

HOW *MYCOPLASMA PNEUMONIAE* MANAGES TO STICK AROUND

Berova I.D. (s5850320), Carl S.A. (s5506506), van der Laan L.D.P. (s5982596), Maire-Richard S. (s5906245), Paingu S. (s5870585), van Veldhuizen R.A. (s6026893)

ABSTRACT

This case study focuses on investigating an unidentified DNA sequence that was blasted and traced back to the pathogen *M. pneumoniae*. To further analyze the genetic and proteomic features of this organisms, KBase, InterPro, AlphaFold, STRING, Pymol and Benchling were utilised.

INTRODUCTION

An unknown DNA fragment was identified as part of the genome from the pathogen *Mycoplasma pneumonia*; the subject of this case study. **Bioinformatics resources** were used to explore the genome and pathogenic gene of the organism and its corresponding protein. In silico analysis was used to design an overexpression vector that could aid in the study of the pathogenic protein and its role in infection.

Mycoplasma pneumoniae is a bacterium that belongs to the genus *Mycoplasma* and the class *Mollicutes*. Its genome has 816,394 bp and 687 genes [1]. *M. pneumoniae* cannot produce a cell wall, the structural support is provided instead by sterols [2]. It is a **parasitic pathogen** that lives on the epithelial surface of its host. It infects only humans, causing atypical community-acquired *pneumonia*, *pharyngitis* and *tracheobronchitis*. The infection is initiated by cytodherence of the organism to the epithelium by proteins like **P1-adhesin** transmembrane proteins. These bacteria attach to the epithelial membranes of the respiratory tract, damaging the cells, recruiting cytokines and producing hydrogen peroxide and superoxide [3]. Antibodies against *M. pneumoniae* may act against human brain cells and red blood cells, causing anaemia, fever, chest pain and coughing [4].

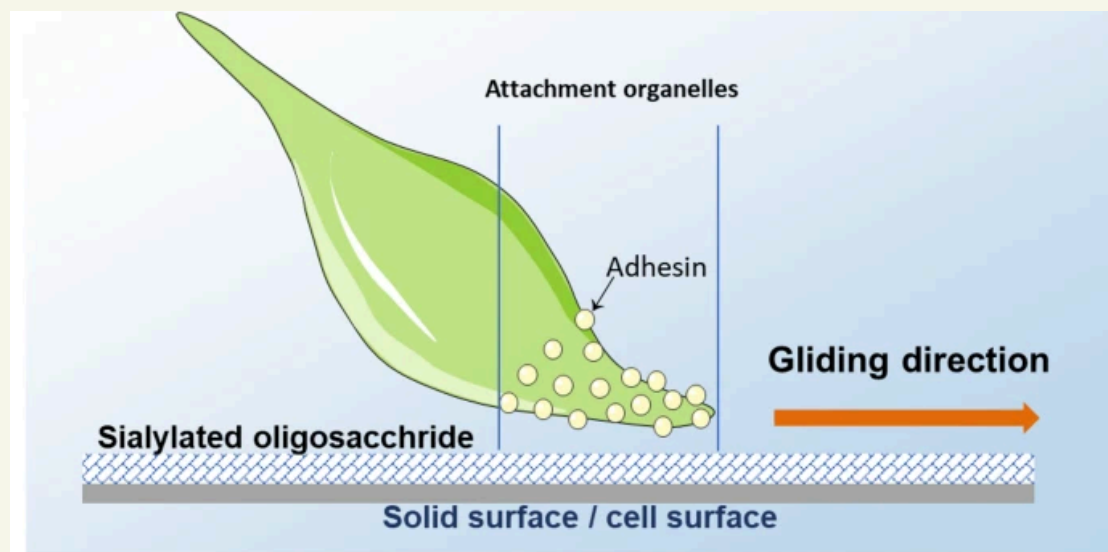


Figure 1: Schematic representation of *M. pneumoniae* and its gliding motility mechanism [5]. The coordinated interaction of adhesin proteins in the bacterium's attachment organelle enables it to bind to sialylated oligosaccharides on a host or solid surface and glide forward in the indicated direction.

