

Functional Annotation of Hypothetical Proteins From the *Enterobacter cloacae* B13 Strain and Its Association With Pathogenicity

Bioinformatics and Biology Insights
Volume 16: 1–16
© The Author(s) 2022
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/11779322221115535



Supantha Dey¹, Sazzad Shahrear¹, Maliha Afroj Zinnia², Ahnaf Tajwar¹ and Abul Bashar Mir Md. Khademul Islam¹

¹Department of Genetic Engineering and Biotechnology, University of Dhaka, Dhaka, Bangladesh. ²Department of Pharmacy, East West University, Dhaka, Bangladesh.

ABSTRACT: *Enterobacter cloacae* B13 strain is a rod-shaped gram-negative bacterium that belongs to the *Enterobacteriaceae* family. It can cause respiratory and urinary tract infections, and is responsible for several outbreaks in hospitals. *E. cloacae* has become an important pathogen and an emerging global threat because of its opportunistic and multidrug resistant ability. However, little knowledge is present about a large portion of its proteins and functions. Therefore, functional annotation of the hypothetical proteins (HPs) can provide an improved understanding of this organism and its virulence activity. The workflow in the study included several bioinformatic tools which were utilized to characterize functions, family and domains, subcellular localization, physicochemical properties, and protein-protein interactions. The *E. cloacae* B13 strain has overall 604 HPs, among which 78 were functionally annotated with high confidence. Several proteins were identified as enzymes, regulatory, binding, and transmembrane proteins with essential functions. Furthermore, 23 HPs were predicted to be virulent factors. These virulent proteins are linked to pathogenesis with their contribution to biofilm formation, quorum sensing, 2-component signal transduction or secretion. Better knowledge about the HPs' characteristics and functions will provide a greater overview of the proteome. Moreover, it will help against *E. cloacae* in neonatal intensive care unit (NICU) outbreaks and nosocomial infections.

KEYWORDS: *Enterobacter cloacae*, hypothetical proteins, functional annotation, virulence factors

RECEIVED: April 1, 2022. **ACCEPTED:** June 11, 2022.

TYPE: Original Research Article

FUNDING: The author(s) received no financial support for the research, authorship, and/or publication of this article.

DECLARATION OF CONFLICTING INTERESTS: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CORRESPONDING AUTHOR: Abul Bashar Mir Md. Khademul Islam, Department of Genetic Engineering and Biotechnology, University of Dhaka, Nilkhet Rd, Dhaka 1000, Bangladesh. Email: khademul@du.ac.bd

Introduction

Enterobacter is a genus of gram-negative, facultative anaerobic, and rod-shaped bacteria. It is a member of the family *Enterobacteriaceae*.¹ Bacteria which belong to the *Enterobacteriaceae* are most often isolated from soil, water, or different clinical specimens.² *Enterobacter cloacae* is an integral part of microflora in the human and animal intestinal tracts.³ *E. cloacae* complex are largely observed in nature, although they can act as pathogens.⁴ *E. cloacae* has emerged as a significant pathogen to study because of its several outbreaks including neonatal intensive care units (NICU).^{5–8} It has become a problematic pathogen for healthcare institutes at a global level as it tends to contaminate various hospital devices.⁴ Up to 5% of all hospital-acquired sepsis and nosocomial pneumonia, 4% of nosocomial urinary tract infections (UTIs), and 10% of postsurgical peritonitis cases are caused by *E. cloacae*.^{9,10} This microorganism's transmission to neonates can be accompanied through contaminated intravenous fluid or medical equipment along with the possibility of inpatients acting as a reservoir.⁵ In general, they contact mostly through the GI tract and skins in humans.^{3,11}

E. cloacae's mechanism of pathogenesis is multifactorial and complicated as it involves a couple of virulence factors, whose role in the disease development is still unclear.⁴ Production of enterotoxins such as cytotoxin similar to Shiga-like toxin II after their adhesion to epithelial cells, type III secretion system (TTSS) with several virulence factors, and phagocytes destroying capability might contribute to *E. cloacae*'s pathogenicity.^{12–14} Furthermore,

induced apoptosis of HEp-2 cells might be a primary strategy of this microorganism to destruct tissues, spread and cause infection or disease.¹⁵ Some studies also pointed to the colony formation ability of *Enterobacteriaceae* which mediates their binding to host proteins, resulting in cell adhesion and invasion.¹⁶ Curli genes expression and curli fimbria demonstrated a correlation between the biofilm formation and morphology of *E. cloacae*.¹⁷ In addition, *E. cloacae* complex has exhibited multidrug-resistance phenotype with their intrinsic β -lactam resistance and genes encoding antibiotic resistance, such as carbapenemase genes.^{18,19} This multidrug-resistance ability of *E. cloacae* can emerge as a global threat.¹⁸

The recent NGS technology can produce a huge amount of genomic data for a wide array of bacteria. However, the lack of complete proteome data because of coding sequences without a proper prediction of functions has made it difficult to understand pathogenesis and virulence determination. These molecules are labeled as hypothetical proteins (HP).^{2,20,21} Nearly 30% to 40% genes of most bacterial genomes are classified as unknown or hypothetical.²² These HPs are the translated nucleic acid sequences based on sequence similarity, but their biochemical and functional characterization evaluation is necessary for the experimental existence.²³ Therefore, the functional annotation of many hypothetical proteins has become an important focus in bioinformatics.²⁴ Homology-dependent gene annotation can assign functions to HPs based on their correlation with known proteins, providing the knowledge of new structures, functions, interactions, and pathways.^{23–27} A well and precise annotation of



E. cloacae HPs can bring additional protein pathways and cascades to our understanding, thus decreasing the gap of knowledge between genome data and protein functions.²⁷ The identification of proteins and their roles in bacterial growth and pathogenesis could be aided by in silico functional annotation of hypothetical proteins. In vitro and in vivo studies have verified the reliability of the in silico functional annotation approach. The proteins of *Pseudomonas* sp. Lz4W involved in cold adaptation, as well as the high arsenic-resistance genes from *Exiguobacterium antarcticum* strain B7, were identified using an integrated in silico and in vivo approach to functionally annotate the hypothetical proteins.^{20,28}

As *E. cloacae*'s pathogenic ability has already helped it to cause multiple NICU outbreaks, deciphering its HP functions in the genome is essential to completely understand the mode of pathogenesis.^{2,5-8} Moreover, novel HPs from *E. cloacae* can be used as markers and pharmacological targets for drug designing and screening, helping to prevent outbreaks.^{29,30}

Many proteins from *E. cloacae* B13 strain are still uncharacterized, which might contain crucial functions in their life. In this study, several in silico approaches were utilized to predict the function of the HPs from this organism. Identifying the functions of the HPs will contribute to a better proteome knowledge of the bacteria and its contribution to virulence. This will help to fight against *E. cloacae* as an emerging threat in nosocomial infections and hospital outbreaks.

Materials and Methods

The methodology overview flowchart is presented in Figure 1.

Extraction of genomic data

E. cloacae B13 genome was used in this study. This strain was isolated from human urine sample in Bangladesh and its genome size is 4963112 bp and encodes 4707 proteins.

The entire sequence of *E. cloacae* strain B13 was downloaded from NCBI sequence set browser with accession PRJNA472680.³¹ Among 4707 proteins, sequences of the HPs were retrieved using the fasta_extract tool in Galaxy.³²

Gene ontology prediction

To determine the HP's functions, Blast2GO with an *E*-value $1e^{-03}$ was used. Blast2GO is a bioinformatics tool which can be used for high-throughput functional annotation of DNA or protein sequences based on the Gene Ontology (GO) vocabulary.³³ The protein sequences with GO IDs were selected and various bioinformatics tools were used for further analysis of their domain and functions.

Family and domain prediction

The conserved domains and protein functions were searched based on the structure of domains. So, Simple Modular Architecture Research Tool (SMART) was used in general mode

to identify and annotate genetically mobile domains of signaling, extracellular and chromatin-associated proteins.³⁴ Furthermore, NCBI Batch CD-Search was applied that allowed to search multiple protein sequences using RPS-BLAST to compare query HP sequences against databases of conserved domain models.³⁵ NCBI Batch CD-Search tool searched against CDD—58235PSSMs database, and threshold was set at 0.01.

The HMMER website, allows protein homology search algorithms within the HMMER 3.3.2 software suite and uses profile hidden Markov model libraries to annotate the HP sequences with protein families and domains.³⁶ The cut off was set at 0.01 for significant *e*-values. Superfamily and Pfam association of the HPs from HMMER was searched. SUPERFAMILY 2.0 database contains superfamily domain annotations for millions of distinct proteins obtained from UniProtKB and NCBI.³⁷ In addition, annotated protein families of Pfam database are represented by Multiple Sequence Alignment (MSA) and hidden Markov Model (HMMs).³⁸ We searched for InterPro IDs for the selected HPs using Blast2GO utilities. InterPro uses predictive models, provided by several databases to provide functional analysis of proteins by classifying families and domains.^{33,39} So far, the functions of HPs with tools like Blast2GO, NCBI Batch CD-Search, SMART, SUPERFAMILY, Pfam, and INTERPRO were predicted and the HPs with predicted functions by 3 or more tools were identified with the help of InteractiVenn.⁴⁰ Finally, the Basic Local Alignment Search Tool (BLAST) against the NCBI nonredundant (nr) database was used to identify the annotated homologous proteins from related organisms, with similarity at $\geq 90\%$.^{20,41}

Subcellular localization determination

In the study, PSORTb v3.0 and CELLO v.2.5: Subcellular Localization Predictor were used to determine the cell locations of the HPs using default parameters for gram-negative bacteria.⁴²⁻⁴⁴ PSORTb database contains information obtained from both laboratory experiments and computational prediction.⁴⁵ A 2-level support vector machine (SVM) is used in CELLO, which involves 4 SVM classifiers and the final assignment is determined by using the jury votes from these classifiers.^{20,43}

Determination of transmembrane proteins

TMHMM 2.0 and HMMTOP 2.0 at default parameters were performed for the prediction of transmembrane helices and topology of the HPs in the study.^{46,47} SignalP 5.0 helped to predict the presence of signal peptides and cleavage site location, which performs through a neural network architecture involving a conditional random field.⁴⁸

Physicochemical prediction parameters

To compute several physical and chemical parameters of the HPs, such as molecular weight, theoretical pI, amino acid composition, instability index, aliphatic index, extinction

