ACCSCIENCE PUBLISHING

Microbes & Immunity

ORIGINAL RESEARCH ARTICLE

Identification and characterization of novel outer membrane proteins of *Brachyspira pilosicoli*

Amisha Panda[1], Jahnvi Kapoor[1], Batchu Hareramadas[2], Ilmas Naqvi[2], Ravindresh Chhabra[3], Sanjiv Kumar[4]*, and Anannya Bandyopadhyay[1]*

[1]Department of Zoology, Faculty of Science, University of Delhi, New Delhi, Delhi, India

[2]Department of Zoology, Zakir Husain Delhi College, University of Delhi, New Delhi, Delhi, India

[3]Department of Biochemistry, School of Basic Sciences, Central University of Punjab, Bathinda, Punjab, India

[4]Independent Researcher, Stockholm, Sweden

Abstract

Brachyspira pilosicoli is a globally prevalent, anaerobic, Gram-negative spirochete that causes intestinal spirochetosis in birds, pigs, and humans. It colonizes the large intestine, causing colitis, diarrhea, and impaired growth. Despite its pathogenic relevance, the outer membrane proteins of *B. pilosicoli* remain largely uncharacterized. In this study, we computationally identified a total of 42 outer membrane β-barrel (OMBB) proteins within the *B. pilosicoli* proteome using a consensus-based computational framework. Structural models generated using AlphaFold 3 confirmed the β-barrel architectures of the predicted proteins. Structure- and sequence-based functional annotations revealed homologs of *β*-barrel assembly machinery BamA protein, lipopolysaccharide-assembly protein LPS-assembly protein D, TolC, transporter proteins, enzymes, diffusion channels, and porins. Notably, seven of the predicted OMBB proteins were previously unannotated in UniProt and the National Center for Biotechnology Information; we report their putative functions here for the 1st time. Sequence variation analysis among the homologs of OMBB proteins across nine *B. pilosicoli* strains revealed that many of the variations were present within surface-exposed loop regions, suggesting roles in host interaction and immune modulation. Our *in silico* study expands the functional repertoire of *B. pilosicoli* outer membrane proteins, highlighting potential targets for diagnostics, vaccine development, and therapeutic interventions.

Keywords: *Brachyspira pilosicoli;* Intestinal spirochetosis; Outer membrane proteins; β-barrel structures; Structural models; Sequence variations; Functional annotations; *In silico*

***Corresponding authors:**
Anannya Bandyopadhyay (anannya@zoology.du.ac.in)
Sanjiv Kumar (drsanjivk@gmail.com)

Citation: Panda A, Kapoor J, Hareramadas B, *et al*. Identification and characterization of novel outer membrane proteins of *Brachyspira pilosicoli*. *Microbes & Immunity*. 2025;2(4):79-109. doi: 10.36922/MI025230050

Received: June 6, 2025

Revised: July 9, 2025

Accepted: August 1, 2025

Published online: August 25, 2025

1. Introduction

Brachyspira pilosicoli, previously known as *Serpulina pilosicoli*, is a zoonotic bacterium belonging to the family Brachyspiraceae, within the order *Spirochaetales* and phylum Spirochaetota.[1] It is a Gram-negative, anaerobic, slow-growing, double-membraned, flagellated bacterium. *B. pilosicoli* causes intestinal spirochetosis (IS) in higher animals, including avian intestinal spirochetosis (AIS) in birds, porcine intestinal spirochetosis

(PIS) in pigs, and human intestinal spirochetosis (HIS) in humans. Spirochetal infections have been reported in the United Kingdom, continental Europe, Scandinavia, North America, Oceania, Iran, Malaysia, and South America.[2-7] *B. pilosicoli* has a broad host range,[8,9] including dogs, monkeys, water birds, game birds, and humans.[8]

In IS, numerous brachyspiral cells penetrate the mucosal layer overlying enterocytes in the small intestine, attaching one end to the luminal surface of the enterocytes, aided by surface lipoproteins. This attachment forms a distinctive layer resembling a "false brush border."[8] *B. pilosicoli* is the sole etiological agent of PIS, which is marked by diarrhea and impaired growth in pigs.[9,10] AIS in chickens is associated with the delayed onset of egg laying, wet and bloody feces, reduced growth rate, and diarrhea.[3,8,11] HIS is associated with a range of non-specific clinical symptoms, including abdominal pain, altered bowel patterns, chronic diarrhea, and rectal bleeding.[3,12-15]

Common risk factors for zoonotic transmission of *B. pilosicoli* to humans include exposure to fecally contaminated water,[9,16-18] rural or animal exposure, overcrowding, socioeconomic depression, travel to less developed countries, immunosuppression due to HIV infection, or being a homosexual male.[8] AIS and PIS are underreported diseases, bearing significant economic consequences for global food production. Although no comprehensive cost analysis for PIS exists, AIS alone is estimated to cost the poultry industry approximately GBP 18 million annually in the United Kingdom.[3] Extrapolating from these figures, combined global economic losses to both industries could reach approximately USD 1–2 billion annually.[19]

Antibiotics are used to treat AIS, PIS, and HIS; however, resistance has been reported.[18] Antibiotics such as co-amoxicillin and metronidazole are used to treat HIS, whereas pleuromutilins, macrolides, and lincosamides are used for AIS and PIS.[18] Although antibiotics are commonly used, no vaccines are currently available to prevent HIS, AIS, or PIS, highlighting the urgent need for vaccine development.

The reference strain (i.e., *B. pilosicoli* strain 95/1000) possesses a single circular chromosome of approximately 2.59 Mb and lacks any extrachromosomal elements. The *B. pilosicoli* genome comprises 2338 genes, with coding regions accounting for approximately 85% of the total genome.[20] Like other Gram-negative bacteria, *B. pilosicoli* consists of a central protoplasmic cylinder enclosed by a membrane sheath known as the outer membrane (OM).[21] The exact composition of *B. pilosicoli*'s OM is not fully understood; however, it is known to be extremely labile due to its high sterol content, which results in low resistance to osmotic stress and destabilisation when exposed to low-ionic-strength buffers.[22]

The *B. pilosicoli* outer envelope contains lipooligosaccharides (LOS) rather than lipopolysaccharides (LPS), exhibiting serological diversity across multiple strains.[23] Bacteria with diderm envelopes possess a diverse family of OM proteins (OMPs), characterized by *β-barrel* structures (OMBBs) and LOS.[24,25] β-barrels are protein structures composed of amphipathic, anti-parallel β-strands that close in on themselves, forming a cylindrical structure. The β-barrels of OMPs are typically composed of an even number of β-strands, typically ranging from 8 to 36.[26] These β-strands are alternately connected on each side of the OM by long loops on the extracellular surface and by shorter turns on the periplasmic side.[27] OMBB proteins are involved in a range of functions, including nutrient acquisition, membrane biogenesis, assembly of OMPs, adhesion, biofilm formation, efflux, proteolysis, and pilus formation.[28] Thus, OMBB proteins represent a crucial area of research and a promising target for developing antibacterial therapies to combat pathogenic microbes.

Notably, few OMPs of *B. pilosicoli* have been studied, including BmpC (a 23 kDa lipoprotein),[22] a 45 kDa surface-exposed lipoprotein,[29] and Bmp72.[30] Christodoulides *et al.*[19] employed an *in silico* reverse vaccinology approach to identify potential vaccine candidates from predicted OMBB proteins. Although a few OMPs and lipoproteins of *B. pilosicoli* have been identified,[22,29-31] the identification and characterization of the complete OM proteome are needed to define their potential roles in disease pathogenesis, particularly in processes such as attachment, virulence, and eliciting host immune responses.

In this study, a comprehensive *in silico* approach was employed to identify novel OMBB proteins in *B. pilosicoli.* A consensus of the outputs from OM localization prediction tools and *β*-barrel conformation prediction tools was considered for OMBB protein prediction. Through stringent screening criteria and manual curation, 42 putative OMBB proteins were selected. In addition, deep-learning-based structural models of the proteins were generated. Structural homologs were identified using the digital addressable lighting interface (DALI) server and Foldseek tool, revealing the functional roles of the proteins. Furthermore, sequence-based annotations were performed using PANNZER and eggNOG-mapper. Amino acid sequence variations in the predicted proteins were obtained from nine strains of *B. pilosicoli* and mapped onto the structural models. This study identified a total of 42 OMBB proteins of *B. pilosicoli*, computationally characterized their structure and function, and identified peptide regions potentially crucial for bacterial pathogenesis.

2. Materials and methods

2.1. OM β-barrel (OMBB) protein prediction

Given that reference genomes provide a streamlined, standardized, and taxonomically diverse representation of the RefSeq collection,[32] we selected the reference strain *B. pilosicoli* 95/1000, a porcine isolate, for our study. Using genome assembly ASM14372v1, the sequences of all proteins from the *B. pilosicoli* 95/1000 genome were downloaded from the National Center for Biotechnology Information (NCBI) (https://www.ncbi.nlm.nih.gov/).[32]

Peptide length, molecular weight, charge, and isoelectric point for all protein sequences were determined using the Pepstats tool from the EMBOSS package (https://www.ebi.ac.uk/jdispatcher/seqstats/emboss_pepstats).[33] The presence of signal peptide was determined using SignalP 5.0 (https://services.healthtech.dtu.dk/services/SignalP-6.0/; accessed on April 20, 2024) and LipoP 1.0 (https://services.healthtech.dtu.dk/services/LipoP-1.0/; accessed on April 17, 2024). SignalP was employed to predict the presence and cleavage position of signal peptides in the protein sequences.

The SignalP server generates output for each protein sequence in the following categories: Secretory signal peptide ("Sec/SPI"), lipoprotein signal peptide ("Sec/SPII"), Tat signal peptide ("Tat/SPI"), Tat lipoprotein signal peptide ("Tat/SPII"), pilin signal peptide ("Sec/SPIII"), or the absence of any signal peptide ("Other").[34]

The LipoP server predicts lipoproteins in Gram-negative bacteria and distinguishes between lipoprotein signal peptides, other signal peptides, and N-terminal transmembrane helices (TMHs). The output is classified into four classes: Secretory signal peptide ("SpI"), lipoprotein signal peptide ("SpII"), N-terminal TMH ("TMH"), and cytoplasmic protein ("Cyt").[35] The N-terminal TMH serves as an anchor, stabilizing the protein within the membrane. Therefore, LipoP was used as a secondary tool to predict signal peptides.

CELLO v.2.5 (http://cello.life.nctu.edu.tw/; accessed on April 19, 2024)[36] and PSORTb 3.0 (https://www.psort.org/psortb/; accessed on April 30, 2024)[37] were utilized to predict the subcellular localization of proteins. Essential proteins from *B. pilosicoli* were predicted by performing BLASTP searches against the Database of Essential Genes (DEG) v15.2 (https://ngdc.cncb.ac.cn/databasecommons/database/id/229; accessed on April 2, 2024).[38] The DEG database is a repository of essential proteins from archaea, bacteria, and eukaryotes, and assumes that proteins essential in one organism are likely to be essential in others. Specifically, proteins with an $E < 1\times10^{-3}$ and a bit score >100 were considered essential.

The computational framework designed to select OMPs is detailed and schematically represented in Figure 1. We employed a consensus-based computational approach to identify OMBB proteins, where the outputs from four OMP prediction tools were considered: One from OMPdb (http://aias.biol.uoa.gr/OMPdb/; accessed on April 21, 2024),[39] MCMBB (http://athina.biol.uoa.gr/bioinformatics/mcmbb/; accessed on May 1, 2024),[40] TMBETADISC-radial basis function (RBF) (http://rbf.bioinfo.tw/~sachen/OMP.html; accessed on April 23, 2024),[41] and TMbed (https://github.com/BernhoferM/TMbed; accessed on April 20, 2024).[42] Protein sequences were searched against those in the OMPdb database to identify homologous proteins (with $E < 1\times10^{-3}$; bit score > 100). OMPdb is a database of integral β-barrel OMPs from Gram-negative bacteria.

MCMBB distinguishes β-barrel OMPs from globular proteins and α-helical membrane proteins. In MCMBB, a score > 0 indicates a higher likelihood of β-barrel conformation, whereas a score lower than zero suggests the protein is not a β-barrel. The TMBETADISC-RBF server predicts OMPs using an RBF network and position-specific scoring matrix profiles. TMbed, based on embeddings from protein language models, predicts the propensity of each residue to form TMHs, transmembrane β-strands, signal peptides, or other structural elements.

Using a consensus-based approach, predictions from the aforementioned tools were used to identify potential OMBB proteins. A final list of 42 OMBB proteins was compiled based on the number of tools predicting β-barrel architecture for each protein. These proteins were categorized based on the number of tools providing positive predictions. Higher confidence was assigned to proteins predicted as OMBBs by a greater number of tools.

2.2. Structural modeling

Structural models of the predicted OMBB proteins were generated using the AlphaFold server, powered by AlphaFold 3 (https://alphafoldserver.com/; accessed on July 21, 2024).[43] The modeling process incorporates physical and chemical constraints to accurately predict protein folding, resulting in atomic coordinates for each OMBB protein. Outputs of AlphaFold 3 include confidence metrics, namely: Predicted local distance difference test, predicted aligned error, predicted template modeling (pTM), and interface predicted template modeling (ipTM) scores. The pTM and ipTM scores assess the accuracy of the overall structure.[44,45] A pTM score above 0.5 and an ipTM score above 0.8 indicate highly reliable predictions. The top-ranked predictions, based on predicted local distance difference test scores, were selected for figure generation

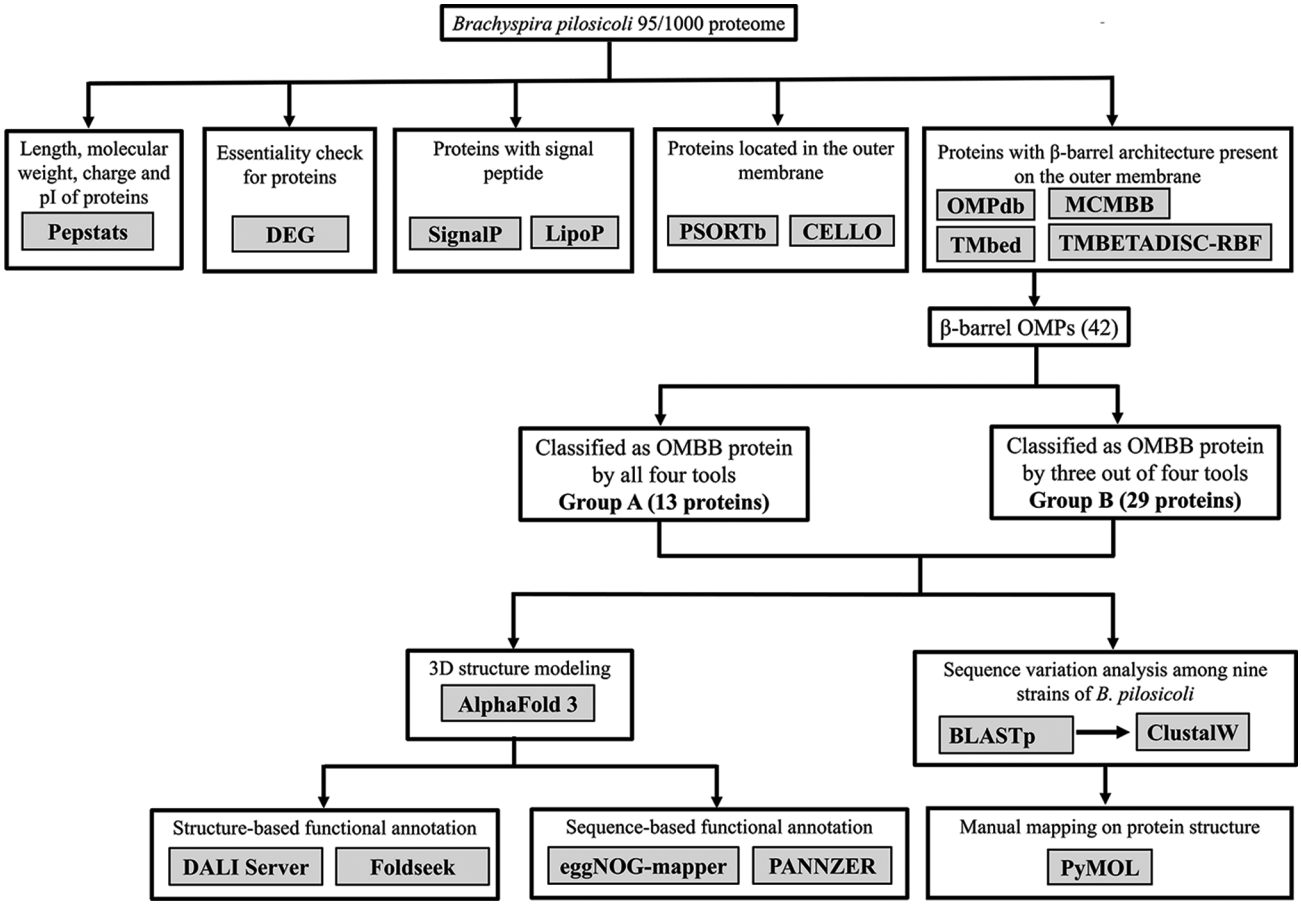

Figure 1. Computational framework for predicting outer membrane-localized β-barrel proteins from *Brachyspira pilosicoli* 95/1000. The *B. pilosicoli* proteome was mined *in silico* using various tools, including SignalP, LipoP, PSORTb, CELLO, OMPdb, MCMBB, TMbed, and TMBETADISC-RBF. Protein essentiality was assessed using the Database of Essential Genes (DEG). Protein size, net charge, and isoelectric point of proteins were predicted using Pepstats. Structural models were generated using the AlphaFold 3 server and used as queries in the DALI server and Foldseek to annotate putative functions. Additionally, sequence-based annotation tools (eggNOG-mapper and PANNZER) were used to predict functional roles. Amino acid sequence variation across nine strains of *B. pilosicoli* was analyzed using ClustalW and mapped onto the structural models using PyMOL.
Abbreviations: OMBB: Outer membrane β-barrel; OMP: Outer membrane protein; pI: Isoelectric point.

and further analysis. The resulting atomic coordinate files were visualized using PyMOL (Warren Lyford DeLano, Schrödinger Inc., USA).[46]

To validate the structures generated by AlphaFold 3, comparative models were also generated using other structure prediction tools: ESMFold (https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/ESMFold.ipynb; accessed on January 23, 2025),[47] SWISS-MODEL (https://swissmodel.expasy.org/; accessed on January 23, 2025),[48] RoseTTA (https://robetta.bakerlab.org/; accessed on January 23, 2025),[49] and TrRosetta (https://yanglab.qd.sdu.edu.cn/trRosetta/; accessed on January 23, 2025).[50]

2.3. Functional annotation of the predicted proteins

Given that many of the OMBB proteins were unannotated hypothetical proteins, a structure-based approach was employed to unravel their functional roles. Atomic coordinates of the structural models were used as queries in the DALI server (http://ekhidna.biocenter.helsinki.fi/dali/; accessed on May 23, 2024)[51] with the full Protein Data Bank (PDB) search option, ensuring that the query was compared against all protein structures available in the PDB. The top-hit protein with the highest Z-score was selected for functional annotation. The Z-score is an optimized similarity score based on the sum of equivalent Cα–Cα distances between two proteins. A score >20 indicates definite homology, 8–20 suggests potential homology, whereas scores <8 indicate insignificant similarity.[52] Functions of the top-hit proteins were retrieved from the available literature to annotate the predicted OMBB proteins.