MATERIALS & METHODS

An NCBI BLASTx was performed on sequence 3, set 21. KBase was used with Prokka annotation to identify the **genes** present. Literature research was done on *M. pneumoniae* and its infection in humans. For the analysis of the **structure** of adhesin P1, NCBI BLASTP with the database (DB) of the protein data bank (PDB) was performed on the amino acid sequence. AlphaFold was used to predict the protein structure, and PyMOL was used to compare the AlphaFold and PDB structures. To see the difference in the encoded domains of the CDS and PDB protein sequence, InterPro was used. The protein adhesin P1 was searched in the STRING DB. To create an **in silico overexpression** of the mgpA gene, Benchling was used. The primers were designed with PrimerBLAST to isolate the gene from the genome. These primers contained restriction sites needed for the ligation in the pBAD-sfGFP plasmid.

RESULTS

GENE ANALYSIS

Description	Total Score	Query Cover	E value	Accession
<i>Mycoplasma pneumoniae</i> FH, complete genome	66719	100%	0	CP002077.1
<i>Mycoplasma pneumoniae</i> strain 10-1385 chromosome, complete genome	66701	100%	0	CP039784.1
<i>Mycoplasma pneumoniae</i> strain NCTC10119 genome assembly, chromosome: 1	66740	100%	0	LR214945.1
<i>Mycoplasma pneumoniae</i> MAC strain Mac, complete genome	67037	100%	0	CP010550.1
<i>Mycoplasma pneumoniae</i> M129, complete genome	66574	100%	0	U00089.2

Table 1. Raw BLASTN results from the unknown sequence 3 from DNA set 21. All the hits belong to the organism *Mycoplasma pneumoniae*.

Type	Function	Aliases	Start	Strand	Length
CDS	Ribosomal large subunit pseudouridine synthase C	rluC	13,003	minus	930
CDS	Lipoprotein signal peptidase	lspA	13,557	minus	555
CDS	Segregation and condensation protein A	scpA	16,734	plus	1,521
CDS	Segregation and condensation protein B	scpB	18,238	plus	627
CDS	ATP-dependent 6-phosphofructokinase isozyme 1	pfkA	18,884	plus	987
CDS	hypothetical protein		3,441	plus	1,290
CDS	Adhesin P1	mgpA	4,833	plus	1,641

Table 2. A summary of the KBase gene annotation of the DNA fragment belonging to *M. pneumoniae*. The gene of interest (adhesin P1, mpaA) is highlighted in bold.

