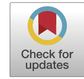
Current Research in Microbial Sciences 7 (2024) 100291

Contents lists available at ScienceDirect

Current Research in Microbial Sciences

journal homepage: www.sciencedirect.com/journal/current-research-in-microbial-sciences



Screening of Novel Drug Targets and Drug Design for *Bordetella pertussis*: A Subtractive Proteomics Approach

Md. Nazmul Islam Bappy [a], Foeaz Ahmed [a,b], Tahera Lasker [a,b], Emran Hossain Sajib [a], Md. Shariful Islam [a,b,*]

[a] *Faculty of Biotechnology and Genetic Engineering, Sylhet Agricultural University, Sylhet 3100, Bangladesh*
[b] *Department of Molecular Biology and Genetic Engineering, Sylhet Agricultural University, Sylhet 3100, Bangladesh*

A R T I C L E I N F O

Keywords:
Bordetella pertussis
Subtractive proteomics approach
Molecular docking
Plant-derived metabolites
ADME properties

A B S T R A C T

Bordetella pertussis causes whooping cough in humans that spreads directly from individual to individual mainly by aerosolized respiratory droplets. Nowadays, it gained the attention of scientific community because it has already been reemerged as one of the major public health threats despite widespread vaccination efforts. Moreover, the growing antibiotic resistance has made it difficult to combat this pathogen with currently available antibiotics. Consequently, screening drug targets and discovering drugs against unique proteins of the pathogen could be a promising alternative. With this view, 3,359 proteins of *B. pertussis* were screened in silico to identify non-duplicate proteins crucial for survival of the bacteria, non-homologous to humans, involved in unique metabolic pathways of the pathogen, and conserved among various bacterial strains. Among these, Chemotaxis protein Mota, Chromosomal replication initiator protein DnaA, Short-chain fatty acids transporter, [protein-PII] uridylyltransferase, Type III secretion protein V, Potassium-transporting ATPase potassium-binding subunit, N-acetylmuramoyl-L-alanine amidase, and RNA polymerase sigma-54 factor fulfilled these criteria. These proteins were further analyzed for qualitative characteristics such as virulence properties and associations with antibiotic resistance, etc. In addition, plant metabolites were screened against these unique proteins utilizing molecular docking to discover putative drugs against them. Four metabolites exhibited superior binding affinity and favorable ADME (Adsorption, distribution, metabolism, and excretion) properties which can further be tested in vivo.

1. Introduction

Bordetella pertussis is the agent behind pertussis, or whooping cough, which is characterized by a ‘whooping’ sound when the person breathes in. Basic features of this pathogen is gram-negative, aerobic, motile coccobacillus that expresses a flagellum-like structure (Hoffman et al., 2019). Airborne droplets are the means of its spreading, with an incubation time that may ranges between 6 and 20 days (7-10 days on average) (David, 2008), and the only known reservoir for this pathogen is humans (Güriş et al., 1999). Children under the age of one are the main sufferers if they are not vaccinated or with faded immunity. However, people of all ages can get infected with it, and it can even be life-threatening, especially in young children.

Currently, vaccine for this pathogen is available in two forms that are acellular pertussis vaccine (aP) and whole-cell pertussis vaccine (wP). Vaccination of baboons with the aP provides protection only against the symptoms, while ineffective against colonization or transmission which is one of the shortcomings of current aP vaccines (Warfel et al., 2014). Pertussis, in any case, started to reappear during the 1990s in a few profoundly vaccinated populaces, and the quantity of pertussis cases is as yet showing the upward patterns around the world (Jakinovich and Sood, 2014). In spite of inclusive immunization policies, pertussis episodes have been accounted for in numerous nations, including United States, Australia, United Kingdom, Brazil, China, and Chile in the last decade (Gohari et al., 2013, Theofiles et al., 2014, Winter et al., 2012). In 2014, 24.1 million cases were reported in youngsters, and 160,700 children under the age of five years died of this disease where the largest portion was from African locale with 7.8 million cases and 92,500 deaths. Moreover, estimated cases in babies younger than one year were 5.1 million and estimated death was 85,900. Potential explanations behind the resurgence of *B. pertussis* disease include expanded surveillance and reporting, mutation of the bacterium, analytical diagnosis

* Corresponding author.
E-mail address: harif.mge@sau.ac.bd (Md.S. Islam).

https://doi.org/10.1016/j.crmicr.2024.100291

2666-5174/© 2024 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

Table 2
List of acronyms, their meanings, and links of the servers.

Acronyms	Full form	Links
NCBI	National Center for Biotechnology Information	https://www.ncbi.nlm.nih.gov/genome
DEG	Database of Essential Genes	www.essentialgene.org
KEGG	Kyoto Encyclopedia of Genes and Genomes	https://www.genome.jp/kegg/pathway.html
KAAS	KEGG Automatic Annotation Server	https://www.genome.jp/kaas-bin/kaas_main
KO	KEGG ORTHOLOGY	https://www.genome.jp/kegg/ko.html
BLAST	Basic Local Alignment Search Tool	https://blast.ncbi.nlm.nih.gov/Blast.cgi
VFDB	Virulence Factor Database	http://www.mgc.ac.cn/VFs/blast/blast.html
ARG-ANNOT	Antibiotic Resistance Gene-ANNOTation	https://ifr48.timone.univ-mrs.fr/blast/arg-annot_v6.html
STRING	Search Tool for the Retrieval of Interacting Genes/Proteins	https://string-db.org
HPIDB	Host-Pathogen Interaction Database	https://hpidb.igbb.msstate.edu/sequence.html
PDB	Protein Data Bank	https://www.rcsb.org/

enhancement, fading immunity after vaccination, and parents rejecting immunization of their youngsters (Yeung et al., 2017). However, vaccination of those children is not possible as according to CDC guidelines; DTaP should be given at 2, 4, and a half year, at 15 through year and a half and at 4 through 6 years. According to Cherry (2015), antibody decaying over time, erroneous antigen balance in vaccine, incomplete packaging of antigen, linked epitope suppression, and genetic alternation of the bacteria contribute to the failure of the vaccine (Cherry, 2015). Findings from some studies suggested that changes in PRN proteins, a protein that aids in adherence of the organism to tracheal epithelial cells, may contribute to the resurgence of pertussis (Lam et al., 2014; Schmidtke et al., 2012). Around 70 % of the time, antibody to PRN is found to be associated with protection, so it seems plausible that PRN deficiency could lead to increased vaccine failures, if deficient mutants became widespread (Storsaeter et al., 1998, Cherry et al., 1998).

Antibiotics are the means of treatment that control symptoms and prevent infected people from spreading the disease. But the uncontrolled utilization of antibiotics has contributed to the growing resistance in the pathogens. Erythromycin and ceftriaxone resistant strain, with MIC range of 0·25 - 6 μg/ml and 1·5 - 256 μg/ml, respectively have been found in a strain of this bacterium (Torres et al., 2015). *B. pertussis* in northern China showed a strikingly high rate (91.9 %) of macrolide resistance (Yang et al., 2015). Again, antibiotics were not sufficient to treat secondary cases of whooping cough. The probable cause may be that they are more effective if initiated within 21 days of the onset of forceful coughing. Various side effects were also found with antibiotics, which differ based on the types of antibiotics (Altunaiji et al., 2012). So, dealing with this fatal pathogen with current vaccines or antibiotics in the near future will be a matter of great concern. Thus, it is imperative to discover new drugs to combat multi-drug resistant bacteria and make available an ideal treatment option to get rid of this deadly pathogen.

The very first step of drug discovery process is the identification of drug targets (Chan et al., 2010). In silico approaches employed to identify novel drug targets and screen metabolites could accelerate the drug discovery process, as the conventional ways of drug discovery require more time, costly experiments, and labor. Plants have been used as a source of medication from the beginning of time and provide the basis for many pharmaceutical applications. Natural compounds derived from plants have been the mainstay of conventional medicine for thousands of years (Ginsburg and Deharo, 2011). Natural products contain complex chemical structures that differ according to their various species in nature, and when the existing high technology methods are applied, it can lead to novel medicine development,

Table 2
Unique metabolic pathway proteins.

