Cell surface processing of the P1 adhesin of *Mycoplasma pneumoniae* identifies novel domains that bind host molecules

Michael Widjaja1, Iain James Berry1, Veronica Maria Jarocki1, Matthew Paul Padula3, Roger Dumke2,4 & Steven Philip Djordjevic1,3,4*

*Mycoplasma pneumoniae* is a genome reduced pathogen and causative agent of community acquired pneumonia. The major cellular adhesin, P1, localises to the tip of the attachment organelle forming a complex with P40 and P90, two cleavage fragments derived by processing Mpn142, and other molecules with adhesive and mobility functions. LC-MS/MS analysis of *M. pneumoniae* M129 proteins derived from whole cell lysates and eluents from affinity matrices coupled with chemically diverse host molecules identified 22 proteoforms of P1. Terminomics was used to characterise 17 cleavage events many of which were independently verified by the identification of semi-tryptic peptides in our proteome studies and by immunoblotting. One cleavage event released 1597-TSAAKPGAPPRPVPKPGAPPKGPPVQPPKKPA1627 from the C-terminus of P1 and this peptide was shown to bind to a range of host molecules. A smaller synthetic peptide comprising the C-terminal 15 amino acids, 1613-PGAPKPPVQPPKKPA1627, selectively bound cytoskeletal intermediate filament proteins cytokeratin 7, cytokeratin 8, cytokeratin 18, and vimentin from a native A549 cell lysate. Collectively, our data suggests that ectodomain shedding occurs on the surface of *M. pneumoniae* where it may alter the functional diversity of P1, Mpn142 and other surface proteins such as elongation factor Tu via a mechanism similar to that described in *Mycoplasma hyopneumoniae*.

The attachment organelle is a structurally and functionally sophisticated component of the *M. pneumoniae* cell that is responsible for the assembly of proteins essential for motility and adherence1–8. An extensive list of host molecules including fibronectin9–13, fibrinogen10–14, plasminogen11–13,15–17, lactoferrin10–12, laminin10–12, and vitronectin10–13 interact with surface accessible adhesins in *M. pneumoniae*. Other less well-defined host molecules include sialylated molecules18, oligosaccharides19, glycolipids20, and glycoproteins21.

The gene *mpn141* encoding the major adhesin P1 is located in the same operon along with *mpn140* and *mpn142* and these three genes constitute a polycistronic transcriptional unit22,23. *mpn140* encodes for a 28 kDa putative phosphoesterase24 and while it has been shown to degrade nanoRNA and dephosphorylate 3′-phosphoadenosine 5′-phosphate to AMP25, no role in adherence has been assigned for this protein. *mpn142* generates a 130 kDa product (Mpn142) that is cleaved into two fragments of 40 kDa (P40) and 90 kDa (P90) immediately after or concurrent with translation26,27. The cleavage event in Mpn142, first described over 25 years ago, was the first in what is now known to be a highly processed molecule on the surface of *M. pneumoniae*28. P1 is a remarkably versatile molecule and the subject of numerous studies over the past 30 years. The only cleavage event that has been accurately assigned to P1 is the removal of the N-terminal 59 amino acids as a leader peptide29. Molecular cross-linking and immunogold-labelling studies indicated that P1 forms a complex with P30, P40, and P9030,31 that colocalise to the tip of the attachment organelle to act in concert to effect different functions32. Cross-linking studies with paraformaldehyde identified P1 complexes containing Mpn309 (P65), Mpn272 (DnaK), C-terminal truncated forms of DnaK and P1, pyruvate dehydrogenase α subunit (Pdh-A), and...
high molecular weight proteins 1 (HMW1) and 3 (HMW3)\textsuperscript{33}. Anti-P1 antibodies reduce adherence of \textit{M. pneumoniae} to abiotic and host cell surfaces\textsuperscript{34,38} and \textit{M. pneumoniae} P1 mutants are also unable to adhere\textsuperscript{35-38,41}. For P1 to translocate to the surface, localise correctly within the attachment organelle and to maintain stability, interactions with accessory proteins P40, P90, HMW1, and Top1 are required\textsuperscript{22,46}. C-terminal regions of P1 have featured in various recombinant vaccines that seek to control infections caused by \textit{M. pneumoniae}.

The P1 adhesin is highly immunogenic and is often detected by sera from \textit{M. pneumoniae}-infected patients\textsuperscript{36,47,48}. Several studies have shown that the carboxyl half of P1 is highly immunogenic and crucial for its function as an adhesin\textsuperscript{37,49-53}. To identify regions within P1 that are recognised by the host humoral immune response, Schurwanz \textit{et al.} generated 15 recombinant fragments spanning the P1 molecule and exposed them to the serum of patients with \textit{M. pneumoniae} infections\textsuperscript{52}. Three recombinant fragments within P1, one in the N-terminus and two spanning C-terminal regions, were strongly immunoreactive with sera from greater than 90% of the patients\textsuperscript{55}. Guinea pig antibodies generated to one of the C-terminal regions significantly reduced binding of \textit{M. pneumoniae} to HBEC (primary bronchial epithelial), MRC-5 (fetal lung fibroblasts), and HeLa (cervical carcinoma) cell lines\textsuperscript{33}. These data informed the creation of a chimeric recombinant protein which included this carboxyl region of P1 and a region in the P30 adhesin. Antibodies raised against this chimeric protein reduced \textit{M. pneumoniae} adherence to human bronchial epithelial cells by more than 95%\textsuperscript{55}, and also successfully reduced \textit{M. pneumoniae} colonisation in animal models\textsuperscript{62}.

Here we sought to determine if P1 is processed on the surface of \textit{M. pneumoniae}. Tryptic peptides that mapped to different regions within P1 were frequently encountered when characterising size-fractionated eluents generated during affinity chromatography using different host molecules as bait. These peptides were mapped to the P1 molecule providing insight into the complex processing events that target this molecule. Precise cleavage sites were determined using an N-terminome approach\textsuperscript{52} and by mapping semi-trypsic peptides identified from our proteome studies. Naturally occurring cleavage fragments of P1 were identified by LC-MS/MS analysis i) because they bound to affinity resins loaded with host proteins, ii) by mapping tryptic peptides derived from proteins spots from 2D-SDS PAGE, and iii) by immunoblotting studies using serum raised against fifteen different regions of P1. These independently acquired, but complementary datasets enabled a rigorous assessment of cleavage events in the P1 adhesin. Finally, microtitre binding assays and microscale thermophoresis showed that the C-terminus of P1 binds various host molecules.

Methods and Materials

Strains. \textit{M. pneumoniae} (M129 strain, ATCC 29342) cells were cultured as described previously\textsuperscript{58}. Cells were grown in modified Hayflick’s medium in tissue culture flasks at 37 °C. Human lung carcinoma (A549, ATCC CCL-185) cells were cultured in RPMI 1640 medium (Invitrogen) supplemented with 10% heat inactivated fetal bovine serum. Cells were grown in tissue culture flasks at 37 °C with 5% CO\textsubscript{2}.

Cell preparation for one dimensional- and two dimensional-SDS polyacrylamide gel electrophoresis. \textit{M. pneumoniae} cells were harvested as described previously\textsuperscript{52}. In brief, cells were lysed with sonication in 7 M urea, 2 M thiourea, 40 mM Tris-HCl, and 1% (w/v) C7BzO detergent (Sigma) after washing with PBS. Proteins were reduced and alkylated with 5 mM tributylphosphine and 20 mM acrylamide monomers before precipitation with aceton. Protein was resuspended in 7 M urea, 2 M thiourea, and 1% (w/v) C7BzO for 1D- and 2D-SDS PAGE.

Gel electrophoresis was performed as described previously\textsuperscript{60,61}. Approximately 80 µg and 250 µg of protein was used for 1D- and 2D-SDS PAGE, respectively. Gels were fixed and stained by either Flamingo fluorescent gel stain (Bio-Rad) or Coomassie Blue G-250 (Sigma).

In-gel trypsin digestion was performed as described previously\textsuperscript{62} for mass spectrometry analysis. Gel pieces were excised, destained, dehydrated, and then incubated with trypsin Gold MS grade (Fromega) in 100 mM NH\textsubscript{4}HCO\textsubscript{3}. Tryptic peptides were extracted by sonication and stored in 4 °C until needed for mass spectrometry.

Liquid chromatography tandem mass spectrometry (LC-MS/MS) and data analysis. LC-MS/MS was performed as described previously\textsuperscript{61}. In brief, 5 µg of peptides in 15 µl was loaded into an Eksigent AS-1 autosampler connected to a Tempo nanolC system (Eksigent, Livermore, CA, USA) and washed onto a PicoFrit column (75 µm × 150 mm) packed with Magic C18AQ resin (Michrom Biosciences, CA). Peptides were eluted from the column into the source of a QSTAR Elite hybrid quadrupole-time-of-flight mass spectrometer (Sciex, Redwood, CA, USA).