PROTEIN ANALYSIS

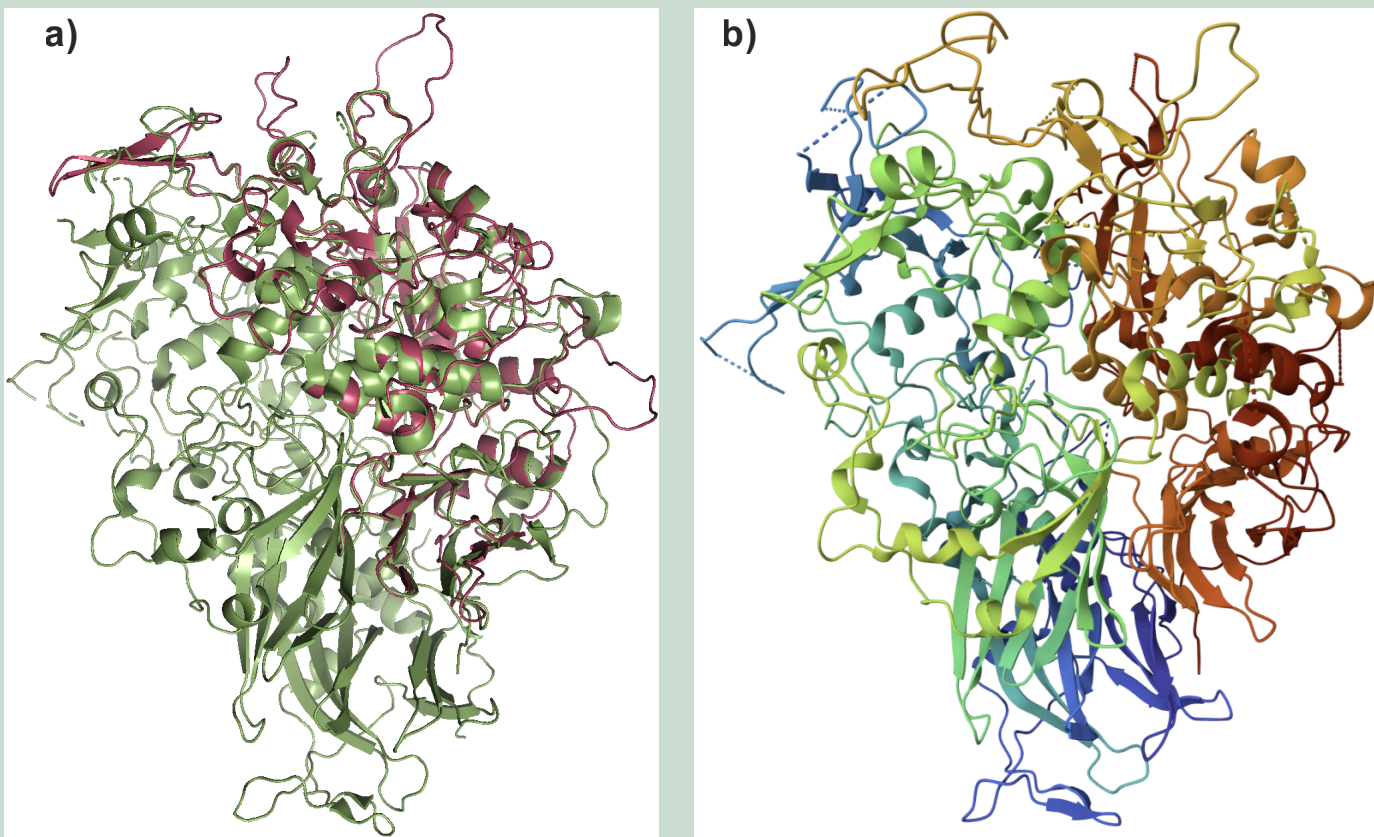


Figure 2: Structural analysis of adhesin P1 in *M. pneumoniae*. NCBI pBLAST (DB: PDB) best hit was Chain A, Adhesin P1 (per. indent. **63.26%**). With InterPro the domains of the PDB structure contained an Adhesin P1 N-terminus domain and C-terminal domain, while the CDS of adhesin P1 only contained a shorter C-terminal domain.

a) PyMOL structural alignment of PDB structure (id: 7BWM) (green) and AlphaFold structure (red).

b) Cryo-EM structure of *M. pneumoniae* Adhesin P1 chain A with resolution of 2.90Å (PDB id: 7BWM). Cartoon representation with color scheme; sequence ID (red= N-terminus, blue= C-terminus).

c) Figure from "Immunodominant proteins P1 and P40/P90 from human pathogen *Mycoplasma pneumoniae*". Crown shown in green comes on top of seven-blade beta-propeller (orange) with the N-terminal domain arranged around it [5].