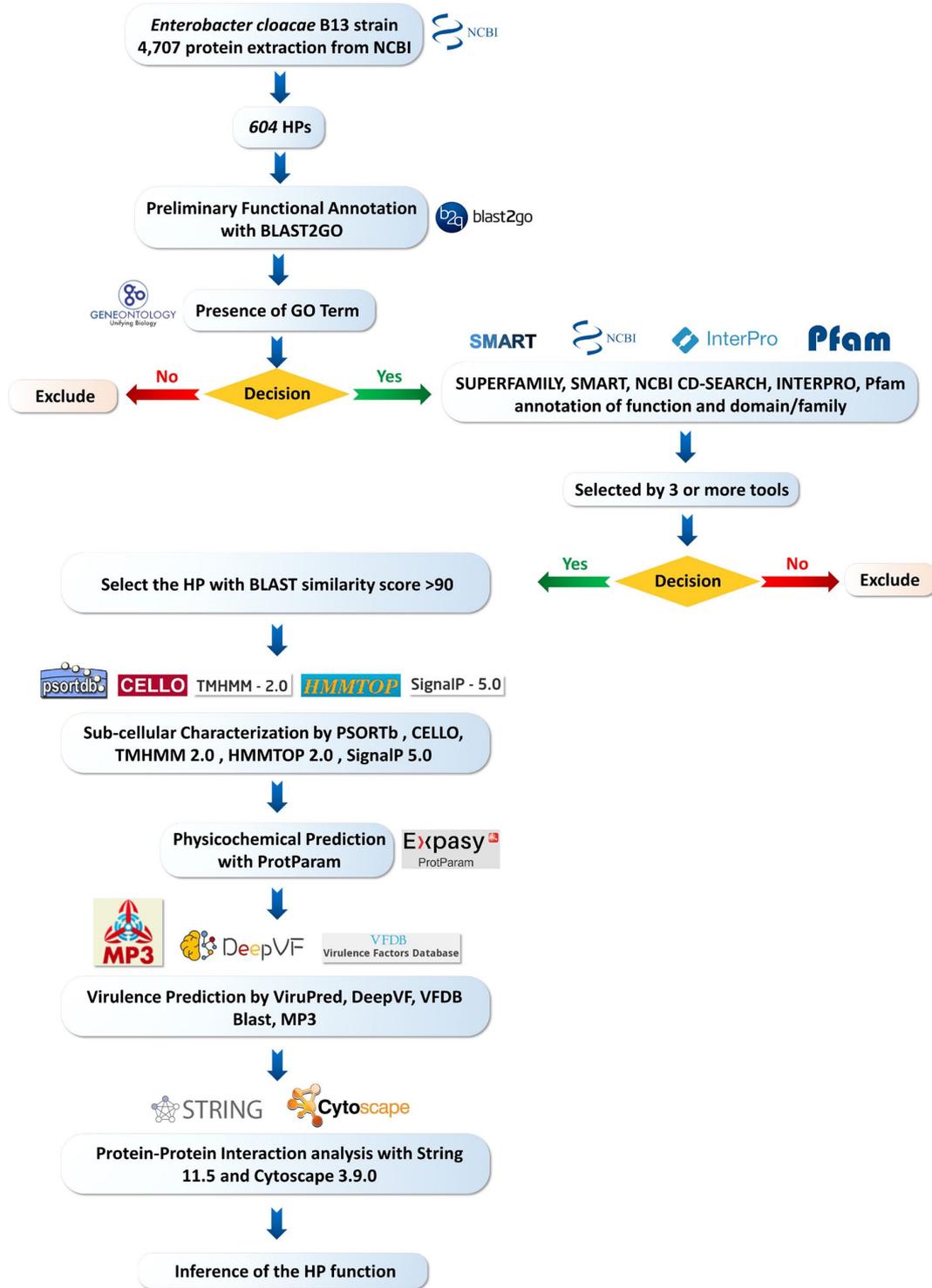


Figure 1. Methodology overview flowchart for the functional annotation and analysis of *E. cloacae* B13 strain hypothetical proteins (HP).

coefficient, and grand average of hydropathicity (GRAVY), ProtParam tool in ExPASy was used.⁴⁹

Virulent HP detection

MP3 is a tool which can accurately predict virulent proteins in genomic and metagenomic data using SVM and HMM approach.⁵⁰ DeepVF uses a deep learning-based hybrid framework to identify virulence factors more accurately by relying on

machine learning.⁵¹ Blast search tool in the Virulence Factor Database (VFDB) identified various virulent factors from the submitted HPs. VFDB contains information about virulent factors from several bacterial pathogens.⁵² Finally, PHI-base was used for virulent factors detection as it contains curated information on pathogen-host interaction affecting genes based on research articles. Only lethal and hypervirulence proteins were selected after completing a blast search (PHI-BLAST) in Phi-base against PHI-base 4.12 protein

sequences.⁵³ Virulent HPs predicted by 2 or more tools were then identified and further analyzed.

Predictions of antigenicity, allergenicity, and toxicity index

The antigenicity of the virulent proteins was predicted using the VaxiJen v2.0 server⁵⁴ and the ANTIGENpro server.⁵⁵ Toxicity and allergenicity of those proteins were predicted using the ToxIBTL server⁵⁶ and the AllerCatPro v. 2.0 server,⁵⁷ respectively.

Protein-protein interaction

String 11.5 database was utilized to predict the protein-protein interactions (PPIs) for the proteins from *E. cloacae* B12 strain.⁵⁸ String database uses all publicly available PPI information to computationally predict direct (physical) and indirect (functional) forecast. To ensure the most reliable PPIs, only the interactions with score values above 0.700 (high confidence) and high FDR stringency (1%).²⁰

String 11.5 search for *E. cloacae* was completed against *E. cloacae* ATCC 13047 as the strain with highest similarities. Here, the identified interactions were transferred to *E. cloacae* by the interolog mapping method, which assumes that when 2 proteins interact, their orthologous pairs will interact too.^{42,59-61} So, hierarchically arranged orthologous groups relations are applied by STRING to transfer association between applicable organisms as described in egglog.⁶²⁻⁶⁴

The network analyzer plugin in Cytoscape 3.9.0 program was utilized for the validation of PPI networks.^{65,66} Cytoscape 3.9.0 was used to obtain a better visualization of the potential virulent HPs with other proteins and among themselves. In Cytoscape, protein molecules are assigned to nodes and molecular interactions to edges. Furthermore, network analyzer tool can compute multiple network topological parameters with details of node degrees, edges, neighbor interactions, and network characteristics.

Results and Discussion

*Functional annotation of *E. cloacae* HPs reveals their association with several biological, molecular, and cellular processes*

A total of 604 proteins out of 4707 (12.83%) were labeled as HP in the *E. cloacae* B13 strain. Functional annotation of this large portion of proteins encoded by the bacterial genome was performed using various bioinformatic tools.

BLAST2GO was utilized to perform a primary prediction of the HPs, which returned 214 HPs with known protein domains or families along with their GO IDs (Supplementary Table 1). Further analysis of the pool of 214 HPs with NCBI Batch CD-Search, SMART, SUPERFAMILY, Pfam, and INTERPRO tools was performed to assign the functions (Supplementary Table 2). Among 214 HPs, functional

characterization with strong confidence was possible for 78 HPs as they demonstrated similar functions predicted by 3 or more tools. NCBI BLASTp tool was used to manually annotate the functions of these 78 HPs according to their homologous proteins (Table 1). Multiple tools increase the reliability of the functional prediction. Moreover, as domains are protein's fundamental unit of structure, folding and function, domain identification is crucial for annotating biological functions of a protein.⁶⁷

GO function analysis

Analysis of predicted GO terms for 78 HPs revealed their association in different GO categories: biological process, cellular components, and molecular functions (Figure 2). For biological process, 34 proteins were identified with distinct GO terms. About 12 of them were involved in protein transport and 18 proteins had functions in metabolic process. The cellular component category had 47 different GO terms, among which 38 were an intrinsic part of the membrane. Finally, among 53 GO terms in molecular functions, 33 were enzymes and 22 proteins were binding proteins.

Enzymes

Enzymes produced by gram-negative bacteria play a significant role in their host as they provide support and nutrients for growth, ensure favorable growth by modifying local environment, conduct the pathogenesis of several infections and help in metabolism.⁶⁸ A total of 33 proteins were characterized as enzymes, among which 15 proteins are hydrolase and 11 proteins are transferase.

Analysis of several infections by gram-negative anaerobes, involving tissue invasion and inflammation, necrosis, or supuration, has revealed that hydrolytic enzymes have roles in pathogenesis of infection.⁶⁸ Furthermore, study of different hydrolases has supported their potential role in pathogenesis.⁶⁹⁻⁷² Four proteins TOZ47235.1, TOZ47607.1, TOZ41437.1 and TOZ48018.1 were identified as the α/β hydrolase that are likely to be involved in the immune system evasion and modulation, detoxification, and metabolic adaptation.⁷³ The α/β hydrolases have also been found to play a major role as virulence factors in *Mycobacterium tuberculosis* and *Staphylococcus aureus*.^{73,74} TOZ40232.1 was predicted as endonuclease protein which functions by stopping the invasion of foreign DNA.⁷⁵ TOZ47235.1 was also predicted to have oxidoreductase activity, which are critical for bacterial virulence and pathogenesis.⁷⁶ Bromoperoxidase A2 protein is a metal-ion-free oxidoreductase from *Streptomyces aureofaciens*, which has both hydrolase and oxidoreductase activity and, has been found to contain an alpha/beta hydrolase fold.⁷⁷

Similarly, 11 proteins were identified as transferase enzymes. They are necessary for lipoprotein biosynthesis, spore germination, and aid the full virulence of bacteria.⁷⁵ TOZ44254.1 was

Table 1. Functionally annotated hypothetical proteins and their homologous accession from *E. cloacae* B13 strain.

HYPOTHETICAL PROTEIN ID	ACCESSION ID	SIMILARITY PERCENTAGE	DESCRIPTION
TOZ50232.1	WP_040020633.1	98.58	5-Oxoprolinase subunit PxpB
TOZ48266.1	WP_139936475.1	99.6	ABC transporter 6-transmembrane domain-containing protein
TOZ47228.1	WP_020689585.1	99.72	Acyltransferase family protein
TOZ47235.1	WP_013097106.1	99.89	Alpha/beta hydrolase
TOZ47607.1	WP_139936593.1	99.31	Alpha/beta hydrolase
TOZ41437.1	KAA3572232.1	99.63	Alpha/beta hydrolase
TOZ48018.1	WP_139936503.1	99.71	Alpha/beta hydrolase
TOZ47572.1	WP_139936577.1	99.61	Alpha-2-macroglobulin family protein
TOZ43519.1	WP_038420663.1	99.69	B3/4 domain-containing protein
TOZ46603.1	WP_094084617.1	99.86	Bestrophin family protein
TOZ50233.1	WP_139936181.1	99.45	Biotin-dependent carboxyltransferase family protein
TOZ46360.1	WP_139936869.1	99.68	CDP-alcohol phosphatidyltransferase family protein
TOZ46208.1	WP_032618344.1	99.7	Cellulase family glycosylhydrolase
TOZ50581.1	WP_023619766.1	99.24	CidA/LrgA family protein
TOZ48809.1	WP_139936311.1	99.8	CNNM family cation transport protein YoaE
TOZ45909.1	WP_001284076.1	99.95	Conjugative transfer system coupling protein TraD
TOZ48835.1	WP_139936314.1	99.53	Copper homeostasis membrane protein CopD
TOZ46383.1	WP_014831497.1	99.7	CTP synthase
TOZ40438.1	WP_045295673.1	100	Cu(+)/Ag(+) efflux RND transporter outer membrane channel SilC
TOZ41378.1	WP_139937455.1	99.72	Cu(+)/Ag(+) efflux RND transporter outer membrane channel SilC
TOZ43300.1	WP_139937218.1	99.62	Diguanylate cyclase
TOZ45347.1	WP_139936994.1	98.96	Rhodanese family protein
TOZ50537.1	WP_139936148.1	99.8	LPS biosynthesis-modulating metalloenzyme YejM
TOZ44410.1	WP_139937108.1	99.91	EAL domain-containing protein
TOZ49766.1	WP_139936297.1	99.91	EAL domain-containing protein
TOZ48059.1	WP_139936525.1	99.46	Efflux transporter outer membrane subunit
TOZ43261.1	TAT64716.1	98.1	Fatty acid desaturase
TOZ46043.1	TOZ46043.1	99.57	Fimbrial biogenesis outer membrane usher protein
TOZ50430.1	WP_013098160.1	99.87	Formate/nitrite transporter family protein
TOZ46589.1	WP_139936830.1	99.57	Four-carbon acid sugar kinase family protein
TOZ46922.1	WP_139936751.1	99.25	FUSC family protein
TOZ48410.1	WP_013099057.1	99.92	FUSC family protein
TOZ46924.1	WP_139936753.1	96.99	Gamma-glutamylcyclotransferase
TOZ48031.1	WP_139936511.1	99.58	Glycoside hydrolase family 10 protein
TOZ44573.1	WP_139937063.1	99.63	Glycoside hydrolase family 127 protein
TOZ48179.1	WP_032663354.1	98.59	Helix-turn-helix domain-containing protein

