. Triskelion structure of the Gli521 protein, involved in the gliding mechanism of *Mycoplasma mobile. J Bacteriol*. 2010 Feb;192(3):636-42.

62. Uenoyama A, Miyata M. Identification of a 123-kilodalton protein (Gli123) involved in machinery for gliding motility of *Mycoplasma mobile*. *J Bacteriol*. 2005;187(16):5578-5584.

63. Ohtani N, Miyata M. Identification of a novel nucleoside triphosphatase from *Mycoplasma mobile*: a prime candidate motor for gliding motility. *Biochem J*. 2007;403(1):71-77.

64. Uenoyama A, Seto S, Nakane D, Miyata M. Regions on Gli349 and Gli521 protein molecules directly involved in movements of *Mycoplasma mobile* gliding machinery, suggested by use of inhibitory antibodies and mutants. *J Bacteriol*. 2009;191(6):1982-1985.

65. Seybert A, Gonzalez-Gonzalez L, Scheffer MP, Lluch-Senar M, Mariscal AM, Querol E, Matthaeus F, Piñol J, Frangakis AS. Cryo-electron tomography analyses of terminal organelle mutants suggest the motility mechanism of *Mycoplasma genitalium. Mol Microbiol*. 2018 May;108(3):319-329.

66. Miyata M, Hamaguchi T. Integrated Information and Prospects for Gliding Mechanism of the Pathogenic Bacterium *Mycoplasma pneumoniae*. *Front Microbiol*. 2016;7:960. Published 2016 Jun 28.

67. Burgos R, Pich OQ, Querol E, Piñol J. Deletion of the *Mycoplasma genitalium* MG_217 gene modifies cell gliding behaviour by altering terminal organelle curvature. *Mol Microbiol*. 2008 Aug;69(4):1029-40.

68. Qin L, Chen Y, You X. Subversion of the Immune Response by Human Pathogenic Mycoplasmas. *Front Microbiol*. 2019;10:1934. Published 2019 Aug 21.

69. Cusick MF, Libbey JE, Fujinami RS. Molecular mimicry as a mechanism of autoimmune disease. *Clin Rev Allergy Immunol*. 2012;42(1):102-111.

70. Jacobs E, Bartl A, Oberle K, Schiltz E. Molecular mimicry by *Mycoplasma pneumoniae* to evade the induction of adherence inhibiting antibodies. *J Med Microbiol*. 1995;43(6):422-429.

71. Deitsch KW, Lukehart SA, Stringer JR. Common strategies for antigenic variation by bacterial, fungal and protozoan pathogens. *Nat Rev Microbiol*. 2009;7(7):493–503.

72. Dorigo-Zetsma JW, Wilbrink B, Dankert J, Zaat SA. *Mycoplasma pneumoniae* P1 type 1- and type 2-specific sequences within the P1 cytadhesin gene of individual strains. *Infect Immun*. 2001;69(9):5612-5618.

73. Kapatais-Zoumbos K, Chandler DK, Barile MF. Survey of immunoglobulin A protease activity among selected species of Ureaplasma and Mycoplasma: specificity for host immunoglobulin A. *Infect Immun*. 1985;47(3):704-709.

74. Yamamoto T, Kida Y, Sakamoto Y, Kuwano K. Mpn491, a secreted nuclease of *Mycoplasma pneumoniae*, plays a critical role in evading killing by neutrophil extracellular traps. *Cell Microbiol*. 2017;19(3):10.1111/cmi.12666.

75. McGowin CL, Popov VL, Pyles RB. Intracellular *Mycoplasma genitalium* infection of human vaginal and cervical epithelial cells elicits distinct patterns of inflammatory cytokine secretion and provides a possible survival niche against macrophage-mediated killing. *BMC Microbiol*. 2009;9:139.

76. Lappin-Scott, H., Burton, S. & Stoodley, P. Revealing a world of biofilms — the pioneering research of Bill Costerton. *Nat Rev Microbiol* **12,** 781–787 (2014).

77. Kornspan JD, Tarshis M, Rottem S. Adhesion and biofilm formation of *Mycoplasma pneumoniae* on an abiotic surface. *Arch Microbiol*. 2011;193(11):833-836.