PROTEIN ANALYSIS

		score
Orf6	Cytadherence protein; MPN142(new), 013(Himmelreich et al, 1996)	0.996
nrnA	DHH family phosphoesterases; Bifunctional enzyme which has both oligoribonuclease and pAp- phosphatase activities.	0.919
hmw1	Cytadherence accessory protein HMW1; component of the cytoskeleton-like structure which stabilizes the shape of the wall-less Mycoplasma.	0.919
hmw3	Cytadherence accessory protein HMW3; component of the cytoskeleton-like structure which stabilizes the shape of the wall-less Mycoplasma.	0.918
p30	30K adhesin-related protein; Adhesin necessary for successful cytodherence and virulence.	0.899
E07_orf175	Hypothetical protein; MPN143(new), 012(Himmelreich et al, 1996)	0.802
P65	Conserved hypothetical protein; MPN143(new), 528(Himmelreich et al, 1996)	0.801
hmw2	Cytadherence high molecular weight protein 2; Component of the cytoskeleton-like structure which stabilizes the shape of the wall-less Mycoplasma	0.76
dnak	Dnak; Acts as a chaperone; Belongs to the heat shock protein 70 family	0.731
E07_orf166	Hypothetical protein; MPN138(new), 017(Himmelreich et al, 1996); see: MPN013; Belongs to the UPF0134 family	0.708

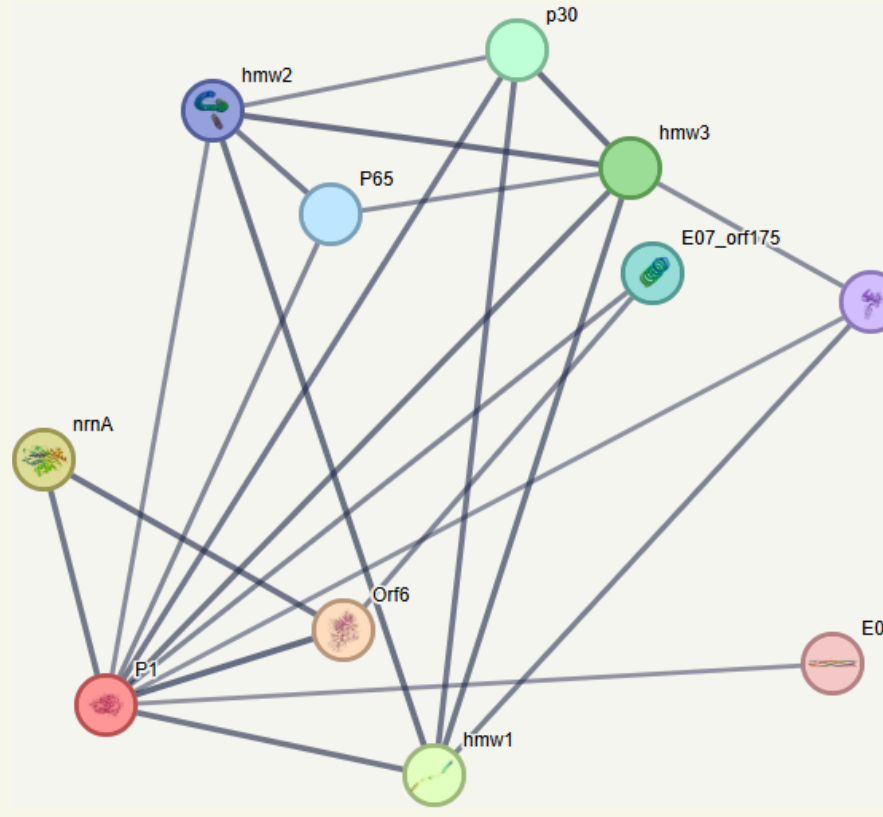


Figure 3: The full STRING network, visualized in STRING.

The string is a prediction of how well other **proteins partner** with the adhesin P1 protein. The **factors** looked at were: gene neighborhoods, gene fusions, gene co-occurrence, textmining, co-expression and protein homology **combined**. Above Figure 3 a table with the score of these prediction are given from the top 10 proteins. The top hits were highly associated with high gene neighborhood scores and also encoded for **adhesion proteins** [6].

