



OMPdb: A Global Hub of Beta-Barrel Outer Membrane Proteins

Ahmed F. Roumia^{1†}, Konstantinos D. Tsirigos^{2†}, Margarita C. Theodoropoulou¹, Ioannis A. Tamposis¹, Stavros J. Hamdrakas³ and Pantelis G. Bagos^{1*}

¹ Department of Computer Science and Biomedical Informatics, University of Thessaly, Lamia, Greece, ² EMBL-EBI, Wellcome Genome Campus, Cambridge, United Kingdom, ³ Section of Cell Biology and Biophysics, Department of Biology, School of Sciences, National and Kapodistrian University of Athens, Athens, Greece

OMPdb (www.ompdb.org) was introduced as a database for β -barrel outer membrane proteins from Gram-negative bacteria in 2011 and then included 69,354 entries classified into 85 families. The database has been updated continuously using a collection of characteristic profile Hidden Markov Models able to discriminate between the different families of prokaryotic transmembrane β -barrels. The number of families has increased ultimately to a total of 129 families in the current, second major version of OMPdb. New additions have been made in parallel with efforts to update existing families and add novel families. Here, we present the upgrade of OMPdb, which from now on aims to become a global repository for all transmembrane β -barrel proteins, both eukaryotic and bacterial.

Keywords: beta barrel membrane proteins, Gram-negative bacteria, mitochondria, chloroplast, β -barrels, Hidden Markov Model, database, sequence analysis

INTRODUCTION

Integral membrane proteins (IMPs) play a vital role in cell tasks and communication. IMPs represent roughly 20–30% of the human genome (von Heijne, 1999). They can be structurally divided into two distinct categories, the α -helical membrane proteins and the β -barrel ones (von Heijne, 1999). While the former are commonly found in the bacterial inner membrane and cell membranes of all eukaryotic cells, the latter are located exclusively in the outer membranes of mitochondria, chloroplasts, and Gram-negative bacteria (Cavalier-Smith, 2000). In contrast to the α -helical proteins, which are the major type of the IMPs, the β -barrel membrane proteins are fewer, comprising <3% of the proteins encoded in bacterial genomes (Casadio et al., 2003) (one order of magnitude lower than that of α -helical membrane proteins). However, the β -barrel membrane proteins participate in crucial biological activities in prokaryotic organisms, as well as in the eukaryotic organelles.

Endosymbiosis of Gram-negative bacteria ancestors with host cells generated eukaryotic organelles, chloroplasts (Kleine et al., 2009) and mitochondria (Gray et al., 1999), whose outer membranes contain pores with β -barrel topology. The eukaryotic β -barrel pores are IMPs and can be classified into three main functional groups. Firstly, the specific diffusion channels group, which consists of the voltage-dependent anion channel (VDAC) (De Pinto et al., 2010), the translocase of the outer mitochondrial membrane (TOM40) (Hill et al., 1998), the outer envelope protein 21 (OEP21) (Hemmler et al., 2006), the outer envelope protein 23 (OEP23) (Goetze et al., 2015), the outer envelope protein (OEP37) (Goetze et al., 2006) and the outer envelope protein 40 (OEP40) (Harsman et al., 2016). Secondly, the non-specific diffusion channels group, that includes only the outer envelope protein 24 (OEP24) (Pohlmeier et al., 1998). Thirdly, the biogenesis/secretion group, that contains the mitochondrial distribution and morphology protein

OPEN ACCESS

Edited by:

Michael Gromiha,
Indian Institute of Technology
Madras, India

Reviewed by:

Bert Van Den Berg,
Newcastle University, United Kingdom
Castrense Savojardo,
University of Bologna, Italy
Joanna Slusky,
University of Kansas, United States

*Correspondence:

Pantelis G. Bagos
pbagos@compngen.org

[†]These authors have contributed
equally to this work and share first
authorship

Specialty section:

This article was submitted to
Protein Bioinformatics,
a section of the journal
Frontiers in Bioinformatics

Received: 27 December 2020

Accepted: 18 March 2021

Published: 09 April 2021

Citation:

Roumia AF, Tsirigos KD,
Theodoropoulou MC, Tamposis IA,
Hamdrakas SJ and Bagos PG (2021)
OMPdb: A Global Hub of Beta-Barrel
Outer Membrane Proteins.
Front. Bioinform. 1:646581.
doi: 10.3389/fbinf.2021.646581

10 (MDM10) (Meisinger et al., 2004), the sorting and assembly machinery (SAM50) (Stojanovski et al., 2007), the trigalactosyldiacylglycerol 4 (TGD4) (Wang et al., 2012), the translocon at the outer membrane of chloroplasts 75 (TOC75) (Hinnah et al., 2002) and the outer envelope protein 80 (OEP80) (Gross et al., 2020).

OMPdb is a database of bacterial β -barrel outer membrane proteins (β OMPs). It was launched in 2011 and is until now the most complete collection of Gram-negative bacteria β OMPs. It has been updated continuously using a collection of characteristic profile Hidden Markov Models (pHMMs) as shown in **Figure 1**.

Up to now, OMPdb (www.ompdb.org) was a repository of only prokaryotic β OMPs and classified them into the following eight functional categories (Tsirigos et al., 2011): receptors, biogenesis/secretion, specific channels, non-specific channels, structural, adhesion, enzymes, and unknown. Specific and non-specific diffusion channels, together with biogenesis/secretion proteins have a crucial role in the bacterial life. Bacterial diffusion channels are the most abundant type of β OMPs in the outer membrane. There are two main classes of bacterial diffusion channels: the non-specific and specific diffusion channels (Schirmer et al., 1995; Koebnik et al., 2000; Schulz, 2002; Nikaido, 2003; Pagès et al., 2008; Wirth et al., 2009; Van Den Berg, 2012). The first class of proteins may have 14 strands, as the outer membrane protein G (OmpG) (Galdiero et al., 2013), 16 strands and permit the flux of hydrophilic substrates smaller than 600 Kda (Ringle and Schulz, 2002) or 26 strands, as the MtrB, the β -barrel member of the outer membrane spanning protein complex, MtrAB (Edwards et al., 2020). The second class members form 12-stranded beta barrels, as in the oligogalacturonate-specific KdgM channel (KdgM) (Wirth et al., 2009), 14-stranded, as the long-chain fatty acid transporter FadL (van den Berg et al., 2004), 16-stranded, like the outer membrane porin B (OprB) (Van Den Berg, 2012) and OprP (Moraes et al., 2007) or 18-stranded, as the outer membrane carboxylate channel (Occ) (Eren et al., 2012; Liu et al., 2012) and maltoporins (Wang et al., 1997; Forst et al., 1998).

In addition to the aforementioned functions of bacterial β OMPs, bacterial adhesion is considered as the first step of biofilm formation and colonization (Dankert et al., 1986). Biofilm, an accumulated biomass of microorganisms, can be harmful to human life where it causes pathogen interaction with host cells (Gristina, 1987; Soto and Hultgren, 1999). Prior to entering inside, pathogenic bacteria adhere to host cells and secrete their products to the host. To this end, bacteria attach to host cells via one of the following three ways: fimbrial (Pili), which is considered the most common proteinous adhesin of Gram-negative bacteria (Soto and Hultgren, 1999) (e.g., *Neisseria gonorrhoeae* Wall and Kaiser, 1999), non-fimbrial adhesions (autotransporters, like e.g., *Haemophilus influenzae* Girard and Mourez, 2006), or using other unique nanofibers like *Caulobacter crescentus* (Corpe, 1972) does. Subsequently, pathogenic bacteria use bacterial secretion as a cellular tool to secrete their virulence factors (predominantly proteins) for invading their host cells. Gram-negative bacteria use eight specific secretion types, with

two of them being the most common, i.e., the general secretion (termed the Sec-pathway) and the twin arginine translocation (termed the Tat-pathway Natale et al., 2008; Green and Meccas, 2016). Despite the high degree of similarity between the targeting signals that direct secretory proteins to their routes, the secretion mechanisms are widely different. The Sec-pathway stimulates the translocation of unfolded proteins, whereupon they fold at the *trans*-side of the membrane into their native structure. On the contrary, the Tat-pathway catalyzes the translocation of folded proteins (Natale et al., 2008).

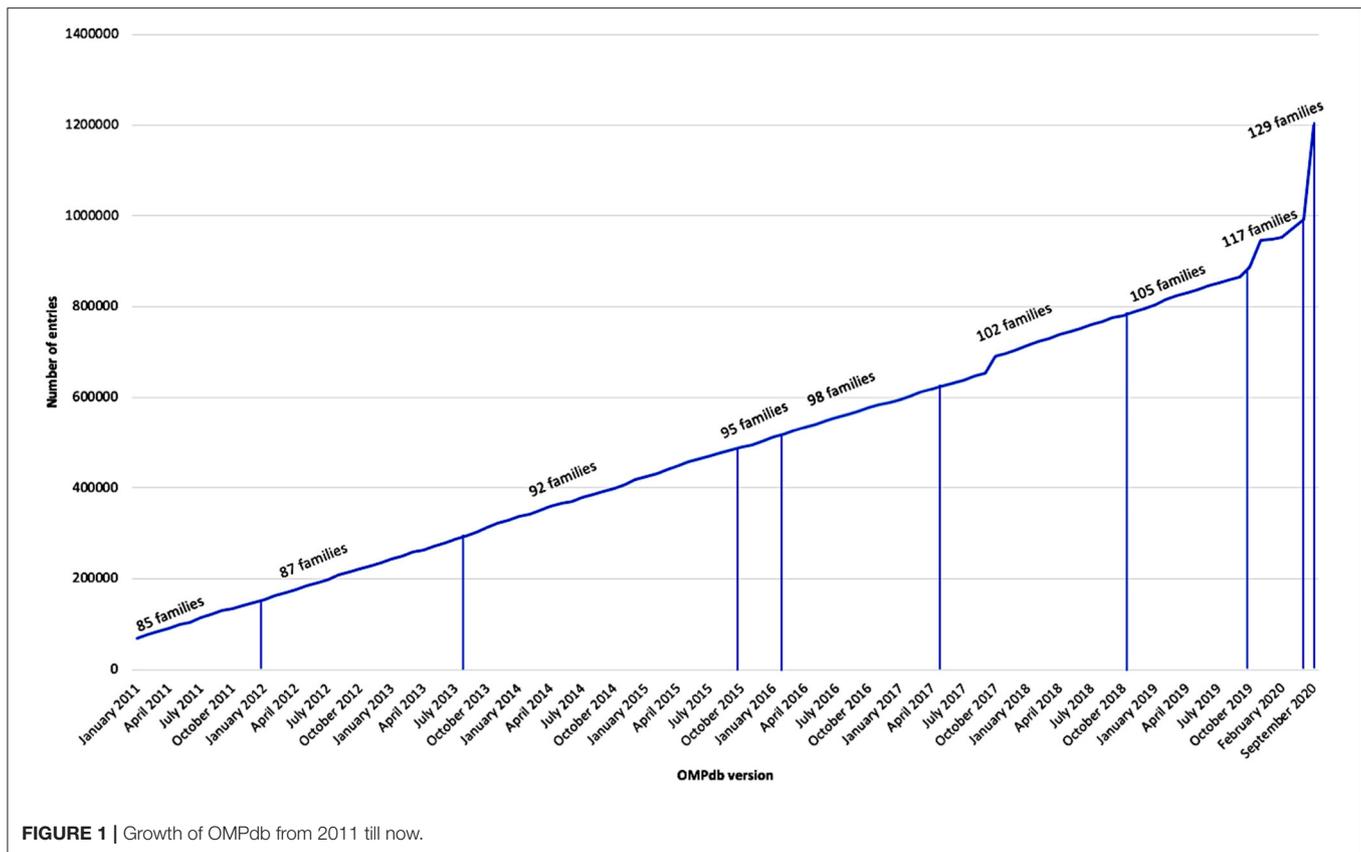
To reach their proper location, the bacterial β OMPs are being accelerated by the five units (BamA, -B, -C, -D, and -E) of the β -barrel assembly machine (Bam) (Tomasek and Kahne, 2021). BamA, the core of the machine, forms a 16-stranded β -barrel in addition to a periplasmic domain with five polypeptide translocation-associated (POTRA) motifs. It is thought that these five POTRA motifs interact with the four lipoproteins of the Bam complex (Jansen et al., 2015). The Omp85 superfamily of outer membrane proteins encompasses the translocation machines of bacterial β OMPs (BamA) (Voulhoux et al., 2003), mitochondrial β OMPs (SAM50) (Höhr et al., 2018), and chloroplastic ones (TOC75/OEP80) (Day et al., 2019; Gross et al., 2020).

Lipopolysaccharide (LPS) is substantial for the vitality of most Gram-negative bacteria. It plays critical roles not only in the survival of them in severe environments by composing a biofilm but also in colonizing the infected host and avoiding attacks from the human immune system (Zhang et al., 2013; Whitfield and Trent, 2014). LptD, an integral membrane protein, is considered as one of the largest reported bacterial β -barrel so far. Its 26 stranded β -barrel together with LptE form the LPS complex translocon “barrel and plug” at the outer membrane of the Gram-negative bacteria (Dong et al., 2014; Gu et al., 2015).

All Gram-negative bacteria are always in a crucial need for large substrates, such as vitamin B12 or iron-siderophore complexes. Therefore, the outer membrane of Gram-negative bacteria contains diverse high-molecular-weight proteins called TonB-dependent transporter (TBDTs) (Braun, 1995) (e.g., FhuA Coulton et al., 1986) and FusA (Grinter et al., 2016). Furthermore, β OMPs may act as enzymes like Lipopolysaccharides 3-O-deacylase enzyme PagL (Bishop, 2008). Also, β OMPs maintain the stability of the outer membrane as outer membrane protein (OmpA) (Ringle and Schulz, 2002). Lastly, due to the ever-growing number of sequenced genomes of Gram-negative bacteria, the number of proteins with unknown function is relatively small (Tsirigos et al., 2011).