78. Meseguer MA, Alvarez A, Rejas MT, Sánchez C, Pérez-Diaz JC, Baquero F. *Mycoplasma pneumoniae*: A reduced-genome intracellular bacterial pathogen. *Infect Genet Evol*. 2003;3:47–55.

79. Yavlovich A, Tarshis M, Rottem S. Internalization and intracellular survival of Mycoplasma pneumoniae by non-phagocytic cells. FEMS Microbiol Lett. 2004 Apr 15;233(2):241-6.

80. Sauteur PM Meyer, Huber BM, Goetschel P. Neuroinvasive *Mycoplasma pneumoniae* infection without intrathecal antibody response. *Pediatr Infect Dis J*. 2012;31:1199–1200.

81. Schomburg J, Vogel M. A 12-year-old boy with severe mucositis: Extrapulmonary manifestation of *Mycoplasma pneumoniae* infection. *Klin Padiatr*. 2012;224:94–95.

82. Paddenberg R, Wulf S, Weber A, Heimann P, Beck LA, Mannherz HG. Internucleosomal DNA fragmentation in cultured cells under conditions reported to induce apoptosis may be caused by mycoplasma endonucleases. *Eur J Cell Biol*. 1996;71(1):105-119.

83. Elkhal CK, Kean KM, Parsonage D, Maenpuen S, Chaiyen P, Claiborne A, Karplus PA. Structure and proposed mechanism of L-α-glycerophosphate oxidase from *Mycoplasma pneumoniae*. *FEBS J*. 2015;282:3030–3042.

84. Almagor M, Yatziv S, Kahane I. Inhibition of host cell catalase by *Mycoplasma pneumoniae*: a possible mechanism for cell injury. *Infect Immun*. 1983;41(1):251-256.

85. Hickman-Davis JM, McNicholas-Bevensee C, Davis IC, Ma HP, Davis GC, Bosworth CA, *et al.* Reactive species mediate inhibition of alveolar type II sodium transport *during Mycoplasma* infection. *Am J Respir Crit Care Med* 2006;173:334-44.

86. Khoury T, Abu Rmeileh A, Kornspan JD, Abel R, Mizrahi M, Nir-Paz R. *Mycoplasma pneumoniae* Pneumonia Associated With Methemoglobinemia and Anemia: An Overlooked Association? [published correction appears in Open Forum Infect Dis. 2015 Jan;2(1):ofv041]. *Open Forum Infect Dis*. 2015;2(1):ofv022.

87. Hardy RD, Coalson JJ, Peters J, Chaparro A, Techasaensiri C, Cantwell AM, Kannan TR, Baseman JB, Dube PH. Analysis of pulmonary inflammation and function in the mouse and baboon after exposure to *Mycoplasma pneumoniae CARDS toxin. PLoS One. 2009;4:e7562.*

88. Techasaensiri C, Tagliabue C, Cagle M, Iranpour P, Katz K, Kannan TR, Coalson JJ, Baseman JB, Hardy RD. Variation in colonization, ADP-ribosylating and vacuolating cytotoxin, and pulmonary disease severity among *Mycoplasma pneumoniae* strains. *Am J Respir Crit Care Med*. 2010;182:797–804.

89. Somarajan SR, Al-Asadi F, Ramasamy K, Pandranki L, Baseman JB, Kannan TR. Annexin A2 mediates *Mycoplasma pneumoniae* community-acquired respiratory distress syndrome toxin binding to eukaryotic cells. *MBio*. 2014;5(pii):e01497-14.

90. Medina JL, Coalson JJ, Brooks EG, Le Saux CJ, Winter VT, Chaparro A, Principe MF, Solis L, Kannan TR, Baseman JB, Dube PH. *Mycoplasma pneumoniae* CARDS toxin exacerbates ovalbumin-induced asthma-like inflammation in BALB/c mice. *PLoS One*. 2014;9:e102613.

91. Johnson C, Kannan TR, Baseman JB. Cellular vacuoles induced by *Mycoplasma pneumoniae* CARDS toxin originate from Rab9-associated compartments. *PLoS One*. 2011;6:e22877.