Files generated from LC-MS/MS were searched against the MSPrn100 database\textsuperscript{61} with the following parameters: Fixed modifications: none; Variable modifications: propionamide, oxidized methionine, deamidation; Enzyme: semi-trypsin; Number of allowed missed cleavages: 3; Peptide mass tolerance: 0.2 Da; and Charge state: 2+. For samples collected from the ‘Surface proteome analysis of \textit{M. pneumoniae} (Biotinylation)’ and ‘Affinity chromatography host binding \textit{M. pneumoniae} complexes (A549)’ listed below, variable modifications also included NHS-LC-Biotin (K) and NHS-LC-Biotin (N-term). ‘Affinity chromatography host binding \textit{M. pneumoniae} complexes (A549)’ was also searched against ‘\textit{homo sapiens}’ entries in MSPrn100 to identify biotinylated surface A549 proteins.

Surface proteome analysis of \textit{M. pneumoniae}. Biotinylation of the \textit{M. pneumoniae} cells was performed as described previously\textsuperscript{68}. The biotinylation reaction was allowed to proceed for 30 seconds on ice. Biotinylated surface proteins were confirmed with western blots using ExtrAvidin-HRP (Sigma).

Tryptic shaving of \textit{M. pneumoniae} cells was carried out as described previously\textsuperscript{52}. Shaving was for 5 minutes at 37°C and released peptides were trypsin digested a second time before analysis by LC-MS/MS.
Affinity chromatography of host binding *M. pneumoniae* complexes. ‘Bait’ host proteins used for affinity chromatography include fibronectin (Code: 341635) and plasminogen (Code: 528175) from human plasma supplied by Merck Millipore. Bovine actin (Code: A3653) and fetuin (Code: F3004) was supplied by Sigma.

Affinity chromatography using host proteins bound to Avidin Agarose (Pierce) as ‘Bait’ was performed as described previously28. *M. pneumoniae* cells were lysed in 1% (w/v) C7BzO (Sigma-Aldrich) in PBS (pH 7.8) to obtain native complexes. The native complex cell lysate was incubated with host proteins bound to Avidin Agarose (‘Bait’). This mixture was washed with PBS and host protein binding complexes (‘Prey’) were eluted 7 M urea, 2 M thiourea, 40 mM Tris-HCl, and 1% (w/v) C7BzO. Elutions were separated by 1D-SDS PAGE and proteins were identified by LC-MS/MS as described above.

Affinity chromatography using human lung carcinoma (A549) surface proteins as ‘Bait’ was performed as described previously28. A549 cells were biotinylated, lysed, and bound to Avidin Agarose (‘Bait’). As above, this mixture was incubated with native *M. pneumoniae* complexes followed by washes and eluents to obtain a fraction of A549 binding complexes (‘Prey’).

Affinity chromatography using heparin HiTrap columns (GE Healthcare) was performed as described previously28. *M. pneumoniae* cells were lysed in 10mM sodium phosphate, 0.1% Triton TX-100 (pH 7.0) to obtain native complexes. Approximately 300 µg of soluble complexes was loaded onto a HiTrap Heparin HP column (GE Healthcare). The column was washed with 10 mM sodium phosphate (pH 7.0) and heparin binding complexes were sequentially eluted in increasing concentrations of sodium chloride (pH 7.0).

Dimethyl labelling of *M. pneumoniae* and LC-MS/MS analysis. Dimethyl labelling of *M. pneumoniae* proteins was carried out as described previously28,64. 1 mg of *M. pneumoniae* protein was labelled in 40 mM formaldehyde (Ultrapur grade, Polysciences), 20 mM sodium cyanoborohydride, 100 mM Hepes (pH 6.7) for 4 hours at 37°C. The reaction was quenched with 100 mM ammonium bicarbonate, precipitated in acetone-methanol (8:1), and digested with trypsin.

Peptides were analysed using both the Sciex 5600 and Thermo Scientific Q Exactive™ mass spectrometers. The methods, protocols, and parameters used have been described previously28.

Bioinformatic analysis of the P1 adhesin. Bioinformatic predictions and analysis was performed as described previously15. The bioinformatic tools used were: ProtParam65, TMpred66, PONDR® (VSL2 predictor)67, and ScanProsite68. Predicted glycosaminoglycan binding motifs searched in ScanProsite included binding sites for heparin (X-[HRK]-[HRK]-X-[HRK]-X motif)69, heparin sulfate (X-[HRK]-X-[HRK]-[HRK]-X)70, or clusters of basic amino acid residues (X-[HRK]-X(0,2)-[HRK]-X(0,2)-[HRK]-X and X-[HRK]-X(1,3)-[HRK]-X(1,3)-[HRK]-X).

Immunoblot of *M. pneumoniae* cell lysates using anti-P1 serum. 60 µg of *M. pneumoniae* cell lysate proteins were separated on 1D-SDS PAGE as described above. Proteins were transferred to PVDF (polyvinylidene fluoride) membranes using a semidry method71. Membranes were blocked with 5% (w/v) skim milk powder in PBS, and 0.1% (v/v) Tween 20 (PBS-Tween) for 1 hour at 25°C. Membranes were cut in to individual lanes and then separately probed with guinea pig sera raised against different regions of the P1 adhesin (guinea pig sera was generated in a previous study55) for 1.5 hours at 25°C in PBS-Tween. Membranes were washed three times over 30 minutes before being probed a second time in peroxidase-conjugated anti-guinea pig antibodies (1:3000, Sigma) for 1 hour at 25°C in PBS-Tween. Membranes were washed again three times over 30 minutes and developed with DAB tablets (3,3′-Diaminobenzidine, Sigma).

Binding of P1 C-terminus to human proteins in ELISA. Human proteins used for ELISA include: plasma fibrinogen (Code: F3879), plasma fibronectin (Code: 11051407001), Glu-plasminogen (Code: P7999), vitronectin (Code: SRP3186), laminin (Code: L6274), and lactoferrin (Code: L1294) which were all supplied by Sigma.

Binding affinity measured by ELISA was performed as described previously17. Recombinant protein RP15 was produced as described55 and both C-terminal peptides were synthesised by Chempeptide Limited (China). P1-30 (1597 TSAAKPGAPRPPVPPKPGAPKPPVQPPKKPA1627) without any tags, but P1-15 (1613 PGAPKPPVQPPKKPA1627) was sequenced with an N-terminal biotin tag.

15 µg/ml of C-terminal P1 fragments were bound to wells and incubated with different host proteins. Wells were then incubated with different antiserum raised against the different host proteins at the following dilutions (all from Sigma): anti-fibrinogen 1:3000, anti-fibronectin 1:1000, anti-plasminogen 1:2500, anti-vitronectin 1:5000, anti-laminin 1:750, and anti-lactoferrin 1:5000. These incubations were followed by incubations with anti-rabbit IgG (Dako) or anti-goat IgG (both 1:2,000). Detection was measured by adding Tetramethylbenzidine (all from Sigma): anti-fibrinogen 1:3000, anti-fibronectin 1:1000, anti-plasminogen: 1:2500, anti-vitronectin.

Binding of the P1 C-terminus to A549 human lung cells. Freshly grown A549 cells were immobilised in 96-well microtitre plates as described in17. Immobilised A549 cells were incubated with 10 µg/ml of either RP15, P1-30, or P1-15 and binding affinity was measured with antiserum raised against RP15 (1:100) as described above. Absorbance detection at 450 nm is the same as described above.

Affinity chromatography of complexes that bind the P1 C-terminus. The C-terminal sequence of P1 (P1-15) was synthesised with an N-terminal biotin tag by Chempeptide Limited (China). Affinity chromatography was performed similar to the section above. In brief, 1 mg of the peptide was added to Avidin Agarose beads for 16 h at 4°C. The beads were washed four times (5 ml per wash) with PBS before being incubated with native A549 cell lysates (harvested in 1% w/v C7BzO in PBS) for 16 h at 4°C. Non-binding proteins were washed from
Figure 1. Cleavage map of the P1 adhesin. The full length proteoform (1627 amino acids) is shown as the black bar with cleavage sites above and fragments below this bar. Cleavage sites identified from dimethyl labelling and semi-tryptic sites are shown as the blue and red arrows, respectively. Sequences where these cleavage sites occur are also shown. Putative heparin binding sites (Hep, blue boxes, motif: X-[HRK]-[HRK]-[HRK]-X), heparan sulfate binding sites (HepS, blue boxes, motif: X-[HRK]-X-[HRK]-[HRK]-X), clusters of basic residues (HepB, blue boxes, motifs: X-[HRK]-X(0,2)-[HRK]-X(0,2)-[HRK]-X or X-[HRK]-X(1,3)-[HRK]-X(1,3)-[HRK]-X), and transmembrane domains (TmD, yellow boxes, TmD1 predicted by TMpred66, TmD2 previously predicted32, TmD3 previously predicted36, and TmD4 previously predicted in a P1 paralog73) are shown within the black bar. Putative transmembrane domains and the location of 15 subregions of P1 (grey ‘RP’ boxes) expressed as recombinant proteins from an earlier study55 are shown. Predicted disordered regions appear as purple boxes in the grey bar. Acidic and basic regions within P1 are identified as yellow and blue bars, respectively. Peptides released from surface shaving experiments and identified by mass spectrometry are shown in the light green boxes within the grey bar. Grey bars represent fragments of P1 identified during SDS-PAGE of whole cell lysates. Red bars represent fragments of P1 recovered from lysates of *M. pneumoniae* that have their surface proteins labelled with biotin (surface exposed fragments of P1). Peptides identified by mass spectrometry of P1
Eluents were concentrated with a Macrosep® 3 kDa cutoff centrifugal device (Pall), precipitated with acetone, and separated by 1D-SDS PAGE. The whole lane was divided into sections, in-gel digested with trypsin, and analysed by LC-MS/MS as described above.