Sl No	Accession Number	KO Number	Protein Name (Gene Name)	Pathways
01	WP_003811927.1	K02556	Chemotaxis protein MotA (motA)	-Bacterial chemotaxis -Two-component system -Flagellar assembly
02	WP_003815032.1	K07789	Multidrug efflux pump (mdtC)	-Two-component system
03	WP_003815201.1	K05515	Penicillin-binding protein 2 (mrdA)	- Peptidoglycan biosynthesis
04	WP_010929554.1	K02313	Chromosomal replication initiator protein (dnaA)	- Two-component system
05	WP_010929693.1	K02106	Short-chain fatty acids transporter (atoE)	- Two-component system
06	WP_010930030.1	K18307	Multidrug efflux pump (mexI)	- Quorum sensing
07	WP_010930146.1	K03412	Two-component system, chemotaxis family, protein-glutamate methylesterase/glutaminase (cheB)	- Two-component system - Bacterial chemotaxis
08	WP_010930331.1	K05874	Methyl-accepting chemotaxis protein I, serine sensor receptor (tsr)	- Two-component system - Bacterial chemotaxis
09	WP_010930348.1	K00990	[protein-PII] uridylyltransferase (glnD)	- Two-component system
10	WP_010930831.1	K03230	Type III secretion protein V (yscV, sctV, hrcV, ssaV, invA)	- Bacterial secretion system
11	WP_010930834.1	K03406	Methyl-accepting chemotaxis protein (mcpA)	- Two-component system - Bacterial chemotaxis
12	WP_010930949.1	K01546	Potassium-transporting ATPase potassium-binding subunit (kdpAs)	- Two-component system
13	WP_019247543.1	K01448	N-acetylmuramoyl-L-alanine amidase (amiABC)	- Cationic antimicrobial peptideresistance
14	WP_023853393.1	K03092	RNA polymerase sigma-54 factor (rpoN)	- Two-component system

Table 3
Subcellular location of unique pathway proteins.

Location	Accession Number
Cytoplasmic	WP_010929554.1 WP_010930146.1 WP_010930348.1 WP_019247543.1 WP_023853393.1
Inner membrane	WP_003811927.1 WP_003815032.1 WP_003815201.1 WP_010929693.1 WP_010930030.1 WP_010930331.1 WP_010930831.1 WP_010930834.1 WP_010930949.1

Table 4
Proteins with higher conservancy.

Sl No	Accession Number	Protein Name (Gene Name)
01	WP_003811927.1	Chemotaxis protein Mota (motA)
02	WP_003815032.1	Multidrug efflux pump (mdtC)
03	WP_003815201.1	Penicillin-binding protein 2 (mrdA)
04	WP_010929554.1	Chromosomal replication initiator protein (dnaA)
05	WP_010929693.1	Short-chain fatty acids transporter (atoE)
06	WP_010930348.1	[protein-PII] uridylyltransferase (glnD)
07	WP_010930831.1	Type III secretion protein V (yscV, sctV, hrcV, ssaV, invA)
08	WP_010930949.1	Potassium-transporting ATPase potassium-binding subunit (kdpAs)
09	WP_019247543.1	N-acetylmuramoyl-L-alanine amidase (amiABC)
10	WP_023853393.1	RNA polymerase sigma-54 factor (rpoN)

Table 5
Unique drug targets for *B. pertussis*.

Sl No	Accession Number	Protein Name (Gene Name)
01	WP_003811927.1	Chemotaxis protein Mota (motA)
02	WP_010929554.1	Chromosomal replication initiator protein (dnaA)
03	WP_010929693.1	Short-chain fatty acids transporter (atoE)
04	WP_010930348.1	[protein-PII] uridylyltransferase (glnD)
05	WP_010930831.1	Type III secretion protein V (yscV, sctV, hrcV, ssaV, invA)
06	WP_010930949.1	Potassium-transporting ATPase potassium-binding subunit (kdpAs)
07	WP_019247543.1	N-acetylmuramoyl-L-alanine amidase (amiABC)
08	WP_023853393.1	RNA polymerase sigma-54 factor (rpoN)

benefiting the entire world (Koparde et al., 2019). Because of the numerous phytochemical properties of plant derived products, traditional people utilized plants to cure numerous diseases (Kumar et al., 2019). Also, the inherent affinity of natural compounds toward biological receptors offer advantages for the development of drugs (Ginsburg and Deharo, 2011), while modern treatment options are associated with various shortcomings like devastating side effects, resistance, toxicity profile, and complicated medication (Cheuka et al., 2016).

In this study, the representative proteome sequence of *B. pertussis* was retrieved to carry out subtractive proteomics approaches. Essential genes were identified as the principal criteria for screening the best therapeutic targets. With a view to avoiding cross reactivity with host proteins and metabolism, host non-homology and metabolic pathway analysis was carried out, respectively. Furthermore, some qualitative characterizations were performed to find out the best suitable ones. To confront this reemerging bacterium, natural products could be utilized as a new remedy to avoid the unwanted outcomes of modern drugs. Therefore, we also aimed to find some plant metabolites as potential therapeutics against *B. pertussis*, utilizing various bioinformatics tools.

2. Materials and methods

2.1. Acquisition of the proteomic data of B. pertussis

The representative proteome of *B. pertussis*, strain Tohama I (assembly GCA_000195715.1) was extracted from National Center for Biotechnology Information (NCBI) Genome database (https://www.ncbi.nlm.nih.gov/genome).

2.2. Exclusion of paralog and mini proteins

The whole Proteome underwent analysis utilizing the CD-HIT server (https://weizhong-lab.ucsd.edu/cdhit_suite/cgi-bin/index.cgi?cmd=cd-hit) to identify paralog proteins (Huang et al., 2010) by applying "sequence identity cut-off" 0.6 (60 % identity) (Dutta et al., 2006) to exclude redundant protein sequences. The server groups similar proteins into clusters based on their sequence similarities. This step was performed to remove duplicate proteins from the datasets (Dutta et al., 2006). In addition, a "length of sequence to skip" of 100 was set to subtract mini proteins (sequence shorter than 100 amino acids), which play important part in various regulatory functions and biological

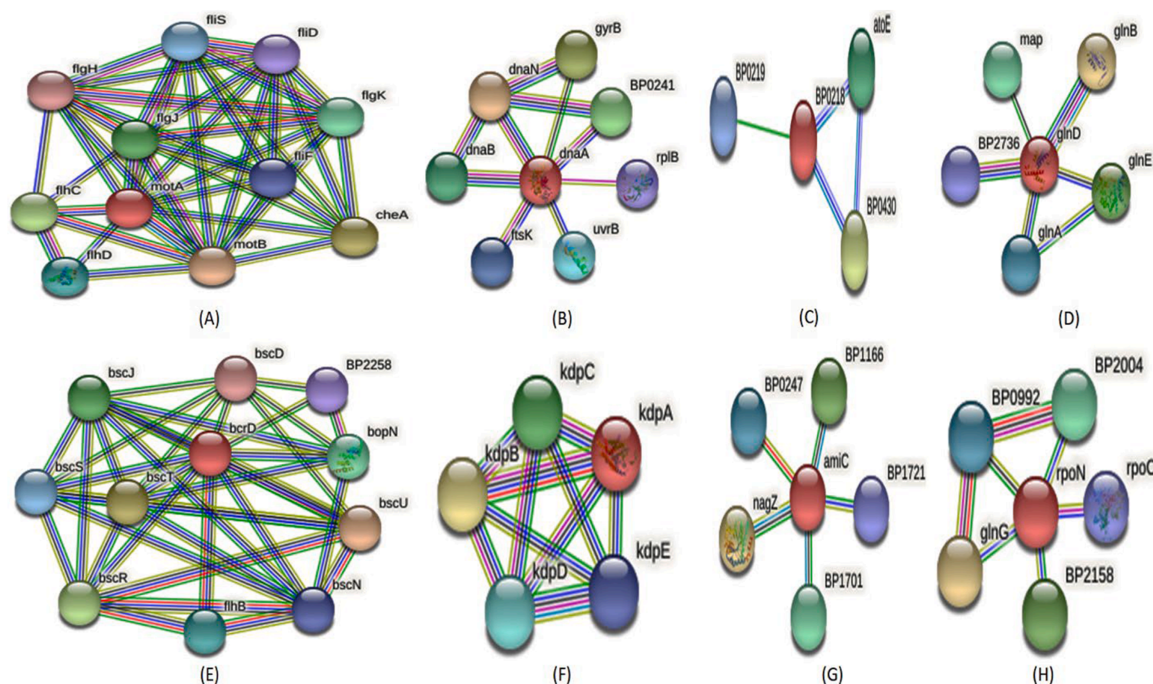


Fig. 1. Protein-Protein Interaction Network. (a) Chemotaxis protein Mota, (b) Chromosomal replication initiator protein, (c) Short-chain fatty acids transporter, (d) [protein-PII] uridylyltransferase, (e) Type III secretion protein, (f) Potassium-transporting ATPase potassium-binding subunit, (g) N-acetylmuramoyl-L-alanine amidase, (h) RNA polymerase sigma-54 factor.