92. Zuo LL, Wu YM, You XX. Mycoplasma lipoproteins and Toll-like receptors. *J Zhejiang Univ Sci B*. 2009;10(1):67-76.

93. Pang HX, Qiao HM, Cheng HJ, Zhang YF, Liu XJ, Li JZ. Levels of TNF-alpha, IL-6 and IL-10 in bronchoalveolar lavage fluid in children with *Mycoplasma pneumoniae* pneumonia. *Zhongguo Dang Dai Er Ke Za Zhi*. 2011;13:808–810.

94. Bose, S., Segovia, J. A., Somarajan, S. R., Chang, T. H., Kannan, T. R., and Baseman, J. B. (2014). ADP-ribosylation of NLRP3 by *Mycoplasma pneumoniae* CARDS toxin regulates inflammasome activity. *MBio* 5:e02186.

95. Shimizu T, Kimura Y, Kida Y, Kuwano K, Tachibana M, Hashino M, Watarai M. Cytadherence of *Mycoplasma pneumoniae* induces inflammatory responses through autophagy and toll-like receptor 4. *Infect Immun*. 2014;82:3076–3086.

96. Yano T, Ichikawa Y, Komatu S, Arai S, Oizumi K. Association of *Mycoplasma pneumoniae* antigen with initial onset of bronchial asthma. Am J Respir Crit Care Med. 1994;149:1348–1353.

97. Mok JY, Waugh PR, Simpson H. *Mycoplasma pneuminiae* infection. A follow-up study of 50 children with respiratory illness. *Arch Dis Child*. 1979;54:506–511.

98. Koh YY, Park Y, Lee HJ, Kim CK. Levels of interleukin-2, interferon-gamma, and interleukin-4 in bronchoalveolar lavage fluid from patients with *Mycoplasma pneumonia*: implication of tendency toward increased immunoglobulin E production. *Pediatrics*. 2001;107:E39.

99. Kraft M, Cassell GH, Henson JE, Watson H, Williamson J, Marmion BP, Gaydos CA, Martin RJ. Detection of *Mycoplasma pneumoniae* in the airways of adults with chronic asthma. *Am J Respir Crit Care Med*. 1998;158:998–1001.

100. Martin RJ, Kraft M, Chu HW, Berns EA, Cassell GH. A link between chronic asthma and chronic infection. *J Allergy Clin Immunol*. 2001;107:595–601.

101. Hong SJ. The Role of *Mycoplasma pneumoniae* Infection in Asthma. *Allergy Asthma Immunol Res*. 2012;4(2):59-61.

102. Xin LH, Wang J, Wang Z, Cheng W, Zhang W. Effect of *Mycoplasma pneumoniae* infection on function of T lymphocytes in bronchoalveolar lavage fluid of asthmatic children. *Zhongguo Dang Dai Er Ke Za Zhi*. 2014;16:277–280.

103. Kang YM, Ding MJ, Han YL, Wang SF, Ma X, Li H. Th1/Th2 immune response in bronchoalveolar lavage fluid in children with severe *Mycoplasma pneumoniae* pneumonia. Zhongguo Dang Dai Er Ke Za Zhi. 2011;13:188–190.

104. Lluch-Senar M, Querol E, Piñol J. Cell division in a minimal bacterium in the absence of ftsZ. *Mol Microbiol*. 2010;78(2):278-289.

105. Hasselbring BM, Jordan JL, Krause RW, Krause DC. Terminal organelle development in the cell wall-less bacterium *Mycoplasma pneumoniae*. *Proc Natl Acad Sci U S A*. 2006 Oct 31;103(44):16478-83.

106. Seto S, Layh-Schmitt G, Kenri T, Miyata M. Visualization of the attachment organelle and cytadherence proteins of *Mycoplasma pneumoniae* by immunofluorescence microscopy. *J Bacteriol*. 2001;183(5):1621-1630.

107. Espeli O, Lee C, Marians KJ. A physical and functional interaction between Escherichia coli FtsK and topoisomerase IV. *J Biol Chem*. 2003;278(45):44639-44644.