(Continued)

Table 1. (Continued)

HYPOTHETICAL PROTEIN ID	ACCESSION ID	SIMILARITY PERCENTAGE	DESCRIPTION
TOZ42186.1	WP_139937356.1	99.52	HlyD family efflux transporter periplasmic adaptor subunit
TOZ40232.1	WP_139937487.1	98.28	HNH endonuclease
TOZ44254.1	WP_205919421.1	99.17	Glycosyltransferase family 8 protein
TOZ44516.1	WP_139937082.1	99.78	Ig-like domain-containing protein
TOZ42967.1	WP_013097375.1	95.65	Inner membrane protein YbjM
TOZ41165.1	WP_007898888.1	99.69	IS66 family insertion sequence element accessory protein TnpB
TOZ42178.1	WP_029881878.1	99.63	Lipocalin family protein
TOZ44728.1	WP_139937026.1	98.92	Lipocalin family protein
TOZ49934.1	WP_130626454.1	99.43	Lysine exporter LysO family protein
TOZ47092.1	WP_063266501.1	99.44	Mannosyl-3-phosphoglycerate phosphatase family
TOZ48860.1	WP_139936319.1	99.08	MAPEG family protein
TOZ49982.1	WP_139936224.1	99.37	MBL fold metallo-hydrolase
TOZ48821.1	WP_013096123.1	98.86	Membrane protein
TOZ48306.1	WP_139936372.1	99.77	MMPL family transporter
TOZ44773.1	WP_023619713.1	99.63	OapA family protein
TOZ40775.1	WP_139937470.1	98.95	OmpA family protein
TOZ48307.1	WP_139936373.1	99.08	Outer membrane lipoprotein carrier protein LolA
TOZ49630.1	WP_013098867.1	99.73	Phage holin family protein
TOZ45869.1	WP_021242119.1	99.97	Predicted P-loop ATPase
TOZ47672.1	WP_020690601.1	97.96	Putative DNA-binding transcriptional regulator
TOZ48919.1	WP_139936331.1	97.15	Putative SAM-dependent methyltransferases/O-linked N-acetylglucosamine transferase
TOZ38888.1	WP_119149238.1	100	Pyridoxal-phosphate dependent enzyme
TOZ42027.1	WP_139937387.1	93.68	Recombinase family protein
TOZ42856.1	WP_013095928.1	98.21	Ribosome-associated protein YbcJ
TOZ48897.1	WP_013096040.1	99.79	RpiB/LacA/LacB family sugar-phosphate isomerase
TOZ48801.1	WP_013096141.1	99.51	Slp family lipoprotein
TOZ42378.1	WP_139937321.1	99.47	Sugar phosphate isomerase/epimerase
TOZ50441.1	WP_013098149.1	99.89	Sulfite exporter TauE/SafE family protein
TOZ49620.1	WP_139936258.1	99.89	TerC family protein
TOZ44280.1	WP_029882175.1	99.07	Threonine/serine exporter ThrE family protein
TOZ42301.1	WP_139937331.1	98.8	Topoisomerase DNA-binding C4 zinc finger domain-containing protein
TOZ45926.1	WP_139936926.1	99.26	Translesion error-prone DNA polymerase V subunit UmuC
TOZ46041.1	WP_139936907.1	99.57	Type 1 fimbrial protein
TOZ42167.1	WP_139937345.1	98.18	Type I addiction module toxin, SymE family
TOZ48195.1	TOZ48195.1	97.5	Zinc ion binding protein
TOZ50295.1	WP_139936192.1	99.49	Uridine diphosphate-N-acetylglucosamine-binding protein YvcK

(Continued)

Table 1. (Continued)

HYPOTHETICAL PROTEIN ID	ACCESSION ID	SIMILARITY PERCENTAGE	DESCRIPTION
TOZ47361.1	WP_139936618.1	99.4	YadA-like family protein
TOZ46379.1	WP_139936875.1	98.92	YbaK/prolyl-tRNA synthetase associated domain-containing protein
TOZ50305.1	WP_014831155.1	99.12	YbhQ family protein
TOZ47583.1	WP_013098275.1	99.71	YfgM family protein
TOZ47636.1	WP_060578689.1	99.92	YpfN family protein
TOZ49631.1	WP_013098868.1	99.49	YqjK-like family protein

predicted as glycosyl transferase protein. Glycosyl transferase family proteins can alter extracellular polysaccharide and lipopolysaccharide synthesis upon mutation, resulting in the reduction in disease symptoms.^{78,79} TOZ46360.1 and TOZ50295.1 were predicted to be CDP-alcohol phosphatidyl-transferase family protein and UDP-GlcNAc. Both of the families are associated with lipid biosynthesis.⁸⁰⁻⁸² Alteration of the synthesized phospholipid has a crucial role in virulence and several human diseases.^{83,84}

TOZ38888.1 and TOZ41165.1 were predicted as lyase enzymes. Lyase enzymes have essential functions for the virulence of pathogenic gram-negative bacteria in host.⁶⁸ TOZ38888.1 is pyridoxal-phosphate (PLP)-dependent enzyme, which are a ubiquitous class of biocatalysts. In several free-living prokaryotes, PLP-dependent enzymes are encoded by almost 1.5% of all genes.⁸⁵ PLP-dependent enzymes with desulphydrase activity help in amino-acid metabolism, adaption to nutrient sources in a new environment, and sometimes can function as virulence factors.^{86,87}

TOZ48897.1 was annotated as RpiB/LacA/LacB family sugar-phosphate isomerase. This family of proteins takes part in the lactose catabolism pathway.⁸⁸

Binding proteins

There are 22 proteins characterized as binding proteins, among which 5 proteins were DNA binding, 3 were RNA binding and 5 were ATP binding ones. HPs with DNA-binding function can contribute to the virulence by altering the expression of virulence factors, which have been observed during *S aureus* infection.⁸⁹ TOZ48179.1 was characterized as helix-turn-helix domain-containing protein. HTH domain containing protein has a large range of functions, such as, DNA repair and replication, RNA metabolism, PPI and 2-component signaling pathway, while 2 component signal transduction system (TCS) are largely used as a target for antimicrobial therapy.⁹⁰⁻⁹² TOZ45926.1 was a translesion error-prone DNA polymerase and TOZ42027.1 was a recombinase family protein, both of which might have functions in DNA repair.^{93,94}

TOZ46383.1 was found to be a CTP synthase, which converts UTP to CTP, a necessary step in pyrimidine metabolic pathway in community-acquired respiratory tract infection (RTI) causing bacteria.⁹⁵ In addition, TOZ48266.1 was identified as an ABC transporter 6-transmembrane domain-containing protein, which are considered to have roles in nutrient uptake and drug resistance. Moreover, evidence of ABC transporters being directly or indirectly involved in the bacterial virulence has been found.⁹⁶ Furthermore, TOZ50233.1 was characterized as a biotin-dependent carboxyltransferase protein. They have roles in fatty acid, amino acid and carbohydrates metabolism.⁹⁷⁻¹⁰⁰ Furthermore, their activity plays important role in the virulence of organisms like *Listeria monocytogenes* and *Candida albicans*.¹⁰¹⁻¹⁰³

Transporter proteins

Eight proteins were characterized to have transmembrane transporter activity. TOZ50430.1 was characterized as formate/nitrite transporter (FNT) protein. Bacterial FNTs monitor the transport of small monoacids.¹⁰⁴ In addition, FNTs can perform as a virulence factor in *Salmonella* species by helping the bacteria to evade killing from activated macrophages in host.¹⁰⁵ TOZ40438.1, TOZ41378.1, and TOZ48059.1 are efflux transmembrane transporter proteins, and the first 2 are Cu(+)/Ag (+) efflux RND transporters. RND transporters are necessary for the multidrug resistance in several pathogens.¹⁰⁶ Furthermore, RND superfamily transporters are organized as tripartite efflux complexes and span inner and outer membrane of cell envelope.¹⁰⁷ Moreover, RND transporters specific to heavy metals in *E coli* have been found to raise resistance to copper(I) and silver(I) ions.¹⁰⁸

Regulatory proteins

Regulatory process is a complex network system in bacteria that helps in various gene expression and maintain bacterial pathogenesis, growth, and survival.² TOZ43300.1 was identified as a diguanylate cyclase which has functions in cellular process regulation and signal transduction. Interestingly,