In addition, the Foldseek tool (https://search.foldseek.com/search; accessed on February 2, 2025) was employed to

identify structural homologs across five protein databases: PDB100, CATH50, AFDB50, AFDB-SWISSPROT, and AFDB-Proteome.[53] The top hit, based on the template modeling score, was selected for functional annotation. Sequence-based functional annotation was performed using PANNZER (http://ekhidna.biocenter.helsinki.fi/pannzer; accessed on February 13, 2025) and eggNOG-mapper v2 (https://eggnog-mapper.embl.de/; accessed on February 2, 2025).[54,55]

2.4. Amino acid sequence variation among different strains of *B. pilosicoli*

Predicted OMBB proteins from the reference genome 95/1000 were searched for similar proteins using BLASTP (E-value < 1×10^{-3}; bit score > 100) across nine completed genomes of *B. pilosicoli* to analyze the amino acid sequence variation in the predicted β-barrel proteins (Table S1). Multiple sequence alignment (MSA) was performed for orthologous sequences of each protein using ClustalW (Thompson JD, Gibson TJ, Higgins DG; EMBL, Heidelberg, Germany).[56] Analysis of the MSA revealed amino acid substitutions among the orthologs. Mapping of these variations onto the structural models was conducted using PyMOL.[46]

In addition, the 16S rRNA gene sequence of *B. pilosicoli* strain P43/6/78 was retrieved from the NCBI database and used as a query for BLASTn against available *B. pilosicoli* genome sequences. From the BLASTn results, strains previously included in the sequence variation analysis were identified, and their corresponding 16S rRNA gene sequences were extracted. A multi-FASTA file was created, followed by MSA using Multiple Sequence Comparison by Log Expectation (Robert C. Edgar, USA). Subsequently, a phylogenetic tree was constructed using the Neighbor-Joining method in MEGA12 software (Kumar S, Tamura K; Temple University, Philadelphia, USA, and Tokyo Metropolitan University, Japan) to determine the evolutionary relationships among the selected strains (Figure S1).[57]

2.5. Structure alignment using the US-align server

Structural models were aligned using the US-align server (https://zhanggroup.org/US-align/; accessed on July 23, 2024)[58] an online web server to assess structural similarities and variations. Structural alignments were visualized, and figures were generated using PyMOL.

3. Results and discussion

3.1. Prediction of OMBB proteins using a consensus-based computational approach

A consensus-based computational framework was applied to the *B. pilosicoli* 95/1000 proteome, consisting of 2,275 proteins, to identify OMBB proteins (Figure 1). Ten computational tools were used for predictions: Pepstats, DEG database, SignalP, LipoP, CELLO, PSORTb, OMPdb, MCMBB, TMBETADISC-RBF, and TMbed (Table S2). Prediction outputs from all tools were combined for each protein. Tools that specifically predict OMBB proteins (e.g., OMPdb, MCMBB, TMBETADISC-RBF, and TMbed) were prioritized for OMBB protein prediction (Table S2).

Through stringent screening criteria and manual curation, a total of 42 OMBB proteins were selected and classified into two groups: Group A (13 proteins, predicted as OMBB by all four tools) and Group B (29 proteins, predicted as OMBB by any three out of four tools) (Table 1).

To gain structural insights into the predicted proteins, we searched for their structures in the PDB but found no experimentally determined models available. Using AlphaFold 3,[43] structural models of the predicted proteins were generated, revealing typical features of OMPs, such as β-barrel architectures with central pores, periplasmic loops, and surface-exposed loops.

Given that most proteins were unannotated, both structure- and sequence-based approaches were employed to assign putative functions. Top-ranking hits were considered for functional annotation. This analysis revealed structural and sequence homologs of well-characterized proteins, including OMP assembly factor BamA, LPS-assembly protein D (LptD), Neisserial surface protein NspA, OM porin F (OmpF), OM phospholipase A (OMPLA), and vitamin B_{12} transporter protein BtuB, thereby providing valuable insights into their possible roles (Tables 2 and S3).

OMPs are located on the bacterial surface, serving as the primary interface between host and pathogen. Due to exposure to the host environment, these proteins are subjected to strong selection pressures, making the analysis of their sequence variability essential for understanding pathogen evolution (Table 3).[59] Building on this, we analyzed the amino acid sequences of the predicted OMBB proteins for residues exhibiting sequence variation across nine *B. pilosicoli* strains (Table S4). Mapping these variations onto the structural models revealed that many variations were located on extracellular loops (ECLs), which are more likely to interact with the host environment (Table 3).

3.2. Identification of OMPs

3.2.1. Group A

Group A comprised of 13 proteins, consisting of three proteins with 16 stranded β-barrel domain (BP951000_RS05730, BP951000_RS10215, and BP951000_RS04760); seven proteins with eight-stranded β-barrel domain

Table 1. Predicted outer membrane β-barrel proteins from *Brachyspira pilosicoli* 95/1000

Locus identifier	Protein name[a]	SignalP	PSORTb	CELLO	TMBETADISC. RBF_AADP	OMPdb Match	MCMBB	Tmbed	No. of β-strands[b]
Group A									
BP951000_RS05730	OMP assembly factor/ BamA	+	+	+	+	+	+	+	16
BP951000_RS10215	Variable surface protein (VspE)	+	U	+	+	+	+	+	16
BP951000_RS04760	Variable surface protein (VspD)	+	U	+	+	+	+	+	16
BP951000_RS01125	CsgG/HfaB family protein	+	U	+	+	+	+	+	8
BP951000_RS03440	OMBB protein	+	U	+	+	+	+	+	8
BP951000_RS05600	TolC family protein	+	–	+	+	+	+	+	18 (trimer of six-stranded protomer)
BP951000_RS09000	TolC family protein	+	U	+	+	+	+	+	18 (trimer of six-stranded protomer)
BP951000_RS06235	TolC family protein	+	+	+	+	+	+	+	12 (trimer of four-stranded protomer)
BP951000_RS04880	Serpentine receptor domain-containing protein	+	+	+	+	+	+	+	8
BP951000_RS02055	Serpentine receptor domain-containing protein	+	+	+	+	+	+	+	8
BP951000_RS02050	Serpentine receptor domain-containing protein	+	+	+	+	+	+	+	8
BP951000_RS07540	Serpentine receptor domain-containing protein	+	+	+	+	+	+	+	8
BP951000_RS00180	Serpentine receptor domain-containing protein	+	U	+	+	+	+	+	8
Group B									
BP951000_RS09575	Lipopolysaccharide-assembly protein (LptD)	+	+	+	+	–	+	+	26
BP951000_RS03215	TonB-dependent siderophore receptor	+	+	+	–	+	+	+	22
BP951000_RS04405	Toxin A	+	+	+	+	–	+	+	18
BP951000_RS09655	DUF5723 domain-containing protein	+	U	+	+	–	+	+	16
BP951000_RS04440	Hypothetical protein	+	U	+	+	–	+	+	16
BP951000_RS08285	Trep protein	–	U	+	+	–	+	+	16
BP951000_RS04505	Variable surface protein (VspH)	–	U	+	–	+	+	+	16
BP951000_RS08455	PorV/PorQ family protein	+	U	+	+	–	+	+	14

(Cont'd...)

Table 1. ***(Continued)***

Locus identifier	Protein name[a]	SignalP	PSORTb	CELLO	TMBETADISC. RBF_AADP	OMPdb Match	MCMBB	Tmbed	No. of β-strands[b]
BP951000_RS01090	Variable surface protein (VspH)	+	+	+	+	−	+	+	14
BP951000_RS06935	Hypothetical protein	+	−	+	+	−	+	+	12
BP951000_RS11380	Toxin A	+	U	+	+	−	+	+	12
BP951000_RS03405	Hypothetical protein	+	U	+	+	−	+	+	12
BP951000_RS00185	Hypothetical protein	+	+	+	+	−	+	+	12
BP951000_RS10320	Hypothetical protein	+	U	+	+	−	+	+	10
BP951000_RS05445	DUF3575 domain-containing protein	+	U	+	+	−	+	+	8
BP951000_RS08300	Tia invasion determinant	+	U	+	−	+	+	+	8
BP951000_RS05490	Tia invasion determinant	+	U	+	+	−	+	+	8
BP951000_RS07500	Hypothetical protein	+	−	+	+	−	+	+	8
BP951000_RS01590	Hypothetical protein	−	−	+	+	−	+	+	8
BP951000_RS08295	Tia invasion determinant	+	U	−	−	+	+	+	8
BP951000_RS08975	TonB-dependent receptor domain-containing protein	+	U	+	+	−	+	+	13
BP951000_RS06930	Serpentine receptor domain-containing protein	−	−	−	+	−	+	+	8
BP951000_RS03290	Serpentine receptor domain-containing protein	+	−	+	+	−	+	+	8
BP951000_RS00765	Serpentine receptor domain-containing protein	+	+	+	+	−	+	+	8
BP951000_RS01280	Serpentine receptor domain-containing protein	+	+	+	+	−	+	+	8
BP951000_RS10445	Serpentine receptor domain-containing protein	+	+	+	+	−	+	+	8
BP951000_RS00365	Serpentine receptor domain-containing protein	+	U	+	+	−	+	+	8
BP951000_RS04620	Serpentine receptor domain-containing protein	+	+	+	+	−	+	+	8
BP951000_RS04220	Serpentine receptor domain-containing protein	+	−	+	+	−	+	+	8

Notes: [a]Protein names follow annotations in the National Center for Biotechnology Information and UniProt databases, retrieved using protein accession numbers (accessed on March 28, 2024). [b]The number of β-strands was predicted using AlphaFold 3. "U" indicates an unknown output from PSORTb, where the tool could not determine the exact cellular localization of the protein. "+" denotes a positive result, and "−" represents a negative result as reported by the respective computational tool.

Abbreviations: DUF: Domain of unknown function; OMBB: Outer membrane β-barrel; OMP: Outer membrane protein; Trep: transcriptional regulating protein.

Table 2. Classification of outer membrane β-barrel proteins based on functional annotation

Functional categories[a]	Subtypes	Protein accession number	Locus identifier	Protein name[a]
Transport and nutrient uptake	Porins	WP_013243854.1	BP951000_RS04405	Toxin A
		WP_013243193.1	BP951000_RS01125	CsgG/HfaB family protein
	Diffusion channels	WP_013244641.1	BP951000_RS08455	PorV/PorQ family protein
		WP_013243185.1	BP951000_RS01090	Variable surface protein (VspH)
		WP_013244339.1	BP951000_RS06935	Hypothetical protein
	TonB-dependent receptors	WP_013244745.1	BP951000_RS08975	TonB-dependent receptor domain-containing protein
		WP_013244607.1	BP951000_RS08285	Trep protein
		WP_041747581.1	BP951000_RS03215	TonB-dependent siderophore receptor
Secretion and export	Secretins	WP_013244995.1	BP951000_RS10215	VspE
		WP_013243917.1	BP951000_RS04760	VspD
		WP_015274839.1	BP951000_RS04505	VspH
	Autotransporters	WP_013244339.1	BP951000_RS06935[d]	Hypothetical protein
	Efflux pumps	WP_013244081.1	BP951000_RS05600	TolC family protein
		WP_013244750.1	BP951000_RS09000	TolC family protein
		WP_041747714.1	BP951000_RS06235	TolC family protein
Structural integrity	OM scaffolding proteins	WP_013243377.1	BP951000_RS02055	Serpentine receptor domain-containing protein[c]
		WP_013243376.1	BP951000_RS02050	Serpentine receptor domain-containing protein[c]
		WP_013244459.1	BP951000_RS07540	Serpentine receptor domain-containing protein[c]
		WP_013242998.1	BP951000_RS00180	Serpentine receptor domain-containing protein[c]
		WP_187287137.1	BP951000_RS07500	Hypothetical protein
		WP_014936494.1	BP951000_RS03290	Serpentine receptor domain-containing protein[c]
		WP_014933009.1	BP951000_RS00765	Serpentine receptor domain-containing protein[c]
		WP_013245039.1	BP951000_RS10445	Serpentine receptor domain-containing protein[c]
		WP_013243896.1	BP951000_RS04620	Serpentine receptor domain-containing protein[c]
		WP_013243815.1	BP951000_RS04220	Serpentine receptor domain-containing protein[c]
	OM biogenesis machinery	WP_013244106.1	BP951000_RS05730	BamA
		WP_041747843.1	BP951000_RS09575	Lipopolysaccharide-assembly protein (LptD)
Adhesion and virulence	Adhesins	WP_013243917.1	BP951000_RS04760[d]	VspD
		WP_013243655.1	BP951000_RS03440	OMBB
		WP_013243940.1	BP951000_RS04880	Serpentine receptor domain-containing protein[c]
		WP_041747873.1	BP951000_RS10320	Hypothetical protein
		WP_013244050.1	BP951000_RS05445	DUF3575 domain-containing protein

(Cont'd...)

Table 2. ***(Continued)***

Functional categories[a]	Subtypes	Protein accession number	Locus identifier	Protein name[a]
		WP_013244610.1	BP951000_RS08300	Tia invasion determinant
		WP_013244059.1	BP951000_RS05490	Tia invasion determinant
		WP_228369485.1	BP951000_RS08295	Tia invasion determinant
		WP_013243225.1	BP951000_RS01280	Serpentine receptor domain-containing protein
		WP_013244338.1	BP951000_RS06930	Serpentine receptor domain-containing protein
		WP_013243037.1	BP951000_RS00365	Serpentine receptor domain-containing protein
	Immune evasion proteins	WP_013244610.1	BP951000_RS08300[d]	Tia invasion determinant
		WP_013242999.1	BP951000_RS00185	Hypothetical protein
		WP_228369485.1	BP951000_RS08295[d]	Tia invasion determinant
Signal transduction	Receptor-like OMPs	WP_181893515.1	BP951000_RS01590	Hypothetical protein
		WP_013243193.1	BP951000_RS01125[d]	CsgG/HfaB family protein
Enzymatic functions	Lipases	WP_013243647.1	BP951000_RS03405	Hypothetical protein

Notes: [a]Protein names follow annotations in the National Center for Biotechnology Information and UniProt databases, retrieved using protein accession numbers (accessed on March 28, 2024). [b]Functional categories were assigned based on consensus predictions from structure- and sequence-based annotation tools (Table S3). [c]As serpentine receptors, or G-protein coupled receptors, are absent in prokaryotes and all tools predicted transmembrane β-barrel structures rather than α-helices, these proteins are likely misannotated as serpentine receptor proteins in UniProt. [d]These proteins were predicted to possess dual roles.
Abbreviations: DUF: Domain of unknown function; OM: Outer membrane; OMBB: Outer membrane β-barrel; OMP: Outer membrane protein; Trep: transcriptional regulating protein.

(BP951000_RS02055, BP951000_RS02055, BP951000_RS07540, BP951000_RS01125, BP951000_RS00180, BP951000_RS03440, and BP951000_RS04880), and three TolC family proteins (BP951000_RS05600, BP951000_RS09000, and BP951000_RS06235) (Table 1). Out of the seven eight-stranded β-barrel proteins, five are annotated as serpentine receptor (SR) domain-containing proteins (BP951000_RS02055, BP951000_RS02055, BP951000_RS07540, BP951000_RS00180, and BP951000_RS04880).

3.2.1.1. BP951000_RS05730

BP951000_RS05730 is annotated as BamA in *B. pilosicoli* strain 95/1000. BamA, along with BamB, BamC, BamD, and BamE, forms the β-barrel assembly machinery complex, which is involved in the assembly and insertion of β-barrel proteins into the OM.[60] BP951000_RS05730 is identified as an essential protein in the DEG database. Its structural model exhibits a characteristic BamA bipartite structure, consisting of a periplasmic N-terminal region and a C-terminal *β*-barrel domain (Figure 2A). The N-terminal segment contains five polypeptide transport-associated (POTRA) domains (P1–P5), each comprising a characteristic β1-α1-α2-β2-β3 motif. In other well-characterized BamA proteins, these domains form a scaffold for the binding of BamB, BamC, BamD, and BamE proteins, and facilitate the folding of OMPs.[61]

Brachyspira pilosicoli BamA consists of 16 antiparallel β-strands, with a characteristic lateral gate between strands 1 and 16. A structural homology search using the DALI server revealed the closest match with BamA of *Escherichia coli* O157:H7 (PDB ID: 7NRE) (Tables 2 and S3). The consensus predictions from other annotation tools (Foldseek, PANNZER, and eggNOG-mapper) validated the functional annotation of BamA in *B. pilosicoli* (Table S3).

Sequence comparison of BP951000_RS05730 across nine strains of *B. pilosicoli* revealed five variations (D60, A184, V465, A467, and F512) (Tables 3 and S4). When mapped onto the structural model, V465, A467, and F512 were present in the β-barrel transmembrane (TM) domain, whereas D60 and V184 were located in the periplasmic region of the protein (Tables 3 and S4).