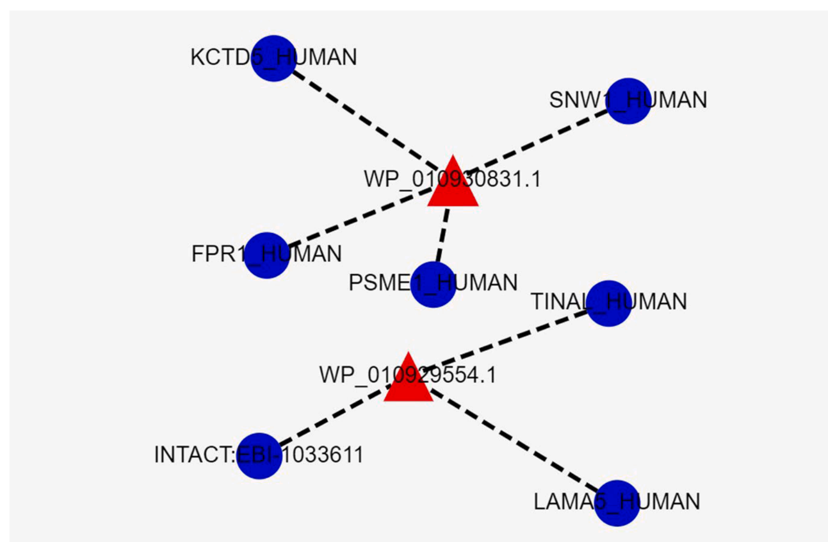


Fig. 2. Interaction of pathogenic proteins with human proteins

Table 6
Refined protein model with their ERRAT and PROCHECK value.

Accession Number	PDB Hit	TM-score	ERRAT VALUE	PROCHECK VALUE	
				Favored Region	Disallowed Region
WP_003811927.1	6ahoA	0.850	94.3262	94.9 %	0.0 %
WP_010929554.1	6ykmA	0.812	79.1757	84.8 %	2.6 %
WP_010929693.1	4r0cA	0.922	93.6508	91.4 %	1.3 %
WP_010930348.1	4c0oA	0.917	79.0179	82.0 %	2.2 %
WP_010930831.1	3mydA	0.487	73.9003	82.5 %	3.6 %
WP_010930949.1	5mrwA	0.937	89.1192	91.4 %	1.4 %
WP_019247543.1	4binA	0.751	86.1915	86.8 %	2.5 %
WP_023853393.1	5byhM	0.832	84.7191	85.5 %	2.2 %

processes (Gupta et al., 2010). On the other hand, larger amino acids sequences are usually found to be involved in critical metabolic pathways (Haag et al., 2012).

2.3. Finding of essential proteins

Essential proteins are defined as proteins that are required for the survival of an organism or a cell. They are of particular interest, not only for their essential biological functions but also in practical applications, such as identifying effective drug targets for pathogenic bacteria and fungi. BLAST analysis was performed by utilizing non-paralog protein sequences against the DEG server database (www.essentialgene.org), known for identifying essential proteins. DEG server includes essential protein coding genes determined by genome-wide gene essentiality analysis and used to find essential proteins (Luo et al., 2014). A cut-off value of 1e-100 for the e-value and a bit score of 100 were considered.

2.4. Screening of orthologs in host

It is desirable to exclude potential targets that have human counterparts to avoid potential toxicity. While toxicity studies must be performed on any new compound intended for human use, comparative analyses of potential bacterial targets and human genome sequences can identify bacterial components that share homology and may thus produce undesired effects on the host (Fields et al., 2017). BLASTp was performed in NCBI database against the human genome for non-paralog essential proteins to screen ortholog in the host. Default settings were utilized, except for specifying *Homo sapiens* (taxid: 9606) in the “organisms” box. Only queries that yielded no significant hits were kept for

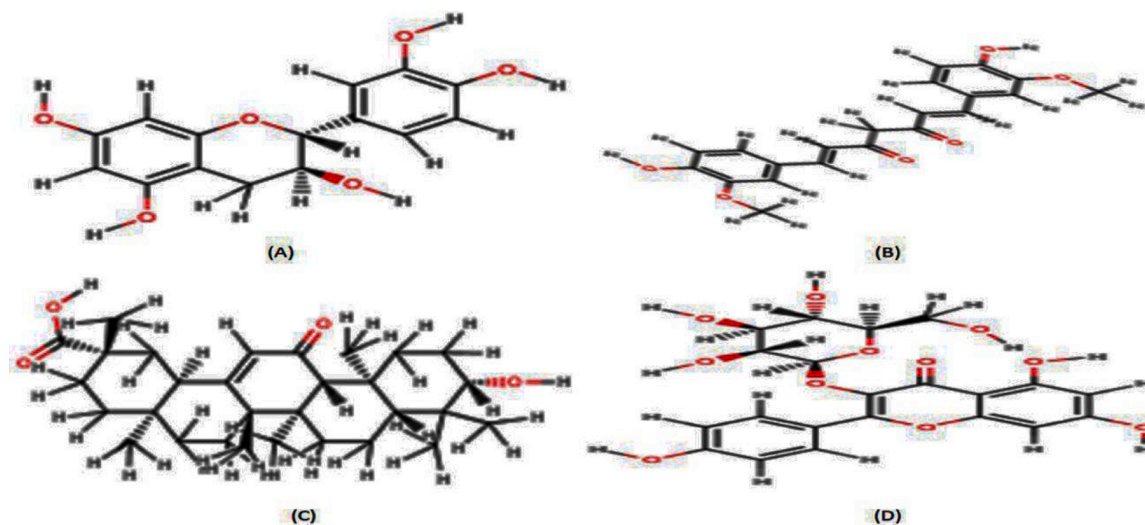


Fig. 3. Chemical structures of selected metabolites. (a) Catechin, (b) Curcumin, (c) 18-beta-glycerrhetenic acid, (d) Astragalin.

Table 7
Binding sites of catechin, curcumin, 18-beta-glycerrhetenic acid, and astragalin.

Chemical Name	Protein ID	Global Binding Energy	Binding Sites
Catechin	WP_003811927.1	-40.57	Glu 63, Arg 66, Ser 68, Tyr 75, Glu 233
	WP_010929554.1	-40.52	Arg 11, Glu 15, Glu 379
	WP_010929693.1	-45.76	Ser 108, Asn 111, Thr 112, Leu 116, Phe 146, Thr 147, Asp 401
	WP_010930348.1	-44.88	Arg 4, Ile 163, Leu 286, His 287, Asp 343,
	WP_010930831.1	-40.28	Leu 462, Arg 653
	WP_010930949.1	-47.02	Ser 414, Tyr 417, Ile 457, Thr 460, Asn 501, Asn 502, Val 456
	WP_019247543.1	-41.37	Thr 136, Ala 195, Gln 201, Val 348,
	WP_023853393.1	-40.43	Ala 189, Pro 196, Leu 197, Gly 200, Gln 201, Asn 349
Curcumin	WP_003811927.1	-40.20	Ser 67, Lys 121, Leu 122, Tyr 154, Glu 157, Ser 234
	WP_010929554.1	-45.45	Arg 11, Gln 20, Thr 377, Gly 432
	WP_010929693.1	-50.66	Gln 62
	WP_010930348.1	-40.51	Asp 299, Arg 456
	WP_010930831.1	-48.58	Arg 667
	WP_010930949.1	-48.27	Ser 126, Gly 381, Gly 504, Ala 506, Asn 513
	WP_019247543.1	-53.96	Gln 187, Pro 196
	WP_023853393.1	-51.87	Pro 196, Leu 197
18-beta-glycerrhetenic acid	WP_003811927.1	-29.28	Arg 230
	WP_010929554.1	-44.09	Arg 11, Gln 20
	WP_010929693.1	-63.50	Tyr 109, Thr 303
	WP_010930348.1	-39.81	Ser 419
	WP_010930831.1	-48.60	Leu 462
	WP_010930949.1	-47.36	Gln 120, Asn 272
	WP_019247543.1	-64.75	Lys 138, Asn 349, Leu 350
	WP_023853393.1	-61.36	Asn 349, Leu 350
Astragalin	WP_003811927.1	-20.25	Arg 66, Ser 67, Ser 68
	WP_010929554.1	-49.98	Arg 11, Cys 8, Leu130, Glu 379, Asp 429,
	WP_010929693.1	-59.33	Ile 66, Tyr 109, Leu 110, Asn 111, Trp 112, Gly 300, Thr 303,
	WP_010930348.1	-37.66	Asp 142, Asp 240, Gly 242, Ile 244, Gly 245, Phe 309, Lys 338
	WP_010930831.1	-41.62	Pro 463, Lys 502
	WP_010930949.1	-42.72	Trp 119, Gln 120, Gly 121, Asn 270, Gly 504,
	WP_019247543.1	-58.15	Gln 133, Ala 195, Pro 196, Leu 197, Gln 201, Asn 349
	WP_023853393.1	-55.26	Pro 188, Ala 189, Ala 195, Pro 196, Gly 200, Gln 201, Leu 350,

further analysis, aiming to develop pathogen-specific remedies (Pourhajibagher and Bahador, 2016).