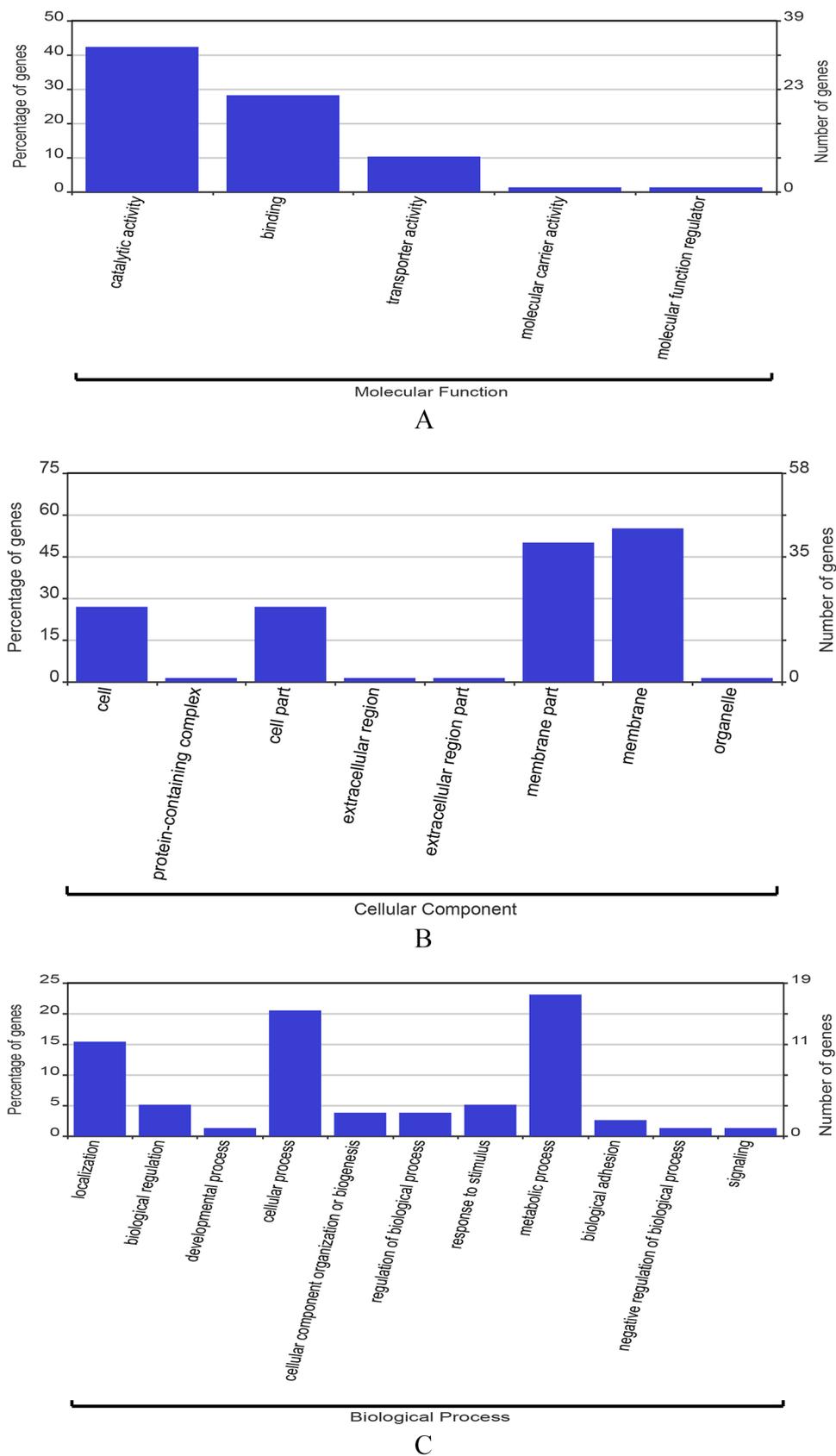


Figure 2. GO categories distribution of the HPs from *E. cloacae* B13 strain. The figure shows the detailed categories, number and percentage of proteins in (A) molecular functions (B) cellular components (C) biological processes. GO indicates Gene Ontology; HP, hypothetical proteins.

diguanylate cyclase is necessary for biofilm development. It also performs as a messenger for bacterial virulence, motility, adhesion, secretion, and community behavior.¹⁰⁹

TOZ47572.1 was predicted as an alpha-2-macroglobulin (A2M) protein, which can structurally mimic proteins of eukaryotic innate immunity in invasive bacteria. Bacterial A2M are located in periplasm where they trap external proteases and provide cellular protection.¹¹⁰ Both pathogenically invasive and saprophytically colonizing species possess A2M and mostly exploit higher eukaryotes as hosts. Therefore, bacterial A2M can be used as useful targets to increase vaccine efficacy in infections.¹¹¹

Membrane protein

A total of 38 proteins were characterized as integral component of the membrane and 1 protein as extrinsic component of the membrane. TOZ40775.1 was annotated as OmpA family protein. This family of proteins is surface-exposed porin proteins with anti-parallel β barrels in the outer membrane.¹¹² HMMTOP and TMHMM also predicted the presence of transmembrane helices for this protein (Supplementary Table 4). Several pathogenic roles including adhesion, invasion, intracellular survival, and host defenses have been assigned to OmpA. In various cases, OmpA proteins are being considered as potential vaccine candidates.¹¹²

TOZ49620.1 was annotated as a TerC family protein. This type of protein is largely found in bacteria species and may influence host-pathogen interaction.¹¹³ Moreover, TerC family proteins in *Bacillus subtilis* have been found to help prevent Manganese (Mn) intoxication. Mn is essential for virulence for many pathogens.¹¹³ Mn detoxification helps in oxidative stress resistance and virulence in *S aureus*.¹¹⁴

TOZ44410.1 and TOZ49766.1 were both characterized as EAL domain-containing proteins. EAL domain is a ubiquitous signal transduction protein domain involved in hydrolysis of second messenger cyclic dimeric GMP (c-di-GMP) as it is the exclusive substrate of EAL.^{114,115} The second messenger c-di-GMP regulates many lifestyle aspects and virulence of several gram-negative bacteria.¹¹⁶ Moreover, EAL domain protein VieA from *Vibrio cholerae* inversely regulate biofilm-specific genes (*vps*) and virulence genes like *ctxA* by decreasing the amount of cellular c-di-GMP. This phenomenon is of particular interest as the shift in gene expression plays a major role in *V. cholerae* life cycle.¹¹⁷ Upon entering to a host, *V. cholerae* tends to undergo a shift in gene expression, where *vps* expression ceases¹¹⁸ and virulence genes are expressed.¹¹⁹⁻¹²³

Virulent protein prediction

MP3, DeepVF, VFDB, and PHI-base were used for virulence factor prediction with high confidence level. A total of 23 HPs were predicted by 2 or more tools to be virulent, and the

remaining HPs were identified by either only one tool or not virulent at all (Supplementary Table 3). As virulence factors help bacteria to colonize and cause disease, the knowledge of biological function and mechanism of the virulence factors is necessary to understand their role in the pathogenesis of bacteria.² Moreover, virulent factors are potential therapeutic targets in case of bacterial infections.¹²⁴ Characterizing virulence factors include several secretion systems (Type I to Type VI secretory systems) 2-component signal transduction systems, quorum sensing, and biofilm formation.^{125,126} Virulent proteins are utilized by a large number of pathogenic bacteria, and therefore identifying inhibitors against essential factors for virulence factors is a new research interest, which is a different molecular approach than traditional drug discovery.¹²⁷ Annotated virulent HPs can obtain a better target-based approach and aid against bacterial infections as a subsidiary therapy to different antibiotics.¹²⁵

Virulent HPs with therapeutic potential

Antigenicity of the virulent HPs was studied, and it was observed that 7 of them have antigenic potential. All of these 7 proteins are likely to be non-allergenic and nontoxic. The subcellular localization of the protein was also explored, and we observed that the 7 antigenic proteins were either membrane bound or periplasmic proteins (Table 2). Our findings suggest that each of these 7 proteins could be a great candidate for vaccine development.¹²⁸⁻¹³¹

Subcellular localization and physiochemical prediction

In the study, amino acid sequences of 78 HPs were analyzed by using various tools, such as PSORTb v3.0, CELLO v.2.5, TMHMM 2.0, HMMTOP 2.0 and ProtParam for assessing their subcellular location along with physiochemical prediction (Supplementary Table 4). However, more attention was paid to the virulent HPs that were predicted to have roles in pathogenesis.

The cellular location along with secretion or signaling ability and transmembrane helices of the 23 HPs were predicted. Nine of them were found to have transmembrane helices predicted by both HMMTOP and TMHMM (TOZ49766.1, TOZ48809.1, TOZ40775.1, TOZ44410.1, TOZ43300.1, TOZ45909.1, TOZ49630.1, TOZ47361.1, and TOZ42186.1). About 19 proteins out of 23 were predicted by CELLO to be an inner or outer membrane and periplasmic proteins. However, pSORTdb predicted 9 proteins as cytoplasmic or cytoplasmic membrane proteins, and 7 proteins as outer membrane proteins. The SignalP 5.0 server predicted 10 proteins out of 23 to contain signal peptides for several secretion pathways. About five of them were predicted to be standard secretory signal peptides and cleaved by Signal Peptidase I. In addition, 5 more

Table 2. Prediction of antigenicity, allergenicity, toxicity, and subcellular localization of the virulent HPs.

VIRULENT HP	ANTIGENICITY	ALLERGENICITY	TOXICITY	SUBCELLULAR LOCALIZATION
TOZ40438.1	Nonantigen	Nonallergen	Nontoxin	Outer membrane
TOZ40775.1	Nonantigen	Allergen	Nontoxin	Inner membrane
TOZ41378.1	Nonantigen	Nonallergen	Nontoxin	Outer membrane
TOZ42178.1	Nonantigen	Nonallergen	Nontoxin	Periplasmic
TOZ42186.1	Antigen	Nonallergen	Nontoxin	Outer membrane
TOZ42301.1	Nonantigen	Nonallergen	Nontoxin	Periplasmic
TOZ43300.1	Nonantigen	Nonallergen	Nontoxin	Cytoplasmic
TOZ44410.1	Nonantigen	Nonallergen	Nontoxin	Inner membrane
TOZ44516.1	Antigen	Nonallergen	Nontoxin	Outer membrane
TOZ44728.1	Nonantigen	Allergen	Nontoxin	Periplasmic
TOZ44773.1	Nonantigen	Nonallergen	Nontoxin	Periplasmic
TOZ45909.1	Nonantigen	Nonallergen	Nontoxin	Outer membrane
TOZ46041.1	Antigen	Nonallergen	Nontoxin	Outer membrane
TOZ46043.1	Antigen	Nonallergen	Nontoxin	Outer membrane
TOZ46589.1	Nonantigen	Nonallergen	Nontoxin	Cytoplasmic
TOZ47361.1	Antigen	Nonallergen	Nontoxin	Periplasmic
TOZ47572.1	Antigen	Nonallergen	Nontoxin	Periplasmic
TOZ48059.1	Nonantigen	Nonallergen	Nontoxin	Outer membrane
TOZ48307.1	Nonantigen	Nonallergen	Nontoxin	Cytoplasmic
TOZ48809.1	Nonantigen	Nonallergen	Nontoxin	Inner membrane
TOZ48919.1	Nonantigen	Nonallergen	Nontoxin	Cytoplasmic
TOZ49630.1	Nonantigen	Nonallergen	Nontoxin	Inner membrane
TOZ49766.1	Antigen	Nonallergen	Nontoxin	Inner membrane

Abbreviation: HPs, hypothetical proteins.

proteins were predicted to be lipoprotein signal peptides and cleaved by Signal Peptidase II. All ten proteins were predicted to be transported by the Sec translocon.