Apart from the typical β OMPs found in the outer membrane of endosymbiotic organelles (e.g., VDAC Bayrhuber et al., 2008, TOC75 Soll and Schleiff, 2004, and OmpG Galdiero et al., 2013), there are also multimeric transmembrane pores formed either from pathogenic microbes in the host's cells, called Pore-forming Toxins (PFTs) (e.g., Aerolysin, which forms heptameric Iacovache et al., 2016 or non-ameris Podobnik et al., 2016 pores and Hemolysin De and Olson, 2011), whose role is not yet clear, or from human cells, called Membrane Attack Complex (MAC), that protect human cells by forming pores in the membrane of pathogenic Gram-negative bacteria (Menny et al., 2018).

Abbreviations: β OMPs, β -barrel outer membrane proteins.



Because of the biomedical importance of the β OMPs, several biological databases have been launched to annotate and organize the β OMPs in several ways. The Protein Data Bank of Transmembrane Proteins (PDBTM) (<http://pdbtm.enzim.hu>) was launched in 2004 as the first comprehensive database for experimentally verified transmembrane proteins either α -helical or β -barrels. The database is being updated continuously using the TMDET algorithm. This algorithm can differentiate between the non-transmembrane proteins and transmembrane proteins based on their 3D coordinates. PDBTM (version 2019-02-22) includes 130 non-redundant β OMPs (Kozma et al., 2013).

TMBB-DB (<http://beta-barrel.tulane.edu/>) was launched in 2012 as a database for transmembrane β OMPs. TMBB-DB includes around 50,000 predicted transmembrane β OMPs (Freeman and Wimley, 2012). To detect the transmembrane β OMPs, the database uses a compilation of two predictors: the Freeman-Wimley algorithm (Freeman and Wimley, 2010) and SignalP5 (Almagro Armenteros et al., 2019).

The Topology Data Bank of Transmembrane Proteins (TOPDB) (<http://topdb.enzim.ttk.mta.hu>) was launched in 2008 as a hub for the topology of the 3D solved structure transmembrane proteins. TOPDB was dependent on the retrieved data from the Protein Data Bank (PDB) and PDBTM in parallel with the generated topology prediction by HMMTOP (Tusnády et al., 2008). In 2015, they improved the generated topology data by incorporating high throughput techniques like

the sequential positions of N- or O-glycosylations. Further, they developed a new algorithm for collecting the scattered topology from various public databases. In addition, they created a new method for evaluating the reliability of the generated topology data. TOPDB provides now the topology information for 201 β OMPs (Dobson et al., 2015).

The Transporter Classification Database (TCDB) (<http://www.tcdb.org>) is a database of all the transporter proteins. TCDB provides comprehensive structural, functional and medical information for all the transporter proteins. Moreover, TCDB has a specialized software that has been designed and integrated into the TCDB to focus on the distinctive characteristics of the transporter proteins, and subsequently their biomedical function. TCDB classifies 1260 β OMPs into 93 different families (Saier et al., 2016).

The MemProtMD database (<http://memprotmd.bioch.ox.ac.uk>) was launched in 2015 acting as a repository for the membrane-embedded protein structures and their lipid interactions. MemProtMD provides the molecular dynamics (MD) simulation results of the protein-lipid interactions for 405 β OMPs (Newport et al., 2019).

The Pfam database (<http://pfam.xfam.org/>) is a widely used repository for protein families. Pfam classifies the protein sequences into clans, families and domains based on specific pHMM for every family. In Pfam, all the β Omp families are gathered together in one clan called MBB (CL0193). This clan

contains 92 families with 1,092,329 protein members (El-Gebali et al., 2019).

Here, we present the upgraded and updated OMPdb, a global repository for all β OMP families with experimentally determined 3D structure or not regardless of their location, function and mechanism of pore formation.

MATERIALS AND METHODS

In order to upgrade OMPdb, new features and novel β OMP families, prokaryotic and eukaryotic ones, are included as shown in **Figure 2**.

New Families

Initially, our efforts were focused on finding as many new transmembrane β OMP families in the literature as possible. To this end, we based our literature search along two main axes, adding novel transmembrane β OMP families and updating current transmembrane β OMP ones, either by adding new functional annotation or by updating their pHMMs as shown in **Figure 3**.

First, we performed an extensive literature search to find novel families of transmembrane β OMPs in Gram-negative bacteria (not included in OMPdb Tsirigos et al., 2011). Furthermore, we retrieved all the 3D structures of transmembrane β OMPs that were not included in OMPdb after a thorough search in both PDBTM (version 2019-02-22) (Kozma et al., 2013) and MemProtMD (version 13-10-2019) (Newport et al., 2019). Additionally, we tried to characterize and assign function to as many of the Domain of Unknown Function (DUF) families, members of the MBB clan of Pfam version 33.1 May 2020 (El-Gebali et al., 2019), as possible. It is important to mention here that the same procedure was used in both the first OMPdb publication (Tsirigos et al., 2011) and our recent eukaryotic transmembrane β OMPs publication (Roumia et al., 2020) to create new pHMMs for those families without known domains in Pfam. In brief, the β -barrel domain of the experimentally verified protein(s) reported in the respective published articles were used to perform search against the UniProt (The UniProt Consortium, 2019) database with protein BLAST (Altschul et al., 1997). Subsequently, Multiple Sequence Alignments (MSAs) of the best-scoring results for each family were made by Clustal Omega (Li et al., 2015) and these alignments were refined with MUSCLE (Edgar, 2004). Finally, we built the corresponding pHMMs using the refined alignments with HMMER version 3.3 (Eddy, 1998).

For the Autotransporter and TBDTs families, we updated their pHMMs by adding the new members to their respective Pfam MSAs, and then, following the above-mentioned method. For the Omp85 family we built a new pHMM corresponding only to its bacterial members (BamA/TamA) and mitochondrial ones (SAM50). To this end, we searched UniProt (The UniProt Consortium, 2019) for mitochondrial SAM50 proteins under Omp85 family, merged the SAM50 search results to the BamA/TamA members, and then, followed the aforementioned steps to building the new pHMM.

Finally, we performed a search with our new pHMMs' collection against the reference proteomes, prokaryotic, and eukaryotic, retrieved from the UniProt database (Release 2020_03) (The UniProt Consortium, 2019).

New Features

As part of the upgrade of OMPdb, we decided to include some new features in the latest version. Providing 3D models for all seed set alignment members for all the families included in the database was the first new feature we added. The models were constructed either via homology modeling using SWISS-MODEL (Biasini et al., 2014) (for the families that contained member(s) with experimentally verified structure) or via *de novo* modeling using PHYRE2 (Kelley et al., 2015) (for the families without any determined 3D structures). As an additional feature, we now provide, in each entry, information about the signal peptide and the topology of each β OMP using the SignalP5.0 (Almagro Armenteros et al., 2019) and the PRED-TMBB2 (Tsirigos et al., 2016) tool, respectively.

RESULTS

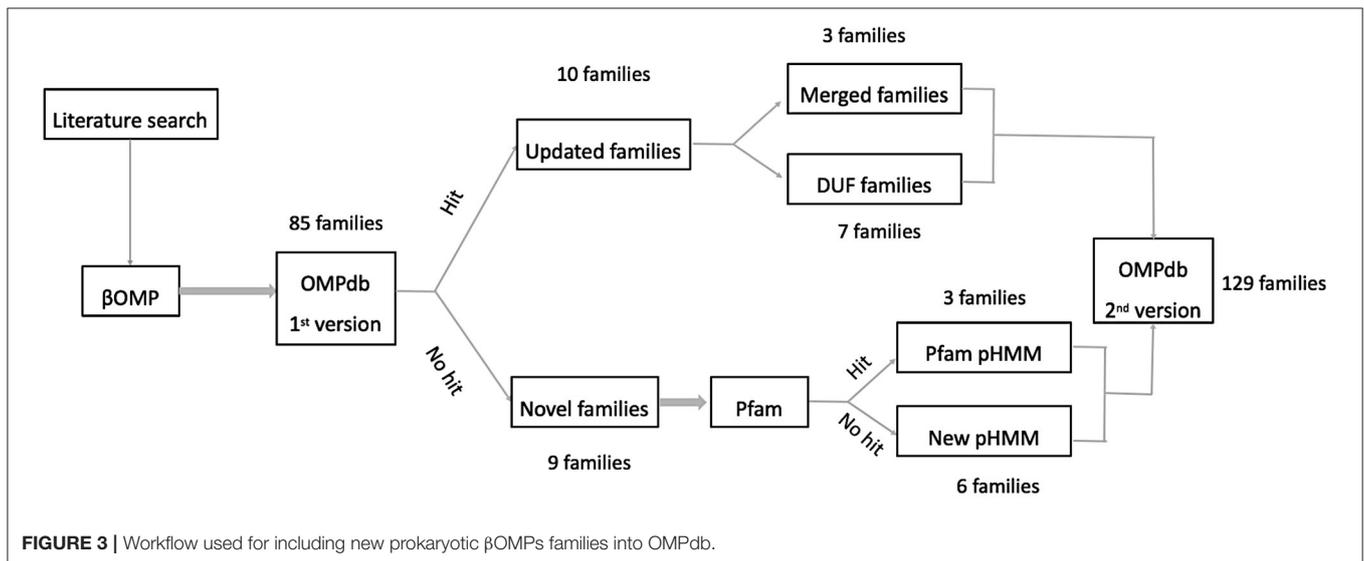
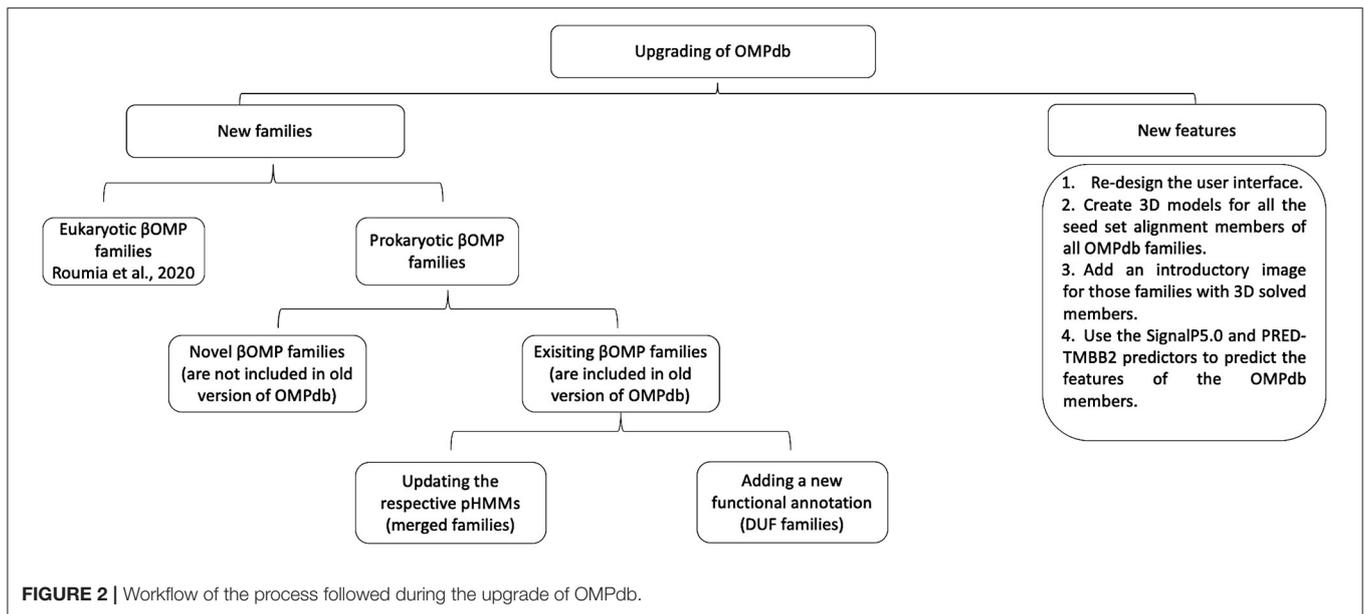
Newly Added Families

In addition to the 11 new eukaryotic transmembrane β OMPs families (Roumia et al., 2020), we included 19 new or updated prokaryotic β OMP families presented in **Table 1**. These families are classified into one of the following three categories: novel β OMP families not included in the old version of OMPdb, known families with new functional annotation or known families with an updated pHMM. Further, for most of the OMPdb families either updated or not, we added 667 recent publications focusing on their respective function.

Novel Families

This first category includes nine novel families presented below.

- (i) TP0969 protein shows structural homology with the TolC protein. TP0969 was predicted as an outer membrane beta barrel by CELLO (Yu et al., 2006), PSORT (Yu et al., 2010), HHomp (Remmert et al., 2009), and PRED-TMBB (Bagos et al., 2004) prediction algorithms and has a signal peptide based on Signal-CF (Chou and Shen, 2007) and SignalP 4.1 (Petersen et al., 2011) prediction algorithms. Furthermore, using mass spectrometry analysis, TP0969 was identified as an outer membrane protein (Cox et al., 2010; Osbak et al., 2016). According to PRED-TMBB2, it forms 12 transmembrane beta strands (Tsirigos et al., 2016).
- (ii) *Vibrio parahaemolyticus* is a significant Gram-negative halophilous pathogen. As a result of the consumption of seafood, it may cause a harmful seafood-borne illness in humans. In addition, it causes a wide range of diseases in aquatic animals (Vora et al., 2005; Kawatsu et al., 2006; Datta et al., 2008). VP1243 is a protective antigen which shows a high effective antimicrobial activity against the mentioned organism. VP1243 is widely distributed and highly conserved among the major *Vibrio* Species. VP1243 is considered as a hopeful candidate against the *Vibrio*



infections. Western blot analysis revealed that VP1243 is an outer membrane protein (Gao et al., 2020). The first 19 amino acids function as a signal peptide according to SignalP 5.0 (Almagro Armenteros et al., 2019). Further, VP1243 forms 10 transmembrane beta strands based on PRED-TMBB2 prediction (Tsirigos et al., 2016).

(iii) PgaA controls the translocation of de-N-acetylated poly- β -1,6-N-acetyl-d-glucosamine (dPNAG) polymer into the outer membrane of *Escherichia coli* k-12. dPNAG is crucial for biofilm adhesion and has an important role in the maintenance and development of biofilms integration for various bacterial species. The PgaA β -barrel domain (513–807) consists of 16 antiparallel transmembrane β -sheets with β 1 and β 16 strands interlocking (Wang et al., 2016).