2.5. Selection of unique metabolic pathway protein

Both the metabolic pathways of *B. pertussis* and humans were collected from KEGG PATHWAY database (https://www.genome.jp/kegg/pathway.html) utilizing their respective three-letter KEGG organism codes: 'bpe' for *B. pertussis* and 'hsa' for *H. sapiens* (Ogata et al., 1999). After that, unique pathways specific to the bacterium were screened via manual comparison. BLASTp analysis through the KAAS server (https://www.genome.jp/kaas-bin/kaas_main) was performed to get the KO numbers of essential host non-homologous proteins. Identified KO numbers were then used to search the KO server (KEGG ORTHOLOGY) (https://www.genome.jp/kegg/ko.html) for the pathways associated with these proteins, and identified the proteins involving in the unique metabolic pathways of the pathogen (Damte et al., 2013). Proteins participating in common pathways were excluded from consideration.

2.6. Analysis of cellular localization

Cytoplasmic proteins have the potential to be used as drug targets, while membrane proteins can serve as both drug targets and vaccine candidates (Mahmud et al., 2019). Four different servers: CELLO v.2.5 (https://cello.life.nctu.edu.tw) (Yu et al., 2006), PSORTb v3.0.2 (https://www.psort.org/psortb) (Yu et al., 2010), ngLOC server (https://genome.unmc.edu/ngLOC) (King and Guda, 2007), and PSLpred (https://webs.iiitd.edu.in/raghava/pslpred/submit.html) (Bhasin et al., 2005) were utilized for predicting the subcellular location of selected proteins. Different servers were utilized to increase the accuracy of the prediction. The final subcellular locations of the proteins were determined based on the consensus of predictions from at least three servers. PSORTb includes new analytical modules designed to capitalize on new discoveries and observations in protein sorting, and benefits from a training dataset of over 11600 proteins of known localization (Yu et al., 2010). CELLO is a multi-class SVM classification system (Yu et al., 2006). ngLOC is a web-based interface for predicting the subcellular localization of the user-supplied protein sequence(s). ngLOC can predict a wide range of subcellular locations including multiple localizations of proteins, and it can be customized to work with a variety of datasets from prokaryotes to eukaryotes, including plant sequences (King and Guda, 2007). PSLpred is a hybrid approach-based method that integrates PSI-BLAST and three SVM modules based on compositions of residues, dipeptides, and physicochemical properties. The prediction accuracies of 90.7, 86.8, 90.3, 95.2, and 90.6 % were attained for cytoplasmic, extracellular, inner-membrane, outer-membrane, and periplasmic proteins, respectively (Bhasin et al., 2005).

2.7. Conservancy analysis of the selected sequences with other strains

The conservancy pattern of the predicted sequences with other strains is essential for determining the range of drug spectrum among the entire homologous bacterial community. Drugs acting against a protein that has a higher conservancy among the strains of the pathogen will work against all other strains of *B. pertussis* (Khan et al., 2020). Therefore, BLASTp was done against 20 randomly chosen strains of *B. pertussis* for all the individual unique pathway protein, determining the percentage identity of the protein. A higher percentage identity indicates greater conservancy. A minimum identity threshold of 99 % was considered during the analysis.

2.8. Finding of novel drug targets using drugbank database

DrugBank 5.1.7 (https://www.drugbank.ca/structures/search/bonds/sequence) was used to find novel drug targets (Wishart et al., 2018). Proteins with higher conservancy were searched against this database. The presence of targets denotes their druggable property, while the absence entitles the uniqueness of the proteins, classifying them as "novel target" (Knox et al., 2011). Default parameters were kept during the action.

2.9. Analyzing virulence factors of the selected proteins

Virulence factors are bacteria-associated molecules that are required for a bacterium to cause disease while infecting eukaryotic hosts such as humans. Virulence factors help bacteria to modulate or degrade host defense mechanisms with the help of adhesion, colonization, and

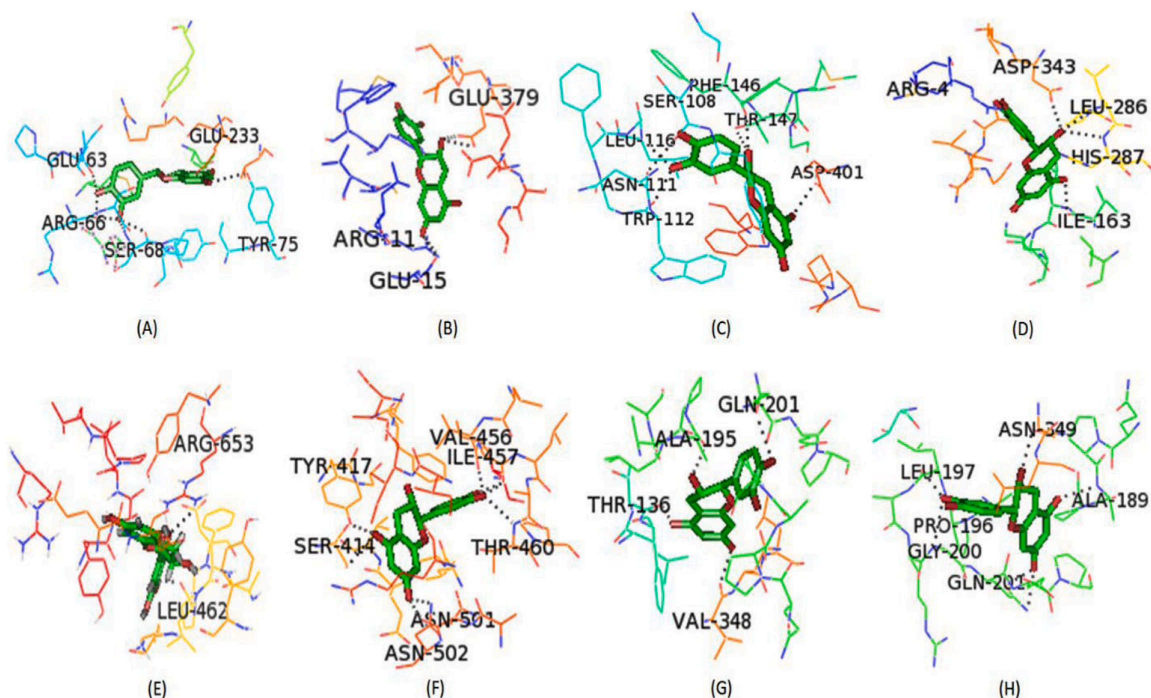


Fig. 4. Binding sites of catechin**.** (a) Chemotaxis protein Mota, (b) Chromosomal replication initiator protein, (c) Short-chain fatty acids transporter, (d) [protein-PII] uridylyltransferase, (e) Type III secretion protein, (f) Potassium-transporting ATPase potassium-binding subunit, (g) N-acetylmuramoyl-L-alanine amidase, (h) RNA polymerase sigma-54 factor.

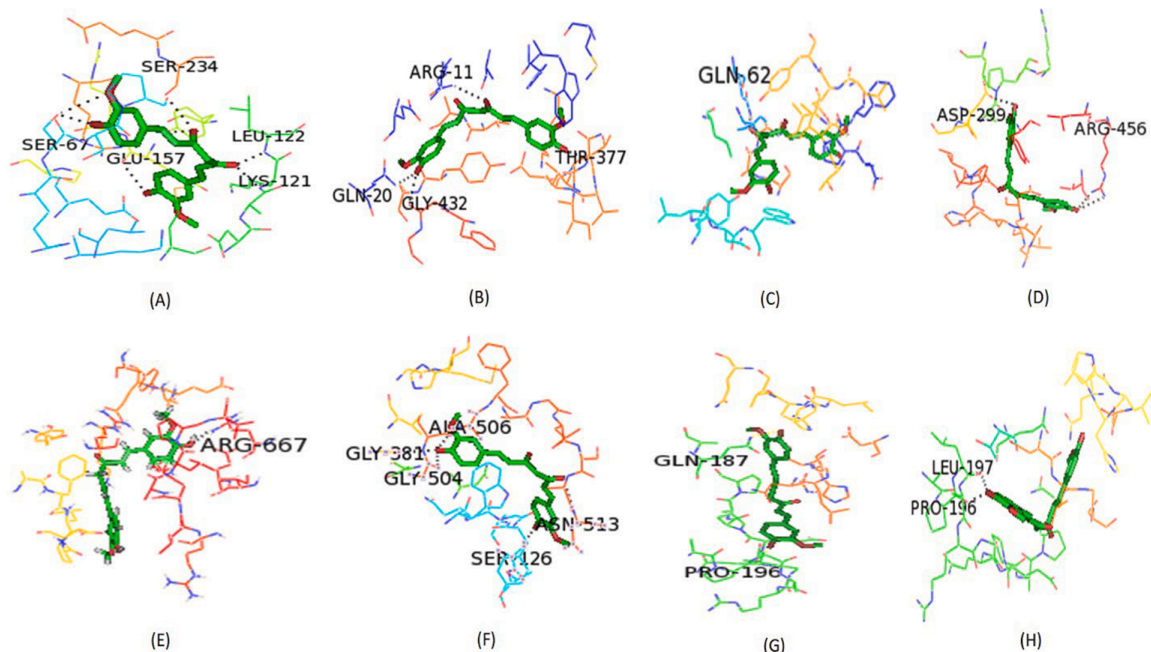


Fig. 5. Binding sites of curcumin**.** (a) Chemotaxis protein Mota, (b) Chromosomal replication initiator protein, (c) Short-chain fatty acids transporter, (d) [protein-PII] uridylyltransferase, (e) Type III secretion protein, (f) Potassium-transporting ATPase potassium-binding subunit, (g) N-acetylmuramoyl-L-alanine amidase, (h) RNA polymerase sigma-54 factor.

invasion. The selected unique drug targets were further analyzed using the protein sequence database from VFDB core dataset (set A) of VFDB server (https://www.mgc.ac.cn/VFs/blast/blast.html) to find whether any of our selected proteins are associated with the virulence of *B. pertussis* (Chen et al., 2005).