3.2.1.2. BP951000_RS10215

BP951000_RS10215 is annotated as a hypothetical protein in NCBI. However, it is annotated as a variable surface protein (Vsp), specifically VspE, in the UniProt database. Vsps are OMPs identified in *Brachyspira hyodysenteriae* and *Mycoplasma bovis*, and are used by these pathogenic bacteria to adapt to host conditions and enhance colonization.[62,63] These proteins can undergo reversible on/off expression

Table 3. Sequence variations among nine strains: Total variations and variations in the extracellular loops of predicted outer membrane β-barrel proteins

Locus identifier	Protein names[a]	Total variations	Variations present on the predicted extracellular loop region
Group A			
BP951000_RS05730	OMP assembly factor BamA	Six variations: D60, A184, V465, A467, and F512	None
BP951000_RS10215[b]	Variable surface protein (VspE)	208 variations	Not determined[c]
BP951000_RS04760[b]	Variable surface protein (VspD)	260 variations	Not determined[c]
BP951000_RS01125	CsgG/HfaB family protein	Five variations: S63, D79, T190, I210, and L380	None
BP951000_RS03440	OMBB protein	Six variations: F24, V47, V64, N110, D169, and A197	V47
BP951000_RS05600	TolC family protein	Two variations: T246 and N499	None
BP951000_RS09000	TolC family protein	Two variations: S90 and S131	S90
BP951000_RS06235	TolC family protein	18 variations: K2, N3, F5, V6, F7, I8, I10, L12, S16, S25, N33, I42, E43, L93, S105, E136, I137, and T210	L93
BP951000_RS04880	Serpentine receptor domain-containing protein	One variation: N69	None
BP951000_RS02055	Serpentine receptor domain-containing protein	Three variations: H101, N163 and M235	N163 and M235
BP951000_RS02050	Serpentine receptor domain-containing protein	Five variations: A27, L28, T108, A228, and I247	T108 and A228
BP951000_RS07540	Serpentine receptor domain-containing protein	Six variations: M1, K2, K3, I4, I5, and L6	None
BP951000_RS00180	Serpentine receptor domain-containing protein	Two variations: I64 and M115	None
Group B			
BP951000_RS09575	Lipopolysaccharide-assembly protein (LptD)	Seven variations: N14, G137, I257, I382, E454, D600, and G944	D600
BP951000_RS03215[b]	TonB-dependent siderophore receptor	59 variations	Not determined[c]
BP951000_RS04405	Toxin A	23 variations: M1, H2, R3, I4, I6, L8, T9, M18, V19, T24, N32, S34, N41, F84, K90, N98, I101, S102, N104, S175, Q183, I264, and T303	N41, F84, K90, N98, S175, and Q183
BP951000_RS09655[b]	DUF5723 domain-containing protein	315 variations	Not determined[c]
BP951000_RS04440	Hypothetical protein	13 variations: K104, K113, S117, Y124, I132, T134, N151, G153, L243, V252, L254, S308, N321	K104 and L243
BP951000_RS08285[b]	Trep protein	43 variations	Not determined[c]
BP951000_RS04505[b]	Variable surface protein (VspH)	247 variations	Not determined[c]
BP951000_RS08455	PorV/PorQ family protein	Six variations: L12, S20, N22, A117, R187, and S253	N22 and A117
BP951000_RS01090	Variable surface protein (VspH)	One variation: E258	E258
BP951000_RS06935	Hypothetical protein	Four variations: S9, I10, V13, and R298	None
BP951000_RS11380	Toxin A	Three variations: M126, M154, and I278	M126

(Cont'd...)

Table 3. ***(Continued)***

Locus identifier	Protein names[a]	Total variations	Variations present on the predicted extracellular loop region
BP951000_RS03405	Hypothetical protein	24 variations: M1, R2, L3, K4, F5, F6, F7, L8, I9, F10, L11, F12, L13, S14, L15, S16, L17, Y18, T19, Q20, D21, N22, E23, and A24	None
BP951000_RS00185[b]	Hypothetical protein	58 variations	Not determined[c]
BP951000_RS10320	Hypothetical protein	25 variations: L18, D48, E55, F251, G285, E400, Y401, G402, I403, F404, T405, K406, Q407, L408, A409, I410, S411, F412, I413, P414, I415, N416, I417, R418, and F419	None
BP951000_RS05445	DUF3575 domain-containing protein	Six variations: K2, I7, A79, N87, H89, and K158	N87 and H89
BP951000_RS08300	Tia invasion determinant	Three variations: L143, N156, and S200	None
BP951000_RS05490[b]	Tia invasion determinant	61 variations	Not determined[c]
BP951000_RS07500[b]	Hypothetical protein	183 variations	Not determined[c]
BP951000_RS01590	Hypothetical protein	Three variations: V205, I215, and V221	None
BP951000_RS08295	Tia invasion determinant	Six variations: N34, I49, V123, S141, I144, and V167	N34
BP951000_RS08975	TonB-dependent receptor domain-containing protein	Two variations: D32 and T371	D32
BP951000_RS06930[b]	Serpentine receptor domain-containing protein	135 variations	Not determined[c]
BP951000_RS03290	Serpentine receptor domain-containing protein	None	None
BP951000_RS00765[b]	Serpentine receptor domain-containing protein	70 variations	Not determined[c]
BP951000_RS01280	Serpentine receptor domain-containing protein	Five variations: V32, A83, V124, E210, and T237	E210
BP951000_RS10445	Serpentine receptor domain-containing protein	G228	None
BP951000_RS00365	Serpentine receptor domain-containing protein	11 variations: V72, Q77, I84, D156, D159, V168, N177, A200, T216, I222, and Y226	D156, D159, and A200
BP951000_RS04620	Serpentine receptor domain-containing protein	Four variations: K2, E95, A140, and V194	None
BP951000_RS04220[b]	Serpentine receptor domain-containing protein	218 variations	Not determined[c]

Notes: [a]Protein names follow annotations in the National Center for Biotechnology Information and UniProt databases, retrieved using protein accession numbers (accessed on March 28, 2024). [b]Sequence comparison of these proteins across nine strains of *Brachyspira pilosicoli* revealed variations at more than 40 positions. Therefore, they have been listed in Table S5. [c]Given that sequence variations were present at more than 40 positions, we have not determined whether these variations are present on the transmembrane region or loop region of the predicted proteins.
Abbreviations: DUF: Domain of unknown function; OMBB: Outer membrane β-barrel; OMP: Outer membrane protein; Trep: transcriptional regulating protein.

switching or antigenic variation by expressing alternative protein phenotypes.[62] They may function as mediators for bacterial attachment to host cells.[64] Vsp-like proteins have been identified in *B. hyodysenteriae*, *B. pilosicoli*, and *Mycoplasma*;[29,62,65] however, homologs in other bacterial genera remain undiscovered, highlighting their unique role in these pathogens.

Brachyspira pilosicoli VspE, BP951000_RS10215, contains a secretory signal peptide. The structural model generated by AlphaFold 3 revealed a β-barrel architecture consisting of 16 β-strands, with the ninth and 10th strands longer than the others, giving an elliptical shape to the extracellular surface of the barrel (Figure 2B). BP951000_RS10215 exhibited the best structural alignment with the

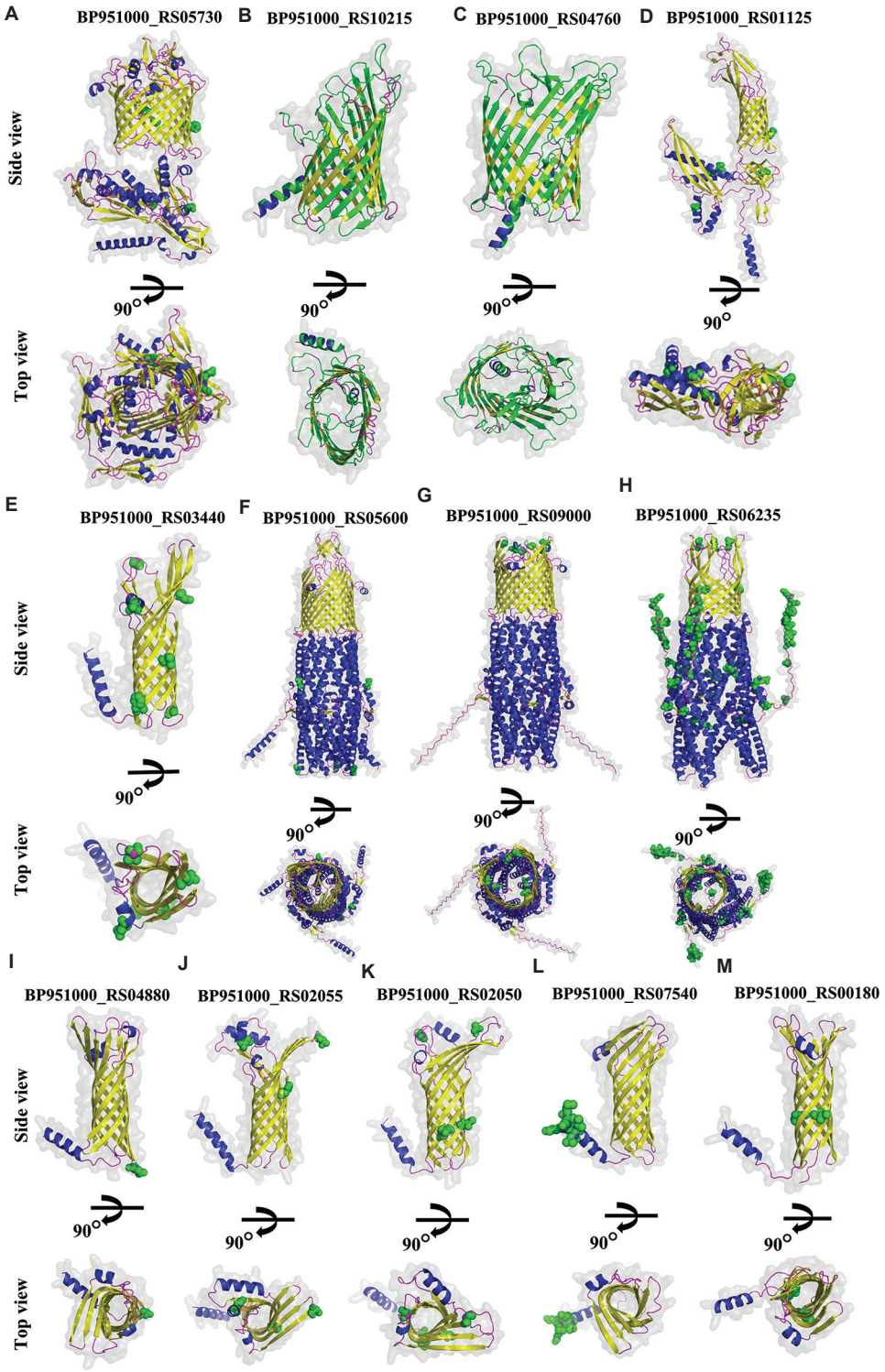

Figure 2. Structural models of β-barrel outer membrane proteins in Group A. Group A consists of 13 proteins. (A-C) Proteins with 16 *β*-strands; (F-H) predicted trimeric structures of TolC family proteins. The remaining proteins exhibit eight-stranded β-barrel structures (D-E, I-M). β-strands, α-helices, and loops are colored yellow, blue, and magenta, respectively. Green spheres indicate amino acid variations identified across nine strains of *Brachyspira pilosicoli*. Proteins with more than 40 variations are shown in green ribbon representation.

C-terminal β-barrel domain of poly-β-1,6-N-acetyl-D-glucosamine (PNAG) export protein PgaA of *E. coli* K-12 (PDB ID: 4Y25) (Tables 2 and S3). *E. coli* PgaA comprises a 16-stranded β-barrel domain at the C-terminal and eight periplasmic tetratricopeptide repeats (TPRs) at the N-terminal. In contrast, BP951000_RS10215 lacks periplasmic TPR domains. PgaA facilitates the translocation of PNAG polymer from the periplasm to the cell surface, a key step in biofilm formation.[66-68] The structural similarity of β-barrels between *E. coli* PgaA and BP951000_RS10215 implies a possible role of the Brachyspiral protein in translocation. Consistently, the Foldseek tool identified its closest match with the bacterial polysaccharide OM secretion of *E. coli* K-12, supporting a potential role in polysaccharide secretion.

In parallel, PANNZER annotated BP951000_RS10215 as VspB, suggesting a possible involvement in surface antigenic variation. Together, these findings suggest a possible dual function in secretion with surface variability (Tables 2 and S2). Given its high sequence and structural homology with *B. hyodysenteriae,* there is a strong likelihood that BP951000_RS10215 plays a role in adherence and host colonization.[62-64] Sequence variation analysis across nine strains of *B. pilosicoli* revealed 208 amino acid substitutions and several deletions (Table S5). These variations are distributed throughout the protein (Figure 2B), suggesting the ability of *B. pilosicoli* to rapidly adapt to changing environments or host immune responses.[62]

3.2.1.3. BP951000_RS04760

BP951000_RS04760 is annotated as VspD in the UniProt database and as a variable surface family protein in NCBI. As discussed in Section 3.2.2, Vsps are involved in bacterial attachment to host cells.[64] *B. hyodysenteriae* VspD is a virulence factor and a potential vaccine development target.[69] Our predictions revealed that BP951000_RS04760 carries a signal peptide and comprises a 16-stranded β-barrel architecture, with varying strand lengths creating an elliptical barrel surface on the extracellular side (Figure 2C). The protein showed the closest structural match with the β-barrel domain of the cellulose synthase operon protein C (BcsC porin; PDB ID: 6TZK) from *E. coli* K-12 (Tables 2 and S3), as determined using the DALI server. BcsC is a 16-stranded β-barrel protein with a periplasmic domain consisting of 19 TPRs, which facilitate the secretion of phosphoethanolamine-cellulose across the OM.[70-74] In contrast, BP951000_RS04760 lacks TPRs. The structural homology between the β-barrel domains of *E. coli* BcsC and BP951000_RS04760 implies a possible role for BP951000_RS04760 in translocation. This structural insight is reinforced by the Foldseek tool, which identified its closest match with the bacterial polysaccharide OM secretin of *E. coli* K-12, further supporting involvement of this protein in secretion. Complementing this, PANNZER annotated BP951000_RS10215 as VspD, which, in *B. hyodysenteriae*, is associated with adhesion and virulence. Together, these findings suggest a dual role for the protein in secretion and adhesion. (Tables 2 and S3). Among nine strains of *B. pilosicoli*, BP951000_RS04760 exhibited 260 amino acid substitutions and multiple deletions, indicating high variability (Table S5). These variations were distributed throughout the protein (Figure 2C).

3.2.1.4. BP951000_RS01125

BP951000_RS01125 is annotated as the curli production assembly/transport component CsgG in both the NCBI and UniProt databases. SignalP predicted a lipoprotein signal peptide, whereas LipoP predicted it as a cytoplasmic protein. The structural model of the protein showed a β-barrel architecture comprising eight β-strands extending into the periplasm via the periplasmic domain (Figure 2D). DALI server results showed that BP951000_RS01125 exhibited the best structural match with OmpF (PDB ID: 4RLC) of *Pseudomonas aeruginosa* (Tables 2 and S3). *P. aeruginosa* OmpF is involved in biofilm formation, OM vesicle production, adhesion, and host immune system modulation.[75-81] The Foldseek tool identified its closest structural match with an uncharacterized protein from the marine metagenome. Sequence-based annotation using PANNZER and eggNOG-mapper confirmed BP951000_RS01125 to be a CsgG homolog (Tables 2 and S3).

The CsgG curli production assembly/transport component OMP is essential for the secretion of curli—functional amyloid fibers that constitute the primary protein component of biofilm extracellular matrices in *Bacteroidetes* and Proteobacteria—and play key roles in pathogenesis.[82] Curli fimbriae are involved in the initial colonization of the host, as well as in bacterial persistence and invasion.[83-85] Considering that both OmpF and CsgG are functionally linked to surface-associated processes and virulence, the combined structural and sequence-level analyses strongly suggest that BP951000_RS01125 may play a similar role in *B. pilosicoli*. Across *B. pilosicoli* strains, BP951000_RS01125 exhibited sequence variations at five positions (S63, D79, T190, I210, and L380) (Tables 3 and S4). Structural mapping revealed that L380 is located within the β-barrel domain, whereas the remaining variations are positioned in the periplasmic region of the protein (Table S4).

3.2.1.5. BP951000_RS03440

BP951000_RS03440 is annotated as a hypothetical protein in the UniProt database, whereas NCBI identifies it as an

"OMBB protein." Our study supports the latter annotation, revealing the presence of a secretory signal peptide. Structural modeling using AlphaFold 3 showed a β-barrel architecture comprising eight β-strands (Figure 2E). Structural similarity assessment using the DALI server identified the closest match with *Neisseria meningitidis* NspA (PDB ID: 1P4T) (Tables 2 and S3). NspA, an eight-stranded β-barrel protein, is involved in bacterial attachment and interaction with the host immune system and has been proposed as a potential vaccine candidate.[86-89] Foldseek analysis identified an uncharacterized protein of *Brachyspira murdochii* as its closest structural homolog, suggesting potential species-specific divergence.

Sequence-based annotation using PANNZER predicted the presence of an OMBB domain (Tables 2 and S3). Together, the structure- and sequence-based data strongly support the classification of BP951000_RS03440 as an OMBB protein, likely involved in host-pathogen interaction. MSA across *B. pilosicoli* strains identified six variations (F24, V47, V64, N110, D169, and A197) (Tables 3 and S4). Structural mapping localized one variation (V47) to the ECL region, four variations (F24, V64, N110, A197) to the TM region, and one variation (D169) to the intracellular region of the protein (Table S4).