(iv) The Gram-negative *Fibrobacteres-Chlorobi-Bacteroidetes* superphylum has a protein export pathway termed the type 9 secretion system (T9SS), which is considered as an important determinant of pathogenicity in serious periodontal disease (Lasica et al., 2017). These bacteroidetes are famous for the quick and unique gliding motility where their cell surface adhesins move on helical tracks (Shrivastava et al., 2016). Cryo-electron microscopy revealed that SprA is the translocon of T9SS. Furthermore, SprA consists of a large transmembrane β -barrel of 36 strands. The SprA barrel has a lateral entry to the external membrane surface because the barrel pore on the extracellular end is closed (Lauber et al., 2018).

In addition to the monomeric β OMPs, we also included the multimeric transmembrane β OMPs to the OMPdb

TABLE 1 | Newly added prokaryotic β -barrel families in OMPdb.

Family ^{Category}	Representative member (Uniprot or PDB Acc.) References	Function	Pfam domain
TP0969 ¹	R9UUH1 (Cox et al., 2010; Osbak et al., 2016)	Biogenesis/Secretion	No domains found
VP1243 ¹	Q87QA7 (Gao et al., 2020)	Active transporter	No domains found
PgaA ¹	4Y25 (Wang et al., 2016)	Biogenesis/Secretion	No domains found
SprA ¹	6H3I (Lauber et al., 2018)	Biogenesis/Secretion	PF14349
Hemolysin ¹	3O44 (De and Olson, 2011)	Non-specific diffusion channel	PF12563; PF07968
Aerolysin ¹	5JZH (Iacovache et al., 2016)	Non-specific diffusion channel	PF01117; PF03440; PF07968
Lysenin ¹	5EC5 (Podobnik et al., 2016)	Non-specific diffusion channel	No domains found
MAC ¹	6H04 (Menny et al., 2018)	Non-specific diffusion channel	PF01823
CsgG ¹	3X2R (Cao et al., 2014)	Biogenesis/Secretion	PF03783
NiiB ²	D3VIY6 (Bhasin et al., 2012)	Adhesion	PF04575
Cj0561c ²	Q0PAV5 (Guo et al., 2008)	Adhesion	PF11059
COG4313 ²	A5W3Z9 (Belchik et al., 2010; van den Berg et al., 2015)	Specific diffusion channel	PF13557
PorP_SprF ²	Q11YQ4 (Zhu and McBride, 2014)	Biogenesis/Secretion	PF11751
OMP_b-brl ²	2LHF (Edrington et al., 2011)	Adhesion	PF13505
TSA ²	P22940 (Lin et al., 2012; Trung et al., 2019)	Adhesion	PF03249
Oms38 ²	O51379 (Thein et al., 2012)	Non-specific diffusion channels	No domains found
Autotransporter ³	FabF (5O67) (Rouse et al., 2017) and Hap β (P45387) (Hendrixson et al., 1997; Meng et al., 2011)	Biogenesis/Secretion	PF03797
TBDTs ³	4ZGV (Grinter et al., 2016)	Active transporter	PF00593
Omp85 ³	POA940 (Kim et al., 2007), O32625 (Flack et al., 1995), Q51930 (Ruffolo and Adler, 1996) and Q9K1H0 (Volokhina et al., 2009)	Biogenesis/Secretion	PF01103

¹This category corresponds to the novel β -barrel protein families which were not included in the previous version of OMPdb.

²This category corresponds to families with a new functional annotation.

³This category corresponds to families for which we built new pHMMs.

families. There are five such families: Hemolysin, Aerolysin, Lysenin, MAC, and CsgG. The first three families act as PFTs (De and Olson, 2011; Iacovache et al., 2016; Podobnik et al., 2016). The MAC family protects the human cells against bacterial attack (Menny et al., 2018) and the CsgG family is involved in protein secretion (Cao et al., 2014).

PFTs attracted the interest of structural biologists and microbiologists early on due to their ability to act as soluble proteins and transmembrane complexes. PFTs are secreted as water soluble monomers, and later on they connect to the membrane of target cells. There they gather into circular oligomers and enable the insertion into the membrane causing the pore forming and finally the resultant cell death (Anderluh and Lakey, 2008).

- (v) Hemolysin, a heptameric porin, is secreted by *Vibrio cholerae*. The heptamer comprises a ring-like structure with an outer diameter of 135. The channel pore consists of an upper large vestibule formed by β -prism and cytolysin domains, and the stem domains which form a 14-strand β -barrel (De and Olson, 2011).
- (vi) Aerolysin is a heptamer where seven monomers form a β -barrel with 14 strands at the outer membrane of *Aeromonas hydrophila*. The aerolysin pore is more stable due its concentric β -barrel fold (Iacovache et al., 2016).
- (vii) Lysenin is a PFT of *Eisena fetida* and a member of the aerolysin family. It is a non-amer and forms an 18-strand smooth-walled tubular β -barrel attached to the outside by the C-terminal domains. It is a mushroom-like structure which has a transmembrane pore with a central stem built of a long β -barrel. Having a uniform and stable β -barrel through the entire length of protein makes lysenin an exception among PFTs (Podobnik et al., 2016).
- (viii) The MAC is a structure formed at the outer membrane of pathogen cells resulting from activation of the host's complement system. It is one of the immune system's first responders. MAC forms damaging transmembrane channels at the cell membrane of pathogens causing cell lysis and death (Charles A Janeway et al., 2001). It is a multiprotein complex, comprised of six polypeptide chains (C5b, C6, C7, C8 α , C8 β , and C8 γ) together with multiple copies of C9 monomer that are arranged in a split-washer configuration, composing a ring in the membrane which permits free flux of molecules in and out of the cell. The cell dies when enough pores are formed (Sharp et al., 2016). Membrane Attack Complex-Perforin (MACPF) domain (PF01823) is a common domain in all C6-C9 (Tschopp et al., 1986). MAC assembly into the membrane starts when C7 connects to C5b6 (a complex formed by C5b and C6) to compose MAC precursor C5b7 (DiScipio et al., 1988). Subsequently, C8 connects irreversibly to the former complex forming the membrane-inserted C5b8 (Steckel et al., 1983). Then, C5b8 binds to 22 C9 monomers to form C5b9 and polymerizes to finish the MAC pore formation (Podack et al., 1982). In the end, the β -barrel pore is composed after transforming the helical bundles in the MACPF domains into transmembrane β -hairpins through

an unknown mechanism (Shepard et al., 1998; Shatursky et al., 1999; Rosado et al., 2007). According to the cryo-EM structure, the formed barrel consists of 88 strands (Dudkina et al., 2016).

- (ix) In *Escherichia coli* Xuzhou21, curli subunits are produced and secreted into the outer membrane through the CsgG secretion channel. Curli are a unique group of functional amyloids, which are crucial for host cell adhesion, biofilm formation and colonization of inert surfaces (Barnhart and Chapman, 2006). They are involved in harmful diseases in humans, since they share similar structural and biochemical characteristics with amyloid fibers (Sunde et al., 1997; Moreno-Gonzalez and Soto, 2011). The crystal structure of CsgG showed that it is a symmetric nonameric channel, composed of monomers each having four strands spanning the outer membrane. A 36-stranded β -barrel is formed from nine CsgG monomers. CsgG could perhaps reduce the biofilm formation by controlling Curli secretion and that is why it is studied as a putative antibiotics target (Cao et al., 2014).

Families With Updated Information

From the literature search, we also were able to assign functional annotation to seven families that had unknown function until now.

- (i) NilB (DUF560) is a surface-exposed outer membrane protein of *Xenorhabdus nematophila*. Expression of NilB is suppressed by NilR and growth in nutrient-rich medium. Members of this family exist in diverse bacteria and are common in the genomes of mucosal pathogens. Bioinformatic analyses reported that NilB is the only characterized member of a family of proteins distinguished by a conserved C-terminal domain of unknown function (DUF560) and N-terminal region tetratricopeptide repeats (TPR). Insertion and deletion mutational analyses revealed that NilB forms a β -barrel with 14 transmembrane strands and seven extracellular surface loops, and an N-terminal globular domain. The globular domain and surface loop 6 play a crucial role in the nematode colonization. Epifluorescence microscopy of these mutants revealed that NilB is necessary at early stages of colonization (Bhasin et al., 2012).
- (ii) Cj0561c (DUF2860) is a probable membrane fusion protein and contributes to intestinal colonization. Cj0561c is the only characterized member of DUF2860. Cj0561c is encoded as a membrane transporter gene whose transcription is inhibited by CmeR in *Campylobacter jejuni* (Guo et al., 2008). Based on PRED-TMBB2, Cj0561c is predicted to have 14 beta strands (Tsirigos et al., 2016).
- (iii) COG4313 (Phenol_MetA_deg) proteins comprise a huge and widespread family of outer membrane channels and have been involved in the uptake of a set of hydrophobic molecules (van den Berg et al., 2015). The first studied member of COG4313 proteins was the TcpY, an outer membrane protein that eases the uptake of polychlorophenols across the outer membrane of *Cupriavidus necator* JMP134 (Belchik et al., 2010). Pput2725 is another member of the family and forms a 12-stranded barrel at the outer membrane of *Pseudomonas putida* with experimentally verified 3D structure. It is suggested that Pput2725 mediates the uptake of hydrophobic aromatic compounds (van den Berg et al., 2015).
- (iv) SprP (PorP_SprF) is associated with the outer membrane of *Cytophaga hutchinsonii*. It is suggested that SprP is involved in protein secretion, a hypothesis made due to its sequence similarity with PorP that is responsible for the secretion of gingipain protease virulence factors in *Porphyromonas gingivalis* type IX secretion system (T9SS) and SprF that is crucial for the delivery of the gliding motility machinery components to the cell surface (Zhu and McBride, 2014). According to PRED-TMBB2, it is predicted to have 14 transmembrane strands (Tsirigos et al., 2016).
- (v) The Outer membrane protein beta-barrel domain (OMP_b-*brl*) domain is found in a wide range of outer membrane proteins, found mainly in Alphaproteobacteria. The Major Outer Membrane Protein (ompP1) is found in various strains of *Coxiella burnetti* with typical porin properties. It is heat-modifiable and has a channel activity. ompP1 is predicted to form a β -barrel; Omp3b from *Brucella abortus* and homologs from various Alphaproteobacteria that are believed to be involved in bacterial surface control and host cell interactions. In *Pseudomonas aeruginosa*, OprH forms a transmembrane β -barrel protein which consisted of eight strands with four extracellular loops of unequal size. Moreover, *in vivo* and *in vitro* biochemical studies revealed that OprH interacts with LPS in *P. aeruginosa* outer membranes (Varghees et al., 2002; Edrington et al., 2011).
- (vi) The Type Specific Antigen (TSA) consists of several antigenic variants in the genera *Rickettsia* and *Orientia* that belong to the family *Rickettsiaceae*. *Rickettsiaceae*, intracellular bacteria, are the causative agent of the Rickettsioses which are an arthropod-borne zoonoses. The TSA of 56-kilodaltons located on the rickettsial surface is responsible for the variation. *Orientia tsutsugamushi* are the causative agent of scrub typhus, is an obligate intracellular pathogen. In *O. tsutsugamushi*, there are three major outer membrane proteins: TSA56, TSA47, and TSA22. The TSA56 is the major outer membrane protein responsible for *O. tsutsugamushi* adhesion (Lin et al., 2012; Trung et al., 2019).
- (vii) Relapsing fever is a worldwide, endemic disease caused by several spirochetal species belonging to the genus *Borrelia*. Oms38 (outer membrane-spanning protein of 38 kDa) is present in the outer membranes of *B. duttonii*, *B. hermsii*, *B. recurrentis*, and *B. turicatae* as well. Characterization of Oms38 was performed using the black lipid bilayer method, which demonstrated that Oms38 forms small, water-filled channels of 80 pS in 1 M KCl that did not exhibit voltage-dependent closure. The Oms38 channel is slightly selective for anions and shows a ratio of permeability for cations over anions of 0.41 in KCl. Analysis of the deduced amino acid sequences revealed that Oms38 contains an N-terminal signal sequence which is processed under *in vivo* conditions.

Oms38 is highly conserved within the four studied relapsing fever species, sharing an overall amino acid identity of 58% and with a strong indication for the presence of a β -barrel formation (Thein et al., 2008). DipA (BB_0418) forms a specific porin for dicarboxylates. It is suggested that DipA plays a crucial role for the flux of specific nutrients toward different *Borrelia* species (Thein et al., 2012).

Families With Updated pHMMs

During the update of OMPdb, we identified proteins with the same biological function as two pre-existing beta barrel families, namely the Autotransporter (PF03797) and the TBDTs (PF00593) families. Interestingly, the Pfam profiles were unable to detect these proteins, so we updated the respective profiles.

For the Autotransporter family, two new members were added in the seed set, the functional amyloid transporter from *Pseudomonas*, FabF (PDB: 5O67) (Rouse et al., 2017), and the *Haemophilus influenzae* Hap β -barrel domain, Hap β (Uniprot: P45387) (Hendrixson et al., 1997; Meng et al., 2011), and the pHMM was rebuilt. FapF is a new class of secretion system similar to type V secretion in the autotransporter proteins (AT-1). FapF is characterized by a C-terminal, 12 β -stranded, helix-blocked pore in the closed state. In contrast to other autotransporters which completely pass the pore, the N terminus of the truncated FapF structure exits the barrel on the periplasmic side (Rouse et al., 2017). Hap β is the β -barrel domain of self-associating autotransporters (SAATs). SAATs is a set of virulence factors that enhance bacterial biofilm formation (Klemm et al., 2006). The *Haemophilus influenzae* Hap SAAT consists of three domains: residues 1–25 correspond to the signal peptide, residues 26–1036 refer to the Hap β s called the passenger domain, and residues 1037–1394 form the 14 β -stranded barrel at the C-terminal (Hendrixson et al., 1997; Meng et al., 2011).

TBDTs pHMM was updated by adding the respective members of FusA in the seed set. FusA is used by the phytopathogenic *Pectobacterium* spp. as a new class of TBDTs. FusA is a 22-stranded transmembrane β -barrel responsible for importing pectocin and ferredoxin in *Pectobacterium* spp. (Grinter et al., 2016).