The pH at which no net electric charge of a molecule remains and does not move in an electric field of direct current is the theoretical pI.^{132,133} For the virulent proteins, the theoretical pI ranged from 4.58 to 9.47. Again, these 23 virulent HPs molecular weight ranged from 11390.68 to 179998.3. 2D gel electrophoresis visualization in laboratorial experiments can be accompanied by the combination of these 2 parameters. The extinction coefficient of the virulent HPs at 280nm ranged from 8450 to 228165 M⁻¹ cm⁻¹ with respect to the Cys (cysteine), Trp (tryptophan), and Tyr (tyrosine) concentration. The extinction coefficient indicates the amount of light absorbent by a protein at a specific wavelength, which is useful for purifying and separating a protein in spectrophotometer. In addition, high extinction coefficient occurred in some HPs

because of the presence of high concentration of Cys, Trp, and Tyr.^{132,134,135} The instability index estimates the stability of a protein in test tubes. Proteins with less than 40 instability index are predicted as stable proteins.¹³⁶ In the study, the 23 predicted virulent HPs instability index ranged from 20.3 to 59.09, and 16 out of 23 proteins were stable. Stable proteins have a longer half-life. The half-lives of several virulent effector proteins are integral to their function. For example, in *Salmonella*, virulent effector proteins' half-life modulations are necessary for the pathogenic cellular functions.¹³⁷ Aliphatic index of a protein determines the relative volume obtained by the aliphatic side chains (alanine, valine, isoleucine, and leucine). Aliphatic index functions as a positive factor for the thermostability increase of globular proteins.¹³⁸ The aliphatic index of the virulent HPs in the study ranged from 60.98 to 124.62. Finally, the Grand Average of Hydropathy (GRAVY) value of a protein is the sum of hydropathy values of all amino acids divided by the

help bacterial survival against antibiotics and in virulence.^{125,145-147} This protein also interacts with at least 5 cus proteins (cusA, cusB, cusF, cusR, cusS). Cus protein complex helps in maintaining copper homeostasis and mediates resistance to copper stress by cation efflux.^{148,149} Toxic properties of copper are often harnessed by the innate immune system, which helps the host to kill bacteria. Bacteria counter this defense by relying on genes for copper tolerance for virulence within the host.¹⁵⁰

The proteins TOZ41378.1 and TOZ40438.1 are Cu(+)/Ag (+) efflux RND transporter outer membrane proteins and demonstrated interactions with 18 and 16 proteins, respectively (Fig S1). They strongly interact with each other along with TOZ48059.1 and most of its interactive proteins. These 3 proteins remain in one cluster. The protein cluster appears to bear the function of 2 component regulatory system with high strength (Log10 observed/expected value is 1.43). Majority of the interacting proteins also contain Histidine kinase domain, and GAF domain, which are associated with osmoregulation, hyphal development and virulence in bacteria like *Agrobacterium tumefaciens* and *Candida albicans*.¹⁵¹⁻¹⁵⁴

TOZ48307.1 is an outer membrane lipoprotein carrier protein, which interacts with 15 other proteins. These proteins also form a cluster (Fig S1). TOZ48307.1 interacts with 4 acyl carrier proteins (ACP) (ECL_04843, ECL_04852, ECL_048550, and ECL_04854). In *Pseudomonas aeruginosa*, Acp3 has been found to be involved in oxidative stress response and Acp1 and Acp3 each contribute to the virulence.¹⁵⁵ Only Acp1 functions in the fatty acid synthesis in *P aeruginosa*¹⁵⁶ and fatty acid plays a multifactorial role in controlling bacterial viability and virulence in this organism.¹⁵⁷⁻¹⁵⁹

Finally, TOZ43300.1, which was predicted as a diguanylate cyclase, interacts with 115 different proteins (Figure 3). Interacting proteins were mostly related to 2-component regulatory system, biofilm formation, diguanylate cyclase activity, and intracellular signal transduction. Environment factors helps to induce bacterial biofilm formation, which are microbial multicellular communities encased within extracellular matrix. Two-component signal transduction system (TCS) strategy is used by bacteria to connect input signals change in environment to changes in physiological output, and coordinate input signals to control biofilm formation.¹⁶⁰ In several *E. cloacae* outbreaks, biofilm formation has been suspected of contributing to its pathogenicity.^{4,161,162} This is alarming because biofilms can show resistance against antibiotics in nosocomial infections and almost 65% of the microbial and 80% of the chronic infections are related to biofilm formation.¹⁶³ In addition, fight against *E. cloacae* in NICU infection and outbreaks is still a major challenge.¹⁶⁴ Moreover, many interacting proteins strengthen the functional prediction of the protein.

Conclusions

Hypothetical proteins form a large portion of a bacterial proteome which play crucial biological roles. Identifying these

proteins and their functional annotation will help us to understand about the organism in a better way. For this study, 78 HPs were from *E. cloacae* B13 strain was functionally annotated with high confidence. The pipeline used in the study obtained a great result and can be followed to assign function to HPs from different organisms. Most proteins were predicted as enzymes, binding proteins, transporter proteins, regulatory proteins or membrane proteins, and their subcellular localization and physiochemical parameters were crucial to the understanding of their characteristics. As *E. cloacae* is responsible for several outbreaks in hospitals and multidrug-resistant *E. cloacae* complex are an emerging global threat, the HPs were analyzed for their role in virulence. We identified several proteins which have potential role in virulence by incorporating antibiotic resistance activity, biofilm formation, quorum sensing, secretion pathway or others. The potential virulent proteins were further investigated for their interaction with other proteins. PPI helped to determine the relationship between these proteins and the known proteome of *E. cloacae*, which also strengthened our prediction of the virulence. Findings from this study will eventually help to fill the gaps in the proteome knowledge of *E. cloacae* and create the possibility to fight against nosocomial infections and NICU outbreaks caused by *E. cloacae*.

Acknowledgements

The authors acknowledge high performance computing facility support from Centre for Bioinformatics Learning Advancement and Systematics Training (cBLAST), University of Dhaka. The authors also acknowledge support of Biomolecular Research Foundation (BMRF), Dhaka, Bangladesh.

Author Contributions

ABMMKI conceived the project. SD, SS, and AT collected the data; SD, SS, and MAZ performed the analyses. SD, SS, MAZ, and ABMMKI wrote the manuscript. The manuscript was reviewed and approved by all authors.

Data and Software Availability

All data added Table, figures and supplementary file and supplementary tables. In this research work publicly available free mostly online and few offline software/tools were used. Necessary link, reference of the software/tools provided in the method section.

Supplemental Material

Supplemental material for this article is available online.

REFERENCES

1. Davin-Regli A, Pagès JM. *Enterobacter aerogenes* and *Enterobacter cloacae*; versatile bacterial pathogens confronting antibiotic treatment. *Front Microbiol.* 2015;6:392. doi:10.3389/fmicb.2015.00392.
2. Prabhu D, Rajamanikandan S, Anusha SB, Chowdary MS, Veerapandiyam M, Jeyakanthan J. In silico functional annotation and characterization of hypothetical proteins from *Serratia marcescens* FG194. *Biol Bull Russ Acad Sci.* 2020;47:319-331. doi:10.1134/S1062359020300019.