3.2.1.6. BP951000_RS05600

BP951000_RS05600 is annotated as a putative OM component of a multidrug efflux system in UniProt and as a TolC family protein in NCBI. Both SignalP and LipoP predicted a secretory signal peptide. The AlphaFold 3 model revealed a trimer 18-stranded β-barrel, with each subunit (protomer) contributing six β-strands (i.e., one-third of the barrel) (Figure 2F). As no known 18-stranded TolC structures have been discovered to date, we validated this model using TrRosetta, RoseTTAFold, ESMFold, and SWISS-MODEL. All models showed high similarity and aligned closely with the AlphaFold 3 model. Alignment of the five monomeric models yielded a root mean square deviation (RMSD) of 3.96 Å, supporting model reliability (Figure S2). Like canonical TolC, the TM barrel extends into the periplasm as an α-helical tunnel, connected to the OM (Figure 2F),[90] with its periplasmic entry blocked, likely to prevent leakage through the OM, as the β-barrel domain remains constantly open.[91]

BP951000_RS05600 showed the best structural match with *E. coli* K-12 TolC (PDB ID: 6WXI) (Tables 2 and S3). TolC functions in hemolysin secretion,[92,93] colicin import,[94,95] antibiotic efflux,[96] and serves as a bacteriophage receptor.[97] PANNZER and eggNOG-mapper respectively annotated the protein as "integral OMP TolC" and "efflux TM transporter" (Tables 2 and S3), suggesting similar roles in *B. pilosicoli*. Sequence comparison across nine strains of *B. pilosicoli* revealed two variations (T246 and N499) (Tables 3 and S4), both mapped to the periplasmic α-helical regions of the protein (Table S4).

3.2.1.7. BP951000_RS09000

BP951000_RS09000 is annotated as an OM efflux protein in UniProt and as a TolC family protein in NCBI. It carries a secretory signal peptide. The AlphaFold 3 model revealed a trimeric 18-stranded β-barrel, with each protomer contributing six β-strands (Figure 2G). Similar to BP951000_RS05600 (Section 3.2.6), we validated this structure using TrRosetta, RoseTTAFold, ESMFold, and SWISS-MODEL. Structural alignment of monomeric models from all five tools yielded an RMSD value of 3.85 Å (Figure S3), confirming strong agreement with the AlphaFold 3 prediction. Consistent with canonical TolC proteins, the barrel extends into the periplasm as an α-helical tunnel (Figure 2G).[90] Structural alignment identified *E. coli* K-12 TolC (PDB ID: 6WXI) as the closest match (Tables 2 and S3). As described in Section 3.2.6, *E. coli* TolC mediates hemolysin secretion,[92,93] colicin import,[94,95] antibiotic efflux,[96] and bacteriophage recognition.[97] The Foldseek tool identified its closest structural match with an uncharacterized protein from a *Spirochaete* bacterium. PANNZER and eggNOG-mapper similarly annotated BP951000_RS09000 as an integral OMP and efflux transporter (Tables 2 and S3). Amino acid sequence comparison across nine *B. pilosicoli* strains revealed two variations: S90 (in the ECL region) (Figure 2G) and S131 (within the β-barrel domain) (Tables 3 and S4).

3.2.1.8. BP951000_RS06235

BP951000_RS06235 is annotated as a TolC family protein in UniProt. The AlphaFold 3 model predicted a trimeric 12-stranded β-barrel, with each subunit contributing four β-strands (Figure 2H). As in Section 3.2.6, we validated this structure using ESMFold, SWISS-MODEL, RoseTTAFold, and TrRosetta. Structural alignment of monomeric models from all tools yielded an RMSD of 3.26 Å (Figure S4), confirming model consistency. The barrel extends into the periplasm as an α-helical tunnel connecting to the OM (Figure 2H).[90] DALI analysis showed the highest similarity to *E. coli* K-12 TolC (PDB ID: 6WXI) (Tables 2 and S3). Foldseek identified the closest homolog as TolC from a *Spirochaete* bacterium. PANNZER and eggNOG-mapper consistently annotated the protein as an OM efflux protein (Tables 2 and S3), suggesting a TolC-like function. Sequence alignment across nine *B. pilosicoli* strains revealed variations at 18 positions: K2, N3, F5, V6, F7, I8, I10, L12, S16, S25, N33, I42, E43, L93, S105, E136, I137, and T210 (Tables 3 and S4). Mapping showed L93

in the ECL region (Figure 2H), S105 in the β-barrel, and the remaining residues within the periplasmic domain (Table S4).

3.2.1.9. Eight-stranded β-barrel proteins annotated as SR domain-containing proteins

Thirteen *B. pilosicoli* proteins annotated in UniProt as "SR domain-containing proteins" were predicted in this study as OMBB proteins. Five proteins—BP951000_RS04880, BP951000_RS02055, BP951000_RS02050, BP951000_RS07540, and BP951000_RS00180—were classified as group A (predicted as OMBB proteins by all tools). Except for BP951000_RS02055 (annotated as an OMBB in NCBI), the rest are listed as hypothetical in NCBI. SRs, or G-protein-coupled receptors (GPCRs), are eukaryote-specific heptahelical membrane proteins.[98,99] However, AlphaFold 3 models of these *B. pilosicoli* proteins displayed eight-stranded β-barrel architectures with four extracellular and three periplasmic loops (Figure 2I–M). These conformations were consistently supported by ESMFold, SWISS-MODEL, RoseTTAFold, and TrRosetta tools. As GPCRs are absent in prokaryotes, and given that all tools predicted TM β-barrel structures rather than α-helices, we conclude that these proteins are misannotated as SR proteins in UniProt. The reason for this misannotation is that UniProtKB/TrEMBL annotates proteins automatically using computational pipelines. These annotations are not manually curated and rely on sequence similarity and automated rule-based systems.[100]

DALI analysis showed that BP951000_RS02055, BP951000_RS07540, and BP951000_RS00180 best matched the N-terminal β-barrel domain of *E. coli* K-12 OM protein A (OmpA) (PDB ID: 9FZC), whereas BP951000_RS02050 exhibited the best structural match with another *E. coli* K-12 OmpA structure (PDB ID: 9FZD) (Tables 2 and S3). Unlike *E. coli* OmpA, which includes both β-barrel and periplasmic domains, these *B. pilosicoli* proteins lack the periplasmic domain. *E. coli* OmpA functions in phage recognition, colicin transport, conjugation, membrane integrity maintenance, solute diffusion, and virulence. It also contributes to the virulence and pathogenicity of *E. coli*, making it a key target in the immune response.[101-109] BP951000_RS04880 aligned best with NspA of *N. meningitidis* (PDB ID: 1P4T) (Tables 2 and S3), a potential vaccine candidate involved in host adhesion and immune interaction.[86-89] This suggests that BP951000_RS04880 might have a role in adhesion and may be explored experimentally as a potential vaccine candidate.

PANNZER annotated all five proteins as OMBB domain-containing proteins. Foldseek identified BP951000_RS02055 and BP951000_RS02050 to be structurally closest to an OMBB protein of *B. hyodysenteriae*, whereas the remaining three proteins matched to an uncharacterized protein of *Brachyspira* spp. (Tables 2 and S3). Functional roles of these five proteins remain uncertain, highlighting the need for experimental validation. Amino acid sequence comparison across nine *B. pilosicoli* strains revealed several variations (Tables 3 and S4). BP951000_RS02055 and BP951000_RS02050 had variations in ECL regions (Figure 2J and 2K); variations in the remaining proteins were located in TM or periplasmic regions (Figure 2I, L, and M, Tables 3 and S4).

3.2.2. Group B

Group B included 29 proteins with diverse β-barrel architectures: a 26-stranded barrel (BP951000_RS09575); a 22-stranded barrel (BP951000_RS03215); an 18-stranded barrel (BP951000_RS04405); four 16-stranded barrels (BP951000_RS09655, BP951000_RS04440, BP951000_RS08285, and BP951000_RS04505); two 14-stranded barrels (BP951000_RS08455 and BP951000_RS01090); one 13-stranded barrel (BP951000_RS08975); four 12-stranded barrels (BP951000_RS06935, BP951000_RS11380, BP951000_RS03405, and BP951000_RS00185); one 10-stranded barrel (BP951000_RS10320); and six 8-stranded barrels (BP951000_RS05445, BP951000_RS08300, BP951000_RS05490, BP951000_RS07500, BP951000_RS01590, and BP951000_RS08295). In addition, eight 8-stranded β-barrel SR domain-containing proteins (BP951000_RS06930, BP951000_RS03290, BP951000_RS00765, BP951000_RS01280, BP951000_RS10445, BP951000_RS00365, BP951000_RS04620, and BP951000_RS04220) were identified (Table 1).

3.2.2.1. BP951000_RS09575

BP951000_RS09575 is annotated as LptD in *B.pilosicoli,*[20] a component of the LPS transport (LPT) system responsible for transporting LPS from the inner OM leaflet to the Gram-negative bacterial surface.[110] It carries a predicted secretory signal peptide. The AlphaFold 3 structural model revealed a 26-stranded β-barrel spanning the OM, with a lateral opening between strands 1 and 26, and a distinctive periplasmic β-jelly roll domain (Figure 3A). DALI analysis showed the best structural alignment with *Yersinia pestis* LptD (PDB ID: 5IXM) (Tables 2 and S3). Functional annotation by PANNZER and eggNOG-mapper confirmed roles in cell envelope biogenesis and LPT function, respectively, further validating its identity as LptD in *B. pilosicoli* (Tables 2 and S3). Sequence variation analysis across nine *B. pilosicoli* strains revealed seven variations: N14, G137, I257, I382, E454, D600, and G944 (Figure 3A, Tables 3 and S1). Structural mapping showed G944 in the ICL region, D600 in the ECL, I382 and E454 in the TM region, and the remaining variations in the β-jelly roll domain (Table S4).

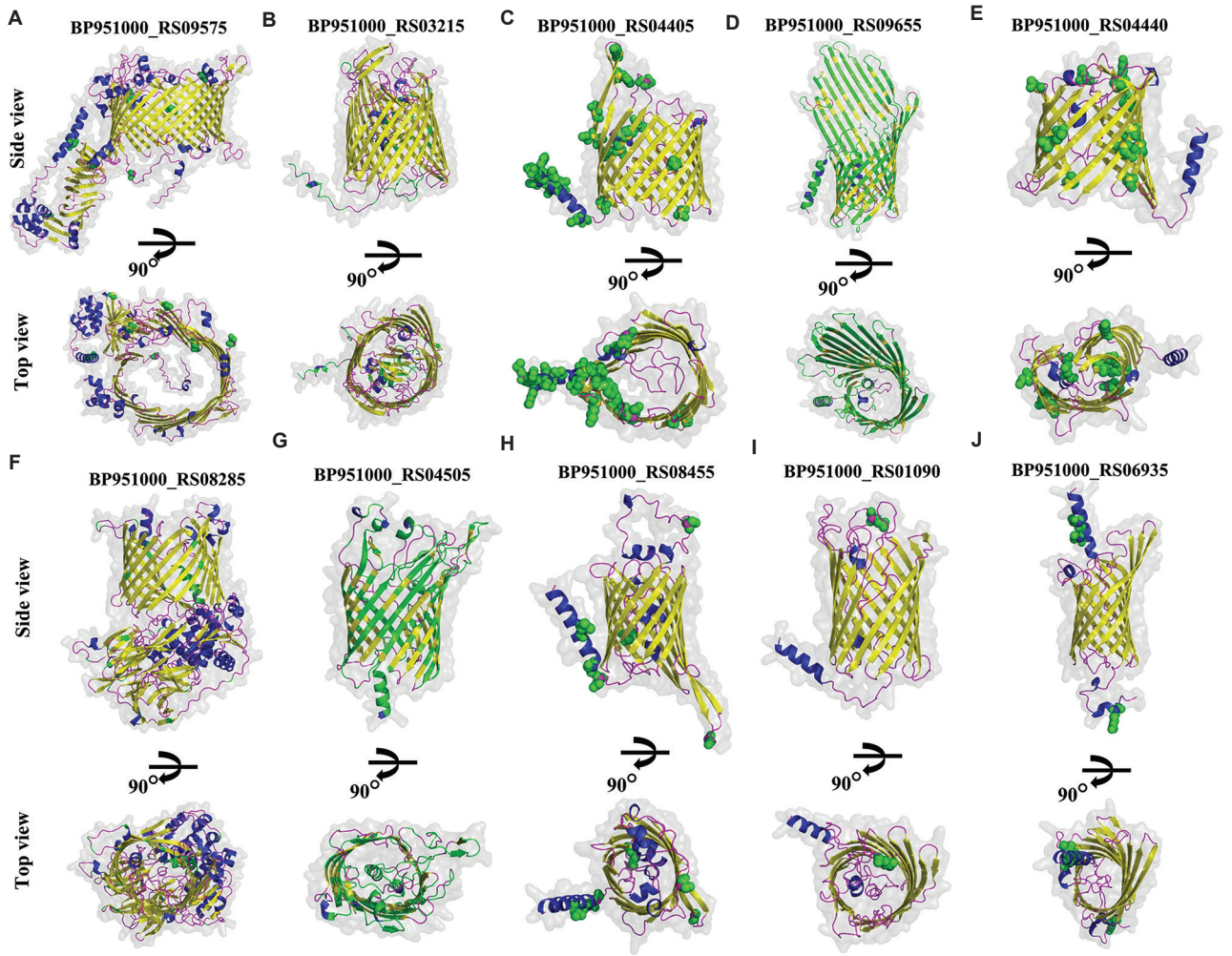

Figure 3. Structural models of β-barrel outer membrane proteins in group B. Group B includes 29 proteins, of which 10 are illustrated here (A-J). β-strands, α-helices, and loops are colored yellow, blue, and magenta, respectively. Green spheres indicate amino acid variations identified across nine strains of *Brachyspira pilosicoli*. Proteins with more than 40 variations are shown in green ribbon representation.

3.2.2.2. BP951000_RS03215

BP951000_RS03215 is annotated as a TonB-dependent siderophore receptor in *B. pilosicoli*. SignalP predicted a secretory signal peptide, while LipoP predicted a cytoplasmic localization. The AlphaFold 3 model revealed a 22-stranded C-terminal β-barrel and an N-terminal plug domain (Figure 3B). DALI analysis showed the best structural alignment with *E. coli* K-12 BtuB (PDB ID: 3RGM) (Tables 2 and S3), a TonB-dependent transporter (TBDT) with a 22-stranded β-barrel and an N-terminal plug domain that occludes the lumen.[111] *E. coli* BtuB mediates vitamin B_{12} (cobalamin) uptake via ECLs.[111-113] TBDTs also transport heme, ferric-siderophores, sucrose, and maltodextrin.[114,115] BP951000_RS03215 shares the same overall architecture as *E. coli* BtuB, and its high Z-score (32.1) confirms strong homology. Foldseek further identified a TonB-dependent receptor (TBDR) from a member of the phylum *Bacteroidetes* as the closest match, supporting a role in nutrient uptake. Sequence-based annotation using PANNZER and eggNOG-mapper, respectively, classified the protein as an OM receptor and linked it to cobalamin transport activity (Tables 2 and S3). Together, these results strongly suggest that BP951000_RS03215 functions as a TonB-dependent OM receptor involved in vitamin B_{12} transport. Sequence variation analysis across nine *B. pilosicoli* strains identified 59 variations (Table S5), which were mapped onto the structural model (Figure 3B).

3.2.2.3. BP951000_RS04405

BP951000_RS04405 is annotated as Toxin A in UniProt[20] but as a hypothetical protein in NCBI. It contains a predicted secretory signal peptide. The AlphaFold 3

structural model revealed a β-barrel comprising 18 β-strands, nine ECLs, and eight periplasmic loops (Figure 3C). DALI analysis showed the highest structural similarity to *P. aeruginosa* PAO1 porin OM channel protein K (OccK)-7 (PDB ID: 4FRT) (Tables 2 and S3). The OccK protein family contains a ladder of basic residues (arginine + lysine = 11%) that form a positively charged channel for the uptake of small, carboxyl-containing substrates.[116-119] BP951000_RS04405 exhibited a lower arginine + lysine content (7.6%), suggesting that while structurally similar to OccK, it may facilitate the transport of alternative substrates. Foldseek identified its closest structural match with an uncharacterized protein from *Treponema vincentii*. PANNZER annotated it as Toxin A, a known virulence factor (Tables 2 and S3).[120]

These findings suggest that BP951000_RS04405 is a porin-like protein with potential functional divergence, warranting further experimental annotation. A total of 23 amino acid sequence variations were identified across nine *B. pilosicoli* strains (Tables 3 and S4). Structural mapping showed six variations (N41, F84, K90, N98, S175, and Q183) in the ECL region, eight (T24, N32, S34, I101, S102, N104, I264, and T303) in the TM β-barrel region, and the remainder in the periplasmic region (Figure 3C and Table S4).

3.2.2.4. BP951000_RS09655

BP951000_RS09655 is annotated as a hypothetical protein in NCBI and as a domain of unknown function (DUF) 5723 domain-containing protein in UniProt. It carries a secretory signal peptide. The AlphaFold model predicted a 16-stranded β-barrel, with strands 3–12 longer than the rest, giving the barrel an asymmetric extracellular profile (Figure 3D).

DALI analysis identified the closest structural match with the tetraheme c-type cytochrome CymA from *Klebsiella oxytoca* (PDB ID: 4V3G) (Tables 2 and S3). CymA, a 14-stranded OMP, facilitates passive diffusion of large molecules such as cyclodextrins and linear maltooligosaccharides.[121,122] It aligned with 14 of the 16 β-strands in BP951000_RS09655. The comparable pore diameter suggests a similar function in passive diffusion channels in *B. pilosicoli*. Structural alignment using the US-Align server yielded an RMSD of 4.28 Å, supporting high structural homology. PANNZER annotated the protein as a cell surface protein (Tables 2 and S3). Sequence comparison across nine strains of *B. pilosicoli* revealed 315 variations and multiple deletions, indicating high variability across the full length of the protein (Figure 3D and Table S5).

3.2.2.5. BP951000_RS04440

BP951000_RS04440, annotated as a hypothetical protein, contains a secretory signal peptide. The AlphaFold 3 model predicted a 16-stranded β-barrel structure with a lateral gate between strands 1 and 16 (Figure 3E). DALI analysis revealed its closest structural match with the translocation and assembly module protein A (TamA) from *E. coli* (PDB ID: 4C00) (Tables 2 and S3), an Omp85 superfamily protein featuring three N-terminal POTRA domains and a C-terminal 16-stranded β-barrel.[123] TamA facilitates autotransporter β-barrel membrane insertion and passenger domain translocation into the extracellular space.[123-125] Foldseek also identified a structural match with an Omp85 domain-containing protein from *Dracunculus medinensis*. While BP951000_RS04440 closely resembles the TamA β-barrel domain, it lacks the POTRA domains, suggesting functional divergence. BP951000_RS04440 is thus predicted to be a structural homolog of TamA, but its specific role in *B. pilosicoli* remains to be clarified. Notably, PANNZER annotated the protein as Toxin A (Tables 2 and S3). Sequence comparison across nine *B. pilosicoli* strains revealed 10 variations (K104, K113, S117, Y124, I132, T134, N151, G153, L243, and V252) (Tables 3 and S4). Structural mapping showed K104 and L243 in the ECL region, with the remaining variations located within the TM β-barrel domain (Figure 3E and Table S4).