Finally, we created a new pHMM for Bama/TamA and SAM50 families. Previously, Bama/TamA and SAM50, along with TOC75 and OEP80 constituted the Omp85 family (PF01103). Since we were able to create two separated chloroplastic pHMMs for TOC75 and OEP80 (Roumia et al., 2020), we decided to build the new pHMM and now we have the ability to easily distinguish members of the three eukaryotic families (TOC75, OEP80, and SAM50) when searching against eukaryotic proteomes.

Figure 4 summarizes the Pfam annotation of the 129 families in OMPdb. 72.9% of the families belong to the outer membrane beta-barrel protein superfamily (MBB clan), while 10.8% of the families are in Pfam but not in MBB clan. Interestingly enough, for 16.3% of OMPdb families, there was no information at all in the Pfam database.

Search Using the OMPdb pHMMs' Collection Against Reference Proteomes

In **Figure 5**, we display the detected proteins by the two pHMM libraries, eukaryotic and prokaryotic, against the retrieved reference proteomes from Uniprot (Release 2020_03). The eukaryotic pHMM library detected at least one protein in 96.97% of the eukaryotic reference proteomes and the number of detected proteins for each family is depicted in **Figure 5A**. As shown in **Figure 5B**, the detected prokaryotic proteins were distributed functionally into eight categories. Most of them belong to the receptor category followed by the biogenesis/secretion category, while the enzymes category shows the lowest number of detected proteins.

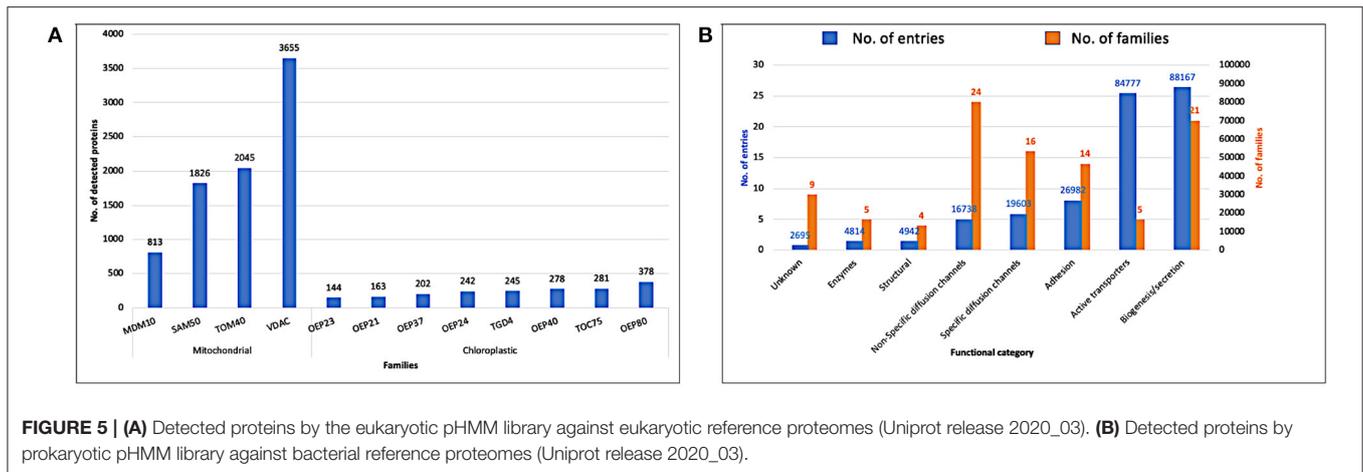
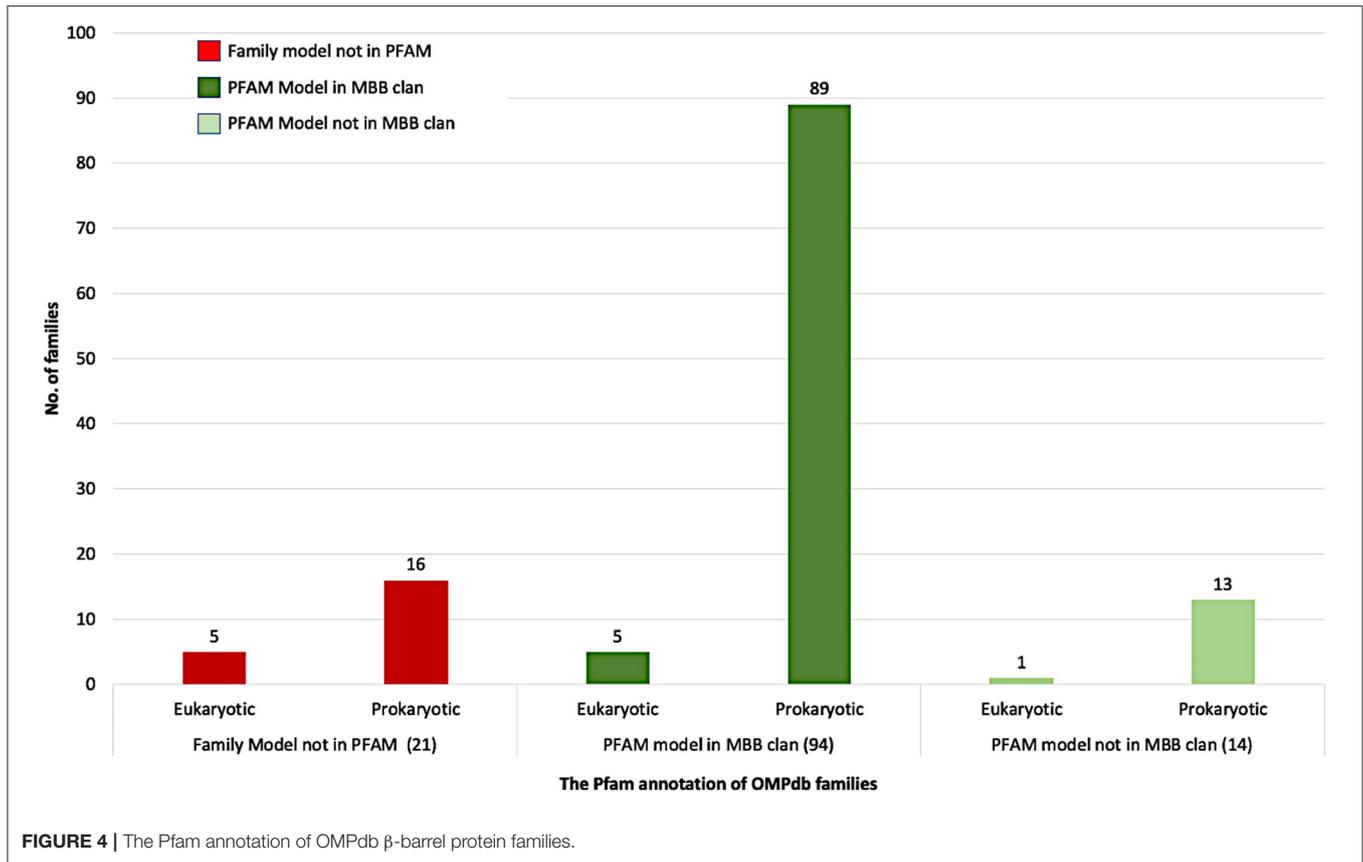
As a next step, we followed the taxonomy to classify the bacterial reference proteomes according to the NCBI Common Taxonomy Tree (Sayers et al., 2009). The bacterial reference proteomes are classified into 16 main taxonomic groups, three groups from Gram-positive bacteria, *Actinobacteria*, *Firmicutes* and *other bacteria*, and the rest 13 groups from Gram-negative bacteria. As shown in **Figure 6**, Gram-negative bacteria are classified into six main groups: *Cytophaga*, *Fusobacterium*, and *Bacteroides* group (CFB group), *Proteobacteria*, *Cyanobacteria*, *Mycoplasmata*, *Negativecutetes*, and *other bacteria*. All taxonomic groups that account for <1% of the total reference proteomes in both classes, Gram-positive or Gram-negative bacteria, are classified under “other bacteria.”

CFB group has the highest average of detected proteins followed by the *Proteobacteria* group of Gram-negative bacteria. In line with what we expected, *Actinobacteria* and *Firmicutes*, the two main Gram-positive bacteria taxonomic groups, have few representatives of the above-mentioned functional categories. Nevertheless, it is clear that the *Firmicutes* group has representatives relatively more than the *Actinobacteria* group as shown in **Figure 6**.

Figure 7 shows in detail the average of detected functional proteins for all bacterial taxonomic groups either Gram-negative or Gram-positive ones. Except for the *Mycoplasmata* group, all Gram-negative bacteria taxonomic groups have a significant number of representatives of the active transporters, biogenesis/secretion, specific diffusion channels, and unknown function categories. For the *Cyanobacteria* group, the non-specific diffusion channels category is almost absent as shown in **Figure 7**.

New Features

Beside re-designing the user interface, we now provide structural information through 3D models for all seed set members of each family and this feature gives deep and concise information about the respective family. In total, we provide 3,264 models for families with at least one member with experimentally verified crystal structure and 2,116 models for families without any previous knowledge about their 3D structure. In addition, we provide an introductory structure image structure at the main page of the respective family, for every family including members with experimentally verified 3D. Moreover, for all the OMPdb



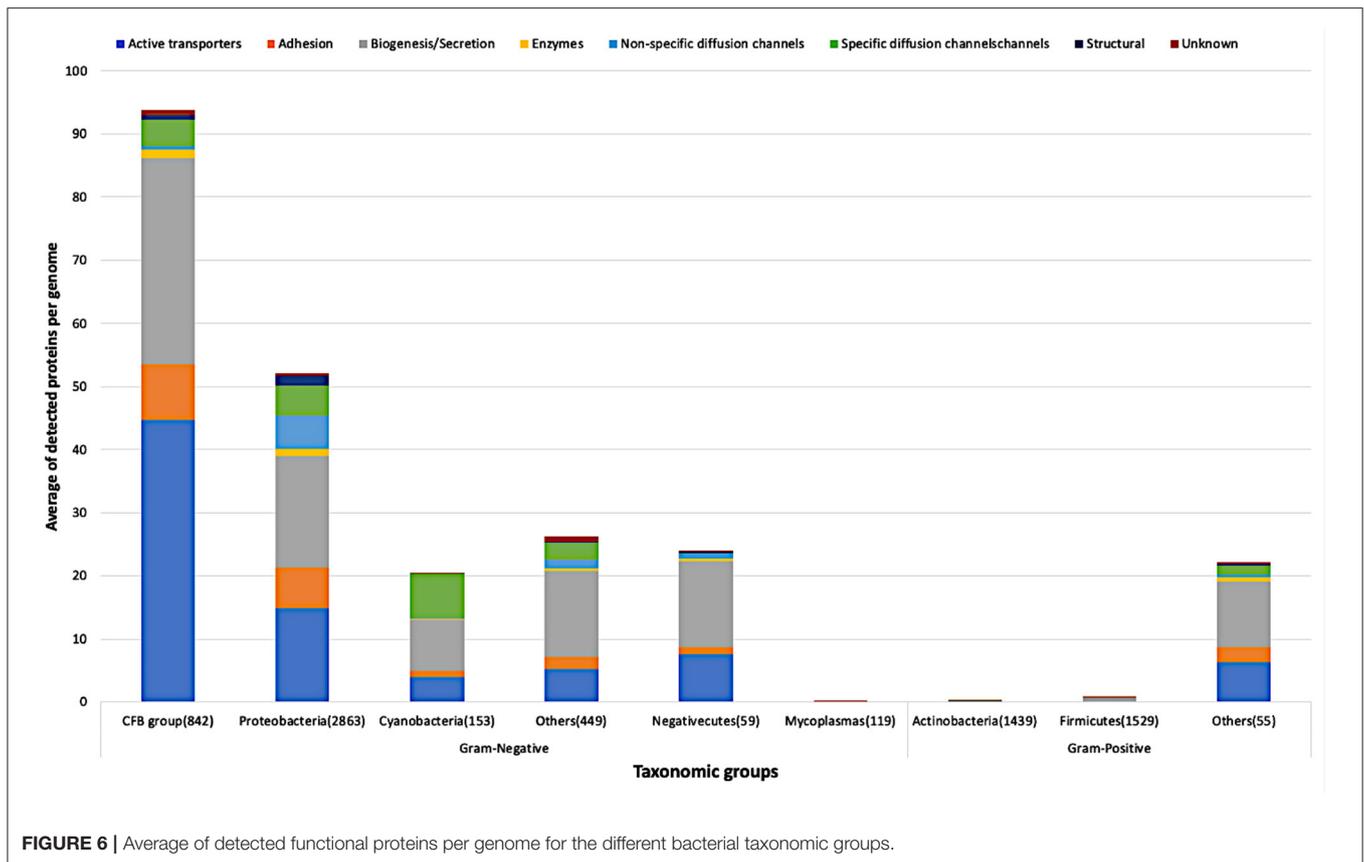
entries, we present the predicted features of the signal peptide and β -barrel topology using SignalP5 and PRED-TMBB2 predictors.

DISCUSSION

In this work, we presented the new content and updated features of OMPdb, a global hub for all β OMPs using 128 characteristic pHMMs. This collection of pHMMs includes 11 eukaryotic

and 117 prokaryotic families in addition to the *Treponema pectinovorum* ATCC 33768 mompA family (Walker et al., 1997) that lacks a pHMM due to missing known homologues of its representative member.

OMPdb in its second version, still has a clear advantage against all the other databases that contain β OMPs. As shown in **Table 2**, OMPdb is currently the most complete resource for transmembrane β -barrels, since it has the highest number of protein and family entries, provides numerous literature



references, offers sequence annotation, cross-references with many related databases and the ability to use search and prediction tools. Due to the discontinuation of TMBETA GENOME (Gromiha et al., 2007), PRNDS (Katta et al., 2004), HHompDB (Remmert et al., 2009), and TMPDB (Ikeda et al., 2003) databases since the first version of OMPdb, we excluded them from the comparison. In contrast, we included in the comparison two new databases, Mpstruc (White, 2009) and MemprotMD (Newport et al., 2019).