**Microscale thermophoresis of P1-15 binding affinity.** Binding affinities to fluorescent labelled host proteins was measured by microscale thermophoresis as described in 72. Microscale thermophoresis was set to 30 s and samples were scanned with 40%, 60% and 80% MST Power. Dissociation constants were determined from generated dissociation curves with set hot/cold or thermophoresis settings. As a control, a scrambled version of the C-terminal P1 peptide (PKPPRAAPPKAPTVPVPGASPVVKPKQAPG) was synthesised by Chempeptide Limited (China) without any tags and binding affinities was measured.

**Ethical approval.** Guinea pig sera used in this study was generated in a previous study55. The animal experiments in that previous study were proved by the ethical board of Landesdirektion Sachsen, Dresden, Germany (no. 24-9168.25-1).

**Results**

**Bioinformatic analysis of the P1 adhesin.** The P1 adhesin has a predicted mass of 176.3 kDa and a pI of 8.53 and contains six predicted transmembrane regions and nine putative glycosaminoglycan binding sites (Fig. 1). The first transmembrane region (spanning the N-terminus), and the last transmembrane region (spanning the C-terminus) have been identified in previous studies of P132,36,66, and a P1 paralog of Mycoplasma genitalium34. The glycosaminoglycan binding sites consist of reiterated copies of positively charged amino acids that are likely to be important in interactions with sulphated derivatives of heparin and heparan sulfate. Analysis of P1 using PONDR® identified seven putative disordered regions that span at least 30 amino acids (Fig. 1). Modules in P1 enriched in acidic (E, K) and basic (K, R, H) amino acids were identified. Disordered region and protein modules enriched in acidic and basic amino acids have been described in adhesin families in the respiratory pathogen *M. hyopneumoniae* and these were influential in the location of a subset of important cleavage sites56-62,74. We confirmed the precise location of 17 cleavage sites in P1 (shown below), 11 of which reside in predicted regions of disorder (Fig. 1). Cleavage sites did not seem to be over-represented in acidic or basic domains.

The **P1 adhesin is processed extensively on the M. pneumoniae cell surface.** P1 peptides identified by LC-MS/MS analyses of size fractionated *M. pneumoniae* lysates identified 23 proteoforms ranging in size from 17 to 176 kDa including the full length proteoform without the N-terminal signal sequence (Fig. 1). The full length and an additional 16 smaller proteoforms of P1 were identified by LC-MS/MS of size fractionated cell lysates separated by SDS-PAGE (grey bars, Fig. 1). The migration behaviour of these 17 proteoforms of P1 was consistent with masses predicted by ProtParam65. Trypsin shaving of the *M. pneumoniae* cell surface released trypsin accessible peptides (green boxes within a grey bar in Fig. 1) that span most of the adhesin indicating that P1 is exposed on the cell surface. This was consistent with LC-MS/MS analysis of size-fractionated biotinylated proteins that were first enriched using avidin chromatography which identified 14 proteoforms (full and fragments 2, 3, 5, 7, 10, 11, 13, 14, 16, 17, 20, 21, and 22) of P1 (red bars in Fig. 2). These data suggest that cleaved P1 proteoforms are surface accessible.

Several other proteoforms of P1 were identified by LC-MS/MS of protein bands digested in-gel from affinity experiments. Two proteoforms of P1 with masses of 149 kDa (fragment 1) and 101 kDa (fragment 3) (orange bars in Fig. 1) were identified from columns coupled with biotinylated A549 surface protein complexes suggesting that large P1 proteoforms with multiple binding domains are required to bind surface receptors on A549 cells. Eluents derived from columns coupled with fetuin and actin were particularly useful for identifying the full length protein and fragments 1, 2, 4, 7, 8, 10, 11, 16 and 18 of P1. Nine fragments (1, 2, 4, 7, 8, 10, 11, 16; yellow bars in Fig. 1) were recovered from columns coupled with fetuin, and six fragments (1, 4, 7, 8, 10, and 16; light blue bars in Fig. 1) were recovered from columns coupled with actin. Six fragments (2, 5, 12, 13, 14, 18) were identified from columns coupled with plasminogen (purple bars in Fig. 1). For the eleven fragments identified from heparin chromatography (blue bars in Fig. 1; fragments: 2, 4, 5, 6, 7, 9, 10, 12, 14, 15, and 16), only two (fragment 14 and 16) did not contain any of the predicted glycosaminoglycan binding motifs identified with ScanProsite. Fifteen fragments (2, 4, 5, 6, 7, 9, 11, 13, 14, 16, 17, 18, 19, 20, and 21) were identified in eluents from columns coupled with fibronectin (green bars in Fig. 1).