2.10. *Resistance Protein Analysis*

Detection of proteins associated with antibiotic resistance mechanisms in the bacteria can be done with the help of ARG-ANNOT server (https://ifr48.timone.univ-mrs.fr/blast/arg-annot_v6.html) (Gupta et al., 2014). All of eight selected proteins were analyzed considering an E-value of 10^{-4} and a bit score of 100.

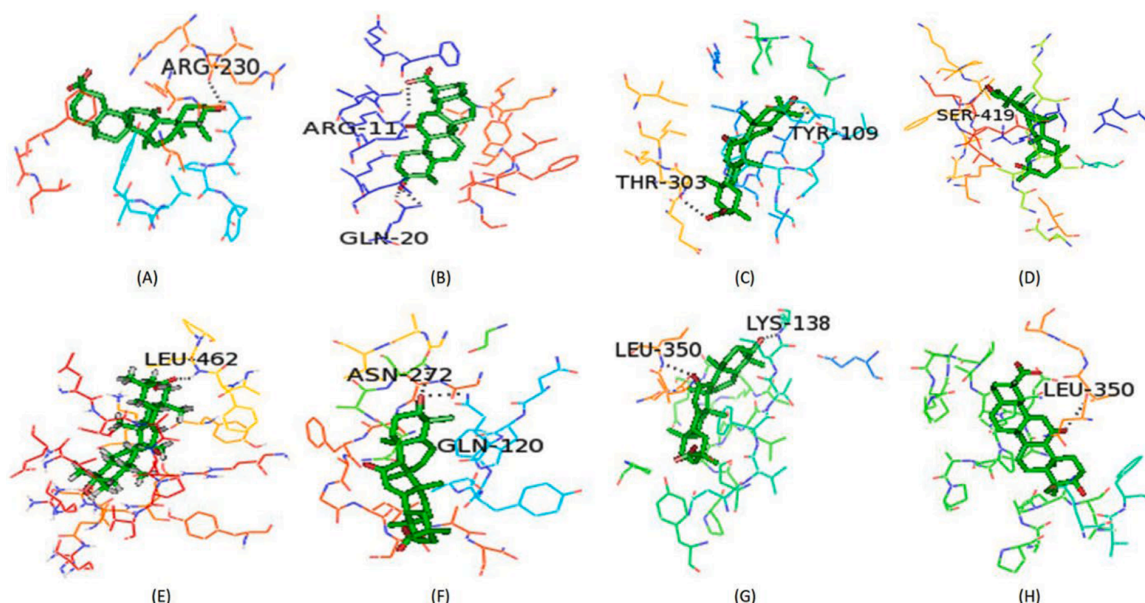


Fig. 6. Binding sites of 18-beta-glycrrhetenic acid. (a) Chemotaxis protein Mota, (b) Chromosomal replication initiator protein, (c) Short-chain fatty acids transporter, (d) [protein-PII] uridylyltransferase, (e) Type III secretion protein, (f) Potassium-transporting ATPase potassium-binding subunit, (g) N-acetylmuramoyl-L-alanine amidase, (h) RNA polymerase sigma-54 factor.

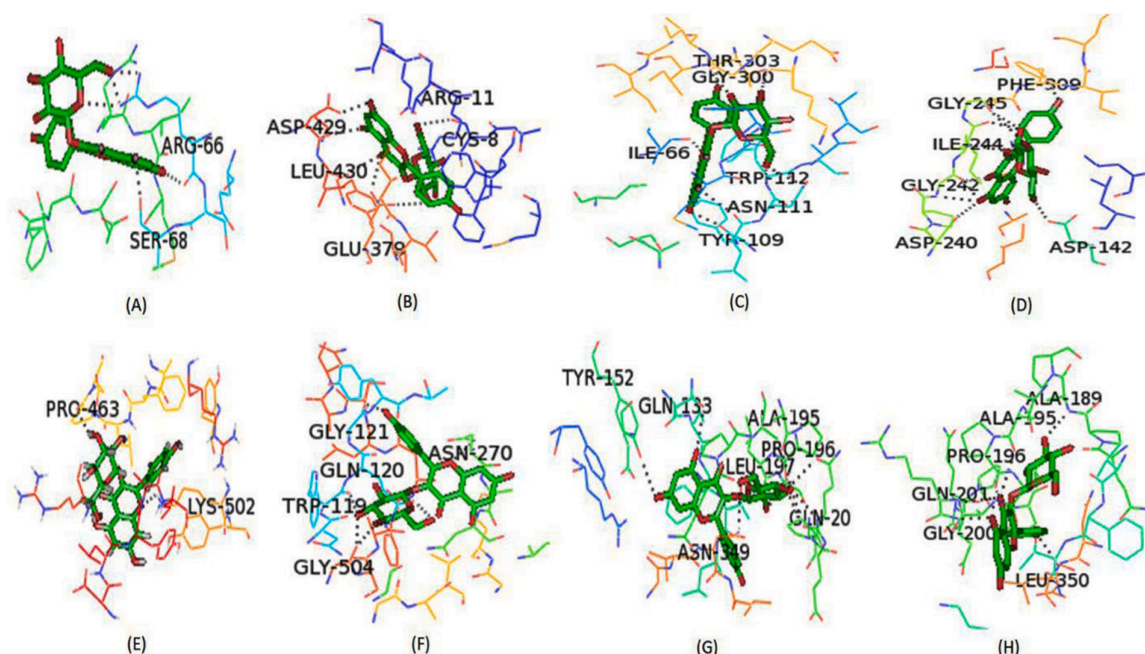


Fig. 7. Binding sites of astragalin. (a) Chemotaxis protein Mota, (b) Chromosomal replication initiator protein, (c) Short-chain fatty acids transporter, (d) [protein-PII] uridylyltransferase, (e) Type III secretion protein, (f) Potassium-transporting aATPase potassium-binding subunit, (g) N-acetylmuramoyl-L-alanine amidase, (h) RNA polymerase sigma-54 factor.

2.11. Interactome Analysis

STRING 11.0 server (https://string-db.org) was utilized to analyze the protein-protein interaction network for the individual selected targets (Szklarczyk et al., 2019). In STRING, one protein interacts with a number of proteins and showed the strength of interaction as score. The interacting score depends on neighborhood in the genome, gene fusions, co-occurrence across genomes, co-expression, association in curated databases and text mining (Jensen et al., 2009). The protein network encompasses high confidence interactors with scores $\geq$ 0.700 to avoid false negative and false positive. Protein sequences were submitted, and the required score was set at "High confidence (0.700)''. The significance of the query protein in the bacterial metabolic system was figured based on the quantity of interacting proteins (nodes) and interactions (edges) interrupted upon its removal (Kushwaha and Shakya, 2010).

2.12. Host-pathogen protein-protein interactions

With a view to exploring the molecular mechanism of pathogenicity and identify pathogenic proteins that interact with host proteins, we analyzed host-pathogen protein-protein interactions (HP-PPIs) (Kumar et al., 2016). HP-PPIs of selected proteins with host proteins were

Table 8
SwissADME properties of top metabolites.