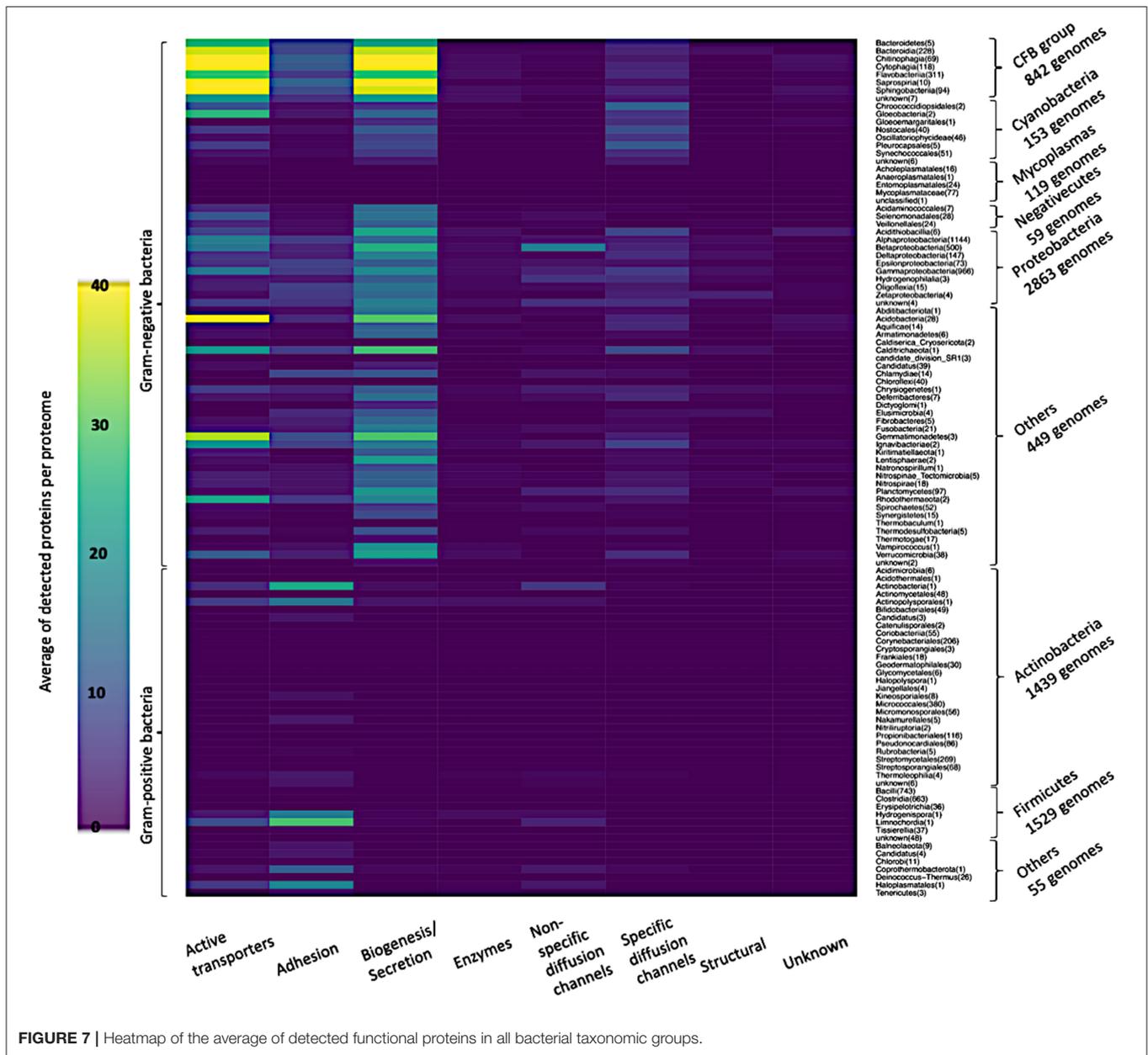
Based on whether a family exists in Pfam, and specifically if it is a member of the MBB clan, the OMPdb families can be classified into three categories as shown in **Figure 4**. The first one includes 94 families belonging to the MBB clan in Pfam. In this category, we updated the pHMMs of MDM10 (PF12519), TBDTs (PF00593) and Autotransporter (PF03797), while the Porin_3 (PF01459) pHMM was replaced with two new pHMMs, VDAC, and TOM40 (Roumia et al., 2020). Furthermore, the pHMM of Omp85 (PF01103) was updated to contain only BamA/TamA and SAM50 proteins due to having two specific pHMMs for TOC75 and OEP80 families (Roumia et al., 2020). The second category refers to 14 families which are in Pfam but not in the MBB clan. Although Hemolysin and SprA families had Pfam pHMMs, we updated their respective pHMMs. For Hemolysin, the respective pHMM (PF12563) did not cover all the transmembrane barrel region, while, in the SprA family (PF14349), the transmembrane barrel part was intermittent. Therefore, we built two novel

pHMMs for Hemolysin and SprA families. The third category (families not included in Pfam), consists of 5 eukaryotic (Roumia et al., 2020) and 16 prokaryotic families. Interestingly, for these 21 families, there is no information in Pfam whatsoever and these novel families can only be found in OMPdb.

In OMPdb we focus on offering as much information about the nature of the transmembrane barrel part in each family to the user as possible. To this end, we re-designed the OMPdb user interface to include various new features. The first one is adding an introductory structure image for 44% of the OMPdb families that presents to the user the β -barrel features of these families at one glance.

The 3D structure dictates the biological function of every protein. Knowing the 3D structure of a protein provides a higher level of understanding of its molecular action. This knowledge allows us to create hypotheses on controlling, modifying or affecting on this protein. We now provide a 3D model for each of the seed set members of every OMPdb family. For the families including members with experimentally verified 3D structure (57 families), we constructed the models using homology modeling via the SWISS-MODEL server (Biasini et al., 2014). We provide the used template as the representative structure for each family. We used *de novo* modeling through PHYRE2 (Kelley et al., 2015) for the rest of the families (71 families).

The nascent β OMPs either eukaryotic (Rapoport, 2007) or prokaryotic (Nouwen et al., 2007) have signal peptides to be



recognized by the Sec or Tat pathways. After or during the assembly of the transmembrane proteins, the signal peptides are being cleaved by signal peptidase (SPase). The Tat substrates can be cleaved by SPase I or II while the Sec substrates can be processed by SPase I, II, or III. SignalP5 is the first predictor that provides a deep neural network-based method for predicting the signal peptide cleaved by (Sec/SPI), (Sec/SPII), and (Tat/SPI) (Almagro Armenteros et al., 2019). PRED-TMBB2 is a predictor for the β -barrel topology based on Hidden Markov Models. It excels the best available β -barrel predictor by 7% (Tsirigos et al., 2016) (Tsirigos et al., 2016). Both predictors are therefore used to provide the topological features of each protein.

For the detection of β OMPs in the reference proteomes, we used the collection of OMPdb's pHMMs after excluding all

the DUF families (19 families) that do not have any literature information regarding their function (Supplementary Table 1). It is worth noting that the updated Omp85 pHMM was used twice, in both the eukaryotic library to detect the SAM50 members and the prokaryotic library to detect the Bama members.

Very few proteins were detected in the *Mycoplasmas* group, although they are Gram-negative bacteria, and the reason is the absence of a cell wall in this taxonomic group (Cimolai, 2001) (Figures 6, 7). On the other hand, the *Firmicutes* group, a group of Gram-positive bacteria, shows a significant number of detected proteins. *Firmicutes* is a bacterial phylum, in which most of its organisms are Gram-positive bacteria. However, *Firmicutes* contains a class called *Negativevecutes*,

TABLE 2 | A comparison between OMPdb and other databases that include β -barrel proteins.

Database Features	OMPdb (current work)	Pfam V. 33.1 (MBB clan) (El-Gebali et al., 2019)	TCDB (Saier et al., 2016)	PDBTM (Kozma et al., 2013)	Mpstruc (White, 2009)	MemprotMD (Newport et al., 2019)	TOPDB (Dobson et al., 2015)	ePSORTdb (Peabody et al., 2016)	OPM (Lomize et al., 2012)	MPdb (Raman et al., 2006)
Total entries	1198558	1092329	1260	130	283	405	201	839	291	153
Total families	129	92	93	-	17	-	-	-	34	29
Data retrieval	Semi-automated	Semi-automated	Manual	Semi-automated	Automated	Automated	Semi-automated	Manual	Semi-automated	Manual
Literature references	+	+	+	-	+	+	+	+	-	+
Signal peptide/TM segments annotation	+/+	-/-	-/-	±	-	-	+/+	-/-	±	-/-
Number of databases cross-referenced	13	5	14	2	3	6	3	2	4	2
Prediction/search tools	BLAST/ HMMER/ PRED-TMBB2/ Text search	Text search/ HMMER	MEMSAT/ HHrepID/ AveHAS/ HMMTop/Text search	Text search	Text search/Fuzzy search	Text search	Text search	Psortb/BLAST	Text search	Text search

which is classified as Gram-negative bacteria (Sutcliffe, 2010). Interestingly enough, around 51.67% of the detected proteins at the *Firmicutes* group (1588 proteomes) belong to the 59 *Negativecutes* proteomes.

Our analysis showed that the biogenesis/secretion category was the largest one in comparison with the other functional categories. The Outer membrane Factor family (OMF) comprises 40.5% of all detected proteins. These proteins are involved in the type I secretion pathway (Sec-independent), which is widespread and permits the secretion of proteins of different sizes and functions using uncleaved secretion signal at the C-terminal (Delepelaire, 2004). Around 9% of all detected proteins belong to each of the following families: Secretin, Autotransporter (AT), *Porphyromonas gingivalis* PorT and Outer Membrane Protein Insertion Porin (OmpIP/Omp85). The members of these four families are involved in type II and type III secretion pathway (Collins et al., 2004), type V secretory pathway (Desvaux et al., 2004), secretion of gingipains from *Porphyromonas gingivalis* (Nguyen et al., 2009) and the translocation of β OMPs into the outer membrane of Gram-negative bacteria (Noinaj et al., 2015), respectively. As for the active transporters category, it consists of five families (Supplementary Figure 2). It should be noted that the TBDTs family detects around 99.83% of all active transporter proteins. The aforementioned family represents the TBDTs transport system, which is involved in the active transport mechanism of nutrients through the outer membrane of Gram-negative bacteria (Higgs et al., 2002). It is noteworthy that the existence of active transporters, biogenesis/secretion, specific diffusion channels, and non-specific diffusion channels categories in all taxonomic groups of Gram-negative bacteria are in line with the findings of Nikaido (Nikaido, 2003), who revealed that the general part of β OMPs is acting as channels and transporters. The only inconsistency is the absence of non-specific diffusion channels in the *Cyanobacteria* group. The adhesion to the host cell is the first step bacteria take to establish a successful infection. The bacterial adhesion may occur directly via monomeric adhesins (Soto and Hultgren, 1999) or via highly advanced machines like the type III secretion systems (Autotransporter) (Pizarro-Cerdá and Cossart, 2006), which belongs to the biogenesis/secretion category, or by using special nanofibers (Corpe, 1972). Although there are multiple ways of bacterial adhesion (e.g., autotransporter and special nanofibers), there are 14 families for which their main function is adhesion. The OMP_b-brl forms around 72.74% of the detected adhesion proteins. The representative member of the aforementioned family, OprH, is involved in the stability of the outer membranes of Gram-negative bacteria by interacting with the LPS (Edrington et al., 2011) (Figure 7 and Supplementary Figure 3).

The widespread existence of specific and non-specific diffusion channels in all Gram-negative bacteria reflects the effective role of porins on the bacterial cell life. In addition to acting as passages for the nutrients into the Gram-negative bacteria, porins are affecting the bacterial pathogenicity due to their high prevalence at the bacterial surface structure (Naumann et al., 1999; Achouak et al., 2006; Galdiero et al., 2013; Choi and Lee, 2019) (Supplementary Figures 4, 5). In *Cyanobacteria*, the absence of the non-specific diffusion channels

category reflects the low permeability of the cyanobacterial outer membrane (Figures 6, 7). The reason why the cyanobacterial outer membranes are highly impermeable is to prevent the flux of toxic compounds into the cell and this is an advantage for the cyanobacterial autotrophic lifestyle (Kowata et al., 2017). In the structural category (Supplementary Figure 6), the OmpA family counts around 71.35% of the proteins. This high percentage gives consideration to the critical effect of this domain on the structural strength and shape of bacterial cells (Höltje, 1998), due to the non-covalent interaction between the respective domain with the peptidoglycan layer (Grizot and Buchanan, 2004).

As expected, the enzyme category has the lowest number of detected proteins, because only a limited number of β OMP act as enzymes (Bishop, 2008). 44.69% of the enzyme proteins belong in the Outer membrane-localized lipid A 3-O-deacylase (PagL) family that is widespread among Gram-negative bacteria (Rutten et al., 2006) (Figure 7, Supplementary Figure 7).

Due to the significant progress that has been made toward the annotation of Gram-negative bacteria genomes, the “unknown” category comprises only around 1.08% of our analysis (Figure 7, Supplementary Figure 8). Around 22% of Pfam proteins belong to DUF families (El-Gebali et al., 2019). Generally, there is an enormous number of proteins with unknown function and the reason is the difficulty of assigning function to proteins, which is a major goal of structural biology, biochemistry and bioinformatics (McKay et al., 2015).

CONCLUSIONS

The OMPdb database is a thorough, up-to date and continuously updated β OMP database. OMPdb, in its second version, contains two different libraries of pHMMs that cover all the kinds of β OMPs. As of September 2020, it has around 1.2 million entries whose domains are classified into 129 different β -barrel families either prokaryotic or eukaryotic ones. OMPdb operates for more than 10 years and will be continuously updated in the future. The database participates in ELIXIR-GR, the Greek National Node of the ESFRI European RI ELIXIR, a distributed infrastructure that will allow the life science research community across Europe

REFERENCES

- Achouak, W., Heulin, T., and Pagès, J.-M. (2006). Multiple facets of bacterial porins. *FEMS Microbiol. Lett.* 199, 1–7. doi: 10.1111/j.1574-6968.2001.tb10642.x
- Almagro Armenteros, J. J., Tsirigos, K. D., Sønderby, C. K., Petersen, T. N., Winther, O., Brunak, S., et al. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* 37, 420–423. doi: 10.1038/s41587-019-0036-z
- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402. doi: 10.1093/nar/25.17.3389
- Anderluh, G., and Lakey, J. H. (2008). Disparate proteins use similar architectures to damage membranes. *Trends Biochem. Sci.* 33, 482–490. doi: 10.1016/j.tibs.2008.07.004
- Babbitt, P. C., Bagos, P. G., Bairoch, A., Bateman, A., Chatonnet, A., Chen, M. J., et al. (2015). Creating a specialist protein resource network: a meeting report

to share and store their research data as part of an organized network, whereas, at the same time is in close collaboration with other specialized protein databases (Babbitt et al., 2015; Holliday et al., 2015).