A global M. pneumoniae dimethyl labelling approach was used to identify internal neo-N termini. Ten cleavage sites were identified in P1 using this approach (Table 1, blue arrows in Fig. 1). Semi-trypptic peptides, defined as peptides with only one tryptic end (Table 1, red arrows in Fig. 1) were also identified, implying seven additional cleavage sites in P1. Four distinct sites in P1 showed evidence that surface accessible amino-peptidases may alter neo-N-termini (Fig. 1 and Table 1): 716NPFVGF, 719GFSFLS, 765NQK, 777TVP, 781TQG, 782ALW, 818PRP, 1179TQW (cleavage site 2), 517NQK, 647L, 758TVP, 765TQG (cleavage site 8), 1146TQW (cleavage site 12), and 1568AGF (cleavage site 16) in a manner that is similar to amino-peptidase processing events reported in the major adhesin families in *M. hyopneumoniae*57,64,74,75. A large predicted disorder region spanning 196 amino acids near the carboxyl terminal fragment of P1 has been identified which may be important for cell surface binding.
| No. | ID | Peptide Sequence                                      | Score | E-value   |
|-----|----|-------------------------------------------------------|-------|-----------|
|     |    | N-terminal dimethyl labelled peptides                  |       |           |
| 1   | N1 | R.73 NTSSSLPGTGENPGAWAVR58.5D                         | 26    | 0.039     |
| 3   | N3 | R.73 TSGQNTSTTGAMGLKGKVEATGSSSTSGSQ7QR157.5G         | 72    | 6.60E-07  |
| 5   | N4 | K.108 AMTANYPSWR158.5T                                 | 67    | 5.70E-13  |
| 6   | N5 | R.108 MAVAGAKFVGR158.5E                                | 62    | 5.00E-04  |
| 8   | N6 | K.135 LTVPQGNTWNHSFPTSLR168.5F                       | 121   | 7.00E-07  |
| 7   | N7 | L.135 TVAPQGNTWNHSFPTSLR168.5F                       | 77    | 8.10E-07  |
| 8   | N8 | T.135 VAPOTGNTWNHSFPTSLR168.5F                       | 64    | 1.00E-05  |
| 9   | N9 | P.135 TQTWNHSFPTSLR168.5F                             | 35    | 5.30E-04  |
| 10  | N10| T.135 RLNLPGYVNGLNPALVETYGNT178.5_A                  | 173   | 7.10E-08  |
| 11  | N11| W.178 LYTFTDFKPR178.5_A                                | 58    | 7.20E-09  |
| 12  | N12| T.178 EFDFKPR178.5_A                                  | 40    | 5.70E-10  |
| 13  | N13| R.182 ALWARPWAFAFR182.5_G                             | 36    | 1.20E-09  |
| 14  | N14| L.186 WAPRPWAFAFR182.5_G                              | 27    | 2.10E-10  |
| 15  | N15| E.186 ALSNQKVDVLTKAVGYSVEIK191.5_SNQKVDVLTKAVGYSVEIK191.5_T  | 160   | 1.70E-14  |
| 16  | N16| A.191 LSNQKVDVLTKAVGYSVEIK191.5_T                     | 158   | 1.30E-13  |
| 17  | N17| L.191 LSNQKVDVLTKAVGYSVEIK191.5_T                     | 184   | 3.90E-08  |
| 18  | N18| S.191 SNQKVDVLTKAVGYSVEIK191.5_T                      | 102   | 4.60E-08  |
| 2   | S1 | F.165 GFGLSGAAPQQWNEVKNKVPVEVAQDSPNYPR197.5_F        | 39    | 2.10E-07  |
| 3   | S2 | G.166 GFGLSGAAPQQWNEVKNKVPVEVAQDSPNYPR197.5_F        | 34    | 3.20E-06  |
| 4   | S3 | F.168 GFLSGAAPQQWNEVKNKVPVEVAQDSPNYPR197.5_F         | 42    | 0.037     |
| 5   | S4 | G.169 LSGAAPQQWNEVKNKVPVEVAQDSPNYPR197.5_F           | 59    | 5.00E-06  |
| 6   | S5 | S.171 GAAPQQWNEVKNKVPVEVAQDSPNYPR197.5_F             | 46    | 3.00E-04  |
| 7   | S6 | D.248 TAKSNEKLQGAETGSTTSSGQS7QR157.5_G                | 126   | 1.00E-05  |
| 8   | S7 | A.250 AKSNEKLQGAETGSTTSSGQS7QR157.5_G                | 80    | 4.20E-07  |
| 9   | S8 | D.396 PKAMTANYPSWR184.5_T                             | 103   | 1.40E-08  |
| 10  | S9 | E.396 ENKDDAKYIYPYR194.5_Y                            | 58    | 1.90E-04  |
| 11  | S10| N.396 GDDAKYIYPYR194.5_Y                              | 61    | 2.00E-04  |
| 12  | S11| L.482 WAPRPWAFAFR182.5_G                              | 56    | 2.00E-04  |
| 13  | S12| L.486 WAPRPWAFAFR182.5_G                              | 46    | 3.90E-04  |
| 14  | S13| W.488 APWPWAFAFR182.5_G                               | 49    | 3.30E-04  |
| 15  | S14| A.492 PRPWAFAFR182.5_G                                | 46    | 3.10E-03  |
| 16  | S15| S.576 DGNTSTSNLAPNTNANTGDVGVR198.5_L                  | 90    | 1.20E-08  |
| 17  | S16| D.576 GNTSTSNLAPNTNANTGDVGVR198.5_L                  | 192   | 8.00E-17  |
| 18  | S17| G.576 NTSTSNLAPNTNANTGDVGVR198.5_L                   | 158   | 1.20E-08  |
| 19  | S18| N.576 NTSTSNLAPNTNANTGDVGVR198.5_L                   | 154   | 6.00E-09  |
| 20  | S19| T.576 STSNLAPNTNANTGDVGVR198.5_L                     | 121   | 7.90E-08  |
| 21  | S20| S.576 STSNLAPNTNANTGDVGVR198.5_L                     | 126   | 1.90E-05  |
| 22  | S21| S.576 TSNLAPNTNANTGDVGVR198.5_L                      | 121   | 4.60E-07  |
| 23  | S22| T.576 TSNLAPNTNANTGDVGVR198.5_L                      | 117   | 4.20E-08  |
| 24  | S23| N.576 NLAPNTNANTGDVGVR198.5_L                        | 132   | 1.00E-05  |
| 25  | S24| N.576 TSNLAPNTNANTGDVGVR198.5_L                      | 118   | 6.30E-06  |
| 26  | S25| L.576 APNANTNANTGDVGVR198.5_L                        | 104   | 4.00E-06  |
| 27  | S26| A.576 APNANTNANTGDVGVR198.5_L                        | 108   | 9.70E-08  |
| 28  | S27| P.576 NTNANTGDVGVR198.5_L                            | 74    | 9.20E-06  |
| 29  | S29| N.576 NTNANTGDVGVR198.5_L                            | 85    | 3.40E-06  |
| 30  | S30| T.576 TSNLAPNTNANTGDVGVR198.5_L                      | 80    | 1.90E-06  |
| 31  | S31| T.576 TSNLAPNTNANTGDVGVR198.5_L                      | 59    | 3.00E-04  |
| 32  | S32| G.576 TDNNTGNDVGVR198.5_L                            | 46    | 2.10E-04  |
| 33  | S33| D.576 DASYNPKMR197.5_L                                | 52    | 4.40E-05  |
| 34  | S34| D.608 EAPNANTGNDVGVR198.5_L                          | 62    | 2.80E-06  |
| 35  | S35| C-terminal semi-tryptic peptides                      |       |           |
| 2   | S2 | R.135 AYLIDLFSKLNQPTQDQTQTFNPEF198.5_G                | 35    | 0.001     |
| 4   | S4 | R.135 TSGQNTSTTGAMFLGKVNAAEPD128.5_D                 | 102   | 3.70E-14  |

Continued
| No. | ID | Peptide Sequence | Score | E-value |
|-----|----|------------------|-------|---------|
| 5   |    | R.386TAIDRVDHLD  | 58    | 6.40E⁻³  |
| 9   |    | R.386PNDKASSGQSDENHTKFTSATGMDQQGTSAGNPDSLKQDNISKSQD  | 57    | 1.10E⁻³  |
| 13  |    | R.1273QSFGTDHSTQPQPQSLKTTPFVTSSQDG  | 68    | 3.50E⁻³  |

Table 1. N-terminal dimethylated peptides identified in P1 adhesin by LC-MS/MS. Exact site within the peptide is indicated by the bold underlined amino acid. Amino acid positions denote the start and end of the peptides. The peptides listed are the highest scores identified from 4 biological replicates analysed separately using Sciex 5600 and Thermo Scientific Q Exactive™ mass spectrometers. All peptides have expectation values < 0.05.

Figure 2. Immunoblots of cell lysates of *M. pneumoniae* probed with sera raised against regions within P1. Sera raised against 15 different regions (‘RP’ boxes) of P1 were a gift from R. Dumke. (Top panel) Simplified cleavage map depicting the P1 adhesin, cleavage sites, and the 15 regions of P1 that have been previously cloned and expressed as recombinant fragments in *E. coli*. The mass of full length and smaller proteoforms of P1 as predicted by ProtParam. (Bottom panel) Immunoblots depicting *M. pneumoniae* cell lysate probed with the panel of anti-recombinant P1 sera. All the immunoblot lanes are part of the same blot. The membrane was blocked and then sliced to separate lanes before incubating with the described P1 sera. (Bottom right) Immunoblots with the intensity adjusted to highlight low abundant bands. Proteins migrating with masses similar to P1 proteoforms identified by LC-MS/MS have been marked on the immunoblot.
of P1 represents a fifth site for high amino-peptidase activity with 18 neo-N-termini residing between amino acid positions 1343 – 1361 (cleavage site 14 in Table 1; sequence: $^{1341}$STS$^{1345}$KQ$^{1350}$T$^{1355}$S$^{1360}$AA$^{1365}$). The C-terminal peptide has a composition comprising five alanine, five lysine, and thirteen proline residues. This C-terminal region also shares sequence identity (53.1%) with the carboxy-terminal 31 residues of Mpn142. Furthermore, the final 15 residues of P1 shares 73.3% sequence identity with the last 14 residues of Mpn142 (11 identical positions). The C-terminal 30 amino acids (named P1-30: $^{1597}$TSAA$^{1602}$KPGAP$^{1607}$RPPV$^{1612}$PPKPGAP$^{1617}$KPPVQ$^{1622}$PPKKPA$^{1627}$), and the C-terminal 15 amino acids (named P1-15: $^{1613}$PGAPK$^{1618}$PPVQ$^{1623}$PPKKPA$^{1627}$) were synthesised chemically (Table 2; Chempeptide Limited, China) and an N-terminal biotin tag was added to the P1-15 peptide. Microtitre binding assays revealed that P1-15, P1-30, and the recombinant protein, RP15$^{25}$, bind a range of host molecules in a dose dependent manner (Fig. 3). M. pneumoniae cells and RP15 bound lactoferin, vitronectin, plasminogen, fibrinectin, and fibrinogen. Only M. pneumoniae cells bound laminin. P1-30 bound fibronectin, fibrinogen and plasminogen in a dose dependent manner but failed to bind laminin. P1-15 only bound vitronectin but failed to bind laminin, lactoferrin, fibrinectin, and fibrinogen (Fig. 3). Compared with P1-30 and P1-15, the C-terminal 106 amino acids of P1 represented by RP15 consistently showed the most consistent and most diverse binding capabilities for the panel of host proteins tested here suggesting that multiple binding domains increase the binding capabilities of P1 proteoforms. Consistent with this hypothesis, RP15 spans two putative glycosaminoglycan binding motifs (underlined motifs in Table 2) that are absent in P1-30 and P1-15.

To investigate whether binding was due to the specific amino acid sequence or to amino acid composition, microscale thermophoresis was performed on P1-30 and a scrambled version of P1-30 (PKPPRAAPPKAPTVPVPGAPSPYVKKPKQAPG). P1-30 had a medium binding affinity for plasminogen ($K_D = 554 \pm 2.1 \text{ nM}$) and a medium/low binding affinity for fetuin ($K_D = 2.4 \pm 0.7 \text{ \mu M}$). No binding affinity could be detected for the scrambled peptide (Fig. 4).

Microtitre binding assays were also employed to determine the binding capabilities of regions spanning the C-terminus of P1 to A549 human epithelial cells (Fig. 5). Recombinant pyruvate dehydrogenase subunit B of M. pneumoniae (rPdhB; positive control$^{24}$) and RP15 bound immobilised A549 cells, but not P1-30. We were not able to determine if P1-15 bound using this assay because we lacked reagents that could detect this peptide.