3.2.2.6. BP951000_RS08285

BP951000_RS08285 is annotated as a transcriptional regulating protein (Trep) in *B. pilosicoli*. In *Pseudomonas fluorescens*, Trep catalyzes trehalose phosphorylation and its translocation across the OM.[126] LipoP predicted BP951000_RS08285 as cytoplasmic, whereas SignalP predicted no signal peptide. The AlphaFold 3 predicted a 16-stranded C-terminal β-barrel with a large N-terminal periplasmic domain and a lateral opening between strands 1 and 16 (Figure 3F). Structure-based annotation showed the closest match with TolB proteins from *E. coli* K-12 (PDB ID: 3IAX) and *Citrobacter freundii* (PDB ID: 2IVZ) (Tables 2 and S3). *E. coli* TolB is a periplasmic protein with a two-domain structure: an α/β N-terminal domain and a six-bladed β-propeller C-terminal domain (Figure S5).[127] The periplasmic domain of BP951000_RS08285 closely resembles the TolB C-terminal domain, suggesting a similar role in porin assembly and cell envelope integrity.[128-130] However, unlike TolB, BP951000_RS08285 includes an additional TM β-barrel domain, indicating potential functions unique to *B. pilosicoli*.

When the β-barrel domain alone was queried, the top DALI hit was filamentous hemagglutinin (FHA) transporter FhaC (PDB ID: 4QL0) from *Bordetella pertussis*, a 16-stranded β-barrel protein that transports FHA (Figure S5).[131] While FhaC includes POTRA domains for substrate recognition,[132] BP951000_RS08285 lacks them, suggesting it may function as a translocation pore

with distinct specificity. Foldseek identified the closest structural match with an uncharacterized protein from a *Spirochaete* bacterium (Tables 2 and S3). PANNZER annotated BP951000_RS08285 as Trep, while eggNOG-mapper linked it to TonB-independent uptake pathways. BLASTp analysis showed no sequence homologs in other spirochetes (e.g., *Treponema*, *Borrelia*, and *Leptospira*), indicating species-specific uniqueness. MSA across nine *B. pilosicoli* strains revealed 43 variations (Table S5).

3.2.2.7. BP951000_RS04505

BP951000_RS04505 is annotated as VspH in *B. pilosicoli*.[20] As described in Section 3.2.3, Vsps mediate bacterial adherence to host cells.[64] SignalP predicted no signal peptide, whereas LipoP predicted the presence of N-terminal TMHs. The AlphaFold 3 model revealed a 16-stranded β-barrel (Figure 3G).

Structural alignment identified the closest match with *E. coli* K-12 PgaA (PDB ID: 4Y25) (Tables 2 and S3). As discussed in Section 3.2.2, PgaA facilitates polysaccharide translocation across the OM. Additionally, Foldseek identified its best alignment with a bacterial polysaccharide OM secretin from *E. coli* K-12, supporting this hypothesis. PANNZER annotated the protein as a cell surface protein. Based on both structural and sequence analyses, BP951000_RS04505 is likely involved in polysaccharide secretion across the OM. Notably, its homolog in *B. hyodysenteriae* (VspH) has been evaluated as a potential vaccine candidate.[69] Sequence comparison across nine *B. pilosicoli* strains revealed 247 amino acid variations and multiple deletions, indicating high variability throughout the protein (Figure 3G and Table S5).

3.2.2.8. BP951000_RS08455

BP951000_RS08455 is annotated as a PorV/PorQ family protein in UniProt and as a hypothetical protein in NCBI. PorV and PorQ are integral components of the type IX secretion system, involved in protein export in Gram-negative members of the Fibrobacteres–Chlorobi–Bacteroidetes superphylum.[133] The AlphaFold 3 model revealed a 14-stranded β-barrel structure with seven ECLs and an interior blocked by an N-terminal hatch domain (Figure 3H). Foldseek identified structural similarity to a PorV/PorQ family protein, and the best DALI match was *E. coli* long-chain fatty acid transporter FadL (PDB ID: 2R88) (Tables 2 and S3). Like FadL, BP951000_RS08455 adopts a monomeric 14-stranded β-barrel,[134] though it lacks the lateral opening formed by the inward bend in a β-strand that is essential for fatty acid transport in FadL.[135] Despite this structural difference, the remaining structural resemblance suggests a potential role in transporting hydrophobic molecules.[136] Although PANZZER annotated the protein as uncharacterized, structure-based evidence strongly supports its function as a transporter of hydrophobic substrates. Sequence comparison across nine *B. pilosicoli* strains revealed an insertion at position 255 and six variations (L12, S20, N22, A117, R187, and S253) (Figure 3H, Tables 3 and S4). Mapping these onto the model placed N22 and A117 in the ECL region, S253 in the ICL region, R187 in the β-barrel TM region, and the remaining variations in the N-terminal domain (Table S4).

3.2.2.9. BP951000_RS01090

BP951000_RS01090 is annotated as a hypothetical protein in NCBI but as a VspH in UniProt. SignalP predicted a secretory signal peptide. The AlphaFold 3 model revealed a 14-stranded β-barrel architecture (Figure 3I). Sequence-based annotation using PANNZER supported the VspH designation. DALI analysis identified *K. oxytoca* CymA (PDB ID: 4V3G) as the closest structural homolog (Tables 2 and S3). As discussed in Section 3.2.13, CymA functions as a diffusion channel for bulky substrates. Foldseek further indicated similarity to an uncharacterized protein from a *Chitinophagaceae* bacterium. Sequence comparison across nine *B. pilosicoli* strains revealed a single variation, E258, located in the ECL region (Figure 3I; Tables 3 and S4). Structural alignment of BP951000_RS04760, BP951000_RS02055, BP951000_RS04505, and BP951000_RS01090 using the US-align server resulted in an RMSD of 4.00 Å (Figure S6).

3.2.2.10. BP951000_RS06935

BP951000_RS06935 is annotated as a hypothetical protein in UniProt and NCBI. It is predicted to contain a secretory signal peptide. The AlphaFold 3 model revealed a 12-stranded β-barrel structure with six ECLs (Figure 3J). Structural homology search identified the closest match as the β-barrel domain of the *E. coli* hemoglobin-binding protease (Hbp) autotransporter (PDB ID: 3AEH) (Tables 2 and S3). Hbp, a member of the serine protease autotransporters of Enterobacteriaceae family, consists of an N-terminal serine protease passenger domain and a C-terminal β-barrel that facilitates its extracellular secretion.[137-139] While BP951000_RS06935 lacks a periplasmic or passenger domain, its β-barrel homology suggests a role in substrate translocation across the OM.

Foldseek identified the best match with a 12-stranded β-barrel OMP from Cluster of Orthologous Groups (COG) 4313 of *Pseudomonas putida* F1, which has a dynamic lateral opening that permits the passage of hydrophobic molecules.[140] Sequence-based annotation by PANNZER linked BP951000_RS06935 to Toxin A. Collectively, structure- and sequence-based analyses suggest that BP951000_RS06935 may act as an autotransporter or facilitate hydrophobic molecule transport. Sequence

variation analysis across nine *B. pilosicoli* strains revealed four variations (S9, I10, V13, and R298) (Figure 3J, Tables 3 and S4). When mapped onto the structural model, these variations were present in the N-terminal and C-terminal regions of the β-barrel structure (Table S4).

3.2.2.11. BP951000_RS11380

BP951000_RS11380, annotated as a hypothetical protein in NCBI, was predicted to contain a secretory signal peptide. UniProt classified the protein as Toxin A. It has an N-terminal TMH, predicted by LipoP. The AlphaFold 3 model revealed a 12-stranded β-barrel (Figure 4A). Structural alignment using DALI showed the closest match to *E. coli* O157:H7 serine protease EspP (PDB ID: 2QOM) (Tables 2 and S3), a member of the serine protease autotransporters of *Enterobacteriaceae* family of autotransporters.[141] These specialized porins contain a C-terminal β-barrel that facilitates the secretion of an N-terminal virulence-associated passenger domain. Some autotransporters consist solely of the β-barrel autotransporter domain, with the passenger or toxin-encoding gene located upstream in the genome.[142] The EspP passenger domain exhibits serine protease activity, cleaving host proteins such as pepsin A and human coagulation factor V.[143-145] Given the structural similarity, BP951000_RS11380 might function as an autotransporter for a serine protease or virulence factor encoded elsewhere in the genome. Foldseek identified the closest structural homolog with an uncharacterized protein from *Candidatus margulis*. Sequence-based analysis by PANNZER also aligned it with Toxin A. Sequence comparison across nine *B. pilosicoli* strains revealed three variations (M126, M154, and I278), with M126 located in the ECL region and the other two (M154 and I278) in the TM β-barrel region (Figure 4A, Tables 3 and S4).

3.2.2.12. BP951000_RS03405

BP951000_RS03405, annotated as a hypothetical protein, contains a secretory signal peptide. The AlphaFold 3 model revealed a 12-stranded β-barrel (Figure 4B). DALI analysis showed *E. coli* OMPLA as the closest structural homolog (PDB ID: 1ILD) (Tables 2 and S3). OMPLA is an acyl hydrolase that cleaves ester bonds in phospholipids and lysophospholipids, and contributes to colicin secretion.[146-148] BP951000_RS03405, as a structural homolog of *E. coli* OMPLA, may serve similar functions in *B. pilosicoli*. Although OMPLA contains a catalytic histidine (His)–serine (Ser)–asparagine (Asn) triad, BP951000_RS03405 lacks this exact motif; however, Ser and Asn residues are present at adjacent positions 224 and 225. Foldseek indicated homology to a DUF1207 domain-containing protein from *Ignavibacteria*, suggesting possible functional diversity, while PANNZER annotated BP951000_RS03405 simply as an OMP. Sequence alignment across nine strains of *B. pilosicoli* revealed 24 variations, all located within the periplasmic region of the β-barrel structure (Figure 4, Tables 3 and S4).

3.2.2.13. BP951000_RS00185

BP951000_RS00185, annotated as a hypothetical protein, contains a secretory signal peptide with a cleavage site between residues 20 and 21. The AlphaFold 3 model predicted a 12-stranded β-barrel (Figure 4C). DALI analysis identified *E. coli* K-12 OMPLA as the closest structural match (PDB ID: 1QD6; Z-score = 18.3, RMSD = 3.1 Å) (Tables 2 and S3). As discussed in Section 3.2.21, OMPLA hydrolyzes acyl ester bonds in phospholipids and lysophospholipids and is involved in colicin secretion.[146-148] Although OMPLA contains a catalytic His–Ser–Asn triad, BP951000_RS00185 lacks this motif. However, adjacent Ser–Asn residues were identified at positions 25–26 and 73–74. Foldseek identified structural homology to a DUF1207 domain-containing protein from *Ignavibacteria*. Sequence-based annotation indicated similarity to Toxin A, a known virulence factor.[120] Sequence variation across nine strains of *B. pilosicoli* revealed 58 variations predominantly present in the signal peptide region (Figure 4C and Table S5).

3.2.2.14. BP951000_RS10320

BP951000_RS10320, annotated as a hypothetical protein, is predicted to contain a secretory signal peptide. The AlphaFold 3 model revealed a 10-stranded β-barrel with several TM β-strands extending toward the periplasmic side, forming an elliptical pore (Figure 4D). Periplasmic loops from the TM barrel are also folded into α-helices and β-strands. Structural homology using Foldseek identified the best match with an uncharacterized protein from *B. murdochii*. DALI analysis returned synthetic TM β-barrels as the top four hits; the fifth hit, *N. meningitidis* opacity-associated protein A (OpcA) (PDB ID: 2VDF), was considered the best structural homolog (Tables 2 and S3). OpcA is a 10-stranded β-barrel OMP that mediates adhesion by binding host cell proteoglycans.[149,150] Sequence analysis using PANNZER also indicated homology with a cell surface protein, supporting a potential role in host cell interaction and adhesion. Sequence analysis across nine *B. pilosicoli* strains revealed 25 variations, none of which were located in the ECL region (Figure 4D, Tables 3 and S4).

3.2.2.15. BP951000_RS05445

BP951000_RS05445, annotated as a DUF3575 domain-containing protein in UniProt and as a hypothetical protein in NCBI, carries a predicted secretory signal peptide (SignalP), although LipoP classified it as a cytoplasmic

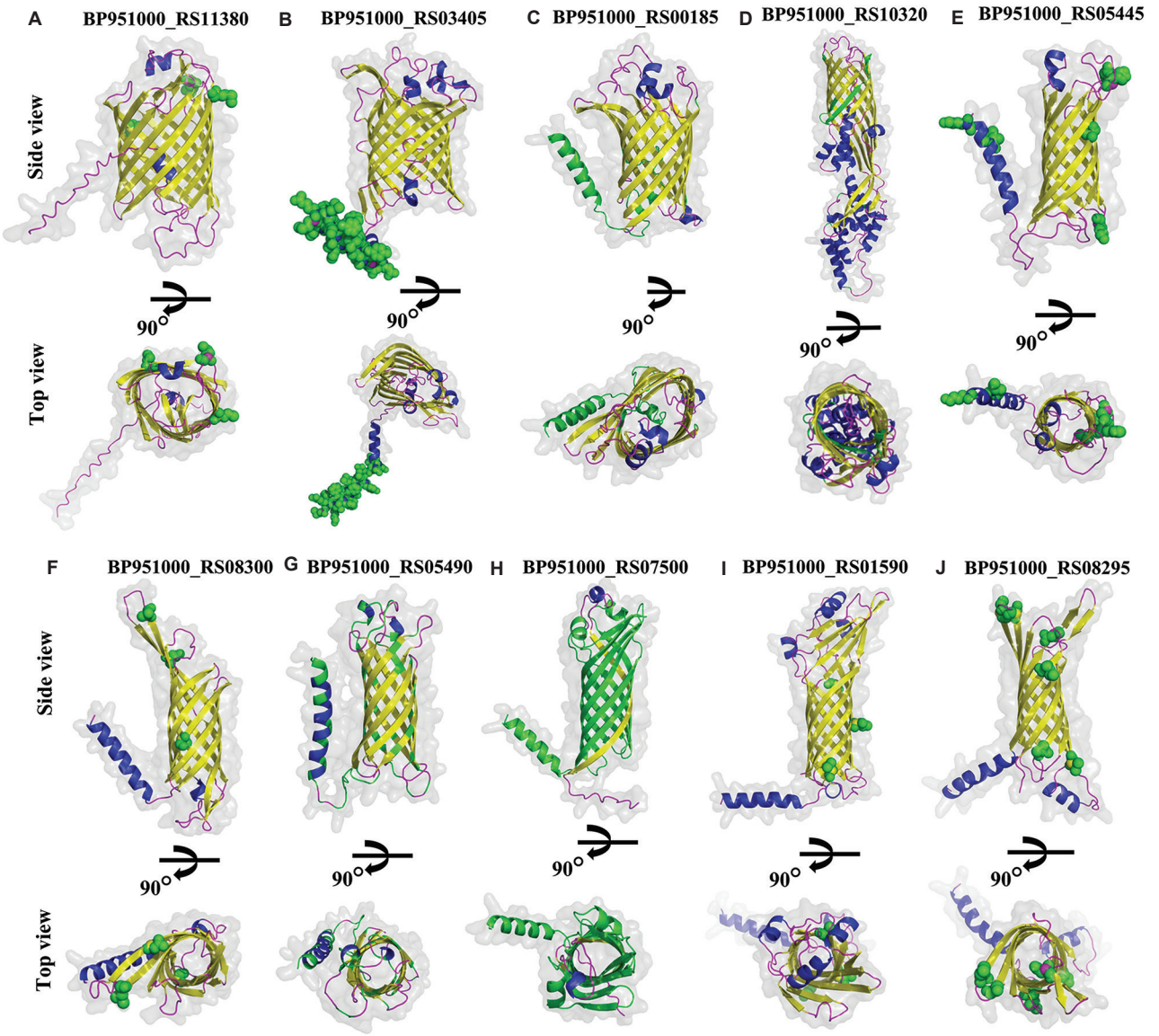

Figure 4. Structural models of β-barrel outer membrane proteins in Group B (continued). Structural models of 10 Group B proteins are shown (A-J), arranged in descending order based on the number of β-strands in their β-barrel architectures. β-strands, α-helices, and loops are colored yellow, blue, and magenta, respectively. Green spheres indicate amino acid variations identified across nine strains of *Brachyspira pilosicoli*. Proteins with more than 40 variations are shown in green ribbon representation.

protein. The AlphaFold 3 model revealed an eight-stranded β-barrel architecture (Figure 4E). DALI analysis for structural homology showed the best matches with two proteins: OmpF (PDB ID: 4RLC) of *P. aeruginosa* PAO1 and NspA (PDB ID: 1P4T) of *N. meningitidis* (Tables 2 and S3). Given the established roles of OmpF and NspA in adhesion, immune modulation, and biofilm formation (as discussed in Sections 3.2.4 and 3.2.5), BP951000_RS05445 may serve similar functions in *B. pilosicoli*.

Sequence-based annotation identified it as Tia invasion protein, supporting its potential role as an adhesin involved in host cell interaction. Comparative sequence analysis across nine *B. pilosicoli* strains revealed six variations (K2, I7, A79, N87, H89, and K158) (Figure 4E, Tables 3 and S4). Mapping onto the structural model showed N87 and H89 in the ECL region, K158 in the ICL region, A79 in the β-barrel domain, and the remaining two in the N-terminal region of the protein (Table S4).