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

PB conceived the presented idea. PB, KT, and MT conceived and planned the experiments. AR carried out the experiments and wrote the manuscript with support from MT and KT. All authors provided critical feedback and helped shape the research, analysis, and manuscript.

FUNDING

AR was financially supported as a PhD scholar from a joint program between the ministry of higher education in Egypt and the Greek ministry of foreign affairs (2016/2017). PB, IT, and MT were partially supported by the project ELIXIR-GR: The Greek Research Infrastructure for Data Management and Analysis in Life Sciences (MIS 5002780), which is implemented under the Action Reinforcement of the Research and Innovation Infrastructure funded by the Operational Programme Competitiveness, Entrepreneurship and Innovation (NSRF 2014–2020) and cofinanced by Greece and the European Union (European Regional Development Fund).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fbinf.2021.646581/full#supplementary-material>

- for the protein bioinformatics and community resources retreat. *Database* 2015:bav063. doi: 10.1093/database/bav063
- Bagos, P. G., Liakopoulos, T. D., Spyropoulos, I. C., and Hamodrakas, S. J. (2004). PRED-TMBB: a web server for predicting the topology of β -barrel outer membrane proteins. *Nucleic Acids Res.* 32, W400–W404. doi: 10.1093/nar/gkh417
- Barnhart, M. M., and Chapman, M. R. (2006). Curli biogenesis and function. *Annu. Rev. Microbiol.* 60, 131–147. doi: 10.1146/annurev.micro.60.080805.142106
- Bayrhuber, M., Meins, T., Habeck, M., Becker, S., Giller, K., Villinger, S., et al. (2008). Structure of the human voltage-dependent anion channel. *Proc. Natl. Acad. Sci. U.S.A.* 105, 15370–15375. doi: 10.1073/pnas.0808115105
- Belchik, S. M., Schaeffer, S. M., Hasenoehrl, S., and Xun, L. (2010). A beta-barrel outer membrane protein facilitates cellular uptake of polychlorophenols in *Cupriavidus necator*. *Biodegradation* 21, 431–439. doi: 10.1007/s10532-009-9313-8
- Bhasin, A., Chaston, J. M., and Goodrich-Blair, H. (2012). Mutational analyses reveal overall topology and functional regions of NilB, a bacterial outer

- membrane protein required for host association in a model of animal-microbe mutualism. *J. Bacteriol.* 194, 1763–1776. doi: 10.1128/JB.06711-11
- Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., et al. (2014). SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* 42, W252–W258. doi: 10.1093/nar/gku340
- Bishop, R. E. (2008). Structural biology of membrane-intrinsic β -barrel enzymes: Sentinels of the bacterial outer membrane. *Biochim. Biophys. Acta - Biomembr.* 1778, 1881–1896. doi: 10.1016/j.bbmem.2007.07.021
- Braun, V. (1995). Energy-coupled transport and signal transduction through the Gram-negative outer membrane via TonB-ExbB-ExbD-dependent receptor proteins. *FEMS Microbiol. Rev.* 16, 295–307. doi: 10.1111/j.1574-6976.1995.tb00177.x
- Cao, B., Zhao, Y., Kou, Y., Ni, D., Zhang, X. C., and Huang, Y. (2014). Structure of the nonameric bacterial amyloid secretion channel. *Proc. Natl. Acad. Sci. U.S.A.* 111, E5439–E5444. doi: 10.1073/pnas.1411942111
- Casadio, R., Faiselli, P., Finocchiaro, G., and Martelli, P. L. (2003). Fishing new proteins in the twilight zone of genomes: the test case of outer membrane proteins in *Escherichia coli* K12, *Escherichia coli* O157:H7, and other Gram-negative bacteria. *Protein Sci.* 12, 1158–1168. doi: 10.1110/ps.0223603
- Cavalier-Smith, T. (2000). Membrane heredity and early chloroplast evolution. *Trends Plant Sci.* 5, 174–182. doi: 10.1016/S1360-1385(00)01598-3
- Charles A Janeway, J., Travers, P., Walport, M., and Shlomchik, M. J. (2001). “The complement system and innate immunity,” in *Immunobiology Interactive*, ed. S. Gibbs (New York, NY: Garland Science), 59–82. Available online at: <https://lib.ub.edu/book/1022278/250ad7?regionChanged=&redirect=29288319> (accessed February 25, 2020).
- Choi, U., and Lee, C.-R. (2019). Distinct roles of outer membrane porins in antibiotic resistance and membrane integrity in *Escherichia coli*. *Front. Microbiol.* 10:953. doi: 10.3389/fmicb.2019.00953
- Chou, K.-C., and Shen, H.-B. (2007). Signal-CF: a subsite-coupled and window-fusing approach for predicting signal peptides. *Biochem. Biophys. Res. Commun.* 357, 633–640. doi: 10.1016/j.bbrc.2007.03.162
- Cimolai, N. (2001). “Mycoplasmas,” in *Laboratory Diagnosis of Bacterial Infections*, ed N. Cimolai (Boca Raton, FL: CRC Press), 861–892. doi: 10.1201/b14028
- Collins, R. F., Frye, S. A., Kitmitto, A., Ford, R. C., Tønnum, T., and Derrick, J. P. (2004). Structure of the *Neisseria meningitidis* outer membrane PilQ secretin complex at 12 Å resolution. *J. Biol. Chem.* 279, 39750–39756. doi: 10.1074/jbc.M405971200
- Corpe, W. A. (1972). “Microfouling: the role of primary film forming marine bacteria,” in *ci.nii.ac.jp*, 598–609. Available online at: <https://ci.nii.ac.jp/naid/10018002926/> (accessed February 12, 2020).
- Coulton, J. W., Mason, P., Cameron, D. R., Carmel, G., Jean, R., and Rode, H. N. (1986). Protein fusions of β -galactosidase to the ferrichrome-iron receptor of *Escherichia coli* K-12. *J. Bacteriol.* 165, 181–192. doi: 10.1128/JB.165.1.181-192.1986
- Cox, D. L., Luthra, A., Dunham-Ems, S., Desrosiers, D. C., Salazar, J. C., Caimano, M. J., et al. (2010). Surface immunolabeling and consensus computational framework to identify candidate rare outer membrane proteins of *Treponema pallidum*. *Infect. Immun.* 78, 5178–5194. doi: 10.1128/IAI.00834-10
- Dankert, J., Hogt, A. H., and Feijen, J. (1986). Biomedical polymers: bacterial adhesion, colonization and infection. *ci.nii.ac.jp* 2, 219–301. Available online at: <https://ci.nii.ac.jp/naid/10018453825/> (accessed February 12, 2020).
- Datta, S., Janes, M. E., and Simonson, J. G. (2008). Immunomagnetic separation and coagglutination of *Vibrio parahaemolyticus* with anti-flagellar protein monoclonal antibody. *Clin. Vaccine Immunol.* 15, 1541–1546. doi: 10.1128/CVI.00141-08
- Day, P. M., Inoue, K., and Theg, S. M. (2019). Chloroplast outer membrane β -barrel proteins use components of the general import apparatus. *Plant Cell* 31, 1845–1855. doi: 10.1105/tpc.19.00001
- De Pinto, V., Messina, A., Lane, D. J. R., and Lawen, A. (2010). Voltage-dependent anion-selective channel (VDAC) in the plasma membrane. *FEBS Lett.* 584, 1793–1799. doi: 10.1016/j.febslet.2010.02.049
- De, S., and Olson, R. (2011). Crystal structure of the *Vibrio cholerae* cytolysin heptamer reveals common features among disparate pore-forming toxins. *Proc. Natl. Acad. Sci. U.S.A.* 108, 7385–7390. doi: 10.1073/pnas.1017442108
- Deleplaire, P. (2004). Type I secretion in gram-negative bacteria. *Biochim. Biophys. Acta - Mol. Cell Res.* 1694, 149–161. doi: 10.1016/j.bbamcr.2004.05.001
- Desvaux, M., Parham, N. J., and Henderson, I. R. (2004). The autotransporter secretion system. *Res. Microbiol.* 155, 53–60. doi: 10.1016/j.resmic.2003.10.002
- DiScipio, R. G., Chakravarti, D. N., Muller-Eberhard, H. J., and Fey, G. H. (1988). The structure of human complement component C7 and the C5b-7 complex. *J. Biol. Chem.* 263, 549–60. doi: 10.1016/S0021-9258(19)57427-0
- Dobson, L., Langó, T., Reményi, I., and Tusnády, G. E. (2015). Expediting topology data gathering for the TOPDB database. *Nucleic Acids Res.* 43, D283–D289. doi: 10.1093/nar/gku1119
- Dong, H., Xiang, Q., Gu, Y., Wang, Z., Paterson, N. G., Stansfeld, P. J., et al. (2014). Structural basis for outer membrane lipopolysaccharide insertion. *Nature* 511, 52–56. doi: 10.1038/nature13464
- Dudkina, N. V., Spicer, B. A., Reboul, C. F., Conroy, P. J., Lukoyanova, N., Elmlund, H., et al. (2016). Structure of the poly-C9 component of the complement membrane attack complex. *Nat. Commun.* 4:10588. doi: 10.1038/ncomms10588
- Eddy, S. R. (1998). Profile hidden Markov models. *Bioinformatics* 14, 755–63. doi: 10.1093/bioinformatics/14.9.755
- Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340
- Edrington, T. C., Kintz, E., Goldberg, J. B., and Tamm, L. K. (2011). Structural basis for the interaction of lipopolysaccharide with outer membrane protein H (OprH) from *Pseudomonas aeruginosa*. *J. Biol. Chem.* 286, 39211–39223. doi: 10.1074/jbc.M111.280933
- Edwards, M. J., White, G. F., Butt, J. N., Richardson, D. J., and Clarke, T. A. (2020). The crystal structure of a biological insulated transmembrane molecular wire. *Cell* 181, 665–673.e10. doi: 10.1016/j.cell.2020.03.032
- El-Gebali, S., Mistry, J., Bateman, A., Eddy, S. R., Luciani, A., Potter, S. C., et al. (2019). The Pfam protein families database in 2019. *Nucleic Acids Res.* 47, D427–D432. doi: 10.1093/nar/gky995
- Eren, E., Vijayaraghavan, J., Liu, J., Cheneke, B. R., Touw, D. S., Lepore, B. W., et al. (2012). Substrate specificity within a family of outer membrane carboxylate channels. *PLoS Biol.* 10:e1001242. doi: 10.1371/journal.pbio.1001242
- Flack, F. S., Loosmore, S., Chong, P., and Thomas, W. R. (1995). The sequencing of the 80-kDa D 15 protective surface antigen of Haemophilus influenzae. *Gene* 156, 97–99. doi: 10.1016/0378-1119(95)00049-C
- Forst, D., Welte, W., Wacker, T., and Diederichs, K. (1998). Structure of the sucrose-specific porin ScrY from *Salmonella typhimurium* and its complex with sucrose. *Nat. Struct. Biol.* 5, 37–46. doi: 10.1038/nsb0198-37
- Freeman, T. C., and Wimley, W. C. (2010). A highly accurate statistical approach for the prediction of transmembrane β -barrels. *Bioinformatics* 26, 1965–1974. doi: 10.1093/bioinformatics/btq308
- Freeman, T. C. J., and Wimley, W. C. (2012). TMDB-DB: a transmembrane β -barrel proteome database. *Bioinformatics* 28, 2425–2430. doi: 10.1093/bioinformatics/bts478
- Galdiero, S., Falanga, A., Cantisani, M., Tarallo, R., Elena Della Pepa, M., D’Orlando, V., et al. (2013). Microbe-host interactions: structure and role of gram-negative bacterial porins. *Curr. Protein Pept. Sci.* 13, 843–854. doi: 10.2174/138920312804871120
- Gao, Z., Ye, C., Zhou, L., Zhang, Y., Ge, Y., Chen, W., et al. (2020). Evaluation of the β -barrel outer membrane protein VP1243 as a candidate antigen for a cross-protective vaccine against *Vibrio* infections. *Microb. Pathog.* 147:104419. doi: 10.1016/j.micpath.2020.104419
- Girard, V., and Mourez, M. (2006). Adhesion mediated by autotransporters of Gram-negative bacteria: structural and functional features. *Res. Microbiol.* 157, 407–416. doi: 10.1016/j.resmic.2006.02.001
- Goetze, T. A., Patil, M., Jeshen, I., Bölter, B., Grahl, S., and Soll, J. (2015). Oep23 forms an ion channel in the chloroplast outer envelope. *BMC Plant Biol.* 15:47. doi: 10.1186/s12870-015-0445-1
- Goetze, T. A., Philippar, K., Ilkavets, I., Soll, J., and Wagner, R. (2006). OEP37 is a new member of the chloroplast outer membrane ion channels. *J. Biol. Chem.* 281, 17989–17998. doi: 10.1074/jbc.M600700200
- Gray, M. W., Burger, G., and Lang, B. F. (1999). Mitochondrial evolution. *Science* 283, 1476–1481. doi: 10.1126/science.283.5407.1476
- Green, E. R., and Mecsas, J. (2016). Bacterial secretion systems: an overview. *Virulence Mech. Bact. Pathog. Fifth Ed.* 4, 215–239. doi: 10.1128/9781555819286.ch8
- Grinter, R., Josts, I., Mosbahi, K., Roszak, A. W., Cogdell, R. J., Bonvin, A. M. J. J., et al. (2016). Structure of the bacterial plant-ferredoxin