Parameter		Metabolites			
		Catechin	Curcumin	18-beta-glycirrhetenic acid	Astragalin
Physicochemical parameters	Formula	$C_{15}H_{14}O_6$	$C_{21}H_{20}O_6$	$C_{30}H_{46}O_4$	$C_{21}H_{20}O_{11}$
	Molecular weight	290.27 g/mol	368.38 g/mol	470.68 g/mol	448.38 g/mol
	Num. heavy atoms	21	27	34	32
	Num. H-bond acceptors	6	6	4	11
	Num. H-bond donors	5	2	2	7
	Molar Refractivity	74.33	102.80	136.85	108.13
	TPSA	110.38 Å^2	93.06 Å^2	74.60 Å^2	190.28 Å^2
Lipophilicity	Log $P_{o/w}$ (iLOGP)	1.33	3.27	3.47	1.29
	Log $P_{o/w}$ (XLOGP3)	0.36	3.20	5.49	0.72
	Log $P_{o/w}$ (WLOGP)	1.22	3.15	6.41	-0.24
	Log $P_{o/w}$ (MLOGP)	0.24	1.47	4.87	-2.10
	Log $P_{o/w}$ (SILICOS-IT)	0.98	4.04	5.55	-0.12
	Consensus Log $P_{o/w}$	0.83	3.03	5.16	-0.09
Pharmacokinetics	GI absorption	High	High	High	Low
	BBB permeant	No	No	No	No
	P-gp substrate	Yes	No	Yes	No
	CYP1A2 inhibitor	No	No	No	No
	CYP2C19 inhibitor	No	No	No	No
	CYP2C9 inhibitor	No	Yes	No	No
	CYP2D6 inhibitor	No	No	No	No
	CYP3A4 inhibitor	No	Yes	No	No
	Log K_p (skin permeation)	-7.82 cm/s	-6.28 cm/s	-5.27 cm/s	-8.52 cm/s
Water solubility	Log S (ESOL)	-2.22	-3.94	-6.15	-3.18
	Solubility	1.74e+00 mg/ml; 5.98e-03 mol/l	4.22e-02 mg/ml; 1.15e-04 mol/l	3.32e-04 mg/ml; 7.06e-07 mol/l	2.97e-01 mg/ml; 6.61e-04 mol/l
	Class	Soluble	Soluble	Poorly soluble	Soluble
	Log S (SILICOS-IT)	-2.14	-4.45	-6.00	-2.10
	Solubility	2.09e+00 mg/ml; 7.19e-03 mol/l	1.31e-02 mg/ml; 3.56e-05 mol/l	4.75e-04 mg/ml; 1.01e-06 mol/l	3.55e+00 mg/ml; 7.91e-03 mol/l
	Class	Soluble	Moderately soluble	Moderately soluble	Soluble
Medicinal chemistry	PAINS	1 alert: catechol_A	0 alert	0 alert	0 alert
	Brenk	1 alert: catechol	2 alerts: beta_keto_anhydride, michael_acceptor_1	0 alert	0 alert
	Leadlikeness	Yes	No; 2 violations: MW>350, Rotors>7	No; 2 violations: MW>350, XLOGP3>3.5	No; 1 violation: MW>350
	Synthetic accessibility	3.50	2.97	6.08	5.29

identified using the Host-Pathogen Interaction Database (HPIDB) (https://hpidb.igbb.msstate.edu/sequence.html). Selected unique protein sequences were submitted, and the predicted network was then analyzed.

2.13. Prediction of 3D structures of the proteins

The structures of the selected proteins did not exist in Protein Data Bank (RCSB PDB). So I-TASSER server was used for predicting the homologous molecular models of the unique proteins (Roy et al., 2010), which were further refined via GalaxyWEB server (http://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE) (Ko et al., 2012). The best model for each protein was then chosen according to their quality at ERRAT server (https://servicesn.mbi.ucla.edu/ERRAT) (Colovos and Yeates, 1993) and Ramachandran plot analysis at PROCHECK server (https://servicesn.mbi.ucla.edu/PROCHECK) (Laskowski et al., 1996).

2.14. Collection of metabolites structure

Currently, different plants are used to treat whopping cough in different regions of the world as home remedies without proven scientific data. By searching literature, active metabolites with antibacterial properties from these plants were identified to check their potency to be acted as drug. The 3D structures of these metabolites were collected from PubChem server (https://pubchem.ncbi.nlm.nih.gov) in SDF (3D) format (Kim et al., 2016), and then converted to PDB format using Open Babel v2.3 software (O'Boyle et al., 2011).

2.15. Docking analysis

In drug discovery, the interaction between small ligands and macromolecules can be modeled by molecular docking (Kitchen et al., 2004). PatchDock Server (https://bioinfo3d.cs.tau.ac.il/PatchDock/php.php) was utilized for docking purposes (Schneidman-Duhovny et al., 2005), as docking provides interactions between drug targets and potential therapeutics (Meng et al., 2011). Here, the proteins were set as the receptor and the chemicals as the ligand. Erythromycin was considered as the standard metabolite for docking analysis, as it is currently used as the treatment of *B. pertussis* (Halperin et al., 1997). FireDock refinement tool (Mashiach et al., 2008) was then used to refine the docked complexes (Mashiach et al., 2008). Finally, PyMOL v2.0 tool was used for visualization and binding site analysis (Wang et al., 2015).

2.16. Pharmacoinformatics studies

Adsorption, distribution, metabolism, and excretion (ADME) properties are mainly associated with the kinetics of drugs exposure to the tissue. Again, analysis of ADME properties during discovery phase will reduce the risk of pharmacokinetics related failure in the clinical phase (Hay et al., 2014). SwissADME server was utilized to assess the ADME properties of top four metabolites (Daina et al., 2017). SDF format of the drugs was uploaded in the server and converted to SMILES, and then run to get the predictions. Additionally, the blood-brain barrier (BBB) permeability of the studied compounds was calculated by the

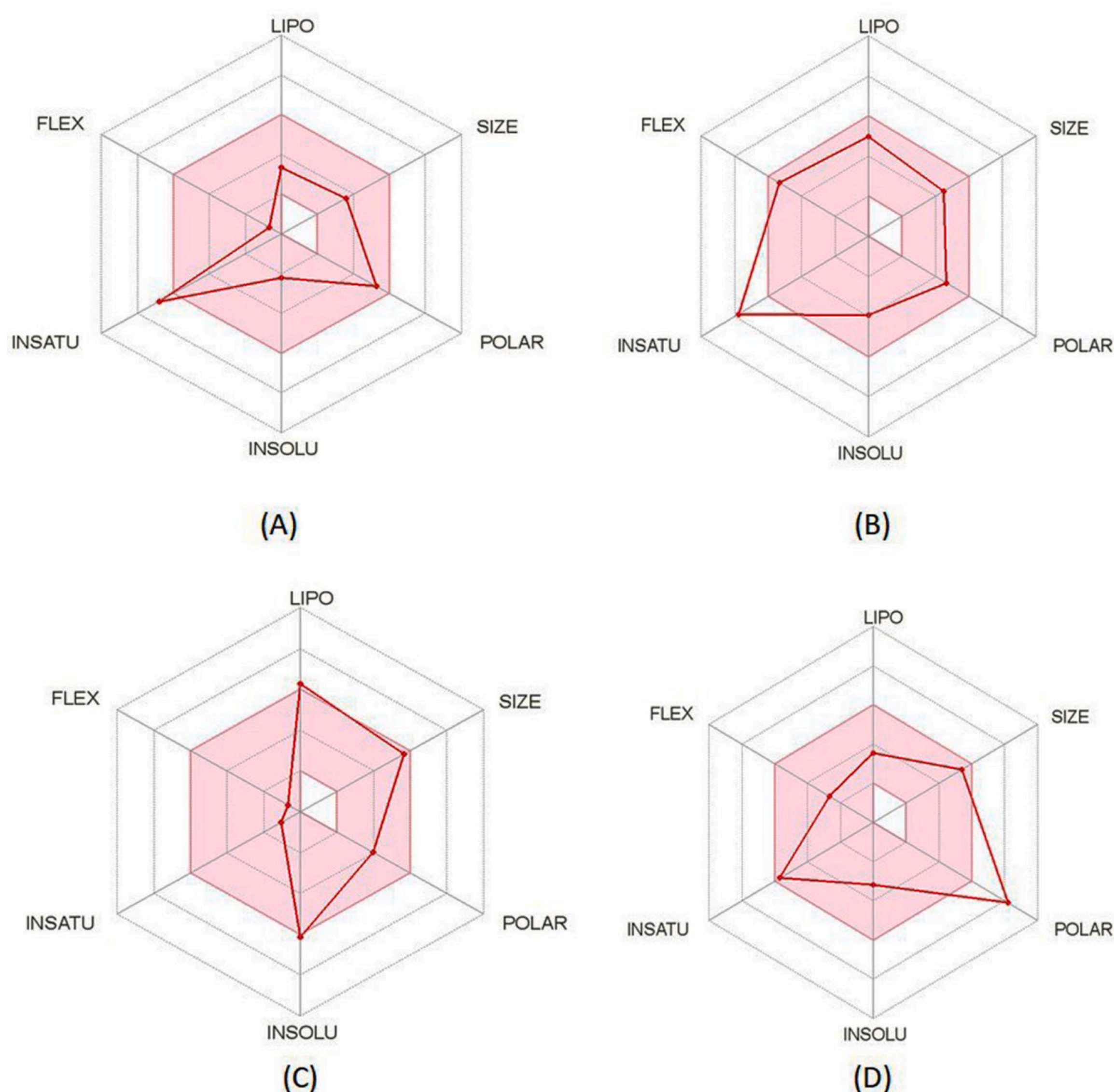


Fig. 8. ADME properties of selected metabolites. (a) Catechin, (b) Curcumin, (c) 18-beta-glycirrhetenic acid, (d) Astragalin.

Table 9
Toxicity parameter of selected chemicals.