IN SILICO

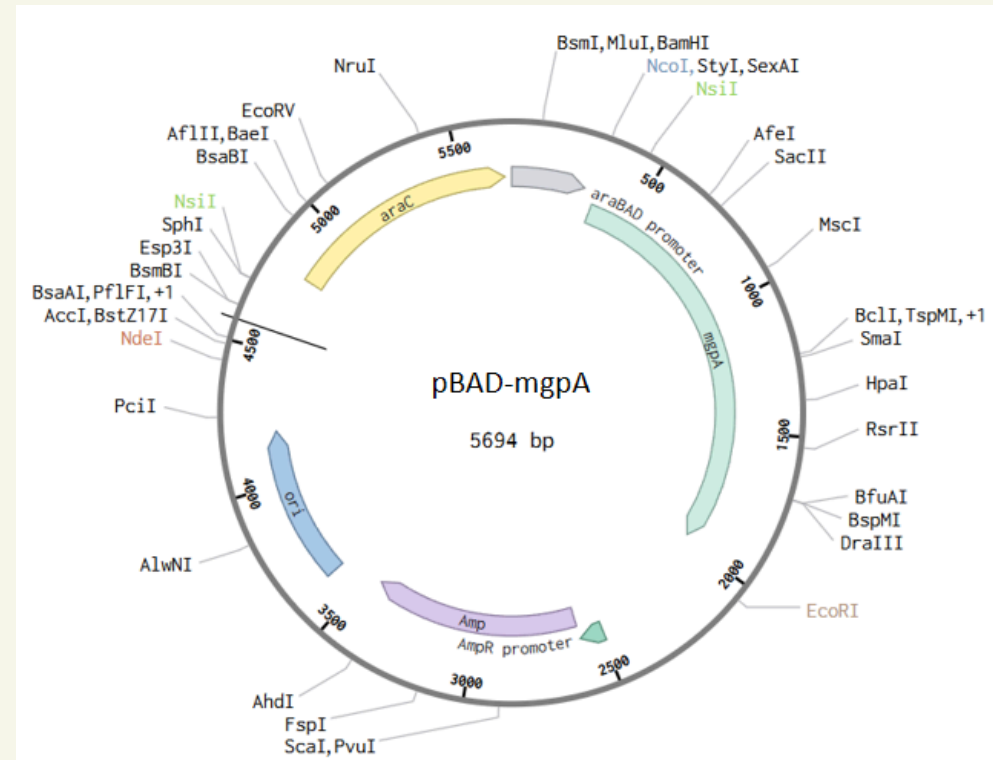


Figure 4: In silico design of the overexpression vector for the mpaA gene. Primers with specific restriction sites were used to isolate the main gene encoding for Adhesin P1, *mgpA*, from the *M. pneumoniae* genome. The **NcoI** and **EcoRI** restriction enzymes were used to introduce sticky in both the vector backbone and GOI in order to create the recombinant **pBAD-mgpA** plasmid. This construct can be transformed into *E.coli* cells, where Adhesin can be expressed.