3. Sanders WE Jr, Sanders CC. Enterobacter spp.: pathogens poised to flourish at the turn of the century. *Clin Microbiol Rev.* 1997;10:220-241. doi:10.3389/fmicb.2015.00392.
4. Mezzatesta ML, Gona F, Stefani S. Enterobacter cloacae complex: clinical impact and emerging antibiotic resistance. *Future Microbiol.* 2012;7:887-902. doi:10.2217/fmb.12.61.
5. Dalben M, Varkulja G, Basso M, et al. Investigation of an outbreak of Enterobacter cloacae in a neonatal unit and review of the literature. *J Hosp Infect.* 2008;70:7-14. doi:10.1016/j.jhin.2008.05.003.
6. Dealler S. Enterobacter cloacae infection and colonization in a neonatal intensive care unit. *J Hosp Infect.* 2002;52:72-73. doi:10.1053/jhin.2002.1265.
7. Pessoa-Silva CL, Richtmann R, Calil R, et al. Healthcare-associated infections among neonates in Brazil. *Infect Control Hosp Epidemiol.* 2004;25:772-777. doi:10.1086/502475.
8. van den Berg RWA, Claahsen HL, Niessen M, Muyltjens HL, Liem K, Voss A. Enterobacter cloacae outbreak in the NICU related to disinfected thermometers. *J Hosp Infect.* 2000;45:29-34. doi:10.1053/jhin.1999.0657.
9. Hoffmann H, Roggenkamp A. Population genetics of the nomenclatures *Enterobacter cloacae*. *Appl Environ Microbiol.* 2003;69:5306-5318. doi:10.1128/AEM.69.9.5306-5318.2003.
10. Paauw A, Caspers MPM, Schuren FHJ, et al. Genomic diversity within the *Enterobacter cloacae* complex. *PLoS ONE.* 2008;3:e3018. doi:10.1371/journal.pone.0003018.
11. Lee SO, Kim YS, Kim BN, Kim MN, Woo JH, Ryu J. Impact of previous use of antibiotics on development of resistance to extended-spectrum cephalosporins in patients with enterobacter bacteremia. *Eur J Clin Microbiol Infect Dis.* 2002;21:577-581. doi:10.1007/s10096-002-0772-7.
12. Barnes AI, Ortiz C, Paraje MG, Balanzino LE, Albesa I. Purification and characterization of a cytotoxin from *Enterobacter cloacae*. *Can J Microbiol.* 1997;43:729-733. doi:10.1139/m97-105.
13. Krzyżmińska S, Mokracka J, Koczura R, Kaznowski A. Cytotoxic activity of *Enterobacter cloacae* human isolates. *FEMS Immunol Med Microbiol.* 2009;56:248-252. doi:10.1111/j.1574-695X.2009.00572.x.
14. Stuber K, Frey J, Burnens AP, Kuhnert P. Detection of type III secretion genes as a general indicator of bacterial virulence. *Mol Cell Probes.* 2003;17:25-32. doi:10.1016/s0890-8508(02)00108-1.
15. Krzyżmińska S, Koczura R, Mokracka J, Puton T, Kaznowski A. Isolates of the *Enterobacter cloacae* complex induce apoptosis of human intestinal epithelial cells. *Microb Pathog.* 2010;49:83-89. doi:10.1016/j.micpath.2010.04.003.
16. Zogaj X, Bokranz W, Nimtz M, Römling U. Production of cellulose and curli fimbriae by members of the family Enterobacteriaceae isolated from the human gastrointestinal tract. *Infect Immun.* 2003;71:4151-4158. doi:10.1128/IAI.71.7.4151-4158.2003.
17. Kim SM, Lee HW, Choi YW, et al. Involvement of curli fimbriae in the biofilm formation of *Enterobacter cloacae*. *J Microbiol.* 2012;50:175-178. doi:10.1007/s12275-012-2044-2.
18. Annavajhala MK, Gomez-Simmonds A, Uhlemann A-C. Multidrug-resistant *Enterobacter cloacae* complex emerging as a global, diversifying threat. *Front Microbiol.* 2019;10:44. <https://www.frontiersin.org/article/10.3389/fmicb.2019.00044>
19. Davin-Regli A, Lavigne J-P, Pagès J-M. *Enterobacter* spp.: update on taxonomy, clinical aspects, and emerging antimicrobial resistance. *Clin Microbiol Rev.* 2019;32:e00002-19. doi:10.1128/CMR.00002-19.
20. da Costa WLO, Araújo CLA, Dias LM, et al. Functional annotation of hypothetical proteins from the *Exiguobacterium antarcticum* strain B7 reveals proteins involved in adaptation to extreme environments, including high arsenic resistance. *PLoS ONE.* 2018;13:e0198965. doi:10.1371/journal.pone.0198965.
21. Singh A, Singal B, Nath O, Singh IK. Functional annotation and classification of the hypothetical proteins of *Neisseria meningitidis* H44/76. *Am J Biosci Bioeng.* 2015;3:57-64. doi:10.11648/J.BIO.20150305.16.
22. Hoskeri HJ, Krishna V, Amruthavalli C. Functional annotation of conserved hypothetical proteins in *Rickettsia massiliae* MTU5. *J Comput Sci Syst Biol.* 2010;3:050-052. doi:10.4172/jcsb.1000055.
23. Shahbaaz M, Imtaiyaz Hassan M, Ahmad F. Functional annotation of conserved hypothetical proteins from *Haemophilus influenzae* Rd KW20. *PLoS ONE.* 2013;8:e84263. doi:10.1371/journal.pone.0084263.
24. Hawkins T, Kihara D. Function prediction of uncharacterized proteins. *J Bioinform Comput Biol.* 2007;5:1-30. doi:10.1142/S0219720007002503.
25. Doerks T, von Mering C, Bork P. Functional clues for hypothetical proteins based on genomic context analysis in prokaryotes. *Nucleic Acids Res.* 2004;32:6321-6326. doi:10.1093/nar/gkh973.
26. Islam MS, Shahik SM, Sohel M, Patwary NI, Hasan MA. In silico structural and functional annotation of hypothetical proteins of *Vibrio cholerae* O139. *Genomics Inform.* 2015;13:53-59. doi:10.5808/GI.2015.13.2.53.
27. Nimrod G, Schushan M, Steinberg DM, Ben-Tal N. Detection of functionally important regions in "hypothetical proteins" of known structure. *Structure.* 2008;16:1755-1763. doi:10.1016/j.str.2008.10.017.
28. Ijaq J, Chandra D, Ray MK, Jagannadham MV. Investigating the functional role of hypothetical proteins from an antarctic bacterium *Pseudomonas* sp. *Lz4W: emphasis on identifying proteins involved in cold adaptation.* *Front Genet.* 2022;13:825269. doi:10.3389/fgene.2022.825269.
29. Lubec G, Afjechi-Sadat L, Yang JW, John JP. Searching for hypothetical proteins: theory and practice based upon original data and literature. *Prog Neurobiol.* 2005;77:90-127. doi:10.1016/j.pneurobio.2005.10.001.
30. Minion FC, Lefkowitz EJ, Madsen ML, Cleary BJ, Swartzell SM, Mahairas GG. The genome sequence of *Mycoplasma hyopneumoniae* strain 232, the agent of swine mycoplasmosis. *J Bacteriol.* 2004;186:7123-7133. doi:10.1128/JB.186.21.7123-7133.2004.
31. Sayers EW, Bolton EE, Brister JR, et al. Database resources of the national center for biotechnology information. *Nucleic Acids Res.* 2022;50:D20-D26. doi:10.1093/nar/gkab1112.
32. Page AJ, Taylor B, Delaney AJ, et al. SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. *Microb Genomics.* 2016;2:e000056. doi:10.1099/MGEN.0.000056.
33. Götz S, García-Gómez JM, Terol J, et al. High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res.* 2008;36:3420-3435. doi:10.1093/nar/gkn176.
34. Letunic I, Khedkar S, Bork P. SMART: recent updates, new developments and status in 2020. *Nucleic Acids Res.* 2021;49:D458. doi:10.1093/nar/gkaa937.
35. Lu S, Wang J, Chitsaz F, et al. CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Res.* 2020;48:D265-D268. doi:10.1093/nar/gkz991.
36. Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD. HMMER web server: 2018 update. *Nucleic Acids Res.* 2018;46:W200-W204. doi:10.1093/nar/gky448.
37. Pandurangan AP, Stahlhacke J, Oates ME, Smithers B, Gough J. The SUPER-FAMILY 2.0 database: a significant proteome update and a new webserver. *Nucleic Acids Res.* 2019;47:D490-D494. doi:10.1093/nar/gky1130.
38. Mistry J, Chuguransky S, Williams L, et al. Pfam: the protein families database in 2021. *Nucleic Acids Res.* 2021;49:D412-D419. doi:10.1093/nar/gkaa913.
39. Blum M, Chang H-Y, Chuguransky S, et al. The InterPro protein families and domains database: 20 years on. *Nucleic Acids Res.* 2021;49:D344-D354. doi:10.1093/nar/gkaa977.
40. Heberle H, Meirelles GV, da Silva FR, Telles GP, Minghim R. InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. *BMC Bioinformatics.* 2015;16:169. doi:10.1186/s12859-015-0611-3.
41. Johnson M, Zaretskaya I, Raytselis Y, Merezukh Y, McGinnis S, Madden TL. NCBI BLAST: a better web interface. *Nucleic Acids Res.* 2008;36:W5-W9. doi:10.1093/nar/gkn201.
42. 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;13:1402-1406. doi:10.1110/ps.03479604.
43. Yu C-S, Chen Y-C, Lu C-H, Hwang J-K. Prediction of protein subcellular localization. *Proteins.* 2006;64:643-651. doi:10.1002/prot.21018.
44. 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;26:1608-1615. doi:10.1093/bioinformatics/btq249.
45. Peabody MA, Laird MR, Vlasschaert C, Lo R, Brinkman FSL. PSORTdb: expanding the bacteria and archaea protein subcellular localization database to better reflect diversity in cell envelope structures. *Nucleic Acids Res.* 2016;44:D663-D668. doi:10.1093/nar/gkv1271.
46. Möller S, Croning MD, Apweiler R. Evaluation of methods for the prediction of membrane spanning regions. *Bioinformatics.* 2001;17:646-653. doi:10.1093/bioinformatics/17.7.646.
47. Tusnády GE, Simon I. The HMMTOP transmembrane topology prediction server. *Bioinformatics.* 2001;17:849-850. doi:10.1093/bioinformatics/17.9.849.
48. Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, et al. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol.* 2019;37:420-423. doi:10.1038/s41587-019-0036-z.
49. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* 2003;31:3784-3788. doi:10.1093/nar/gkg563.
50. Gupta A, Kapil R, Dhakan DB, Sharma VK. MP3: a software tool for the prediction of pathogenic proteins in genomic and metagenomic data. *PLoS ONE.* 2014;9:e93907. doi:10.1371/journal.pone.0093907.
51. Xie R, Li J, Wang J, et al. DeepVF: a deep learning-based hybrid framework for identifying virulence factors using the stacking strategy. *Brief Bioinform.* 2021;22:bbaa125. doi:10.1093/bib/bbaa125.
52. Chen L, Yang J, Yu J, et al. VFDB: a reference database for bacterial virulence factors. *Nucleic Acids Res.* 2005;33:D325-D328. doi:10.1093/nar/gki008.
53. Urban M, Cuzick A, Seager J, et al. PHI-base: the pathogen-host interactions database. *Nucleic Acids Res.* 2020;48:D613. doi:10.1093/nar/gkz904.
54. Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics.* 2007;8:1-7. doi:10.1186/1471-2105-8-4.