To overcome this experimental limitation and to attempt to identify potential binding partners for P1-15, we designed an affinity bait-prey experiment. The biotinylated P1-15 was coupled to avidin agarose and, in parallel with uncoupled avidin-agarose (negative control), were exposed to a native A549 cell lysate as described in Methods, washed and eluants were characterised by SDS-PAGE and LC-MS/MS (Fig. S1). Three protein bands identified in eluents from avidin-agarose coupled with biotinylated P1-15 that were absent in the control were analysed by LC-MS/MS (Fig. S1). LC-MS/MS analysis of slice 1 and 2 identified tryptic peptides that mapped to the intermediate filament cytokeratin 7 (Mascot score = 1157), cytokeratin 8 (Mascot score = 2737 & 1486), cytokeratin 18 (Mascot score = 2592), and vimentin (Mascot score = 617) (Fig. S1). Tryptic peptides to these filament proteins were not identified in the control experiment. Tryptic peptides identified in slice 3 identified glyceraldehyde-3-phosphate dehydrogenase, however, this protein was also identified in the eluents from the control and was not considered further as a potential binding partner with P1-15.

**Discussion.**

*M. pneumoniae* binds diverse host cell proteins including plasminogen, fibronectin, vitronectin, fibrinogen, lactoferrin, glycosaminoglycans, and sialoglyconjugates$^{9-17}$. The P1 adhesin and proteins it associates with at the tip of the attachment organelle are central to binding interactions that enable *M. pneumoniae* to target host cell receptors and is likely to contain binding domains
for some or all of these host molecules. Here we show that Mpn141 is processed extensively generating 23 proteoforms and that many proteoforms are retained on affinity matrices loaded with different host molecules and mimics of regions of host proteins including fetuin, fibronectin, actin, heparin, and plasminogen. Microtitre plate binding assays and microscale thermophoresis assays confirmed several of these preliminary findings and showed that the C-terminal region of P1 binds vitronectin, fibrinogen and fibronectin. Apart from removal of a 59 amino acid N-terminal leader peptide, only a ~40 kDa carboxyl terminal truncated fragment of P1 (potentially representing fragment 18 from this study), that forms a complex with full length P1 protein, and other accessory proteins has been reported previously33 but earlier immunoblotting studies with anti-P1 monospecific antisera identified numerous smaller proteoforms of P1 that were not characterised 2. Dimethyl labelling experiments enabled us to map the precise location of cleavage events in P1 (Table 1). P1 proteoforms are likely generated by proteases on the cell surface of *M. pneumoniae* or associated with the protein translocation machinery but their identities have not been confirmed. Biotinylation studies identified 13 proteoforms of P1 that were accessible on the surface of *M. pneumoniae* and our surface labelling and trypsin shaving experiments indicate that the

![Figure 3.](image)

Concentration-dependent binding of the C-terminus of P1 to different human proteins. Microtitre plate binding assays were used to measure the binding abilities of RP15, P1-30, and P1-15 to human plasminogen and to different components of the human extracellular matrix. Bovine serum albumin (BSA) and whole cell lysate proteins of *M. pneumoniae* (Mpn) were used as a negative and positive control, respectively. Results are shown from a single experiment with a mean and standard deviation of eight replicates. The experiment was independently repeated twice.
proteoforms remain attached to the extracellular side of \textit{M. pneumoniae} cell membranes. Our data is consistent with electron micrographs of \textit{M. pneumoniae} immunostained with ferretin-labelled anti-P1 antibodies that depict gold particles at: i) the tip of the attachment organelle; ii) along the shaft of this structure; iii) at sites along the cell body; and iv) at sites distant from the \textit{M. pneumoniae} membrane\textsuperscript{2}. It is not known if some proteoforms are excreted into the extracellular milieu but it is conceivable that processing of P1 occurs after translocation and the fragments may remain anchored to the surface via the predicted C-terminal transmembrane domain similarly seen in P40 and P90 of \textit{M. pneumoniae}\textsuperscript{27}. Consistent with this view, we were unable to find tryptic peptides that mapped to the putative leader peptide residing in the N-terminus of P1 or in the bioinformatically predicted transmembrane domains, or the well characterised C-terminal transmembrane domain. However, we did find tryptic peptides in the bioinformatically predicted transmembrane domain located around residue 1294.

Regions in P1 have been extensively characterised in an earlier study\textsuperscript{55}. Highly immunogenic regions and adherence mediating regions were found distributed throughout P1 particularly in the carboxy-terminal half of the molecule\textsuperscript{55}. Sera from patients infected with \textit{M. pneumoniae} bound to regions in P1 that were not responsible for adherence\textsuperscript{55}. It is conceivable that P1-derived proteoforms divert the binding of host antibodies away from regions in P1 required for adherence. We hypothesise that post-translational processing events release a portion of P1-derived proteoforms into the extracellular milieu, a process that may represent an immune decoy mechanism that seeks to bind and direct host antibodies away from \textit{M. pneumoniae}. A similar scenario has been hypothesised for Protein M of \textit{Mycoplasma genitalium}; a close relative of \textit{M. pneumoniae}\textsuperscript{79}.

**Figure 4.** Plasminogen and fetuin binding by P1-30 using microscale thermophoresis. Left: Thermophoretic output representing P1-30 (triangles) binding to plasminogen with a \( K_D \) of 554 nM. A scrambled version of P1-30 (circles) could not be assigned a \( K_D \) value. Right: Thermophoretic output representing P1-30 binding to fetuin with a \( K_D \) of 2 \( \mu \)M. The scrambled peptide could not be assigned a \( K_D \) value.

**Figure 5.** Binding of the C-terminus of P1 to immobilized A549 cells. Proteins, peptides, or A549 cells were immobilised in wells of a 96-well plate and binding was measured in a microtitre plate binding assay. 'Without A549' cells: protein and peptides were immobilised and detected by corresponding antisera (peptides with anti-RP15). 'With A549 cells': A549 cells were immobilised first before incubating with proteins and peptides. Whole antigen of \textit{M. pneumoniae} (Mpn), recombinant PdhB, and RP8 served as positive and negative controls, respectively\textsuperscript{17}. Bars represent mean and standard deviation of eight replicates from a single experiment. The experiment was independently repeated twice.
Our affinity studies suggest that the different proteoforms retain the ability to bind to different host proteins, glycosaminoglycans and sialoglyconjugates. RP15 was observed to bind immobilised A549 cells in microtitre plate assays (Fig. 4). This was surprising as no adherence regions have been previously identified within RP15. Anti-RP15 antibodies were reported to be unable to inhibit M. pneumoniae adherence to primary human bronchial epithelial (HBEC) cells, human fetal lung fibroblasts (MRC-5), and human cervical carcinoma cells (HeLa) suggesting that RP15 may bind to specific receptors only present on the A549 cell surface. We were unable to determine binding activity to A549 cells for P1-30 or P1-15 (Fig. 4) because anti-RP-15 antibodies did not detect these peptides. To investigate the binding capabilities of the C-terminal peptide P1-15, it was bound to avidin agarose and incubated with A549 cell lysates. This strategy selectively recovered cytoskeletal proteins, vimentin, cytokeratin 7, cytokeratin 8, and cytokeratin 18 (Fig. S1) from P1-15-avidin agarose but not from avidin agarose control experiments. Although preliminary, these observations are worthy of further study. Cytokeratin 7 is found in epithelia of lungs and other tissues, and has been shown to be involved in stabilising cytokeratin 18. Both cytokeratin 8 and 18 are major structural proteins of epithelial cells and are found in the intermediate filaments of A549 cells. Cytokeratin 8 has been identified to reside on the cellular surface of carcinogenic keratinocyte cells (HaCaT) and carcinogenic mammary cells, and carcinogenic hepatocytes suggesting they may be surface accessible on many cells. Cytokeratin 8 and 18 are co-expressed and frequently found associated together. Vimentin forms filament and is primarily expressed when epithelial cells transition into mesenchymal cells and function to induce changes in cell shape, motility and adhesin during this transition. Vimentin has also been observed to be secreted to the extracellular matrix and on the surface of activated macrophages. Cytokeratin 8, 18, and vimentin are suggested to be targeted by different pathogens after successfully invading host cells or after inducing cytoskeletal rearrangements. Pathogenic bacteria are known to interact with these cytoskeletal proteins during infection. Although mycoplasma have long been considered to be cell surface-associated parasitic bacteria, this dogma has been challenged with numerous reports citing phylogenetically-divergent mycoplasma residing within eukaryote cells and possessing the molecular machinery for selective uptake into survival within, and release from phagosomes.