MECHANISM OF ADHESION

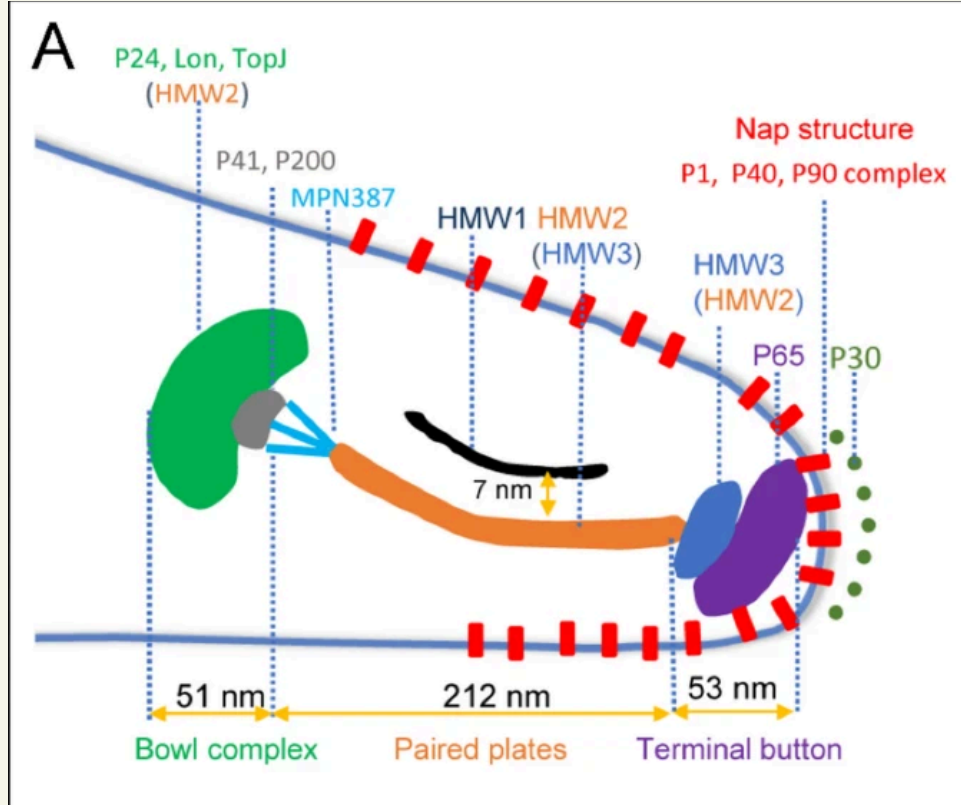
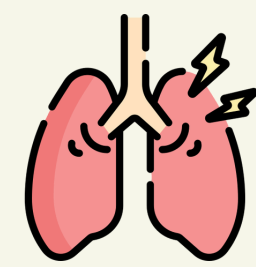


Figure 5: Attachment organelle in *M. pneumoniae* [7]. The C-terminal of P1 plays a crucial role in the attachment process. It is mostly clustered at the tip of the attachment organelle, where **P1**, attached to **P40/90 complex**, recognises and attaches to **sialylated oligosaccharides**. This regulates the P40/90 binding site, making P1 essential for gliding motility and adhesion [7]. **Proteins A, B, and C** act as **scaffolds**, ensuring the correct localisation of P1 in the mycoplasma membrane. Also, the proteins **HMW1, 2, and 3** are critical for the structural integrity of the entire attachment organelle. Regarding the motility in *M. pneumoniae*, it is **ATP-dependent** as well as **unidirectional** thanks to the Nap machinery [7, 8]. Moreover, the energy requirements of *mycoplasma* are directly related to its overall ability to attach to different surfaces via the relationship between the **phosphorylation rate of P1** and **HMWs** and **glucose content** [8].

CONCLUSION

The fragment was identified as part of *M. pneumoniae* by BLASTN. **Adhesin P1** was found in the DNA fragment by KBase. Other proteins found in the same gene neighborhood, also had cell adhesion properties. The structure of the amino acid sequence (from KBase) only **partially covers** the structure (7BWM) from blastp, but they do align. There was found that the KBase sequence only contains a shorter C-terminal domain of adhesin P1 and no N-terminal domain. In addition, its cooperation with P40/90 is essential for inducing adhesion and gliding motility. Adhesin P1 was overexpressed in silico in the plasmid pBAD-sfGFP. This enables further study of the protein structure and functionality, host interactions, immune recognition and potential vaccine production. Since P1 exhibits such a central role in adhesion and inhibition of the immune response, targeting its C-terminal is a leading and sensible strategy for development of **mRNA vaccines**.

REFERENCES