55. Magnan CN, Zeller M, Kayala MA, et al. High-throughput prediction of protein antigenicity using protein microarray data. *Bioinformatics*. 2010;26:2936-2943. doi:10.1093/bioinformatics/btq551.
56. Wei L, Ye X, Sakurai T, Mu Z, Wei L. ToxIBTL: prediction of peptide toxicity based on information bottleneck and transfer learning. *Bioinformatics*. 2022;38:1514-1524. doi:10.1093/bioinformatics/btac006.
57. Maurer-Stroh S, Krutz NL, Kern PS, et al. AllerCatPro—prediction of protein allergenicity potential from the protein sequence. *Bioinformatics*. 2019;35:3020-3027. doi:10.1093/bioinformatics/btz029.
58. Szklarczyk D, Gable AL, Lyon D, et al. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res*. 2019;47:D607-D613. doi:10.1093/nar/gky1131.
59. Fodor EL, de Carvalho PVSD, Silva WM, et al. In silico identification of essential proteins in *Corynebacterium pseudotuberculosis* based on protein-protein interaction networks. *BMC Syst Biol*. 2016;10:103. doi:10.1186/s12918-016-0346-4.
60. Ngounou Wetie AG, Sokolowska I, Woods AG, Roy U, Deinhardt K, Darie CC. Protein–protein interactions: switch from classical methods to proteomics and bioinformatics-based approaches. *Cell Mol Life Sci*. 2014;71:205-228. doi:10.1007/s00018-013-1333-1.
61. Sharan R, Suthram S, Kelley RM, et al. Conserved patterns of protein interaction in multiple species. *Proc Natl Acad Sci USA*. 2005;102:1974-1979. doi:10.1073/pnas.0409522102.
62. Franceschini A, Szklarczyk D, Frankild S, et al. STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res*. 2013;41:D808-D815. doi:10.1093/nar/gks1094.
63. Huerta-Cepas J, Szklarczyk D, Forslund K, et al. eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res*. 2016;44:D286-D293. doi:10.1093/nar/gkv1248.
64. Yu H, Luscombe NM, Lu HX, et al. Annotation transfer between genomes: protein-protein interologs and protein-DNA regulogs. *Genome Res*. 2004;14:1107-1118. doi:10.1101/gr.1774904.
65. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13:2498-2504. doi:10.1101/gr.1239303.
66. Assenov Y, Ramírez F, Schelhorn S-E, Lengauer T, Albrecht M. Computing topological parameters of biological networks. *Bioinformatics*. 2008;24:282-284. doi:10.1093/bioinformatics/btm554.
67. Wang Y, Zhang H, Zhong H, Xue Z. Protein domain identification methods and online resources. *Comput Struct Biotechnol J*. 2021;19:1145-1153. doi:10.1016/j.csbj.2021.01.041.
68. Bjornson HS. Enzymes associated with the survival and virulence of gram-negative anaerobes. *Rev Infect Dis*. 1984;6(Suppl_1):s21-s24. doi:10.1093/clinids/6.Supplement_1.S21.
69. Bender J, Fliieger A. Lipases as pathogenicity factors of bacterial pathogens of humans. In: Timmis KN, ed. *Handbook of Hydrocarbon and Lipid Microbiology*. Berlin; Heidelberg: Springer; 2010:3241-3258. doi:10.1007/978-3-540-77587-4_246.
70. Kim B, Espinosa J, Hang HC. Biochemical analysis of NlpC/p60 peptidoglycan hydrolase activity. *Methods Enzymol*. 2020;638:109-127. doi:10.1016/b.mie.2020.02.017.
71. Jan S, Baron F, Coat R, Gonçalves O. Spoilage of egg products. In: Gonçalves O, Legrand JBT, eds. *Alteration of Oviproducts*. Elsevier; 2018:51-156. doi:10.1016/B978-1-78548-271-7.50002-X.
72. Iversen L, Kragballe K, Ziboh VA. Significance of leukotriene-A4 hydrolase in the pathogenesis of psoriasis. *Skin Pharmacol*. 1997;10:169-177. doi:10.1159/000211501.
73. Johnson G. The α/β hydrolase fold proteins of mycobacterium tuberculosis, with reference to their contribution to virulence. *Curr Protein Pept Sci*. 2017;18:190-210. doi:10.2174/1389203717666160729093515.
74. Fellner M, Lentz CS, Jamieson SA, et al. Structural basis for the inhibitor and substrate specificity of the unique Fph serine hydrolases of *Staphylococcus aureus*. *ACS Infect Dis*. 2020;6:2771-2782. doi:10.1021/acinfecdis.0c00503.
75. van den Broek B, Noom MC, Wuite GJ. DNA-tension dependence of restriction enzyme activity reveals mechanochemical properties of the reaction pathway. *Nucleic Acids Res*. 2005;33:2676-2684. doi:10.1093/nar/gki565.
76. Davey L, Ng CKW, Halperin SA, Lee SF. Functional analysis of paralogous thiol-disulfide oxidoreductases in *Streptococcus gordonii*. *J Biol Chem*. 2013;288:16416-16429.
77. Hecht HJ, Sobek H, Haag T, Pfeifer O, van Pée KH. The metal-ion-free oxidoreductase from *Streptomyces aureofaciens* has an alpha/beta hydrolase fold. *Nat Struct Biol*. 1994;1:532-537. doi:10.1038/nsb0894-532.
78. Li Q, Zhang Y, Sheng Y, et al. Large T-antigen up-regulates Kv4.3 K⁺ channels through Sp1, and Kv4.3 K⁺ channels contribute to cell apoptosis and necrosis through activation of calcium/calmodulin-dependent protein kinase II. *Biochem J*. 2012;441:859-869. doi:10.1042/BJ20111604.
79. McQuiston JR, Vemulapalli R, Inzana TJ, et al. Genetic characterization of a Tn5-disrupted glycosyltransferase gene homolog in *Brucella abortus* and its effect on lipopolysaccharide composition and virulence. *Infect Immun*. 1999;67:3830-3835. doi:10.1128/IAI.67.8.3830-3835.1999.
80. Valvano MA. Genetics and biosynthesis of lipopolysaccharide. In: Tang Y-W, Sussman M, Liu D, Poxton I, Schwartzman J, eds. *Molecular Medical Microbiology* (Second Edition). Academic Press; 2015:55-89. doi:10.1016/B978-0-12-397169-2.00004-4.
81. Wang X, Quinn PJ, Yan A. Kdo2-lipid A: structural diversity and impact on immunopharmacology. *Biol Rev Camb Philos Soc*. 2015;90:408-427. doi:10.1111/brv.12114.
82. Nogly P, Gushchin I, Remeeva A, et al. X-ray structure of a CDP-alcohol phosphatidyltransferase membrane enzyme and insights into its catalytic mechanism. *Nat Commun*. 2014;5:4169. doi:10.1038/ncomms5169.
83. Gibellini F, Smith TK. The Kennedy pathway—De novo synthesis of phosphatidylethanolamine and phosphatidylcholine. *IUBMB Life*. 2010;62:414-428. doi:10.1002/iub.337.
84. Heath RJ, White SW, Rock CO. Lipid biosynthesis as a target for antibacterial agents. *Prog Lipid Res*. 2001;40:467-497. doi:10.1016/S0163-7827(01)00012-1.
85. Percudani R, Peracchi A. A genomic overview of pyridoxal-phosphate-dependent enzymes. *EMBO Rep*. 2003;4:850-854. doi:10.1038/sj.embor.embor914.
86. Fukamachi H, Nakano Y, Yoshimura M, Koga T. Cloning and characterization of the L-cysteine desulfhydrase gene of *Fusobacterium nucleatum*. *FEMS Microbiol Lett*. 2002;215:75-80. doi:10.1111/j.1574-6968.2002.tb11373.x.
87. Krupka HI, Huber R, Holt SC, Clausen T. Crystal structure of cystalysin from *Treponema denticola*: a pyridoxal 5'-phosphate-dependent protein acting as a haemolytic enzyme. *EMBO J*. 2000;19:3168-3178. doi:10.1093/emboj/19.13.3168.
88. Rosey EL, Stewart GC. Nucleotide and deduced amino acid sequences of the lacR, lacABCD, and lacFE genes encoding the repressor, tagatose 6-phosphate gene cluster, and sugar-specific phosphotransferase system components of the lactose operon of *Streptococcus mutans*. *J Bacteriol*. 1992;174:6159-6170. doi:10.1128/jb.174.19.6159-6170.1992.
89. Kaito C, Morishita D, Matsumoto Y, Kurokawa K, Sekimizu K. Novel DNA binding protein SarZ contributes to virulence in *Staphylococcus aureus*. *Mol Microbiol*. 2006;62:1601-1617.
90. Barrett JF, Hoch JA. Two-component signal transduction as a target for microbial anti-infective therapy. *Antimicrob Agents Chemother*. 1998;42:1529-1536. doi:10.1128/AAC.42.7.1529.
91. Aravind L, Anantharaman V, Balaji S, Babu MM, Iyer LM. The many faces of the helix-turn-helix domain: transcription regulation and beyond. *FEMS Microbiol Rev*. 2005;29:231-262. doi:10.1016/j.fmrre.2004.12.008.
92. Tiwari S, Jamal SB, Hassan SS, et al. Two-component signal transduction systems of pathogenic bacteria as targets for antimicrobial therapy: an overview. *Front Microbiol*. 2017;8:1878. <https://www.frontiersin.org/article/10.3389/fmicb.2017.01878>
93. Hofstatter PG, Tice AK, Kang S, Brown MW, Lahr DJG. Evolution of bacterial recombinase A (recA) in eukaryotes explained by addition of genomic data of key microbial lineages. *Proceedings Biol Sci*. 2016;283:20161453. doi:10.1098/rspb.2016.1453.
94. Shen X, Woodgate R, Goodman MF. *Escherichia coli* DNA polymerase V subunit exchange: a post-SOS mechanism to curtail error-prone DNA synthesis. *J Biol Chem*. 2003;278:52546-52550. doi:10.1074/jbc.M310127200.
95. Yoshida T, Nasu H, Namba E, Ubukata O, Yamashita M. Discovery of a compound that acts as a bacterial PyrG (CTP synthase) inhibitor. *J Med Microbiol*. 2012;61:1280-1285. doi:10.1099/jmm.0.046052-0.
96. Garmory HS, Titball RW. ATP-binding cassette transporters are targets for the development of antibacterial vaccines and therapies. *Infect Immun*. 2004;72:6757-6763. doi:10.1128/IAI.72.12.6757-6763.2004.
97. Cronan JE Jr, Waldrop GL. Multi-subunit acetyl-CoA carboxylases. *Prog Lipid Res*. 2002;41:407-435. doi:10.1016/s0163-7827(02)00007-3.
98. Jitrapakdee S, St Maurice M, Rayment I, Cleland WW, Wallace JC, Attwood PV. Structure, mechanism and regulation of pyruvate carboxylase. *Biochem J*. 2008;413:369-387. doi:10.1042/BJ20080709.
99. Wakil SJ, Stoops JK, Joshi VC. Fatty acid synthesis and its regulation. *Annu Rev Biochem*. 1983;52:537-579. doi:10.1146/annurev.bi.52.070183.002541.
100. Tong L. Acetyl-coenzyme A carboxylase: crucial metabolic enzyme and attractive target for drug discovery. *Cell Mol Life Sci*. 2005;62:1784-1803. doi:10.1007/s00018-005-5121-4.
101. Ghosh S, Navarathna DH, Roberts DD, et al. Arginine-induced germ tube formation in *Candida albicans* is essential for escape from murine macrophage line RAW 264.7. *Infect Immun*. 2009;77:1596-1605. doi:10.1128/IAI.01452-08.
102. Schär J, Stoll R, Schauer K, et al. Pyruvate carboxylase plays a crucial role in carbon metabolism of extra- and intracellularly replicating *Listeria monocytogenes*. *J Bacteriol*. 2010;192:1774-1784. doi:10.1128/JB.01132-09.
103. Tong L. Structure and function of biotin-dependent carboxylases. *Cell Mol Life Sci*. 2013;70:863-891. doi:10.1007/s00018-012-1096-0.