We recently showed that Mnp142, a member of the same operon that houses the P1 gene (mnp141), and the surface accessible moonlighting adhesin, elongation factor Tu (Ef-Tu), are cleaved extensively. Post-translational processing of adhesins has been well characterised in M. hypopneumoniae where cleavage fragments have been shown to adhere to porcine cilia, porcine kidney epithelial cells, and a range of host molecules such as the glycosaminoglycan mimic heparin, plasminogen, fibronectin, and heparin. Processing of adhesin molecules is not confined to M. hypopneumoniae but has been described in Mycoplasma gallisepticum, Mycoplasma fermentans, M. genitalium, and Spiroplasma citri. We show that major adhesion molecules in M. pneumoniae, a phylogenetically distinct human pathogen, are processed. All these studies suggest that the processing of surface accessible proteins is widespread in Mollicutes. It is notable that all the P1 fragments that were recovered during heparin affinity chromatography contained putative glycosaminoglycan binding motifs except an N-terminal and a central fragment (Fig. 1, fragments 14 and 16). These motifs consist of clustered, positively charged amino acids that have been shown to have a role in binding to glycosaminoglycans, actin, and plasminogen. Heparin mimics the glycosaminoglycans found in the extracellular matrix and on the surface of host cells. M. hypopneumoniae, M. gallisepticum, and M. fermentans have been shown to bind heparin to the bacterial cell surface and employ bound heparin to bind other host molecules. Finally, heparin has also been implicated in biofilm formation by increasing cell-cell interactions in the Gram-positive pathogens, S. aureus and Lactobacillus rhamnosus forms large, complex biofilms on ablomatic surfaces. Heparin affinity chromatography of M. pneumoniae has been performed identifying only nine proteins, none of which was P1. Recently, we showed that Ef-Tu in M. pneumoniae displays a strong affinity to heparin. Collectively, our studies suggest that the ability to bind heparin is a universal strategy in microbial pathogenesis.

In several instances, we observed multiple cleavage sites within P1 that clustered within a defined region of P1. For example, 18 cleavage sites clustered between amino acids 1343–1361 in the C-terminus of P1 (Table 1). Sequential cleavage patterns similar to this were also reported in Mpn142 and in Mhp493, a paralog of the major adhesin P97 (Mhp183) in M. hypopneumoniae. Surfaceome studies of M. pneumoniae revealed the presence of surface accessible aminopetidases that may target a neo-N-terminal cleavage event and sequentially clip amino acids subsequent to the initial cleavage event. The function of these clipping events remains unknown but could be a mechanism to alter function and localisation of cleavage fragments, or represent a mechanism to recycle amino acids. Cleavage site 14 in P1 (Fig. 1) occurs within a large predicted disorder region (amino acid range 1187–1382). The inherent flexibility of disordered regions make them accessible to protease activity. Many major cleavage events identified in M. hypopneumoniae adhesin molecules reside with large disordered regions.

The C-terminus of the P1 tail is homologous to the C-terminus of Mnp142 and the C-terminal 15 amino acids of P1 has 73.3% sequence identity with the same region in Mnp142. Almost half of this sequence consists of proline residues while lysine is also heavily represented in this region. Proline-rich regions in proteins have been implicated in protein-protein interactions and has been suggested that proline residues could anchor the C-terminus of P1 in the cell membrane. Lysine-rich regions are associated with binding plasminogen, actin, and DNA. While P1-15 and P1-30 bound plasminogen in a dose-responsive manner, it was notable that RP-15 bound it more strongly. RP-15 also bound fibronectin and fibrinogen more strongly than P1-30 (Fig. 3). These data suggest that extra binding sites for these host molecules are located upstream of the C-terminal 30 amino acids of P1. Previous work suggests that sialic acid is the dominant host receptor for the P1 adhesin. Consistent with these earlier studies the P1 tail has a strong affinity to the sialic acid rich protein, fetuin. Our data indicates that the mature P1 proteoform and a
further nine smaller proteoforms of P1 bind fetuin. The ability to bind fetuin has been linked with biofilm formation in *M. pneumoniae*.

### Conclusion

In summary, this study reports that the P1 adhesin is subject to extensive post-translational processing forming twenty-two proteoforms from seventeen cleavage sites. Each of the proteoforms retain the ability to bind to host molecules or their structural mimics and are surface accessible. Processing has been described in *M. hypopneumoniae*, *M. gallisepticum*, and *S. citri* and is likely to be a widespread mechanism to generate surface protein diversity and promote protein:protein interactions. Specifically we show that the C-terminus of P1 plays a role in adhering to a range of host molecules including cytoskeletal proteins. This study expands on our knowledge of the role that the P1 adhesin plays in interactions between *M. pneumoniae* and host cells.

### Data availability

Data for this study is available on request from the corresponding author.

Received: 17 January 2019; Accepted: 20 March 2020; Published online: 14 April 2020

### References

1. Krause, D. C. & Balish, M. F. Structure, function, and assembly of the terminal organelle of Mycoplasma pneumoniae. *FEMS Microbiol Lett* 198, 1–7 (2001).
2. 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).
3. Feldner, J., Gobel, U. & Bredd, W. Mycoplasma pneumoniae adhesin localized to tip structure by monoclonal antibody. *Nature* 289, 765–767 (1982).
4. Krause, D. C. Mycoplasma pneumoniae cytheradherence: unravelling the tie that binds. *Mol Microbiol* 20, 247–253 (1996).
5. Seto, S., Layh-Schmitt, G., Kenri, T. & Miyata, M. Visualization of the attachment organelle and cytheradherence proteins of *Mycoplasma pneumoniae* by immunofluorescence microscopy. *J Bacteriol* 183, 1621–1630, https://doi.org/10.1128/JB.183.6.1621-1630.2001 (2001).
6. Seto, S. & Miyata, M. Attachment organelle formation represented by localization of cytheradherence proteins and formation of the electron-dense core in wild-type and mutant strains of *Mycoplasma pneumoniae*. *J Bacteriol* 185, 1062–1091 (2003).
7. Nakane, D., Kenri, T., Matsuo, L. & Miyata, M. Systematic Structural Analyses of Attachment Organelle in *Mycoplasma pneumoniae*. *PLoS Pathog* 11, e1005299, https://doi.org/10.1371/journal.ppat.1005299 (2015).
8. Baseman, J. B. et al. Identification of a 32-kilodalton protein of *Mycoplasma pneumoniae* associated with hemadsorption. *Int J Med Sci* 23, 474–479 (1987).
9. Dallo, S. F., Kannan, T. R., Blaylock, M. W. & Baseman, J. B. Elongation factor Tu and E1 beta subunit of pyruvate dehydrogenase complex act as fibronectin binding proteins in *Mycoplasma pneumoniae*. *Mol Microbiol* 46, 1041–1051 (2002).
10. Gründel, A., Jacobs, E. & Dumke, R. Interactions of surface-displayed glycolytic enzymes of *Mycoplasma pneumoniae* with components of the human extracellular matrix. *Int J Med Microbiol*, https://doi.org/10.1016/j.ijmm.2016.09.001 (2016).
11. Hagemann, L., Gründel, A., Jacobs, E. & Dumke, R. The surface-displayed chaperones GroEL and DnaK of *Mycoplasma pneumoniae* interact with human plasminogen and components of the extracellular matrix. *Pathog Dis*, https://doi.org/10.1093/femspd/bfx017 (2017).
12. Widjaja, M. et al. Elongation factor Tu is a multifunctional and processed moonlighting protein. *Sci Rep* 7, 11227, https://doi.org/10.1038/s41598-017-10644-z (2017).
13. Grimmer, J. & Dumke, R. Organization of multi-binding to host proteins: The glycolaldehyde-3-phosphate dehydrogenase (GAPDH) of *Mycoplasma pneumoniae*. *Microbiol Res* 218, 22–31, https://doi.org/10.1016/j.mires.2018.09.006 (2019).
14. Dumke, R., Hausner, M. & Jacobs, E. Role of *Mycoplasma pneumoniae* glyceraldehyde-3-phosphate dehydrogenase (GAPDH) in mediating interactions with the human extracellular matrix. *Microbiology* 157, 2328–2338, https://doi.org/10.1099/mic.0.048298-0 (2011).
15. Thomas, C., Jacobs, E. & Dumke, R. Characterization of pyruvate dehydrogenase subunit B and enolase as plasminogen-binding proteins in *Mycoplasma pneumoniae*. *Microbiology* 159, 352–365, https://doi.org/10.1099/mic.0.061184-0 (2013).
16. Gründel, A., Friedrich, R., Pfeiffer, M., Jacobs, E. & Dumke, R. Subunits of the Pyruvate Dehydrogenase Cluster of *Mycoplasma pneumoniae* Are Surface-Display Proteins that Bind and Activate Human Plasminogen. *PLoS One* 10, e0126600, https://doi.org/10.1371/journal.pone.0126600 (2015).
17. Gründel, A., Pfeiffer, M., Jacobs, E. & Dumke, R. Network of Surface-Displayed Glycolytic Enzymes in *Mycoplasma pneumoniae* and Their Interactions with Human Plasminogen. *Infect Immun* 84, 666–676, https://doi.org/10.1128/IAI.01037-15 (2016).
18. Kahane, I., Banai, M., Razin, S. & Feldner, I. Attachment of mycoplasmas to host cell membranes. *Rev Infect Dis* 4(Suppl), S183–192 (1982).
19. Loomes, L. M. et al. Erythrocyte receptors for *Mycoplasma pneumoniae* are sialylated oligosaccharides of II antigen type. *Nature* 307, 560–563 (1984).
20. Loomes, L. M., Uemura, K. & Feizi, T. Interaction of *Mycoplasma pneumoniae* with erythrocyte glycoproteins of I and I antigen types. *Infect Immun* 47, 15–20 (1985).
21. Roberts, D. D., Olson, L. D., Barile, M. F., Ginsburg, V. & Krivan, H. C. Sialic acid-dependent adhesion of *Mycoplasma pneumoniae* to purified glycoproteins. *J Biol Chem* 264, 9289–9293 (1989).
22. Inamine, J. M., Loechel, S. & Hu, P. C. Analysis of the nucleotide sequence of the P1 operon of *Mycoplasma pneumoniae*. *Gene* 73, 175–183 (1988).
23. Waldo, R. H. II & Krause, D. C. Synthesis, stability, and function of cytosdishin P1 and accessory protein B/C complex of *Mycoplasma pneumoniae*. *J Bacteriol* 188, 569–575, https://doi.org/10.1128/JB.188.2.569-575.2006 (2006).
24. Aravind, L. & Koonin, E. V. A novel family of predicted phosphoesterases includes Drosophila prune protein and bacterial Rec exonuclease. *Trends Biochem Sci* 23, 17–19 (1998).
25. Postic, G., Danchin, A. & Mechold, U. Characterization of NrnA homologs from *Mycobacterium tuberculosis* and *Mycoplasma pneumoniae*. *RNA* 18, 155–165, https://doi.org/10.1261/rna.029132.111 (2012).
26. Speker, B., Hu, P. & Herrmann, R. Identification of gene products of the P1 operon of *Mycoplasma pneumoniae*. *Mol Microbiol* 5, 299–306 (1991).
27. 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).
104. Tsiodras, S., Kelesidis, I., Kelesidis, T., Stamboulis, E. & Giamarellou, H. Central nervous system manifestations of Mycoplasma pneumoniae. J Infect 33, 29–32, https://doi.org/10.1016/j.jinf.2005.07.006 (2014).