3.2.2.16. BP951000_RS08300

BP951000_RS08300 is annotated as a Tia invasion determinant. Tia proteins, known in enterotoxigenic *E.*

coli, function as both adhesins and invasins.[151] Tia proteins consist of eight TM β-sheets with four surface-exposed loops and bind specific receptors on HCT8 human ileocecal epithelial cells.[152]

BP951000_RS08300 is predicted to contain a secretory signal peptide, and its AlphaFold 3 model revealed an elliptical eight-stranded β-barrel, with two extended β-strands projecting extracellularly (Figure 4F). Structural alignment showed the highest similarity to *N. meningitidis* NspA (PDB ID: 1P4T) (Tables 2 and S3). Foldseek analysis identified structural homology with an OmpA family protein of *Brachyspira hampsonii* 30446, suggesting potential involvement in adhesion, invasion, intracellular survival, or immune modulation. EggNOG mapper further predicted lipid A 3-O-deacylase activity, which modifies lipid A structure to facilitate immune evasion. Collectively, structure- and sequence-based evidence suggests that BP951000_RS08300 is a multifunctional OMP, potentially involved in host interaction and immune modulation. Sequence comparison across nine *B. pilosicoli* strains revealed three variations (L143, N156, and S200) (Figure 4F; Tables 3 and S4), all located in the β-barrel domain (Table S4).

3.2.2.17. BP951000_RS05490

BP951000_RS05490 is annotated as a hypothetical protein in NCBI and as a Tia invasion determinant in UniProt. SignalP predicted a secretory signal peptide, whereas LipoP classified it as cytoplasmic. The AlphaFold 3 model revealed an eight-stranded β-barrel architecture (Figure 4G). DALI analysis identified *N. meningitidis* NspA (PDB ID: 1P4T) as the closest structural match (Tables 2 and S3). As a structural homolog of NspA, BP951000_RS05490 may play similar roles in host interaction and immune modulation, as discussed in Section 3.2.5. PANNZER annotation also identified the protein as a Tia invasion determinant, aligning with its UniProt classification. Comparative sequence analysis across nine *B. pilosicoli* strains revealed 61 variations distributed throughout the sequence and two deletions (Figure 4G and Table S5).

3.2.2.18. BP951000_RS07500

BP951000_RS07500, annotated as a hypothetical protein, is predicted to contain a secretory signal peptide. The AlphaFold 3 model predicted an eight-stranded β-barrel structure (Figure 4H). Structural alignment identified the closest match with the β-barrel domain of *E. coli* K-12 OmpA (PDB ID: 9FZC) (Tables 2 and S3). Unlike *E. coli* OmpA, BP951000_RS07500 lacks the periplasmic domain. Given that OmpA is involved in phage binding, vesicle transport, conjugation, and membrane integrity, BP951000_RS07500 may perform similar functions.[101-109] Sequence-based annotation using PANNZER identified it as an OMBB protein. Sequence comparison across nine *B. pilosicoli* strains revealed 183 variations within the first seven strands of the β-barrel (Figure 4H and Table S5), indicating high variability across its 232-residue sequence.

3.2.2.19. BP951000_RS01590

BP951000_RS01590, annotated as a hypothetical protein, is predicted to contain a secretory signal peptide. LipoP predicted a TMH at the N-terminal. The structural model revealed an eight-stranded β-barrel architecture (Figure 4I) and additional short β-strands and α-helices extending extracellularly. As identified by DALI, structural homology was highest with *P. aeruginosa* OprG (PDB ID: 2X27) (Tables 2 and S3), an OmpW family protein involved in catabolism and uptake of hydrophobic molecules, including hydrocarbons.[153] Foldseek identified the closest structural match with an uncharacterized protein of *B. pilosicoli* P43/6/78. PANNZER analysis predicted the protein to be a cell surface protein, suggesting a possible role in host–pathogen interaction. Sequence variation analysis across nine *B. pilosicoli* strains revealed three variations (V205, I215, and V221), all located in the TM β-barrel domain (Figure 4I, Tables 3 and S4).

3.2.2.20. BP951000_RS08295

BP951000_RS08295, annotated as a Tia invasion determinant, is predicted to carry a secretory signal peptide. The AlphaFold 3 model revealed an eight-stranded β-barrel with two elongated strands, forming an elliptical pore on the extracellular side (Figure 4J). DALI analysis identified the best structural alignment with *N. meningitidis* NspA (PDB ID: 1P4T) (Tables 2 and S3), suggesting potential roles in adhesion and immune evasion, similar to NspA (Section 3.2.5). Sequence-based annotation indicated adhesive function and potential lipid A 3-O-deacylase activity, potentially contributing to immune evasion. Among nine *B. pilosicoli* strains, six variations were detected (N34, I49, V123, S141, I144, and V167), with N34 located in the ECL region and the others in the β-barrel domain (Figure 4J, Tables 3 and S4).

3.2.2.21. BP951000_RS08975

BP951000_RS08975 is annotated as a TonB-dependent receptor (TBDR) domain-containing protein in NCBI,[20] and as a Ser/Threonine protein kinase in UniProt. TBDR domain-containing proteins transport substrates across the OM using the proton motive force, transmitted via the TonB–ExbB–ExbD complex.[114,154] SignalP predicted a secretory signal peptide. AlphaFold 3 revealed a 13-stranded incomplete β-barrel, likely forming a dimer or a higher-order oligomer (Figure 5A). Attempts to model a

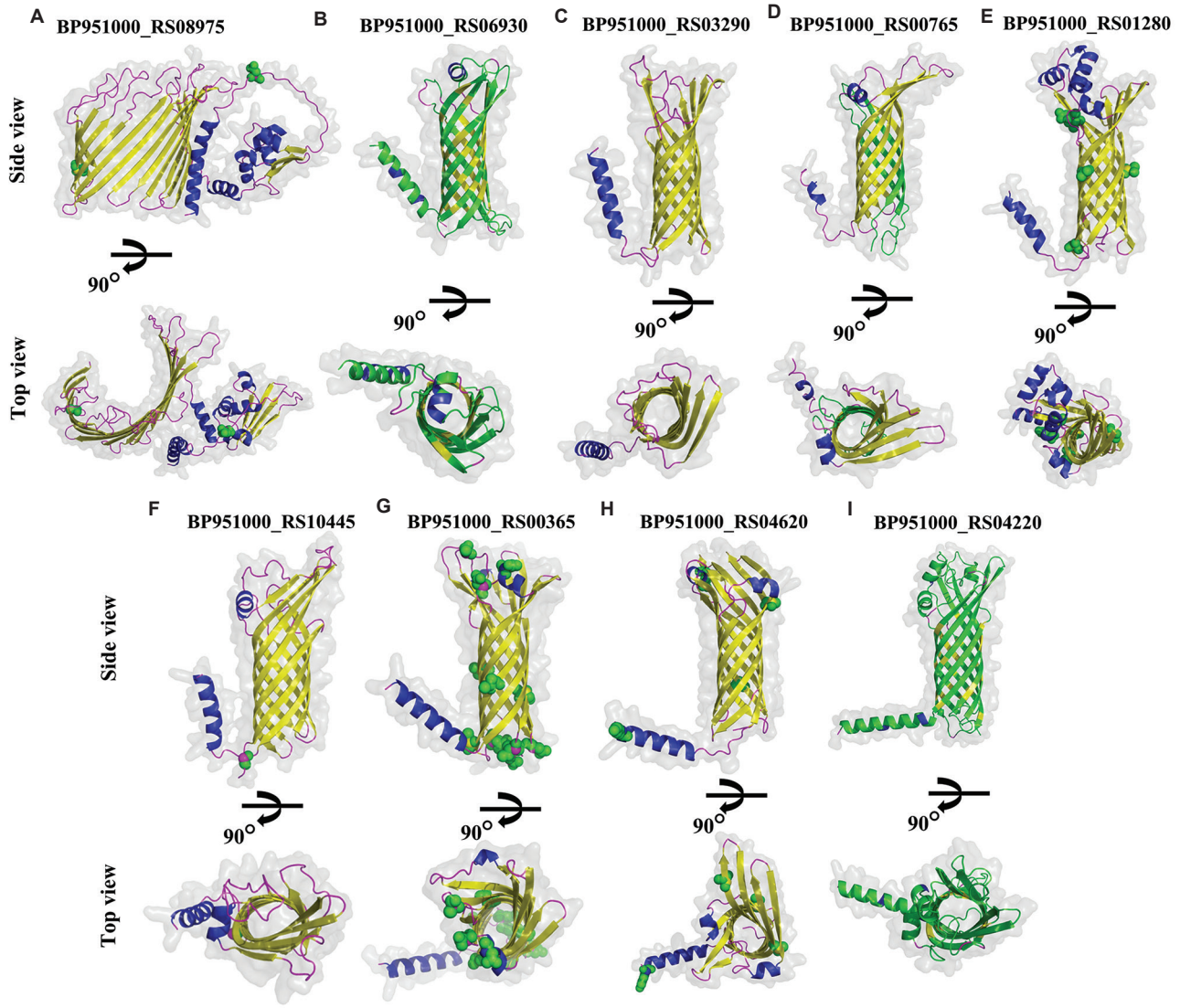

Figure 5. Structural models of β-barrel outer membrane proteins in Group B (continued). Structural models of the remaining nine proteins of Group B are shown (A-I). Proteins are arranged based on the decreasing number of β-strands. Structural model of BP951000_RS08975 revealed an incomplete β-barrel structure consisting of 13 β-strands. β-strands, α-helices, and loops are colored yellow, blue, and magenta, respectively. Green spheres indicate amino acid variations identified across nine strains of *Brachyspira pilosicoli*. Proteins with more than 40 variations are shown in green ribbon representation.

dimer using AlphaFold 3 were inconclusive (pTM = 0.35; ipTM = 0.19).

Structural models generated using ESMFold, SWISS-MODEL, RoseTTAFold, and TrRosetta aligned well with the AlphaFold 3 prediction (RMSD = 4.61 Å). Foldseek identified structural homology with an uncharacterized protein from *B. hampsonii*, whereas the DALI server matched it to *E. coli* BtuB (PDB ID: 2GSK), a vitamin B_{12} transporter (Tables 2 and S3). However, the high RMSD suggests significant structural differences, and functional BtuB homology is unlikely. Unlike canonical BtuB, BP951000_RS08975 does not possess a full β-barrel. Furthermore, BP951000_RS08975 showed no homologs in the spirochete genera—*Treponema*, *Borrelia*, or *Leptospira*—by BLASTp. PANNZER annotation also identified the protein as a Ser/Threonine protein kinase, aligning with its UniProt classification. Sequence comparison across nine *B. pilosicoli* strains revealed two variations: D32 in the ECL and T371 in the TM region (Figure 5A, Tables 3 and S4).

3.2.2.22. Eight-stranded β-barrel proteins annotated as Serpentine receptor domain-containing proteins

Eight *B. pilosicoli* proteins of Group B—BP951000_RS06930, BP951000_RS03290, BP951000_RS00765,

BP951000_RS01280, BP951000_RS10445, BP951000_RS00365, BP951000_RS04620, and BP951000_RS04220—are annotated in UniProt as SR domain-containing proteins, but listed as hypothetical in NCBI. All are predicted to contain a signal peptide. Five similarly annotated proteins (Group A) were previously discussed in Section 3.2.9.

The AlphaFold 3 models revealed a conserved eight-stranded β-barrel architecture (Figure 5B-I). Structural alignment analysis using DALI showed that BP951000_RS01280, BP951000_RS00365, and BP951000_RS06930 closely matched with *N. meningitidis* NspA (PDB ID: 1P4T), whereas BP951000_RS04620, BP951000_RS00765, BP951000_RS03290, BP951000_RS04220, and BP951000_RS10445 aligned best with the N-terminal β-barrel domain of *E. coli* K-12 OmpA (PDB ID: 9FZC) (Tables 2 and S3).

As described in Sections 3.2.5 and 3.2.9, NspA contributes to host colonization and immune evasion and is a potential vaccine candidate,[86-89] whereas OmpA plays roles in phage recognition, conjugation, membrane integrity, diffusion, and virulence.[101-109] Foldseek analysis further revealed that seven of the eight proteins showed the closest structural match with uncharacterized proteins from the *Brachyspira* genus. Notably, BP951000_RS01280 aligned with an OMBB protein from *B. hyodysenteriae*, supporting its role as a conserved OMP component.

Consistent with structural data, PANNZER annotated BP951000_RS06930, BP951000_RS00765, BP951000_RS01280, BP951000_RS10445, and BP951000_RS04620 as OMBB proteins, whereas BP951000_RS03290, BP951000_RS00365, and BP951000_RS04220 remained uncharacterized. The convergence of structure- and sequence-based annotations supports their classification as OMBB proteins, although experimental validation is needed to confirm their functional roles. Sequence comparison across nine *B. pilosicoli* strains revealed variations at several positions (Tables 3, S4, and S5).

This study did not include experimental validation of OMBBs, which is a limitation. Future research should focus on validating these predictions through experimental characterization and functional analysis.

4. Conclusion

Brachyspira pilosicoli is a globally prevalent enteric spirochete associated with IS in both animals and humans. Despite its clinical significance, the molecular mechanisms underlying its pathogenesis remain poorly understood, particularly the role of OMPs, which are critical for nutrient uptake, adhesion, immune evasion, and virulence. In this study, we addressed this knowledge gap by systematically identifying and characterizing the OMBB proteome of *B. pilosicoli* using a consensus-based computational approach.

We predicted 42 OMBB proteins and validated their β-barrel architectures using AlphaFold 3. These proteins exhibited diverse topologies, ranging from 8 to 26 strands. Structural homology-based functional annotation revealed putative homologs of BamA, LptD, TolC, TBDRs, NspA, OmpA, FadL, and others, suggesting roles in membrane biogenesis, LPT, efflux activity, substrate uptake, host colonization, and immune modulation.

Among the 42 OMBB proteins, seven lacked annotation in both UniProt and NCBI. Combined structure- and sequence-based analyses enabled putative functional assignment for these hypothetical proteins. Comparative sequence analysis across nine *B. pilosicoli* strains revealed extensive polymorphisms in 12 proteins, each containing 40 or more variations, suggesting potential roles in immune evasion and host adaptation.

This study expands current knowledge of the *B. pilosicoli* OMP repertoire and provides a framework for identifying potential targets for diagnostic, prophylactic, and therapeutic development. The functional predictions and structural insights reported here lay the foundation for future experimental work aimed at elucidating the precise roles of these OMPs in IS pathogenesis.

Acknowledgements

None.

Funding

Amisha Panda is a recipient of a junior research fellowship from the Department of Biotechnology, Government of India (grant number: DBT/2023-24/UOD/2326). Jahnvi Kapoor is a recipient of a junior research fellowship from the University Grants Commission, Government of India (grant number: 231620067077).

Conflict of interest

The authors declare that they have no competing interests.

Author contributions

Conceptualization: Sanjiv Kumar
Data curation: Amisha Panda
Formal analysis: Amisha Panda, Jahnvi Kapoor, Ravindresh Chhabra, Anannya Bandyopadhyay
Investigation: Amisha Panda
Methodology: Amisha Panda, Sanjiv Kumar, Anannya Bandyopadhyay
Software: Amisha Panda, Sanjiv Kumar
Supervision: Sanjiv Kumar

Visualization: Jahnvi Kapoor, Batchu Hareramadas, Ilmas Naqvi

Writing–original draft: Amisha Panda, Ravindresh Chhabra, Anannya Bandyopadhyay

Writing–review & editing: Jahnvi Kapoor, Batchu Hareramadas, Ilmas Naqvi, Ravindresh Chhabra, Sanjiv Kumar, Anannya Bandyopadhyay

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

References

1. Ludwig W, Euzéby J, Whitman WB. Taxonomic outlines of the phyla bacteroidetes, spirochaetes, tenericutes (mollicutes), acidobacteria, fibrobacteres, fusobacteria, dictyoglomi, gemmatimonadetes, lentisphaerae, verrucomicrobia, chlamydiae, and planctomycetes. In: Krieg NR, Staley JT, Brown DR, *et al*, editors. *Bergey's Manual® of Systematic Bacteriology: Volume Four The Bacteroidetes, Spirochaetes, Tenericutes (Mollicutes), Acidobacteria, Fibrobacteres, Fusobacteria, Dictyoglomi, Gemmatimonadetes, Lentisphaerae, Verrucomicrobia, Chlamydiae, and Planctomycetes*. New York: Springer; 2010. p. 21-24.
2. Hampson DJ, Oxberry SL, La T. Potential for zoonotic transmission of *Brachyspira pilosicoli*. *Emerg Infect Dis*. 2006;12(5):869-870.

   doi: 10.3201/eid1205.051180
3. Le Roy CI, Mappley LJ, La Ragione RM, Woodward MJ, Claus SP. *Brachyspira pilosicoli*-induced avian intestinal spirochaetosis. *Microb Ecol Health Dis*. 2015;26(1):28853.

   doi: 10.3402/mehd.v26.28853
4. Bano L, Merialdi G, Bonilauri P, *et al*. Prevalence, disease associations and risk factors for colonization with intestinal spirochaetes (*Brachyspira* spp.) in flocks of laying hens in North-Eastern Italy. *Avian Pathol*. 2008;37(3):281-286.

   doi: 10.1080/03079450802043726
5. Myers SE, Dunn PA, Phillips ND, La T, Hampson DJ. *Brachyspira* intermedia and *Brachyspira pilosicoli* are commonly found in older laying flocks in Pennsylvania. *Avian diseases*. 2009;53(4):533-537.

   doi: 10.1637/8900-042709-Reg.1
6. Amin MM, Phillips ND, La T, Robertson ID, Hampson DJ. Intestinal spirochaetes (*Brachyspira* spp.) colonizing flocks of layer and breeder chickens in Malaysia. *Avian Pathol*. 2014;43(6):501-505.

   doi: 10.1080/03079457.2014.966056
7. Zarabi M, Jamshidi A, Khanzadi S, Razmyar J. Intestinal colonization of different *Brachyspira* spp. In laying hens. *Iran J Vet Med*. 2014;8:213-218.
8. Hampson DJ. The spirochete *Brachyspira pilosicoli*, enteric pathogen of animals and humans. *Clin Microbiol Rev*. 2018;31(1):e00087-17.

   doi: 10.1128/cmr.00087-17
9. Smith JL. Colonic spirochetosis in animals and humans. *J Food Prot*. 2005;68(7):1525-1534.