108. Lluch. M, Piñol. J, Querol. Enrique. Estudio de la expresión génica y la división celular en *Mycoplasma genitalium*. Universidad Autónoma de Barcelona, 2011.

109. Ferwerda, A., Moll, H.A. & de Groot, R. Respiratory tract infections by *Mycoplasma pneumoniae* in children: a review of diagnostic and therapeutic measures. *Eur J Pediatr* 160, 483-91 (2001).

110. Bajantri B, Venkatram S, Diaz-Fuentes G. *Mycoplasma pneumoniae*: A Potentially Severe Infection. *J Clin Med Re*s. 2018 Jul;10(7):535-544.

111. Khoury T, Sviri S, Rmeileh AA, Nubani A, Abutbul A, Hoss S, van Heerden PV. et al. Increased rates of intensive care unit admission in patients with *Mycoplasma pneumoniae*: a retrospective study. *Clin Microbiol Infect*. 2016;22(8):711–714.

112. Marrie TJ. *Mycoplasma pneumoniae* pneumonia requiring hospitalization, with emphasis on infection in the elderly. *Arch Intern Med*. 1993;153(4):488–494.

113. Amin D, McKitish K, Shah PS. Association of Mortality and Recent *Mycoplasma Pneumoniae* Infection in COVID-19 Patients [published online ahead of print, 2020 Aug 27]. *J Med Virol*. 2020;10.1002/jmv.26467.

114. Kenri T, Okazaki N, Yamazaki T, Narita M, Izumikawa K, Matsuoka M, et al. Genotyping analysis of *Mycoplasma pneumoniae* clinical strains in Japan between 1995 and 2005: type shift phenomenon of M. pneumoniae clinical strains. *J Med Microbiol*. 2008;57(Pt 4):469-475.

115. Onozuka D, Chaves LF. Climate variability and nonstationary dynamics of *Mycoplasma pneumoniae* pneumonia in Japan. *PLoS One*. 2014;9(4):e95447. Published 2014 Apr 16.

116. McGowin CL, Anderson-Smits C. *Mycoplasma genitalium*: an emerging cause of sexually transmitted disease in women. *PLoS Pathog*. 2011;7:e1001324.

117. Asenjo A, Kusters JG, Severs TT, Alós JI. *Mycoplasma genitalium* in Spain: prevalence of genital infection and frequency of resistance to macrolides. *Enferm Infecc Microbiol Clin*. 2018;36(3):169-171.

118. Anagrius C, Loré B, Jensen JS. *Mycoplasma genitalium*: prevalence, clinical significance, and transmission. *Sex Transm Infect*. 2005;81(6):458-462.

119. Gatski M, Martin DH, Theall K, et al. *Mycoplasma genitalium* infection among HIV-positive women: prevalence, risk factors and association with vaginal shedding. *Int J STD AIDS*. 2011;22(3):155-159.

120. Morozumi M, Takahashi T, Ubukata K. Macrolide-resistant *Mycoplasma pneumoniae*: characteristics of isolates and clinical aspects of community-acquired pneumonia. *J Infect Chemother*. 2010;16(2):78–86.

121. Morrison-Plummer, J., Lazzell, A. & Baseman, J.B. Shared epitopes between *Mycoplasma pneumoniae* major adhesin protein P1 and a 140-kilodalton protein of *Mycoplasma genitalium*. *Infection and Immunity* 55, 49-56 (1987).

122. Feizi T, Taylor-Robinson D. Cold agglutinin anti-I and *Mycoplasma pneumoniae*. *Immunology*. 1967;13(4):405-409.

123. Guo, H., Suzuki, T. & Rubinstein, J. L. Structure of a bacterial ATP synthase. *Elife* 8, e43128 (2019).

124. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol*. 372, 774–797 (2007).

125. Nishikawa MS, Nakane D, Toyonaga T, Kawamoto A, Kato T, Namba K, Miyata M.
126. Refined Mechanism of *Mycoplasma mobile* Gliding Based on Structure, ATPase Activity, and Sialic Acid Binding of Machinery. *mBio*. 2019 Dec 24;10(6):e02846-19.