Toxicity Parameter	Chemical Name			
	Catechin	Curcumin	18-beta-glycirrhetenic acid	Astragalin
AMES Toxicity	No	No	No	No
Max. Tolerated Dose (log mg/kg/day)	0.438	0.081	0.196	0.582
hERG I inhibitor	No	No	No	No
hERG II inhibitors	No	No	No	No
Oral Rat Acute Toxicity, LD50 (mol/kg)	2.428	1.833	2.735	2.546
Oral Rat Chronic Toxicity, LOAEL (log mg/kg_bw/day)	2.5	2.228	1.664	4.53
Hepatotoxicity	No	No	No	No
Carcinogen	No	No	No	No
Skin Sensitization	No	No	No	No
Minnow Toxicity (log mM)	3.585	-0.081	1.026	6.735

BOILED-Egg model (Daina and Zoete, 2016).

2.17. Toxicity analysis

pkCSM, an online tool, was utilized to predict the comparative toxic effects of top drugs (Pires et al., 2015). Moreover, admetSAR was used to analyze the carcinogenicity of these selected drugs (Cheng et al., 2012). Both of these servers require SMILES that were collected from PubChem for the top four metabolites. Table 1 represents acronyms stated above with their meanings and links of the servers.

3. Results

3.1. Removal of duplicate proteins

NCBI contains 889 genomes for *B. pertussis* in total. From these, we collected the proteome of Tohama I strain, which contains a total of 3,359 proteins. At 60 % identity, CD-HIT server found 3054 clusters among which 34 groups contained paralogous proteins that are greater than 100 amino acids (Supplementary file 1). The paralogous sequences were eliminated leaving 3054 non-duplicate large proteins.

3.2. Screening of essential proteins

After BLAST analysis against the DEG server, proteins for which significant hit was found under the selected conditions were considered as essential proteins for *B. pertussis*. Among 3,054 non-paralog proteins, 509 proteins were identified as essential for the organism, and these were selected for further analysis (Supplementary file 2). These proteins were kept for further analysis as blocking those proteins will kill or weekend the bacteria and the non-essential ones were eliminated as blocking those proteins will not affect the pathogen significantly.

3.3. Removal of orthologs in host

Drugs and therapeutic compounds should be such that they don't cause the unintentional blocking of host proteins. This step is done for diminishing undesirable binding of the drugs to the active sites of the host homologous proteins (Sarkar et al., 2012). Therefore, orthologs were excluded from the list. In this step, 159 non-paralog essential proteins showed no hits with human proteins during BLASTp analysis in NCBI database (Supplementary file 3). That's why these proteins were considered as non-ortholog to the host. These proteins were considered for further analysis as the drugs targeting these proteins will not block the human proteins.

3.4. Finding of unique metabolic pathway proteins

The KEGG server contained 115 *B. pertussis* metabolic pathways, alongside 337 human metabolic pathways, among which 38 metabolic pathways were found to be unique for *B. pertussis* (Supplementary file 4). Proteins within these unique pathways could be potential targets for therapeutics. After BLASTp analysis at KAAS server, 143 out of 159 non-homolog essential proteins were found to have KO orthology and participate in metabolic pathways. These 143 proteins play key roles in metabolism for bacterial survival, among which 14 proteins were exclusively associated with *B. pertussis* unique pathways (Table 2). Targeting these proteins will ensure that the provided drug may not affect the human metabolic pathway as the listed pathways in Table 4 is absent in human.

3.5. Identification of subcellular location

All unique pathway proteins (14) identified were found to be located either in the cytoplasm or the inner membrane (Supplementary file 5). More specifically, five proteins were located in the cytoplasm, while the rest were located in the inner membrane (Table 3). We kept all of them for analysis in the next steps, as both cytoplasmic and membrane proteins can act as a putative drug target.

3.6. Conservancy among the species

Ten out of the fourteen unique pathway proteins showed identity over 99 % across all selected strains, indicating a higher conservancy (Supplementary file 6). Therefore, they were nominated for the succeeding steps. The designed drugs targeting these proteins will be effective against broad ranges of *B. pertussis* strains. On the other hand, four proteins were excluded from further analysis, as they showed identity lower than 99 % against any of the selected strains. If the drugs are designed against these proteins, the drugs may not work against all strains as the target proteins evolve fast. Proteins showing higher conservancy is provided in Table 4.

3.7. Druggability analysis

Among our selected proteins, two proteins showed similarity with approved, investigational, and experimental drug targets listed in the Drugbank Database (Supplementary file 7). These proteins were eliminated from the list, as we aimed to find novel targets. However, eight proteins didn't show any similarity to known drugs targets, allowing them to be marked as a novel drug target. Therefore, they were finally selected for further analysis (Table 5).

3.8. Analyzing virulence factors and resistance protein

Our analysis showed that Type III secretion protein V (WP_010930831.1) was associated with the virulence of the pathogen. Designing drugs against this protein offers some extra facilities over the other selected proteins, as blocking this protein will disarm the bacterium instead of killing (Totsika, 2017). On contrary, results of analysis via ARG-ANNOT server confirmed that none of the selected proteins were associated with antibiotic resistance mechanism.

3.9. Protein-protein interaction network analysis

Proteins that interact with more proteins are considered metabolically active and they are suitable to be considered as drug target (Cui et al., 2009, Kushwaha and Shakya, 2010). STRING analysis revealed that chemotaxis protein MotA, chromosomal replication initiator protein, short-chain fatty acids transporter, [protein-PII] uridylyltransferase, type III secretion protein V, potassium-transporting ATPase potassium-binding subunit, N-acetylmuramoyl-L-alanine amidase, and RNA polymerase sigma-54 factor (rpoN), exhibits interactions with 10, 7, 3, 5, 10, 4, 5, and 5 proteins, respectively (Fig. 1). The result indicated that blocking these proteins will interrupt the interacting proteins indirectly and will hamper the associated pathways of the pathogen.

3.10. Host-pathogen protein-protein interactions

Among the eight unique drug targets, only chromosomal replication initiator protein (WP_010929554.1) and Type III secretion protein V (WP_010930831.1) were found to interact with human proteins (Fig. 2). The chromosomal replication initiator protein interacts with Laminin subunit alpha-5 (LAMA5) and Tubulointerstitial nephritis antigen-like (TINAGL1) proteins. On the other hand, Type III secretion protein V interacts with fMet-Leu-Phe receptor (FPR1), Proteasome activator complex subunit 1 (PSME1), SNW domain-containing protein 1 (SNW1), and BTB/POZ domain-containing protein KCTD5 (KCTD5) proteins.

3.11. Molecular modeling and quality assessment

The I-TASSER server provided five models for each protein. I-TASSER modeling started from the structure templates identification by LOMETS from the PDB library, using the templates of the highest significance in threading alignments. TM-score (template modeling score), and accession no of the PDB hits are listed in Table 6. After initial modeling, the ERRAT values and Ramachandan plots were considered for selecting the best model. Then, the refinement was done via GalaxyWEB server, resulting in ten refined models for each protein. Finally, the best model was selected by analyzing ERRAT quality scores and Ramachandran plots from the PROCHECK server. The ERRAT and PROCHECK scores are provided in Table 6.

3.12. Molecular docking and binding site analysis

The list of selected plant metabolites is provided in supplementary file 8. All of our selected unique proteins (macromolecules) were docked against these plant metabolites (ligands) (Supplementary file 9). Among them, catechin, curcumin, 18-beta-glycrrhetenic acid, and astragalin (Fig. 3) showed superior or nearly equal global binding energy compared to erythromycin across all selected proteins. Consequently, we further analyzed the binding residues of these metabolites using Pymol tool.

To elucidate the drug surface hotspots of the targeted unique

proteins, the structural conformations of the docked complexes was investigated. The ligand binding patterns and interacting residues along with their respective positions were examined (Table 7, Figs. 4, 5, 6, and 7). It has been found that amino acids spanning positions 63-75, 121-157, and 230-234 were vital for the binding interactions of the chemotaxis protein MotA. Similarly, residues at positions 8-15, 377-379, and 429-432 were identified as binding residues for the chromosomal replication initiator protein. Furthermore, the ligands exhibited maximum binding affinity for regions spanning positions 108-116 and 303-306 of the short-chain fatty acids transporter. Similarly, all the binding residues of RNA polymerase sigma-54 factor were located between positions 188-201 and 349-350. Amino acids spanning positions 133-201 were the hotspot for N-acetylmuramoyl-L-alanine amidase.

3.13. Pharmacoinformatics study

The ADME properties of top drugs were evaluated to characterize drug profiles of top antibacterial drugs (Table 8, Fig. 8). All metabolites except astragalin showed robust gastrointestinal absorption, while none showed BBB permeability among the selected top drugs. They did not show interaction probability with cytochromes P450 isoforms when their inhibitory effects were analyzed with different CYP isoforms, and all drugs showed poor to high-level water solubility. Furthermore, only catechin showed a single alert related to potential pain (Table 8).