   doi: 10.4315/0362-028X-68.7.1525
10. Trott DJ, Stanton TB, Jensen NS, Duhamel GE, Johnson JL, Hampson DJ. *Serpulina pilosicoli* sp. Nov., the agent of porcine intestinal spirochetosis. *Int J Syst Bacteriol*. 1996;46(1):206-215.

    doi: 10.1099/00207713-46-1-206
11. Stephens CP, Hampson DJ. Intestinal spirochete infections of chickens: A review of disease associations, epidemiology and control. *Anim Health Res Rev*. 2001;2(1):83-91.
12. Gad A, Willén R, Furugård K, Fors B, Hradsky M. Intestinal spirochaetosis as a cause of longstanding diarrhoea. *Ups J Med Sci*. 1977;82(1):49-54.

    doi: 10.3109/03009737709179059
13. Gan J, Bryant C, Arul D, Parmar C. Intestinal spirochaetosis mimicking acute appendicitis with review of the literature. *BMJ Case Rep*. 2017;2017:bcr2017221574.

    doi: 10.1136/bcr-2017-221574
14. Nishii S, Higashiyama M, Ogata S, *et al*. Human intestinal spirochetosis mimicking ulcerative colitis. *Clin J Gastroenterol*. 2018;11:145-149.

    doi: 10.1007/s12328-017-0807-3
15. Marthinsen L, Willén R, Carlén B, Lindberg E, Värendh G. Intestinal spirochetosis in eight pediatric patients from Southern Sweden. *APMIS*. 2002;110(7-8):571-579.

    doi: 10.1034/j.1600-0463.2002.11007809.x
16. Oxberry SL, Trott DJ, Hampson DJ. *Serpulina pilosicoli*, waterbirds and water: Potential sources of infection for humans and other animals. *Epidemiol Infect*. 1998;121(1):219-225.

    doi: 10.1017/S0950268898008863
17. Trott DJ, Mikosza AS, Combs BG, Oxberry SL, Hampson DJ. Population genetic analysis of *Serpulina pilosicoli* and its molecular epidemiology in villages in the Eastern Highlands of Papua New Guinea. *Int J Syst Bacteriol*. 1998;48(3):659-668.

    doi: 10.1099/00207713-48-3-659

18. Hampson DJ, Lugsomya K, La T, Phillips ND, Trott DJ, Abraham S. Antimicrobial resistance in *Brachyspira* - an increasing problem for disease control. *Vet Microbiol*. 2019;229:59-71.

doi: 10.1016/j.vetmic.2018.12.019

19. Christodoulides M, De Oliveira D, Cleary DW, Humbert MV, Machado-De-Ávila RA, La Ragione RM. An *in silico* reverse vaccinology study of *Brachyspira pilosicoli*, the causative organism of intestinal spirochaetosis, to identify putative vaccine candidates. *Process Biochem*. 2022;122:128-148.

doi: 10.1016/j.procbio.2022.08.014

20. Wanchanthuek P, Bellgard MI, La T, *et al*. The complete genome sequence of the pathogenic intestinal spirochete *Brachyspira pilosicoli* and comparison with other *Brachyspira* genomes. *PLoS One*. 2010;5(7):e11455.

doi: 10.1371/journal.pone.0011455

21. Erlandson KM, Klingler ET. Intestinal spirochetosis: Epidemiology, microbiology, and clinical significance. *Clin Microbiol Newsl*. 2005;27(12):91-96.

doi: 10.1016/j.clinmicnews.2005.05.002

22. Trott DJ, Alt DP, Zuerner RL, *et al*. Identification and cloning of the gene encoding BmpC: an outer-membrane lipoprotein associated with *Brachyspira pilosicoli* membrane vesicles. *Microbiology (Reading)*. 2004;150(4):1041-1053.

doi: 10.1099/mic.0.26755-0

23. Lee BJ, Hampson DJ. Lipo-oligosaccharide profiles of *Serpulina pilosicoli* strains and their serological cross-reactivities. *J Med Microbiol*. 1999;48(4):411-415.

doi: 10.1099/00222615-48-4-411

24. Koebnik R, Locher KP, Van Gelder P. Structure and function of bacterial outer membrane proteins: Barrels in a nutshell. *Mol Microbiol*. 2000;37(2):239-253.

doi: 10.1046/j.1365-2958.2000.01983.x

25. Fairman JW, Noinaj N, Buchanan SK. The structural biology of β-barrel membrane proteins: A summary of recent reports. *Curr Opin Struct Biol*. 2011;21(4):523-531.

doi: 10.1016/j.sbi.2011.05.005

26. Horne JE, Brockwell DJ, Radford SE. Role of the lipid bilayer in outer membrane protein folding in gram-negative bacteria. *J Biol Chem*. 2020;295(30):10340-10367.

doi: 10.1074/jbc.REV120.011473

27. Kleinschmidt JH. Folding of β-barrel membrane proteins in lipid bilayers - unassisted and assisted folding and insertion. *Biochim Biophys Acta*. 2015;1848(9):1927-1943.

doi: 10.1016/j.bbamem.2015.05.004

28. Solan R, Pereira J, Lupas AN, Kolodny R, Ben-Tal N. Gram-negative outer-membrane proteins with multiple β-barrel domains. *Proc Natl Acad Sci U S A*. 2021;118(31):e2104059118.

doi: 10.1073/pnas.2104059118

29. Trott DJ, Alt DP, Zuerner RL, Wannemuehler MJ, Stanton TB. The search for *Brachyspira* outer membrane proteins that interact with the host. *Anim Health Res Rev*. 2001;2(1):19-30.

doi: 10.1079/AHRR200112

30. Tenaya IW, Penhale WJ, Hampson DJ. Preparation of diagnostic polyclonal and monoclonal antibodies against outer envelope proteins of *Serpulina pilosicoli*. *J Med Microbiol*. 1998;47(4):317-324.

doi: 10.1099/00222615-47-4-317

31. La T, Phillips ND, Hampson DJ. Vaccination of chickens with the 34 kDa carboxy-terminus of Bpmp72 reduces colonization with *Brachyspira pilosicoli* following experimental infection. *Avian Pathol*. 2019;48(1):80-85.

doi: 10.1080/03079457.2018.1546377

32. Medicine NLo. *National Center for Biotechnology Information (NCBI)*. Bethesda (MD): Medicine NLo; 1988.

doi: 10.1093/nar/gkaa892

33. Rice P, Longden I, Bleasby A. EMBOSS: The European molecular biology open software suite. *Trends Genet*. 2000;16(6):276-277.

doi: 10.1016/S0168-9525(00)02024-2

34. Teufel F, Almagro Armenteros JJ, Johansen AR, *et al*. SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nat Biotechnol*. 2022;40(7):1023-1025.

doi: 10.1038/s41587-021-01156-3

35. Juncker AS, Willenbrock H, Von Heijne G, Brunak S, Nielsen H, Krogh A. Prediction of lipoprotein signal peptides in gram-negative bacteria. *Protein Sci*. 2003;12(8):1652-1662.

doi: 10.1110/ps.0303703

36. Yu CS, Lin CJ, Hwang JK. Predicting subcellular localization of proteins for gram-negative bacteria by support vector machines based on n-peptide compositions. *Protein Sci*. 2004;13(5):1402-1406.

doi: 10.1110/ps.03479604

37. Yu NY, Wagner JR, Laird MR, *et al*. PSORTb 3.0: Improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics*. 2010;26(13):1608-1615.

doi: 10.1093/bioinformatics/btq249

38. Luo H, Lin Y, Liu T, *et al*. DEG 15, an update of the database of essential genes that includes built-in analysis tools. *Nucleic Acids Res*. 2021;49(D1):D677-D686.

doi: 10.1093/nar/gkaa917

39. Roumia AF, Tsirigos KD, Theodoropoulou MC, Tamposis IA, Hamodrakas SJ, Bagos PG. OMPdb: A global hub of

beta-barrel outer membrane proteins. *Front Bioinform*. 2021;1:646581.

doi: 10.3389/fbinf.2021.646581

40. Bagos PG, LiakopoulipopTD, Hamodrakas SJ. Finding beta-barrel outer membrane proteins with a markov chain model. *WSEAS Trans Biol Biomed*. 2004;2(1):186-189.

41. Ou YY, Gromiha MM, Chen SA, Suwa M. TMBETADISC-RBF: Discrimination of beta-barrel membrane proteins using RBF networks and PSSM profiles. *Comput Biol Chem*. 2008;32(3):227-231.

doi: 10.1016/j.compbiolchem.2008.03.002

42. Bernhofer M, Rost B. TMbed: Transmembrane proteins predicted through language model embeddings. *BMC Bioinformatics*. 2022;23(1):326.

doi: 10.1186/s12859-022-04873-x

43. Abramson J, Adler J, Dunger J, *et al*. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*. 2024;630:493-500.

doi: 10.1038/s41586-024-07487-w

44. Xu J, Zhang Y. How significant is a protein structure similarity with TM-score= 0.5? *Bioinformatics*. 2010;26(7):889-895.

doi: 10.1093/bioinformatics/btq066

45. Zhang Y, Skolnick J. Scoring function for automated assessment of protein structure template quality. *Proteins*. 2004;57(4):702-710.

doi: 10.1002/prot.20264

46. DeLano WL. *The PyMOL Molecular Graphics System*; 2002. https://www/pymol/org [Last accessed on 2025 Jan 25].

47. Lin Z, Akin H, Rao R, *et al*. Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science*. 2023;379(6637):1123-1130.

doi: 10.1126/science.ade2574

48. Waterhouse A, Bertoni M, Bienert S, *et al*. SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res*. 2018;46(W1):W296-W303.

doi: 10.1093/nar/gky427

49. Baek M, DiMaio F, Anishchenko I, *et al*. Accurate prediction of protein structures and interactions using a three-track neural network. *Science*. 2021;373(6557):871-876.

doi: 10.1126/science.abj8754

50. Du Z, Su H, Wang W, *et al*. The trRosetta server for fast and accurate protein structure prediction. *Nat Protoc*. 2021;16(12):5634-5651.

doi: 10.1038/s41596-021-00628-9

51. Holm L. Dali server: Structural unification of protein families. *Nucleic Acids Res*. 2022;50(W1):W210-W215.

doi: 10.1093/nar/gkac387

52. Holm L. Using dali for protein structure comparison. *Methods Mol Biol*. 2020;2112:29-42.

doi: 10.1007/978-1-0716-0270-6_3

53. Van Kempen M, Kim SS, Tumescheit C, *et al*. Fast and accurate protein structure search with foldseek. *Nat Biotechnol*. 2024;42(2):243-246.

doi: 10.1038/s41587-023-01773-0

54. Törönen P, Holm L. PANNZER-a practical tool for protein function prediction. *Protein Sci*. 2022;31(1):118-128.

doi: 10.1002/pro.4193

55. Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J. EggNOG-mapper v2: Functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol Biol Evol*. 2021;38(12):5825-5829.

doi: 10.1093/molbev/msab293

56. Larkin MA, Blackshields G, Brown NP, *et al*. Clustal W and clustal X version 2.0. *Bioinformatics*. 2007;23(21):2947-2948.

doi: 10.1093/bioinformatics/btm404

57. Kumar S, Stecher G, Suleski M, Sanderford M, Sharma S, Tamura K. MEGA12: Molecular evolutionary genetic analysis version 12 for adaptive and green computing. *Mol Biol Evol*. 2024;41(12):msae263.

doi: 10.1093/molbev/msae263

58. Zhang C, Shine M, Pyle AM, Zhang Y. US-align: Universal structure alignments of proteins, nucleic acids, and macromolecular complexes. *Nat Methods*. 2022;19(9):1109-1115.

doi: 10.1038/s41592-022-01585-1

59. Derbyshire MC. Bioinformatic detection of positive selection pressure in plant pathogens: The neutral theory of molecular sequence evolution in action. *Front Microbiol*. 2020;11:644.

doi: 10.3389/fmicb.2020.00644

60. Bakelar J, Buchanan SK, Noinaj N. The structure of the β-barrel assembly machinery complex. *Science*. 2016;351(6269):180-186.

doi: 10.1126/science.aad3460

61. Konovalova A, Kahne DE, Silhavy TJ. Outer membrane biogenesis. *Annu Rev Microbiol*. 2017;71(1):539-556.

doi: 10.1146/annurev-micro-090816-093754

62. Lysnyansky I, Ron Y, Sachse K, Yogev D. Intrachromosomal recombination within the vsp locus of *Mycoplasma bovis* generates a chimeric variable surface lipoprotein antigen. *Infect Immun*. 2001;69(6):3703-3712.

doi: 10.1128/iai.69.6.3703-3712.2001

63. Gabe JD, Dragon E, Chang RJ, McCaman MT. Identification of a linked set of genes in Serpulina hyodysenteriae (B204) predicted to encode closely related 39-kilodalton

extracytoplasmic proteins. *J Bacteriol*. 1998;180(2):444-448.

doi: 10.1128/jb.180.2.444-448.1998

64. Gömmel M, Barth S, Heydel C, Baljer G, Herbst W. Adherence of *Brachyspira* hyodysenteriae to porcine intestinal epithelial cells is inhibited by antibodies against outer membrane proteins. *Curr Microbiol*. 2013;66:286-292.

doi: 10.1007/s00284-012-0267-4

65. Bellgard MI, Wanchanthuek P, La T, *et al*. Genome sequence of the pathogenic intestinal spirochete *Brachyspira* hyodysenteriae reveals adaptations to its lifestyle in the porcine large intestine. *PLoS One*. 2009;4(3):e4641.

doi: 10.1371/journal.pone.0004641

66. Wang Y, Pannuri AA, Ni D, *et al*. Structural basis for translocation of a biofilm-supporting exopolysaccharide across the bacterial outer membrane. *J Biol Chem*. 2016;291(19):10046-10057.

doi: 10.1074/jbc.M115.711762

67. Itoh Y, Rice JD, Goller C, *et al*. Roles of pgaABCD genes in synthesis, modification, and export of the *Escherichia coli* biofilm adhesin poly-beta-1,6-N-acetyl-D-glucosamine. *J Bacteriol*. 2008;190(10):3670-3680.

doi: 10.1128/jb.01920-07

68. Wang X, Preston JF 3rd, Romeo T. The pgaABCD locus of *Escherichia coli* promotes the synthesis of a polysaccharide adhesin required for biofilm formation. *J Bacteriol*. 2004;186(9):2724-2734.

doi: 10.1128/jb.186.9.2724-2734.2004

69. Bellgard M, Hampson DJ, La T. *Genes and Proteins of Brachyspira Hyodysenteriae and uses Thereof*. Google Patents; 2015.

70. Zimmer J. A molecular description of cellulose biosynthesis. *Biophys J*. 2015;108(2):499a.

doi: 10.1016/j.bpj.2014.11.2734

71. Römling U, Galperin MY. Bacterial cellulose biosynthesis: Diversity of operons, subunits, products, and functions. *Trends Microbiol*. 2015;23(9):545-557.

doi: 10.1016/j.tim.2015.05.005

72. Acheson JF, Derewenda ZS, Zimmer J. Architecture of the cellulose synthase outer membrane channel and its association with the periplasmic TPR domain. *Structure*. 2019;27(12):1855-1861.e3.

doi: 10.1016/j.str.2019.09.008

73. Mathes A, Engelhardt H. Nonlinear and asymmetric open channel characteristics of an ion-selective porin in planar membranes. *Biophys J*. 1998;75(3):1255-1262.

doi: 10.1016/S0006-3495(98)74045-7

74. Zeth K, Diederichs K, Welte W, Engelhardt H. Crystal structure of Omp32, the anion-selective porin from comamonas acidovorans, in complex with a periplasmic peptide at 2.1 Å resolution. *Structure*. 2000;8(9):981-992.

doi: 10.1016/S0969-2126(00)00189-1

75. Azghani AO, Idell S, Bains M, Hancock RE. *Pseudomonas aeruginosa* outer membrane protein F is an adhesin in bacterial binding to lung epithelial cells in culture. *Microb Pathog*. 2002;33(3):109-114.

doi: 10.1006/mpat.2002.0514

76. Wu L, Estrada O, Zaborina O, *et al*. Recognition of host immune activation by *Pseudomonas aeruginosa*. *Science*. 2005;309(5735):774-777.

doi: 10.1126/science.1112422

77. Fito-Boncompte L, Chapalain A, Bouffartigues E, *et al*. Full virulence of *Pseudomonas aeruginosa* requires OprF. *Infect Immunity*. 2011;79(3):1176-1186.

doi: 10.1128/iai.00850-10

78. Wessel AK, Liew J, Kwon T, Marcotte EM, Whiteley M. Role of *Pseudomonas aeruginosa* peptidoglycan-associated outer membrane proteins in vesicle formation. *J Bacteriol*. 2013;195(2):213-219.

doi: 10.1128/jb.01253-12

79. Abellón-Ruiz J, Zahn M, Baslé A, Van Den Berg B. Crystal structure of the *Acinetobacter baumannii* outer membrane protein Omp33. *Acta Crystallogr D Struct Biol*. 2018;74(9):852-860.

doi: 10.1107/S205979831800904X

80. Alhede M, Bjarnsholt T, Givskov M, Alhede M. *Pseudomonas aeruginosa* biofilms: Mechanisms of immune evasion. *Adv Appl Microbiol*. 2014;86:1-40.

doi: 10.1016/B978-0-12-800262-9.00001-9

81. Mayeux G, Gayet L, Liguori L, *et al*. Cell-free expression of the outer membrane protein OprF of *Pseudomonas aeruginosa* for vaccine purposes. *Life Sci Alliance*. 2021;4(6):e202000958.

doi: 10.26508/lsa.202000958

82. Goyal P, Krasteva PV, Van Gerven N, *et al*. Structural and mechanistic insights into the bacterial amyloid secretion channel CsgG. *Nature*. 2014;516(7530):250-253.

doi: 10.1038/nature13768

83. La Ragione RM, Sayers AR, Woodward MJ. The role of fimbriae and flagella in the colonization, invasion and persistence of *Escherichia coli* O78:K80 in the day-old-chick model. *Epidemiol Infect*. 2000;124(3):351-363.