105. Burki, S. & Negroni, Z. Role of tyrosine kinases and the tyrosine phosphatase SptP in the interaction of Salmonella with eukaryotic cells. J Biol Chem 282, 21481–21491, https://doi.org/10.1074/jbc.M112.370791 (2017).

106. Ossendorp, F. et al. Electroblotting of multiple gels: a simple apparatus without buffer tank for rapid transfer of proteins from polyacrylamide to nitrocellulose. J Biochem Biophys Methods 10, 203–209 (1984).

107. Peng, K., Radivojac, P., Vucetic, S., Dunker, A. K. & Obradovic, Z. Length-dependent prediction of protein intrinsic disorder. BMC Bioinformatics 7, 208, https://doi.org/10.1186/1471-2105-7-208 (2006).

108. de Castro, E. et al. ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. Nucleic Acids Res 34, W362–365, https://doi.org/10.1093/nar/gkl124 (2006).

109. Cardin, A. D. & Weintraub, H. J. Molecular modeling of protein-glycosaminoglycan interactions. Arterioscler Thromb Vasc Biol 11, 21–28 (1991).

110. Kumar, Y. & Valdivia, R. H. Actin and intermediate filaments stabilize the Chlamydia trachomatis vacuole by forming dynamic filamentous complexes in urea reveals different polypeptide separation characteristics. Proc Natl Acad Sci USA 80, 7113–7117 (1983).

111. Ovenden, D. W. & Lane, E. B. The quest for the function of simple epithelial keratins. Bioessays 25, 748–758, https://doi.org/10.1002/bies.10316 (2003).

112. Dobashi, N. et al. Detection of anti-cytokeratin 8 antibody in the serum of patients with cryptogenic fibrosing alveolitis and pulmonary fibrosis associated with collagen vascular disorders. Thorax 53, 969–974 (1998).

113. Haim, M. et al. Cytokeratin 8 interacts with clumping factor B: a new possible virulence factor target. Microbiology 156, 3710–3721, https://doi.org/10.1099/mic.0.34414-0 (2010).

114. Godfroid, E., Geuskens, M., Dupressoir, T., Parent, I. & Szpirer, C. Cytokeratins are exposed on the outer surface of established human mammary carcinoma cells. J Cell Sci 99 (Pt 3), 595–607 (1991).

115. Hamborg, T. A., Vasudevan, J., Alliotta, M. M., Glass, W. A. 2nd & Giovannucci, S. L. A cytokeratin 8-like protein with plasminogen-binding activity is present on the external surfaces of hepatocytes, HepG2 cells and breast carcinoma cell lines. J Cell Sci 108 (Pt 3), 1071–1082 (1995).

116. Lu, X. & Lane, E. B. Retrovirus-mediated transgenic keratin expression in cultured fibroblasts: specific domain functions in keratin stabilization and filament formation. Cell 62, 681–696 (1990).

117. Moll, R., Franke, W. W., Schiller, D. L., Geiger, B. & Krepler, R. The catalog of human cytokeratins: patterns of expression in normal and pathological situations. J Histochem Cytochem 39 (Pt 3), 108 (1991).

118. Mor-Vaknin, N., Punturieri, A., Sitwala, K. & Markovitz, D. M. Vimentin is secreted by activated macrophages. Nat Cell Biol 31, 2920–2930, https://doi.org/10.1038/ncb698 (2015).

119. Carlson, S. A., Omary, M. B. & Jones, B. D. Identification of cytokeratins as accessory mediators of Salmonella entry into eukaryotic cells. J Infect Dis 187, 1537–1540, https://doi.org/10.1086/343167 (2003).

120. Viswanathan, V. K. et al. Antibodies specific for vimentin filaments. J Immunol 165, 145–152, https://doi.org/10.4049/jimmunol.165.1.145 (1998).

121. Klimstra, W. B., Heidner, H. W. & Johnston, R. E. The furin protease cleavage recognition sequence of Sindbis virus PE2 can mediate virion attachment to cell surface heparan sulfate. J Biol Chem 262, 6399–6406 (1987).

122. Kyhse-Andersen, J. Electroblotting of multiple gels: a simple apparatus without buffer tank for rapid transfer of proteins from polyacrylamide to nitrocellulose. J Biochem Biophys Methods 10, 203–209 (1984).

123. Raymond, B. B. et al. Proteolytic processing of the ciliain adhesin MHJ_0194 (P123J) in Mycoplasma hyopneumoniae generates a functionally diverse array of cleavage fragments that bind multiple host molecules. Cell Microbiol 17, 425–444, https://doi.org/10.1111/cmi.12377 (2015).

124. Aparicio, D. et al. Mycoplasma genitalium adhesin P110 binds sialic-acid human receptors. Nat Commun 9, 4471, https://doi.org/10.1038/s41467-018-06963-y (2018).

125. Escobedo, K. et al. Cytokine-related expression of enolase 1 in keratinocytes and in different epidermal layers. J Investig Dermatol Symp Proc 14, 15–22, https://doi.org/10.1016/j.jidsp.2019.09.004 (2019).

126. Saberi, S. et al. Cytokeratin 18 interacts with the enteropathogenic Escherichia coli secreted protein F (EspF) and is redistributed after infection. EMBO Rep 17, 595–607, https://doi.org/10.15252/embr.201543069 (2016).

127. Viswanathan, V. K. et al. Cytokeratin 18 interacts with the enteropathogenic Escherichia coli secreted protein F (EspF) and is redistributed after infection. Cell Microbiol 6, 987–997, https://doi.org/10.1111/j.1462-5822.2004.00416.x (2004).

128. Icenogle, L. M. et al. Molecular and biological characterization of Streplococcal SpyA-mediated ADP-ribosylation of intermediate filament protein vimentin. J Biol Chem 283, 21481–21491, https://doi.org/10.1074/jbc.M112.370791 (2017).

129. Murf, S., Watson, R. O. & Galan, J. E. Role of tyrosine kinases and the tyrosine phosphatase Stif in the interaction of Salmonella with host cells. Cell Microbiol 3, 795–810 (2001).

130. Guignot, J. & Servin, A. L. Maintenance of the Salmonella-containing vacuole in the juxtanuclear area: a role for intermediate filaments. Microbiol Pathog 45, 415–422, https://doi.org/10.1016/j.micpath.2008.09.007 (2008).

131. Saberi, S. et al. A potential association between Helicobacter pylori CagA EPITa and multimerization motifs with cytokeratin 18 cleavage rate during early apoptosis. Helicobacter 17, 350–357, https://doi.org/10.1111/j.1572-9223.2012.12095.x (2012).

132. Jahren, A. H. & Goldschmidt, C. R. Zooming into the future of the cell: telomeres are dynamic and telomeres are conserved. Nature 466, 386–392, https://doi.org/10.1038/nature09118 (2010).