104. Wiechert M, Beitz E. Mechanism of formate-nitrite transporters by dielectric shift of substrate acidity. *EMBO J*. 2017;36:949-958. doi:10.15252/embj.201695776.
105. Das P, Lahiri A, Lahiri A, Chakravorty D. Novel role of the nitrite transporter NirC in Salmonella pathogenesis: SPI2-dependent suppression of inducible nitric oxide synthase in activated macrophages. *Microbiology (Reading)*. 2009;155:2476-2489. doi:10.1099/mic.0.029611-0.
106. Routh MD, Zalucki Y, Su CC, Zhang Q, Shafer WM, Yu EW. Efflux pumps of the resistance-nodulation-division family: a perspective of their structure, function, and regulation in gram-negative bacteria. *Adv Enzymol Relat Areas Mol Biol*. 2011;77:109-146. doi:10.1002/9780470920541.ch3.
107. Delmar JA, Su C-C, Yu EW. Bacterial multidrug efflux transporters. *Annu Rev Biophys*. 2014;43:93-117. doi:10.1146/annurev-biophys-051013-022855.
108. Mealman TD, Blackburn NJ, McEvoy MM. Metal export by CusCFBA, the periplasmic Cu(I)/Ag(I) transport system of Escherichia coli. *Curr Top Membr*. 2012;69:163-196. doi:10.1016/B978-0-12-394390-3.00007-0.
109. Whiteley CG, Lee DJ. Bacterial diguanylate cyclases: structure, function and mechanism in exopolysaccharide biofilm development. *Biotechnol Adv*. 2015;33:124-141. doi:10.1016/j.biotechadv.2014.11.010.
110. Wong SG, Dessen A. Structure of a bacterial $\alpha 2$ -macroglobulin reveals mimicry of eukaryotic innate immunity. *Nat Commun*. 2014;5:4917. doi:10.1038/ncomms5917.
111. Budd A, Blandin S, Levshina EA, Gibson TJ. Bacterial alpha2-macroglobulins: colonization factors acquired by horizontal gene transfer from the metazoan genome. *Genome Biol*. 2004;5:R38. doi:10.1186/gb-2004-5-6-r38.
112. Confer AW, Ayalew S. The OmpA family of proteins: roles in bacterial pathogenesis and immunity. *Vet Microbiol*. 2013;163:207-222. doi:10.1016/j.vetmic.2012.08.019.
113. Paruthiyil S, Pinochet-Barros A, Huang X, Helmann JD. Bacillus subtilis TerC family proteins help prevent manganese intoxication. *J Bacteriol*. 2020;202:e00624-19. doi:10.1128/JB.00624-19.
114. Grunenwald CM, Choby JE, Juttukonda LJ, et al. Manganese detoxification by MntE is critical for resistance to oxidative stress and virulence of *Staphylococcus aureus*. *mBio*. 2019;10. doi:10.1128/mBio.02915-18.
115. Schmidt AJ, Ryjenkov DA, Gomelsky M. The ubiquitous protein domain EAL is a cyclic diguanylate-specific phosphodiesterase: enzymatically active and inactive EAL domains. *J Bacteriol*. 2005;187:4774-4781. doi:10.1128/JB.187.14.4774-4781.2005.
116. Wei C, Jiang W, Zhao M, et al. A systematic analysis of the role of GGDEF-EAL domain proteins in virulence and motility in *Xanthomonas oryzae* pv. *Oryzicola*. *Sci Rep*. 2016;6:23769. doi:10.1038/srep23769.
117. Tamayo R, Tischler AD, Camilli A. The EAL domain protein VieA is a cyclic diguanylate phosphodiesterase. *J Biol Chem*. 2005;280:33324-33330. doi:10.1074/jbc.M506500200.
118. Schoolnik GK, Voskuil MI, Schnappinger D, et al. [1] Whole genome DNA microarray expression analysis of biofilm development by *Vibrio cholerae* O1 El Tor. *Meth Enzymol*. 2001;336:3-18. doi:10.1016/S0076-6879(01)36573-4.
119. Herrington DA, Hall RH, Losonsky G, Mekalanos JJ, Taylor RK, Levine MM. Toxin, toxin-coregulated pilus, and the toxR regulon are essential for *Vibrio cholerae* pathogenesis in humans. *J Exp Med*. 1988;168:1487-1492. doi:10.1084/jem.168.4.1487.
120. Taylor RK, Miller VL, Furlong DB, Mekalanos JJ. Use of phoA gene fusions to identify a pilus colonization factor coordinately regulated with cholera toxin. *Proc Natl Acad Sci USA*. 1987;84:2833-2837. doi:10.1073/pnas.84.9.2833.
121. Lee SH, Hava DL, Waldor MK, Camilli A. Regulation and temporal expression patterns of *Vibrio cholerae* virulence genes during infection. *Cell*. 1999;99:625-634. doi:10.1016/S0092-8674(00)81551-2.
122. Miller VL, Mekalanos JJ. Genetic analysis of the cholera toxin-positive regulatory gene toxR. *J Bacteriol*. 1985;163:580-585. doi:10.1128/jb.163.2.580-585.1985.
123. Lee SH, Butler SM, Camilli A. Selection for in vivo regulators of bacterial virulence. *Proc Natl Acad Sci USA*. 2001;98:6889-6894. doi:10.1073/pnas.111581598.
124. Clatworthy AE, Pierson E, Hung DT. Targeting virulence: a new paradigm for antimicrobial therapy. *Nat Chem Biol*. 2007;3:541-548. doi:10.1038/nchembio.2007.24.
125. Marra A. Targeting virulence for antibacterial chemotherapy: identifying and characterising virulence factors for lead discovery. *Drugs R D*. 2006;7:1-16. doi:10.2165/00126839-200607010-00001.
126. Depluvere S, Devos S, Devreese B. The role of bacterial secretion systems in the virulence of gram-negative airway pathogens associated with cystic fibrosis. *Front Microbiol*. 2016;7:1336. doi:10.3389/fmicb.2016.01336.
127. Marra A. Can virulence factors be viable antibacterial targets. *Expert Rev Anti Infect Ther*. 2004;2:61-72. doi:10.1586/14787210.2.1.61.
128. Zhang B-Z, Hu D, Dou Y, et al. Identification and evaluation of recombinant outer membrane proteins as vaccine candidates against *Klebsiella pneumoniae*. *Front Immunol*. 2021;12:730116. doi:10.3389/fimmu.2021.730116.
129. Qian R-H, Xiao Z-H, Zhang C-W, et al. A conserved outer membrane protein as an effective vaccine candidate from *Vibrio alginolyticus*. *Aquaculture*. 2008;278:5-9. doi:10.1016/j.aquaculture.2008.03.010.
130. Pan J, Li C, Ye Z. Immunoproteomic approach for screening vaccine candidates from bacterial outer membrane proteins. *Methods Mol Biol*. 2016;1404:519-528. doi:10.1007/978-1-4939-3389-1_34.
131. Elhossary MA, Bahey-El-Din M, AbdelBary A, El Guink N, Aboushleib HM. Immunization with the ferric iron-binding periplasmic protein HitA provides protection against *Pseudomonas aeruginosa* in the murine infection model. *Microb Pathog*. 2019;131:181-185. doi:10.1016/j.micpath.2019.04.014.
132. Enany S. Structural and functional analysis of hypothetical and conserved proteins of Clostridium tetani. *J Infect Public Health*. 2014;7:296-307. doi:10.1016/j.jiph.2014.02.002.
133. Audain E, Ramos Y, Hermjakob H, et al. Accurate estimation of isoelectric point of protein and peptide based on amino acid sequences. *Bioinformatics*. 2016;32:821-27. doi:10.1093/bioinformatics/btv674.
134. Gasteiger E, Gattiker A, Hoogland C, et al. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res*. 2003;31:3784-88. doi:10.1093/nar/gkg563.
135. Gill SC, von Hippel PH. Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem*. 1989;182:319-326. doi:10.1016/0003-2697(89)90602-7.
136. Guruprasad K, Reddy BV, Pandit MW. Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. *Protein Eng*. 1990;4:155-161. doi:10.1093/protein/4.2.155.
137. Kubori T, Galán JE. Temporal regulation of salmonella virulence effector function by proteasome-dependent protein degradation. *Cell*. 2003;115:333-342. doi:10.1016/s0092-8674(03)00849-3.
138. Ikai A. Thermostability and aliphatic index of globular proteins. *J Biochem*. 1980;88:1895-1898.
139. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. *J Mol Biol*. 1982;157:105-132. doi:10.1016/0022-2836(82)90515-0.
140. Magdeldin S, Yoshida Y, Li H, et al. Murine colon proteome and characterization of the protein pathways. *Biodata Min*. 2012;5:11. doi:10.1186/1756-0381-5-11.
141. Ezaj MMA, Haque MS, Syed SB, et al. Comparative proteomic analysis to annotate the structural and functional association of the hypothetical proteins of *S. PLoS ONE*. 2021;16:e0252295. doi:10.1371/journal.pone.0252295.
142. Zahiri J, Bozorgmehr JH, Masoudi-Nejad A. Computational prediction of protein-protein interaction networks: algo-rithms and resources. *Curr Genomics*. 2013;14:397-414. doi:10.2174/1389202911314060004.
143. Andreani J, Guerois R. Evolution of protein interactions: from interactomes to interfaces. *Arch Biochem Biophys*. 2014;554:65-75. doi:10.1016/j.abb.2014.05.010.
144. Brun C, Chevenet F, Martin D, Wojcik J, Guénoche A, Jacq B. Functional classification of proteins for the prediction of cellular function from a protein-protein interaction network. *Genome Biol*. 2003;5:R6. doi:10.1186/gb-2003-5-1-r6.
145. Lu S, Zgurskaya HI. Role of ATP binding and hydrolysis in assembly of MacAB-TolC macrolide transporter. *Mol Microbiol*. 2012;86:1132-1143. doi:10.1111/mmi.12046.
146. Nikaido H. Antibiotic resistance caused by gram-negative multidrug efflux pumps. *Clin Infect Dis*. 1998;27:S32-S41. http://www.jstor.org/stable/4460548
147. Nikaido H, Takatsuka Y. Mechanisms of RND multidrug efflux pumps. *Biochim Biophys Acta*. 2009;1794:769-781. doi:10.1016/j.bbapap.2008.10.004.
148. Franke S, Grass G, Rensing C, Nies DH. Molecular analysis of the copper-transporting efflux system CusCFBA of Escherichia coli. *J Bacteriol*. 2003;185:3804-3812. doi:10.1128/JB.185.13.3804-3812.2003.
149. Magnani D, Solioz M. How bacteria handle copper BT—molecular microbiology of heavy metals. In: Nies DH, Silver S, eds. Berlin and Heidelberg, Germany: Springer; 2007:259-285. doi:10.1007/7171_2006_081.
150. Ladomersky E, Petris MJ. Copper tolerance and virulence in bacteria. *Metallo-mics*. 2015;7:957-964. doi:10.1039/c4mt00327f.
151. Catlett NL, Yoder OC, Turgeon BG. Whole-genome analysis of two-component signal transduction genes in fungal pathogens. *Eukaryot Cell*. 2003;2:1151-1161. doi:10.1128/EC.2.6.1151-1161.2003.
152. Jacob S, Foster AJ, Yemelin A, Thines E. Histidine kinases mediate differentiation, stress response, and pathogenicity in Magnaporthe oryzae. *Microbiologyopen*. 2014;3:668-687. doi:10.1002/mbo3.197.
153. Lin YH, Pierce BD, Fang F, Wise A, Binns AN, Lynn DG. Role of the VirA histidine autokinase of *Agrobacterium tumefaciens* in the initial steps of pathogenesis. *Front Plant Sci*. 2014;5:195. https://www.frontiersin.org/article/10.3389/fpls.2014.00195
154. Nagahashi S, Mio T, Ono N, et al. Isolation of CaSLN1 and CaNIK1, the genes for osmosensing histidine kinase homologues, from the pathogenic fungus *Candida albicans*. *Microbiology (Reading)*. 1998;144(Pt2):425-432. doi:10.1099/00221287-144-2-425.
155. Chen W, Wang B, Gruber JD, Zhang Y-M, Davies C. Acyl carrier protein 3 is involved in oxidative stress response in *Pseudomonas aeruginosa*. *Front Microbiol*. 2018;9:2244. doi:10.3389/fmicb.2018.02244.
156. Ma J-C, Wu Y-Q, Cao D, Zhang W-B, Wang H-H. Only acyl carrier protein 1 (AcpP1) functions in *Pseudomonas aeruginosa* fatty acid synthesis. *Front Microbiol*. 2017;8:2186. https://www.frontiersin.org/article/10.3389/fmicb.2017.02186

157. Hoang TT, Sullivan SA, Cusick JK, Schweizer HP. β -Ketoacyl acyl carrier protein reductase (FabG) activity of the fatty acid biosynthetic pathway is a determining factor of 3-oxo-homoserine lactone acyl chain lengths. *Microbiology (Reading)*. 2002;148:3849-3856.
158. Yuan Y, Sachdeva M, Leeds JA, Meredith TC. Fatty acid biosynthesis in *Pseudomonas aeruginosa* is initiated by the FabY class of β -ketoacyl acyl carrier protein synthases. *J Bacteriol*. 2012;194:5171-5184.
159. Zhu K, Rock CO. RhIA converts β -hydroxyacyl-acyl carrier protein intermediates in fatty acid synthesis to the β -hydroxydecanoyl- β -hydroxydecanoate component of rhamnolipids in *Pseudomonas aeruginosa*. *J Bacteriol*. 2008;190:3147-3154.
160. Liu C, Sun D, Zhu J, Liu W. Two-component signal transduction systems: a major strategy for connecting input stimuli to biofilm formation. *Front Microbiol*. 2019;9:3279. doi:10.3389/fmicb.2018.03279.
161. Hurrell E, Kucerova E, Loughlin M, et al. Neonatal enteral feeding tubes as loci for colonisation by members of the Enterobacteriaceae. *BMC Infect Dis*. 2009;9:146. doi:10.1186/1471-2334-9-146.
162. Rahal A, Andreo A, Le Gallou F, et al. Enterobacter cloacae complex outbreak in a neonatal intensive care unit: multifaceted investigations and preventive measures are needed. *J Hosp Infect*. 2021;116:87-90. doi:10.1016/j.jhin.2021.07.012.
163. Jamal M, Ahmad W, Andleeb S, et al. Bacterial biofilm and associated infections. *J Chin Med Assoc*. 2018;81:7-11. doi:10.1016/j.jcma.2017.07.012.
164. Ferry A, Plaisant F, Ginevra C, et al. Enterobacter cloacae colonisation and infection in a neonatal intensive care unit: retrospective investigation of preventive measures implemented after a multiclonal outbreak. *BMC Infect Dis*. 2020;20:682. doi:10.1186/s12879-020-05406-8.