doi: 10.1017/S0950268899004045

84. La Ragione RM, Cooley WA, Woodward MJ. The role of fimbriae and flagella in the adherence of avian strains of *Escherichia coli* O78:K80 to tissue culture cells and tracheal and gut explants. *J Med Microbiol*. 2000;49(4):327-338.

doi: 10.1099/0022-1317-49-4-327

85. La Ragione RM, Woodward MJ. Virulence factors of *Escherichia coli* serotypes associated with avian colisepticaemia. *Res Vet Sci*. 2002;73(1):27-35.

doi: 10.1016/S0034-5288(02)00075-9

86. Vandeputte-Rutten L, Bos MP, Tommassen J, Gros P. Crystal structure of neisserial surface protein A (NspA), a conserved outer membrane protein with vaccine potential. *J Biol Chem*. 2003;278(27):24825-24830.

doi: 10.1074/jbc.M302803200

87. Martin D, Cadieux N, Hamel J, Brodeur BR. Highly conserved *Neisseria meningitidis* surface protein confers protection against experimental infection. *J Exp Med*. 1997;185(7):1173-1184.

doi: 10.1084/jem.185.7.1173

88. Moe GR, Tan S, Granoff DM. Differences in surface expression of NspA among *Neisseria meningitidis* group B strains. *Infect Immun*. 1999;67(11):5664-5675.

doi: 10.1128/iai.67.11.5664-5675.1999

89. Moe GR, Zuno-Mitchell P, Lee SS, Lucas AH, Granoff DM. Functional activity of anti-neisserial surface protein A monoclonal antibodies against strains of *Neisseria meningitidis* serogroup B. *Infect Immun*. 2001;69(6):3762-3771.

doi: 10.1128/iai.69.6.3762-3771.2001

90. Koronakis V, Eswaran J, Hughes C. Structure and function of TolC: The bacterial exit duct for proteins and drugs. *Annu Rev Biochem*. 2004;73(1):467-489.

doi: 10.1146/annurev.biochem.73.011303.074104

91. Koronakis V, Sharff A, Koronakis E, Luisi B, Hughes C. Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export. *Nature*. 2000;405(6789):914-919.

doi: 10.1038/35016007

92. Vakharia H, German GJ, Misra R. Isolation and characterization of *Escherichia coli* tolC mutants defective in secreting enzymatically active alpha-hemolysin. *J Bacteriol*. 2001;183(23):6908-6916.

doi: 10.1128/jb.183.23.6908-6916.2001

93. Wandersman C, Delepelaire P. TolC, an *Escherichia coli* outer membrane protein required for hemolysin secretion. *Proc Natl Acad Sci*. 1990;87(12):4776-4780.

doi: 10.1073/pnas.87.12.4776

94. Clowes R. Transmission and elimination of colicin factors and some aspects of immunity to colicin E1 in *Escherichia coli*. *Zentrabl Bakteriol Parasiten Abt I Orig*. 1965;196:152-157.

95. Davies JK, Reeves P. Genetics of resistance to colicins in *Escherichia coli* K-12: Cross-resistance among colicins of group A. *J Bacteriol*. 1975;123(1):102-117.

doi: 10.1128/jb.123.1.102-117.1975

96. Fralick JA. Evidence that TolC is required for functioning of the Mar/AcrAB efflux pump of *Escherichia coli*. *J Bacteriol*. 1996;178(19):5803-5805.

doi: 10.1128/jb.178.19.5803-5805.1996

97. German GJ, Misra R. The TolC protein of *Escherichia coli* serves as a cell-surface receptor for the newly characterized TLS bacteriophage. *J Mol Biol*. 2001;308(4):579-585.

doi: 10.1006/jmbi.2001.4578

98. Fredriksson R, Lagerström MC, Lundin LG, Schiöth HB. The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Mol Pharmacol*. 2003;63(6):1256-1272.

doi: 10.1124/mol.63.6.1256

99. Pierce KL, Premont RT, Lefkowitz RJ. Seven-transmembrane receptors. *Nat Rev Mol Cell Biol*. 2002;3(9):639-650.

doi: 10.1038/nrm908

100. Consortium TU. UniProt: The universal protein knowledgebase in 2025. *Nucleic Acids Res*. 2024;53(D1):D609-D617.

doi: 10.1093/nar/gkae1010

101. Van Alphen L, Havekes L, Lugtenberg B. Major outer membrane protein d of *Escherichia coli* K12. Purification and *in vitro* activity of bacteriophage k3 and f-pilus mediated conjugation. *FEBS Lett*. 1977;75:285-290.

doi: 10.1016/0014-5793(77)80104-x

102. Datta DB, Arden B, Henning U. Major proteins of the *Escherichia coli* outer cell envelope membrane as bacteriophage receptors. *J Bacteriol*. 1977;131(3):821-829.

doi: 10.1128/jb.131.3.821-829.1977

103. Chai T, Foulds J. Demonstration of a missing outer membrane protein in tolG mutants of *Escherichia* cell. *J Mol Biol*. 1974;85(3):465-474.

doi: 10.1016/0022-2836(74)90445-8

104. Skurray RA, Hancock RE, Reeves P. Con--mutants: Class of mutants in *Escherichia coli* K-12 lacking a major cell wall protein and defective in conjugation and adsorption of a bacteriophage. *J Bacteriol*. 1974;119(3):726-735.

doi: 10.1128/jb.119.3.726-735.1974

105. Schweizer M, Henning U. Action of a major outer cell envelope membrane protein in conjugation of *Escherichia coli* K-12. *J Bacteriol*. 1977;129(3):1651-1652.

doi: 10.1128/jb.129.3.1651-1652.1977

106. Braun V. Covalent lipoprotein from the outer membrane of *Escherichia coli*. *Biochim Biophys Acta*. 1975;415(3):335-377.

doi: 10.1016/0304-4157(75)90013-1

107. Sonntag I, Schwarz H, Hirota Y, Henning U. Cell envelope and shape of *Escherichia coli*: Multiple mutants missing

the outer membrane lipoprotein and other major outer membrane proteins. *J Bacteriol*. 1978;136(1):280-285.

doi: 10.1128/jb.136.1.280-285.1978

108. Sugawara E, Nikaido H. Pore-forming activity of OmpA protein of *Escherichia coli*. *J Biol Chem*. 1992;267(4):2507-2511.

doi: 10.1016/S0021-9258(18)45908-X

109. Krishnan S, Prasadarao NV. Outer membrane protein A and OprF: Versatile roles in gram-negative bacterial infections. *FEBS J*. 2012;279(6):919-931.

doi: 10.1111/j.1742-4658.2012.08482.x

110. Laguri C, Sperandeo P, Pounot K, *et al*. Interaction of lipopolysaccharides at intermolecular sites of the periplasmic Lpt transport assembly. *Sci Rep.* 2017;7(1):9715.

doi: 10.1038/s41598-017-10136-0

111. Pieńko T, Trylska J. Extracellular loops of BtuB facilitate transport of vitamin B12 through the outer membrane of *E. Coli*. *PLoS Comput Biol*. 2020;16(7):e1008024.

doi: 10.1371/journal.pcbi.1008024

112. Di Girolamo PM, Kadner RJ, Bradbeer C. Isolation of vitamin B 12 transport mutants of *Escherichia coli*. *J Bacteriol*. 1971;106(3):751-757.

doi: 10.1128/jb.106.3.751-757.1971

113. White JC, Di Girolamo PM, Fu ML, Preston YA, Bradbeer C. Transport of vitamin B 12 in *Escherichia coli*. Location and properties of the initial B 12 -binding site. *J Biol Chem*. 1973;248(11):3978-3986.

doi: 10.1016/S0021-9258(19)43828-3

114. Noinaj N, Guillier M, Barnard TJ, Buchanan SK. TonB-dependent transporters: Regulation, structure, and function. *Annu Rev Microbiol*. 2010;64(1):43-60.

doi: 10.1146/annurev.micro.112408.134247

115. Schauer K, Rodionov DA, De Reuse H. New substrates for TonB-dependent transport: Do we only see the 'tip of the iceberg'? *Trends Biochem Sci*. 2008;33(7):330-338.

doi: 10.1016/j.tibs.2008.04.012

116. Biswas S, Mohammad MM, Patel DR, Movileanu L, Van Den Berg B. Structural insight into OprD substrate specificity. *Nat Struct Mol Biol*. 2007;14(11):1108-1109.

doi: 10.1038/nsmb1304

117. Moraes TF, Bains M, Hancock RE, Strynadka NC. An arginine ladder in OprP mediates phosphate-specific transfer across the outer membrane. *Nat Struct Mol Biol*. 2007;14(1):85-87.

doi: 10.1038/nsmb1189

118. Eren E, Vijayaraghavan J, Liu J, *et al*. Substrate specificity within a family of outer membrane carboxylate channels. *PLoS Biol*. 2012;10(1):e1001242.

doi: 10.1371/journal.pbio.1001242

119. Biswas S, Mohammad MM, Movileanu L, Van Den Berg B. Crystal structure of the outer membrane protein OpdK from *Pseudomonas aeruginosa*. *Structure*. 2008;16(7):1027-1035.

doi: 10.1016/j.str.2008.04.009

120. Kuehne SA, Cartman ST, Heap JT, Kelly ML, Cockayne A, Minton NP. The role of toxin A and toxin B in *Clostridium* difficile infection. *Nature*. 2010;467(7316):711-713.

doi: 10.1038/nature09397

121. Van Den Berg B, Prathyusha Bhamidimarri S, Dahyabhai Prajapati J, Kleinekathöfer U, Winterhalter M. Outer-membrane translocation of bulky small molecules by passive diffusion. *Proc Natl Acad Sci*. 2015;112(23):E2991-E2999.

doi: 10.1073/pnas.1424835112

122. Pajatsch M, Andersen C, Mathes A, Böck A, Benz R, Engelhardt H. Properties of a cyclodextrin-specific, unusual porin from *Klebsiella oxytoca*. *J Biol Chem*. 1999;274(35):25159-25166.

doi: 10.1074/jbc.274.35.25159

123. Gruss F, Zähringer F, Jakob RP, Burmann BM, Hiller S, Maier T. The structural basis of autotransporter translocation by TamA. *Nat Struct Mol Biol*. 2013;20(11):1318-1320.

doi: 10.1038/nsmb.2689

124. Selkrig J, Mosbahi K, Webb CT, *et al*. Discovery of an archetypal protein transport system in bacterial outer membranes. *Nat Struct Mol Biol*. 2012;19(5):506-510.

doi: 10.1038/nsmb.2261

125. Pavlova O, Peterson JH, Ieva R, Bernstein HD. Mechanistic link between β barrel assembly and the initiation of autotransporter secretion. *Proc Natl Acad Sci*. 2013;110(10):E938-E947.

doi: 10.1073/pnas.1219076110

126. Matthijs S, Koedam N, Cornelis P, De Greve H. The trehalose operon of *Pseudomonas* fluorescens ATCC 17400. *Res Microbiol*. 2000;151(10):845-851.

doi: 10.1016/S0923-2508(00)01151-7

127. Abergel C, Bouveret E, Claverie JM, *et al*. Structure of the *Escherichia coli* TolB protein determined by MAD methods at 1.95 Å resolution. *Structure*. 1999;7(10):1291-1300.

doi: 10.1016/S0969-2126(00)80062-3

128. Rigal A, Bouveret E, Lloubes R, Lazdunski C, Benedetti H. The TolB protein interacts with the porins of *Escherichia coli*. *J Bacteriol*. 1997;179(23):7274-7279.

doi: 10.1128/jb.179.23.7274-7279.1997

129. Fognini-Lefebvre N, Lazzaroni JC, Portalier R. tolA, tolB and excC, three cistrons involved in the control of pleiotropic release of periplasmic proteins by *Escherichia coli* K12. *Mol Gen Genet*. 1987;209:391-395.

doi: 10.1007/BF00329670

130. Lazzaroni JC, Fognini-Lefebvre N, Portalier R. Cloning of the excC and excD genes involved in the release of periplasmic proteins by *Escherichia coli* K12. *Mol Gen Genet*. 1989;218:460-464.

doi: 10.1007/BF00332410

131. Maier T, Clantin B, Gruss F, *et al*. Conserved Omp85 lid-lock structure and substrate recognition in FhaC. *Nat Commun*. 2015;6(1):7452.

doi: 10.1038/ncomms8452

132. Delattre AS, Saint N, Clantin B, *et al*. Substrate recognition by the POTRA domains of TpsB transporter FhaC. *Mol Microbiol*. 2011;81(1):99-112.

doi: 10.1111/j.1365-2958.2011.07680.x

133. Lauber F, Deme JC, Lea SM, Berks BC. Type 9 secretion system structures reveal a new protein transport mechanism. *Nature*. 2018;564(7734):77-82.

doi: 10.1038/s41586-018-0693-y

134. Van Den Berg B, Black PN, Clemons WM Jr., Rapoport TA. Crystal structure of the long-chain fatty acid transporter FadL. *Science*. 2004;304(5676):1506-1509.

doi: 10.1126/science.1097524

135. Black PN, Said B, Ghosn CR, Beach JV, Nunn WD. Purification and characterization of an outer membrane-bound protein involved in long-chain fatty acid transport in *Escherichia coli*. *J Biol Chem*. 1987;262(3):1412-1419.

doi: 10.1016/S0021-9258(19)75801-3

136. Lauber F, Deme JC, Liu X, *et al*. Structural insights into the mechanism of protein transport by the type 9 secretion system translocon. *Nat Microbiol*. 2024;9(4):1089-1102.

doi: 10.1038/s41564-024-01644-7

137. Otto BR, Sijbrandi R, Luirink J, *et al*. Crystal structure of hemoglobin protease, a heme binding autotransporter protein from pathogenic *Escherichia coli*. *J Biol Chem*. 2005;280(17):17339-17345.

doi: 10.1074/jbc.M412885200

138. Otto BR, Van Dooren SJ, Dozois CM, Luirink J, Oudega B. *Escherichia coli* hemoglobin protease autotransporter contributes to synergistic abscess formation and heme-dependent growth of *Bacteroides fragilis*. *Infect Immun*. 2002;70(1):5-10.

doi: 10.1128/iai.70.1.5-10.2002

139. Otto BR, Van Dooren SJ, Nuijens JH, Luirink J, Oudega B. Characterization of a hemoglobin protease secreted by the pathogenic *Escherichia coli* strain EB1. *J Exp Med*. 1998;188(6):1091-1103.

doi: 10.1084/jem.188.6.1091

140. Berg BVD, Bhamidimarri SP, Winterhalter M. Crystal structure of a COG4313 outer membrane channel. *Sci Rep*. 2015;5(1):11927.

doi: 10.1038/srep11927

141. Henderson IR, Nataro JP. Virulence functions of autotransporter proteins. *Infect Immun*. 2001;69(3):1231-1243.

doi: 10.1128/iai.69.3.1231-1243.2001

142. Yen MR, Peabody CR, Partovi SM, Zhai Y, Tseng YH, Saier MH Jr. Protein-translocating outer membrane porins of gram-negative bacteria. *Biochim Biophys Acta*. 2002;1562(1-2):6-31.

doi: 10.1016/S0005-2736(02)00359-0

143. Henderson IR, Navarro-Garcia F, Nataro JP. The great escape: Structure and function of the autotransporter proteins. *Trends Microbiol*. 1998;6(9):370-378.

doi: 10.1016/S0966-842X(98)01318-3

144. Dutta PR, Cappello R, Navarro-García F, Nataro JP. Functional comparison of serine protease autotransporters of Enterobacteriaceae. *Infect Immun*. 2002;70(12):7105-7113.

doi: 10.1128/iai.70.12.7105-7113.2002

145. Brunder W, Schmidt H, Karch H. EspP, a novel extracellular serine protease of enterohaemorrhagic *Escherichia coli* O157:H7 cleaves human coagulation factor V. *Mol Microbiol*. 1997;24(4):767-778.

doi: 10.1046/j.1365-2958.1997.3871751.x

146. Scandella CJ, Kornberg A. Membrane-bound phospholipase A1 purified from *Escherichia coli*. *Biochemistry*. 1971;10(24):4447-4456.

doi: 10.1021/bi00800a015

147. Pugsley AP, Schwartz M. Colicin E2 release: Lysis, leakage or secretion? Possible role of a phospholipase. *EMBO J*. 1984;3(10):2393-2397.

doi: 10.1002/j.1460-2075.1984.tb02145.x

148. Van Der Wal FJ, Luirink J, Oudega B. Bacteriocin release proteins: Mode of action, structure, and biotechnological application. *FEMS Microbiol Rev*. 1995;17(4):381-399.

doi: 10.1111/j.1574-6976.1995.tb00221.x

149. De Vries FP, Cole R, Dankert J, Frosch M, Van Putten JP. *Neisseria meningitidis* producing the Opc adhesin binds epithelial cell proteoglycan receptors. *Mol Microbiol*. 1998;27(6):1203-1212.

doi: 10.1046/j.1365-2958.1998.00763.x

150. Virji M, Makepeace K, Moxon ER. Distinct mechanisms of interactions of Opc-expressing meningococci at apical and basolateral surfaces of human endothelial cells; the role of integrins in apical interactions. *Mol Microbiol*. 1994;14(1):173-184.

doi: 10.1111/j.1365-2958.1994.tb01277.x

151. Mammarappallil JG, Elsinghorst EA. Epithelial cell adherence mediated by the enterotoxigenic *Escherichia coli* Tia protein. *Infect Immun*. 2000;68(12):6595-6601.

doi: 10.1128/iai.68.12.6595-6601.2000

152. Fleckenstein JM, Kopecko DJ, Warren RL, Elsinghorst EA. Molecular characterization of the tia invasion locus from enterotoxigenic *Escherichia coli*. *Infect Immun*. 1996;64(6):2256-2265.

doi: 10.1128/iai.64.6.2256-2265.1996

153. Gensberg K, Smith AW, Brinkman FS, Hancock RE. Identification of *oprG*, a gene encoding a major outer membrane protein of *Pseudomonas aeruginosa*. *J Antimicrob Chemother*. 1999;43(4):607-608.

doi: 10.1093/jac/43.4.607

154. Celia H, Noinaj N, Zakharov SD, *et al*. Structural insight into the role of the ton complex in energy transduction. *Nature*. 2016;538(7623):60-65.

doi: 10.1038/nature19757