133. Murphy, K. A., Saiki, R. K. & Kleckner, N. The role of cellular DNA replication in chromosome structure and stability. Annu Rev Genet 21, 1–32 (1987).
107. Suleman, M. et al. Mycoplasma bovis isolates recovered from cattle and bison (Bison bison) show differential in vitro effects on PRMC proliferation, alveolar macrophage apoptosis and invasion of epithelial and immune cells. Vet Microbiol 186, 28–36, https://doi.org/10.1016/j.vetmic.2016.02.016 (2016).

108. Josi, C. et al. Bovine Epithelial in vitro Infection Models for Mycoplasma bovis. Frontiers in cellular and infection microbiology 8, 329, https://doi.org/10.3389/fcimb.2018.00329 (2018).

109. Raymond, B. B. et al. Mycoplasma hypnepneumoniae resides intracellularly within porcine epithelial cells. Sci Rep 8, 17697, https://doi.org/10.1038/s41598-018-36054-3 (2018).

110. Burnett, T. A. et al. P159 is a proteolytically processed, surface adhesin of Mycoplasma hypnepneumoniae: defined domains of P159 bind heparin and promote adherence to eukaryote cells. Mol Microbiol 60, 669–686, https://doi.org/10.1111/j.1365-2958.2006.05139.x (2006).

111. Wilton, L. M. et al. Mhp693 (P216) is a proteolytically processed, cilia and heparin binding protein of Mycoplasma hypnepneumoniae. Mol Microbiol 71, 566–582, https://doi.org/10.1111/j.1365-2958.2008.06546.x (2009).

112. Seymour, L. M. et al. A processed multidomain mycoplasma hypnepneumoniae adhesin binds fibronectin, plasminogen, and swine respiratory cilia. J Biol Chem 285, 33971–33978, https://doi.org/10.1074/jbc.M110.10463 (2010).

113. Seymour, L. M. et al. Mhp107 is a member of the multifunctional adhesin family of Mycoplasma hypnepneumoniae. J Biol Chem 286, 10097–10104, https://doi.org/10.1074/jbc.M108108140 (2011).

114. Seymour, L. M. et al. Mhp182 (P102) binds fibronectin and contributes to the recruitment of plasmin(ogen) to the Mycoplasma hypnepneumoniae cell surface. Cell Microbiol 14, 81–94, https://doi.org/10.1111/j.1462-5822.2011.01702.x (2012).

115. Deutscher, A. T. et al. Mycoplasma hypnepneumoniae Surface proteins Mhp385 and Mhp384 bind host cilia and glycosaminoglycans and are endoproteolytically processed by proteases that recognize different cleavage motifs. J Proteome Res 11, 1924–1936, https://doi.org/10.1021/pr201115v (2012).

116. Raymond, B. B. et al. Extracellular Actin Is A Receptor for Mycoplasma hypnepneumoniae. Frontiers in cellular and infection microbiology 8, 54, https://doi.org/10.3389/fcimb.2018.00054 (2018).

117. Szczepanek, S. M. et al. Identification of lipoprotein MlA as a neoteric virulence factor of Mycoplasma gallisepticum. Infect Immun 78, 3473–3483, https://doi.org/10.1128/IAI.00154-10 (2010).

118. Calcutt, M. J., Kim, M. F., Karpas, A. B., Muhlradt, P. F. & Wise, K. S. Differential posttranslational processing confers intraspecies variation of a major surface lipoprotein and a macrophage-activating lipoprotein of Mycoplasma fermentans. Infect Immun 67, 760–771 (1999).

119. Davis, K. L. & Wise, K. S. Site-specific proteolysis of the MALP-40 lipoprotein determines the release of a soluble selective lipoprotein-associated motif-containing fragment and alteration of the surface phenotype of Mycoplasma fermentans. Infect Immun 70, 1129–1135 (2002).

120. Muhlradt, P. F., Kiess, M., Meyer, H., Sussmuth, R. & Jung, G. Isolation, structure elucidation, and synthesis of a macrophage stimulatory lipopeptide from Mycoplasma fermentans acting at picomolar concentration. J Exp Med 185, 1951–1958 (1997).

121. Shimizu, T., Kida, Y. & Kuwano, K. A triacylated lipoprotein from Mycoplasma genitalium activates NF-kappaB through Toll-like receptor 1 (TLR1) and TLR2. Infect Immun 76, 3672–3678, https://doi.org/10.1128/IAI.00257-08 (2008).

122. Dubrana, M. P. et al. Proteolytic post-translational processing of adhesins in the plant pathogen Sprolopa citri. J Mol Biol 400, 772–788, https://doi.org/10.1016/j.jmb.2010.05.004 (2010).

123. Boone, T. J. & Tyrrell, G. J. Identification of the actin and plasminogen binding regions of group B streptococcal phosphoglycerate kinase. J Biol Chem 287, 29035–29044, https://doi.org/10.1074/jbc.M112.361261 (2012).

124. Kamhi, E., Joo, E. J., Dordick, J. S. & Linhardt, R. J. Glycosaminoglycans in infectious disease. Biol Rev Camb Philos Soc 88, 928–943, https://doi.org/10.1111/j.1462-5822.2013.01204.x (2013).

125. Jenkins, C., Geary, S. J., Gladd, M. & Djordjevic, S. P. The Mycoplasma gallisepticum OsmC-like protein MG1142 resides on the cell surface and binds heparin. Microbiology 153, 1455–1463, https://doi.org/10.1099/mic.0.004937-0 (2007).

126. Duensing, T. D., Wing, J. S. & van Putten, J. P. Sulfated polysaccharide-directed recruitment of mammalian host proteins: a novel strategy in microbial pathogenesis. Infect Immun 67, 4463–4468 (1999).

127. Shanks, R. M. et al. Heparin stimulates Staphylococcus aureus biofilm formation. Infect Immun 73, 4596–4606, https://doi.org/10.1128/IAI.01349-09 (2005).

128. Chen, X. E., Ling, P., Duan, R. & Zhang, T. Effects of heparosan and heparin on the adhesion and biofilm formation of several bacteria in vitro. Carbohydrate polymers 88, 1288–1292 (2012).

129. Ueberle, B., Frank, R. & Herrmann, R. The proteome of the bacterium Mycoplasma pneumoniae: comparing predicted open reading frames to identified gene products. Proteomics 2, 754–764, https://doi.org/10.1002/pmic.200020035 (2000).

130. Williamson, M. P. The structure and function of proline-rich regions in proteins. Biochem J 297( Pt 2), 249–260 (1994).

131. Kay, B. K., Williamson, M. P. & Sudol, M. The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. J Biol Chem 279, 231–241 (2004).

132. Cork, A. J. et al. Defining the structural basis of human plasminogen binding by streptococcal surface enolase. J Biol Chem 284, 17129–17137, https://doi.org/10.1074/jbc.M109.04317 (2009).

133. Hantke, K., Pohlmann, B. B. & Djordjevic, S. P. Exploitation of plasmin(ogen) by bacterial pathogens of veterinary significance. Vet Microbiol 178, 1–13, https://doi.org/10.1016/j.vetmic.2015.04.008 (2015).

134. Robinson, M. W. et al. MHP1,025 is an M42 glutamyl aminopeptidase that moonlights as a multifunctional adhesin on the surface of Mycoplasma hypnepneumoniae. Open Biol 3, 130017, https://doi.org/10.1098/rsob.130017 (2013).

135. Lee, A. Y., Hsu, C. H. & Wu, S. H. Functional domains of Brevibacillus thermodiuberon lip protease for oligomerization and DNA binding: role of N-terminal and sensor and substrate discrimination domains. J Biol Chem 279, 34903–34912, https://doi.org/10.1074/jbc.M3043562200 (2004).

Acknowledgements
This research is supported by an Australian Government Research Training Program Scholarship provided to I.J.B. The authors would like to thank Mark Raftery and the Bioanalytical Mass Spectrometry Facility (BMSF) for access to the Scielx 5600 and Thermo Scientific Q Exactive™ Plus mass spectrometers purchased with the ARC grant LE130100096 entitled ‘Advanced high resolution mass spectrometer for collaborative proteomic and lipidomics research’. The authors would like to thank both the University of Technology Sydney and the Technische Universität Dresden for funding this research.
Author contributions
M.W. acquired and analysed the data except those listed for I.J.B., V.M.J. and R.D. Dimethyl labelling experiments was performed by I.J.B. Results from dimethyl labelling experiments was analysed by both M.W. and I.J.B. The experiments investigating the binding of the P1 C-terminus to human proteins and A549 human lung cells were conducted and analysed by R.D. Microscale thermophoresis of the scrambled P1 peptide was performed and data analysed by V.M.J. Acquisition of mass spectrometry data was performed by M.P.P. who also assisted with mass spectrometry data interpretation. R.D. created the figures for microtitre plate binding assays and all other figures were created by M.W. Both M.W. and S.P.D. wrote the manuscript. S.P.D. initiated the study. Both S.P.D. and R.D. secured funding for this study. The manuscript was reviewed and approved by all authors before submission.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-63136-y.

Correspondence and requests for materials should be addressed to S.P.D.

Reprints and permissions information is available at www.nature.com/reprints.

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2020