- receptor FusA. *Nat. Commun.* 31:13308. doi: 10.1038/ncomms13308
- Gristina, A. (1987). Biomaterial-centered infection: microbial adhesion versus tissue integration. *Sci. Sciencemag Org.* 237, 1588–1595. doi: 10.1126/science.3629258
- Grizot, S., and Buchanan, S. K. (2004). Structure of the OmpA-like domain of RmpM from *Neisseria meningitidis*. *Mol. Microbiol.* 51, 1027–1037. doi: 10.1111/j.1365-2958.2003.03903.x
- Gromiha, M. M., Yabuki, Y., Kundu, S., Suharnan, S., and Suwa, M. (2007). TMβETA-GENOME: database for annotated beta-barrel membrane proteins in genomic sequences. *Nucleic Acids Res.* 35, D314–D316. doi: 10.1093/nar/gkl805
- Gross, L. E., Spies, N., Simm, S., and Schleiff, E. (2020). Toc75-V/OEP80 is processed during translocation into chloroplasts and the membrane embedded form exposes its POTRA domain to the intermembrane space. *FEBS Open Bio.* 10, 444–454. doi: 10.1002/2211-5463.12791
- Gu, Y., Stansfeld, P. J., Zeng, Y., Dong, H., Wang, W., and Dong, C. (2015). Lipopolysaccharide is inserted into the outer membrane through an intramembrane hole, a lumen gate, and the lateral opening of LptD. *Structure* 23, 496–504. doi: 10.1016/j.str.2015.01.001
- Guo, B., Wang, Y., Shi, F., Barton, Y. W., Plummer, P., Reynolds, D. L., et al. (2008). CmeR functions as a pleiotropic regulator and is required for optimal colonization of *Campylobacter jejuni* in vivo. *J. Bacteriol.* 190, 1879–1890. doi: 10.1128/JB.01796-07
- Harsman, A., Schock, A., Hemmis, B., Wahl, V., Jeshen, I., Bartsch, P., et al. (2016). OEP40, a regulated glucose-permeable β-barrel solute channel in the chloroplast outer envelope membrane. *J. Biol. Chem.* 291, 17848–17860. doi: 10.1074/jbc.M115.712398
- Hemmler, R., Becker, T., Schleiff, E., Bölter, B., Stahl, T., Soll, J., et al. (2006). Molecular properties of Oep21, an ATP-regulated anion-selective solute channel from the outer chloroplast membrane. *J. Biol. Chem.* 281, 12020–12029. doi: 10.1074/jbc.M513586200
- Hendrixon, D. R., De La Morena, M. L., Stathopoulos, C., and St Geme, J. W. III. (1997). Structural determinants of processing and secretion of the Haemophilus influenzae Hap protein. *Mol. Microbiol.* 26, 505–518. doi: 10.1046/j.1365-2958.1997.5921965.x
- Higgs, P. I., Larsen, R. A., and Postle, K. (2002). Quantification of known components of the *Escherichia coli* TonB energy transduction system: TonB, ExbB, ExbD and FepA. *Mol. Microbiol.* 44, 271–281. doi: 10.1046/j.1365-2958.2002.02880.x
- Hill, K., Model, K., Ryan, M. T., Dietmeier, K., Martint, F., Wagner, R., et al. (1998). Tom40 forms the hydrophilic channel of the mitochondrial import pore for preproteins. *Nature* 395, 516–521. doi: 10.1038/26780
- Hinnah, S. C., Wagner, R., Sveshnikova, N., Harrer, R., and Soll, J. (2002). The chloroplast protein import channel Toc75: pore properties and interaction with transit peptides. *Biophys. J.* 83, 899–911. doi: 10.1016/S0006-3495(02)75216-8
- Höhr, A. I. C., Lindau, C., Wirth, C., Qiu, J., Stroud, D. A., Kutik, S., et al. (2018). Membrane protein insertion through a mitochondrial β-barrel gate. *Science* 359:eaah6834. doi: 10.1126/science.aah6834
- Holliday, G. L., Bairoch, A., Bagos, P. G., Chatonnet, A., Craik, D. J., Finn, R. D., et al. (2015). Key challenges for the creation and maintenance of specialist protein resources. *Proteins* 83, 1005–1013. doi: 10.1002/prot.24803
- Höltje, J.-V. (1998). Growth of the stress-bearing and shape-maintaining murein sacculus of *Escherichia coli*. *Microbiol. Mol. Biol. Rev.* 62, 181–203. doi: 10.1128/MMBR.62.1.181-203.1998
- Iacovache, I., De Carlo, S., Cirauqui, N., Dal Peraro, M., Van Der Goot, F. G., and Zuber, B. (2016). Cryo-EM structure of aerolysin variants reveals a novel protein fold and the pore-formation process. *Nat. Commun.* 13:12062. doi: 10.1038/ncomms12062
- Ikeda, M., Arai, M., Okuno, T., and Shimizu, T. (2003). TMPDB: a database of experimentally-characterized transmembrane topologies. *Nucleic Acids Res.* 31, 406–409. doi: 10.1093/nar/gkg020
- Jansen, K. B., Baker, S. L., and Sousa, M. C. (2015). Crystal structure of BamB bound to a periplasmic domain fragment of BamA, the central component of the β-barrel assembly machine. *J. Biol. Chem.* 290, 2126–2136. doi: 10.1074/jbc.M114.584524
- Katta, A. V. S. K. M., Marikkannu, R., Basaiwmoit, R. V., and Krishnaswamy, S. (2004). Consensus based validation of membrane porins. *In Silico Biol.* 4, 549–561.
- Kawatsu, K., Ishibashi, M., and Tsukamoto, T. (2006). Development and evaluation of a rapid, simple, and sensitive immunochromatographic assay to detect thermostable direct hemolysin produced by *Vibrio parahaemolyticus* in enrichment cultures of stool specimens. *J. Clin. Microbiol.* 44, 1821–1827. doi: 10.1128/JCM.44.5.1821-1827.2006
- Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* 10, 845–858. doi: 10.1038/nprot.2015.053
- Kim, S., Malinverni, J. C., Sliz, P., Silhavy, T. J., Harrison, S. C., and Kahne, D. (2007). Structure and function of an essential component of the outer membrane protein assembly machine. *Science* 317, 961–964. doi: 10.1126/science.1143993
- Kleine, T., Maier, U. G., and Leister, D. (2009). DNA transfer from organelles to the nucleus: the idiosyncratic genetics of endosymbiosis. *Annu. Rev. Plant Biol.* 60, 115–138. doi: 10.1146/annurev.arplant.043008.092119
- Klemm, P., Vejborg, R. M., and Sherlock, O. (2006). Self-associating autotransporters, SAATs: functional and structural similarities. *Int. J. Med. Microbiol.* 296, 187–195. doi: 10.1016/j.ijmm.2005.10.002
- Koebnik, R., Locher, K. P., and Van Gelder, P. (2000). Structure and function of bacterial outer membrane proteins: barrels in a nutshell. *Mol. Microbiol.* 37, 239–253. doi: 10.1046/j.1365-2958.2000.01983.x
- Kowata, H., Tochigi, S., Takahashi, H., and Kojima, S. (2017). Outer membrane permeability of cyanobacterium *Synechocystis* sp. strain PCC 6803: studies of passive diffusion of small organic nutrients reveal the absence of classical porins and intrinsically low permeability. *J. Bacteriol.* 199:e00371-17. doi: 10.1128/JB.00371-17
- Kozma, D., Simon, I., and Tusnády, G. E. (2013). PDBTM: Protein data bank of transmembrane proteins after 8 years. *Nucleic Acids Res.* 41, D524–D529. doi: 10.1093/nar/gks1169
- Lasica, A. M., Ksiazek, M., Madej, M., and Potempa, J. (2017). The Type IX Secretion System (T9SS): highlights and recent insights into its structure and function. *Front. Cell. Infect. Microbiol.* 7:215. doi: 10.3389/fcimb.2017.00215
- Lauber, F., Deme, J. C., Lea, S. M., and Berks, B. C. (2018). Type 9 secretion system structures reveal a new protein transport mechanism. *Nature* 564, 77–82. doi: 10.1038/s41586-018-0693-y
- Li, W., Cowley, A., Uludag, M., Gur, T., McWilliam, H., Squizzato, S., et al. (2015). The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic Acids Res.* 43, W580–W584. doi: 10.1093/nar/gkv279
- Lin, C.-C., Chou, C.-H., Lin, T.-C., Yang, M.-C., Cho, C.-L., Chang, C.-H., et al. (2012). Molecular characterization of three major outer membrane proteins, TSA56, TSA47 and TSA22, in *Orientia tsutsugamushi*. *Int. J. Mol. Med.* 30, 75–84. doi: 10.3892/ijmm.2012.967
- Liu, J., Eren, E., Vijayaraghavan, J., Cheneke, B. R., Indic, M., Van Den Berg, B., et al. (2012). OccK channels from *Pseudomonas aeruginosa* exhibit diverse single-channel electrical signatures but conserved anion selectivity. *Biochemistry* 51, 2319–2330. doi: 10.1021/bi300066w
- Lomize, M. A., Pogozheva, I. D., Joo, H., Mosberg, H. I., and Lomize, A. L. (2012). OPM database and PPM web server: resources for positioning of proteins in membranes. *Nucleic Acids Res.* 40, D370–D376. doi: 10.1093/nar/gkr703
- McKay, T., Hart, K., Horn, A., Kessler, H., Dodge, G., Bardhi, K., et al. (2015). Annotation of proteins of unknown function: initial evidence results. *J. Struct. Funct. Genomics* 16, 43–54. doi: 10.1007/s10969-015-9194-5
- Meisinger, C., Rissler, M., Chacinska, A., Sanjuán Szklarz, L. K., Milenkovic, D., Kozjak, V., et al. (2004). The mitochondrial morphology protein Mdm10 functions in assembly of the preprotein translocase of the outer membrane. *Dev. Cell* 7, 61–71. doi: 10.1016/j.devcel.2004.06.003
- Meng, G., Spahich, N., Kenjale, R., Waksman, G., and St Geme, J. W. (2011). Crystal structure of the Haemophilus influenzae Hap adhesin reveals an intercellular oligomerization mechanism for bacterial aggregation. *EMBO J.* 30, 3864–3874. doi: 10.1038/emboj.2011.279
- Menny, A., Serna, M., Boyd, C. M., Gardner, S., Joseph, A. P., Morgan, B. P., et al. (2018). CryoEM reveals how the complement membrane attack complex ruptures lipid bilayers. *Nat. Commun.* 9:5316. doi: 10.1038/s41467-018-07653-5
- Moraes, T. F., Bains, M., Hancock, R. E. W., and Strynadka, N. C. J. (2007). An arginine ladder in OprP mediates phosphate-specific transfer across the outer membrane. *Nat. Struct. Mol. Biol.* 14, 85–87. doi: 10.1038/nsmb1189

- Moreno-Gonzalez, I., and Soto, C. (2011). Misfolded protein aggregates: mechanisms, structures and potential for disease transmission. *Semin. Cell Dev. Biol.* 22, 482–487. doi: 10.1016/j.semcdb.2011.04.002
- Natale, P., Brüser, T., and Driessen, A. J. M. (2008). Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane-Distinct translocases and mechanisms. *Biochim. Biophys. Acta - Biomembr.* 1778, 1735–1756. doi: 10.1016/j.bbamem.2007.07.015
- Naumann, M., Rudel, T., and Meyer, T. F. (1999). Host cell interactions and signalling with *Neisseria gonorrhoeae*. *Curr. Opin. Microbiol.* 2, 62–70. doi: 10.1016/S1369-5274(99)80011-3
- Newport, T. D., Sansom, M. S. P., and Stansfeld, P. J. (2019). The MemProtMD database: a resource for membrane-embedded protein structures and their lipid interactions. *Nucleic Acids Res.* 47, D390–D397. doi: 10.1093/nar/gky1047
- Nguyen, K. A., Zylicz, J., Szczesny, P., Sroka, A., Hunter, N., and Potempa, J. (2009). Verification of a topology model of PorT as an integral outer-membrane protein in *Porphyromonas gingivalis*. *Microbiology* 155, 328–337. doi: 10.1099/mic.0.024323-0
- Nikaido, H. (2003). Molecular basis of bacterial outer membrane permeability revisited. *Microbiol. Mol. Biol. Rev.* 67, 593–656. doi: 10.1128/MMBR.67.4.593-656.2003
- Noiraj, N., Rollauer, S. E., and Buchanan, S. K. (2015). The β -barrel membrane protein insertase machinery from Gram-negative bacteria. *Curr. Opin. Struct. Biol.* 31, 35–42. doi: 10.1016/j.sbi.2015.02.012
- Nouwen, N., Berrelkamp, G., and Driessen, A. J. M. (2007). Bacterial sec-translocase unfolds and translocates a class of folded protein domains. *J. Mol. Biol.* 372, 422–433. doi: 10.1016/j.jmb.2007.07.003
- Osbak, K. K., Houston, S., Lithgow, K. V., Meehan, C. J., and Strouhal, M., Šmajš, D., et al. (2016). Characterizing the syphilis-causing *Treponema pallidum* ssp. *pallidum* proteome using complementary mass spectrometry. *PLoS Negl. Trop. Dis.* 10:e0004988. doi: 10.1371/journal.pntd.0004988
- Pagès, J. M., James, C. E., and Winterhalter, M. (2008). The porin and the permeating antibiotic: a selective diffusion barrier in Gram-negative bacteria. *Nat. Rev. Microbiol.* 6, 893–903. doi: 10.1038/nrmicro1994
- Peabody, M. A., Laird, M. R., Vlasschaert, C., Lo, R., and Brinkman, F. S. L. (2016). PSORTdb: expanding the bacteria and archaea protein subcellular localization database to better reflect diversity in cell envelope structures. *Nucleic Acids Res.* 44, D663–D668. doi: 10.1093/nar/gkv1271
- Petersen, T. N., Brunak, S., von Heijne, G., and Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* 8, 785–786. doi: 10.1038/nmeth.1701
- Pizarro-Cerdá, J., and Cossart, P. (2006). Bacterial adhesion and entry into host cells. *Cell* 124, 715–727. doi: 10.1016/j.cell.2006.02.012
- Podack, E., Tschopp, J., and Müller-Eberhard, H. J. (1982). Molecular organization of C9 within the membrane attack complex of complement. induction of circular C9 polymerization by the C5b-8 assembly. *J. Exp. Med.* 156, 268–282. doi: 10.1084/jem.156.1.268
- Podobnik, M., Savory, P., Rojko, N., Kisovec, M., Wood, N., Hambley, R., et al. (2016). Crystal structure of an invertebrate cytolysin pore reveals unique properties and mechanism of assembly. *Nat. Commun.* 7:11598. doi: 10.1038/ncomms11598
- Pohlmeier, K., Soll, J., Grimm, R., Hill, K., and Wagner, R. (1998). A high-conductance solute channel in the chloroplastic outer envelope from Pea. *Plant Cell* 10, 1207–1216. doi: 10.1105/tpc.10.7.1207
- Raman, P., Cherezov, V., and Caffrey, M. (2006). The membrane protein data bank. *Cell. Mol. Life Sci.* 63, 36–51. doi: 10.1007/s00018-005-5350-6
- Rapoport, T. A. (2007). Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes. *Nature* 450, 663–669. doi: 10.1038/nature06384
- Remmert, M., Linke, D., Lupas, A. N., and Söding, J. (2009). HHomp—prediction and classification of outer membrane proteins. *Nucleic Acids Res.* 37, W446–W451. doi: 10.1093/nar/gkp325
- Ringler, P., and Schulz, G. E. (2002). OmpA membrane domain as a tight-binding anchor for lipid bilayers. *ChemBioChem* 3, 463–466. doi: 10.1002/1439-7633(20020503)3:5<463::AID-CBIC463>3.0.CO;2-P
- Rosado, C. J., Buckle, A. M., Law, R. H. P., Butcher, R. E., Kan, W.-T., Bird, C. H., et al. (2007). A common fold mediates vertebrate defense and bacterial attack. *Science* 317, 1548–1551. doi: 10.1126/science.1144706
- Roumia, A. F., Theodoropoulou, M. C., Tsigiris, K. D., Nielsen, H., and Bagos, P. G. (2020). Landscape of eukaryotic transmembrane beta barrel proteins. *J. Proteome Res.* 19, 1209–1221. doi: 10.1021/acs.jproteome.9b00740
- Rouse, S. L., Hawthorne, W. J., Berry, J. L., Chorev, D. S., Ionescu, S. A., Lambert, S., et al. (2017). A new class of hybrid secretion system is employed in *Pseudomonas amyloid* biogenesis. *Nat. Commun.* 8:263. doi: 10.1038/s41467-017-00361-6
- Ruffolo, C. G., and Adler, B. (1996). Cloning, sequencing, expression, and protective capacity of the oma87 gene encoding the *Pasteurella multocida* 87-kilodalton outer membrane antigen. *Infect. Immun.* 64, 3161–3167. doi: 10.1128/IAI.64.8.3161-3167.1996
- Rutten, L., Geurtsen, J., Lambert, W., Smolenaers, J. J. M., Bonvin, A. M., De Haan, A., et al. (2006). Crystal structure and catalytic mechanism of the LPS 3-O-deacetylase PagL from *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. U.S.A.* 103, 7071–7076. doi: 10.1073/pnas.0509392103
- Saier, M. H. J., Reddy, V. S., Tsu, B. V., Ahmed, M. S., Li, C., and Moreno-Hagelsieb, G. (2016). The Transporter Classification Database (TCDB): recent advances. *Nucleic Acids Res.* 44, D372–D379. doi: 10.1093/nar/gkv1103
- Sayers, E. W., Barrett, T., Benson, D. A., Bryant, S. H., Canese, K., Chetvernin, V., et al. (2009). Database resources of the national center for biotechnology information. *Nucleic Acids Res.* 37, D5–D15. doi: 10.1093/nar/gkp382
- Schirmer, T., Keller, T. A., Wang, Y. F., and Rosenbusch, J. P. (1995). Structural basis for sugar translocation through maltoporin channels at 3.1 Å resolution. *Science* 267, 512–514. doi: 10.1126/science.7824948
- Schulz, G. E. (2002). The structure of bacterial outer membrane proteins. *Biochim. Biophys. Acta* 1565, 308–317. doi: 10.1016/S0005-2736(02)00577-1
- Sharp, T. H., Koster, A. J., and Gros, P. (2016). Heterogeneous MAC Initiator and pore structures in a lipid bilayer by phase-plate cryo-electron tomography. *Cell Rep.* 15, 1–8. doi: 10.1016/j.celrep.2016.03.002
- Shatursky, O., Heuck, A. P., Shepard, L. A., Rossjohn, J., Parker, M. W., Johnson, A. E., et al. (1999). The mechanism of membrane insertion for a cholesterol-dependent cytolysin: a novel paradigm for pore-forming toxins. *Cell* 99, 293–299. doi: 10.1016/S0092-8674(00)81660-8
- Shepard, L. A., Heuck, A. P., Hamman, B. D., Rossjohn, J., Parker, M. W., Ryan, K. R., et al. (1998). Identification of a membrane-spanning domain of the thiol-activated pore-forming toxin *Clostridium perfringens* perfringolysin O: an alpha-helical to beta-sheet transition identified by fluorescence spectroscopy. *Biochemistry* 37, 14563–14574. doi: 10.1021/bi981452f
- Shrivastava, A., Roland, T., and Berg, H. C. (2016). The screw-like movement of a gliding bacterium is powered by spiral motion of cell-surface adhesins. *Biophys. J.* 111, 1008–1013. doi: 10.1016/j.bpj.2016.07.043
- Soll, J., and Schleiff, E. (2004). Protein import into chloroplasts. *Nat. Rev. Mol. Cell Biol.* 9, 222–227. doi: 10.1038/nrm1333
- Soto, G. E., and Hultgren, S. J. (1999). MINIREVIEW bacterial adhesins: common themes and variations in architecture and assembly. *J. Bacteriol.* 181, 1059–1071. doi: 10.1128/JB.181.4.1059-1071.1999
- Steckel, E. W., Welbaum, B. E., and Sodetz, J. M. (1983). Evidence of direct insertion of terminal complement proteins into cell membrane bilayers during cytolysis. labeling by a photosensitive membrane probe reveals a major role for the eighth and ninth components. *J. Biol. Chem.* 258, 4318–4324. doi: 10.1016/S0021-9258(18)32625-5
- Stojanovski, D., Guiard, B., Kozjak-Pavlovic, V., Pfanner, N., and Meisinger, C. (2007). Alternative function for the mitochondrial SAM complex in biogenesis of α -helical TOM proteins. *J. Cell Biol.* 179, 881–893. doi: 10.1083/jcb.200706043
- Sunde, M., Serpell, L. C., Bartlam, M., Fraser, P. E., Pepys, M. B., and Blake, C. C. F. (1997). Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *J. Mol. Biol.* 273, 729–739. doi: 10.1006/jmbi.1997.1348
- Sutcliffe, I. C. (2010). A phylum level perspective on bacterial cell envelope architecture. *Trends Microbiol.* 18, 464–470. doi: 10.1016/j.tim.2010.06.005
- The UniProt Consortium (2019). UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* 47, D506–D515. doi: 10.1093/nar/gky1049
- Thein, M., Bonde, M., Bunikis, I., Denker, K., Sickmann, A., Bergström, S., et al. (2012). DipA, a pore-forming protein in the outer membrane of Lyme disease spirochetes exhibits specificity for the permeation of dicarboxylates. *PLoS ONE* 7:e36523. doi: 10.1371/journal.pone.0036523
- Thein, M., Bunikis, I., Denker, K., Larsson, C., Cutler, S., Drancourt, M., et al. (2008). Oms38 is the first identified pore-forming protein in the

- outer membrane of relapsing fever spirochetes. *J. Bacteriol.* 190, 7035–7042. doi: 10.1128/JB.00818-08
- Tomasek, D., and Kahne, D. (2021). The assembly of β -barrel outer membrane proteins. *Curr. Opin. Microbiol.* 60, 16–23. doi: 10.1016/j.mib.2021.01.009
- Trung, N. V., Hoi, L. T., Cuong, D. D., Ha, D. T., Hoa, T. M., Lien, V. N., et al. (2019). Analysis of the 56-kDa type specific antigen gene of *Orientia tsutsugamushi* from northern Vietnam. *PLoS ONE* 14:e0221588. doi: 10.1371/journal.pone.0221588
- Tschopp, J., Masson, D., and Stanley, K. K. (1986). Structural/functional similarity between proteins involved in complement- and cytotoxic T-lymphocyte-mediated cytotoxicity. *Nature* 322, 831–834. doi: 10.1038/322831a0
- Tsirigos, K. D., Bagos, P. G., and Hamodrakas, S. J. (2011). OMPdb: a database of β -barrel outer membrane proteins from Gram-negative bacteria. *Nucleic Acids Res.* 39, D324–D331. doi: 10.1093/nar/gkq863
- Tsirigos, K. D., Elofsson, A., and Bagos, P. G. (2016). PRED-TMBB2: improved topology prediction and detection of beta-barrel outer membrane proteins. *Bioinformatics* 32, i665–i671. doi: 10.1093/bioinformatics/btw444
- Tusnady, G. E., Kalmar, L., and Simon, I. (2008). TOPDB: topology data bank of transmembrane proteins. *Nucleic Acids Res.* 36, D234–D239. doi: 10.1093/nar/gkm751
- Van Den Berg, B. (2012). Structural basis for outer membrane sugar uptake in pseudomonads. *J. Biol. Chem.* 287, 41044–41052. doi: 10.1074/jbc.M112.408518
- van den Berg, B., Bhamidimarri, S. P., and Winterhalter, M. (2015). Crystal structure of a COG4313 outer membrane channel. *Sci. Rep.* 5:11927. doi: 10.1038/srep11927
- van den Berg, B., Black, P. N., Clemons, W. M. J., and Rapoport, T. A. (2004). Crystal structure of the long-chain fatty acid transporter FadL. *Science* 304, 1506–1509. doi: 10.1126/science.1097524
- Varghees, S., Kiss, K., Frans, G., Braha, O., and Samuel, J. E. (2002). Cloning and porin activity of the major outer membrane protein P1 from *Coxiella burnetii*. *Infect. Immun.* 70, 6741–6750. doi: 10.1128/IAI.70.12.6741-6750.2002
- Volokhina, E., Beckers, F., Tommassen, J., and Bos, M. (2009). The beta-barrel outer membrane protein assembly complex of *Neisseria Meningitidis*. *J. Bacteriol.* 191, 7074–7085. doi: 10.1128/JB.00737-09
- von Heijne, G. (1999). Recent advances in the understanding of membrane protein assembly and structure. *Q. Rev. Biophys.* 32, 285–307. doi: 10.1017/S0033583500003541
- Vora, G. J., Meador, C. E., Bird, M. M., Bopp, C. A., Andreadis, J. D., and Stenger, D. A. (2005). Microarray-based detection of genetic heterogeneity, antimicrobial resistance, and the viable but nonculturable state in human pathogenic *Vibrio* spp. *Proc. Natl. Acad. Sci. U.S.A.* 102, 19109–19114. doi: 10.1073/pnas.0505033102
- Voulhoux, R., Bos, M. P., Geurtsen, J., Mols, M., and Tommassen, J. (2003). Role of a highly conserved bacterial protein in outer membrane protein assembly. *Science* 299, 262–5. doi: 10.1126/science.1078973
- Walker, S. G., Ebersole, J. L., and Holt, S. C. (1997). Identification, isolation, and characterization of the 42-kilodalton major outer membrane protein (MompA) from *Treponema pectinovorum* ATCC 33768. *J. Bacteriol.* 179, 6441–6447. doi: 10.1128/JB.179.20.6441-6447.1997
- Wall, D., and Kaiser, D. (1999). Type IV pili and cell motility. *Mol. Microbiol.* 32, 01–10. doi: 10.1046/j.1365-2958.1999.01339.x
- Wang, Y., Pannuri, A. A., Ni, D., Zhou, H., Cao, X., Lu, X., et al. (2016). Structural basis for translocation of a biofilm-supporting exopolysaccharide across the bacterial outer membrane. *J. Biol. Chem.* 291, 10046–10057. doi: 10.1074/jbc.M115.711762
- Wang, Y. F., Dutzler, R., Rizkallah, P. J., Rosenbusch, J. P., and Schirmer, T. (1997). Channel specificity: Structural basis for sugar discrimination and differential flux rates in maltoporin. *J. Mol. Biol.* 272, 56–63. doi: 10.1006/jmbi.1997.1224
- Wang, Z., Xu, C., and Benning, C. (2012). TGD4 involved in endoplasmic reticulum-to-chloroplast lipid trafficking is a phosphatidic acid binding protein. *Plant J.* 70, 614–623. doi: 10.1111/j.1365-313X.2012.04900.x
- White, S. H. (2009). Biophysical dissection of membrane proteins. *Nature* 459, 344–346. doi: 10.1038/nature08142
- Whitfield, C., and Trent, M. S. (2014). Biosynthesis and export of bacterial lipopolysaccharides. *Annu. Rev. Biochem.* 83, 99–128. doi: 10.1146/annurev-biochem-060713-035600
- Wirth, C., Condemine, G., Boiteux, C., Berneche, S., Schirmer, T., and Peneff, C. M. (2009). NanC crystal structure, a model for outer-membrane channels of the acidic sugar-specific KdgM porin family. *J. Mol. Biol.* 394, 718–731. doi: 10.1016/j.jmb.2009.09.054
- Yu, C.-S., Chen, Y.-C., Lu, C.-H., and Hwang, J.-K. (2006). Prediction of protein subcellular localization. *Proteins* 64, 643–651. doi: 10.1002/prot.21018
- Yu, N. Y., Wagner, J. R., Laird, M. R., Melli, G., Rey, S., Lo, R., et al. (2010). PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* 26, 1608–1615. doi: 10.1093/bioinformatics/btq249
- Zhang, G., Meredith, T. C., and Kahne, D. (2013). On the essentiality of lipopolysaccharide to Gram-negative bacteria. *Curr. Opin. Microbiol.* 16, 779–785. doi: 10.1016/j.mib.2013.09.007
- Zhu, Y., and McBride, M. J. (2014). Deletion of the *Cytophaga hutchinsonii* type IX secretion system gene sprP results in defects in gliding motility and cellulose utilization. *Appl. Microbiol. Biotechnol.* 98, 763–775. doi: 10.1007/s00253-013-5355-2

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Roumia, Tsirigos, Theodoropoulou, Tamposis, Hamodrakas and Bagos. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.