3.14. Toxicity prediction

Several toxicity criteria of top metabolites are outlined in Table 9. The negative results from the AMES test revealed that all drugs were non-mutagenic, and none of them were identified as hERG I and hERG II inhibitor. In addition, the LD50 values for these top drugs ranged from 1.833 to 2.735 mol/kg, with negative results observed for skin sensitization and oral rat acute toxicity. Minnow Toxicity values of all drugs were more than -0.3 log mM, except curcumin, proving them non-toxic. Besides, these drugs were non-carcinogenic as they provided negative results on admetSAR carcinogenicity prediction. Additionally, negative hepatotoxicity results of all drugs indicate that the normal function of the liver will not be disrupted through these drugs.

4. Discussion

In many developed and developing countries throughout the world, whopping cough still considered as a significant public health concern. Shortcomings of available aP vaccine as well as the development of rapid antibiotic resistance making it difficult to combat this pathogen. Therefore, discovering new drugs is becoming an essential task to fight with it. Here, we explored the proteome of *B. pertussis,* Tohama I strain through subtractive proteomics approach to identify novel drug targets, and molecular docking was performed to screen probable drugs against the essential non homologous unique pathway proteins of *B. pertussis.*

Essential proteins are considered as most suitable antimicrobial therapeutic targets, as most drugs tend to dock with these essential gene products. Our target pathogen has 509 such essential proteins, indicating that drugs could act against these proteins to kill the bacteria. Subtracting host homolog proteins is a crucial step in the in silico selection of drug targets (Anishetty et al., 2005, Sarkar et al., 2012). After analysis, 159 such non-orthologous proteins were found, against which drugs could be designed without risking cross reactivity (Sarkar et al., 2012). Among these non-orthologs, 143 proteins were predicted to be associated with the metabolism of the bacteria, but only fourteen of them were found to be involved with the unique metabolic pathways of *B. pertussis*. Proteins localized in either cytoplasm or membrane are considered as the best drug targets (Michael et al., 2014). Interestingly, all of the unique pathway proteins were located either in the cytoplasm or membrane, indicating their potential as drug targets. Although the research was started with the proteome analysis of a specific strain, Tohama I, ultimately conservancy of all unique proteins among other strains was checked with rigid criteria of 99 % similarity to find out universal targets applicable to all strains. Subsequently, ten proteins meeting these criteria were shortlisted, and DrugBank databases were screened for avoiding mutational changes and to inhibit the development of resistant bacteria by the broad-spectrum drugs. Two out of ten highly conserved proteins were druggable, and potentially susceptible to already approved and available drugs. Therefore, the remaining eight proteins were selected as the novel drug targets.

Virulence factors and proteins associated with antibiotic resistance provide extra facilities if they are targeted for drug action. Following advantages can be gained by targeting virulence factors: 1) antimicrobial development with novel working mechanisms 2) diminished selective pressure will cause reduced resistance emergence, and 3) potential preservation of intestinal biomass (Heras et al., 2015). Importantly, a 'disarm-don't kill' strategy enables the delivery of antimicrobials which can replace the available failing antibiotics, while avoiding two main shortcomings: destruction of gut micro biota and antibiotic resistance (Totsika, 2017). Only type III secretion protein V was found to be associated with the virulence of the pathogen, but none were associated with the development of antibiotic resistance in *B. pertussis*. Protein-protein interactions (PPIs) for all unique proteins were revealed, as it allows us to understand the role of a protein in a particular pathway and alterations in the protein-protein network disrupt the natural flow of events of the cell (Genengnews, 2005, March 5). Overall, Type III secretion protein V could be considered as the best as it is associated with the virulence of the pathogen, which interacts with the highest number of proteins and also interacts with the host proteins.

Metabolites from plants play an important role by being a lead molecule in finding suitable drug candidates (Joseph et al., 2017). Therefore, we assessed some inhibitory plant metabolites for *B. pertussis* based on their affinity of binding to the selected unique drug targets. The Docking result revealed that four drug molecules i.e. curcumin, catechin, 18-beta-glycrrhetenic acid, and astragalin showed high affinity for each of the eight proteins with the lowest global binding energy. Both curcumin and 18-beta-glycrrhetenic acid had maximum binding affinity with N-acetylmuramoyl-L-alanine amidase (-53.96 kcal/mol and -64.75 kcal/mol, respectively). Catechin had the highest interacting affinity with potassium-transporting ATPase potassium-binding subunit (-47.02 kcal/mol), whereas astragalin showed a peak with short-chain fatty acids transporter (-59.33 kcal/mol).

Poor ADME data is often associated with failure of clinical trials during many of the drug development projects (Shin et al., 2016). Therefore, ADME analysis is critical for such projects which can be performed by in vitro, in vivo, or in silico methods. The top four drug candidates displayed no undesirable consequences in ADME study, which could reduce the drug related properties. All the potential drugs exhibited water solubility and absorption in GI, while catechin showed highest solubility in water. As whooping cough caused by *B. pertussis* is associated with the respiratory system, the four drugs would not create any problem despite of their non-permeate nature to BBB. Toxicity prediction showed that all four drug candidates are non-carcinogenic, non-mutagenic, and insensitive to skin and non-hepatotoxic. Overall, the toxicity analysis revealed that the predicted drugs are safe to administer and can be used as therapeutics to treat *B. pertussis*.

5. Conclusion

Availability of numerous bioinformatics tools has helped to revolutionize the drug discovery process by subtractive proteomics analysis. Our findings from this research could aid in developing suitable therapeutics with less trials and error repeats of assays. Furthermore, this research may also help to save time and money for in vitro research in future, and thus will help to decrease the morbidity and mortality associated with *B. pertussis*. However, in vivo studies are highly

recommended to validate the predicted results in model organisms.

CRediT authorship contribution statement

Md. Nazmul Islam Bappy: Conceptualization, Methodology, Formal analysis, Investigation, Writing – original draft. **Foeaz Ahmed:** Investigation, Formal analysis, Writing – original draft. **Tahera Lasker:** Investigation, Formal analysis, Writing – original draft. **Emran Hossain Sajib:** Formal analysis, Writing – original draft. **Md. Shariful Islam:** Conceptualization, Methodology, Supervision, Project administration, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Not Applicable.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.crmicr.2024.100291.

Data availability

All data generated or analyzed during this study are included in this published article and its supplementary information files.

References

Altunaiji, S.M., Kukuruzovic, R.H., CURTIS, N.C., Massie, J., 2012. Cochrane review: antibiotics for whooping cough (pertussis). Evid.-Based Child Health 7, 893–956.
Bhasin, M., Garg, A., Raghava, G.P., 2005. PSLpred: prediction of subcellular localization of bacterial proteins. Bioinformatics. 21, 2522–2524.
Chan, J.N., Nislow, C., Emili, A., 2010. Recent advances and method development for drug target identification. Trends. Pharmacol. Sci. 31, 82–88.
Chen, L., Yang, J., Yu, J., Yao, Z., Sun, L., Shen, Y., Jin, Q., 2005. VFDB: a reference database for bacterial virulence factors. Nucleic. Acids. Res. 33, D325–D328.
Cheng, F., Li, W., Zhou, Y., Shen, J., Wu, Z., Liu, G., Lee, P.W., Tang, Y., 2012. admetSAR: a comprehensive source and free tool for assessment of chemical ADMET properties. J. Chem. Inf. Model. 52, 3099–3105.
Cherry, J.D., Gornbein, J., Heininger, U., Stehr, K., 1998. A search for serologic correlates of immunity to *Bordetella pertussis* cough illnesses. Vaccine 16, 1901–1906.
Cherry, J.D., 2015. The history of pertussis (whooping cough); 1906–2015: facts, myths, and misconceptions. Curr. Epidemiol. Rep. 2, 120–130.
Cheuka, P.M., Mayoka, G., Mutai, P., Chibale, K., 2016. The role of natural products in drug discovery and development against neglected tropical diseases. Molecules 22.
Colovos, C., Yeates, T.O., 1993. Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci. 2, 1511–1519.
Cui, T., Zhang, L., Wang, X., He, Z.G., 2009. Uncovering new signaling proteins and potential drug targets through the interactome analysis of *Mycobacterium tuberculosis*. BMC. Genomics. 10, 118.
Daina, A., Michielin, O., Zoete, V., 2017. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Sci. Rep. 7, 42717.
Daina, A., Zoete, V., 2016. A boiled-egg to predict gastrointestinal absorption and brain penetration of small molecules. Chem Med Chem 11, 1117–1121.
Damte, D., Suh, J.W., Lee, S.J., Yohannes, S.B., Hossain, M.A., Park, S.C., 2013. Putative drug and vaccine target protein identification using comparative genomic analysis of KEGG annotated metabolic pathways of mycoplasma hyopneumoniae. Genomics 102, 47–56.
David, H., 2008. Pertussis. Control of Communicable Diseases Manual. American Public Health Association, Washington DC.
Dutta, A., Singh, S.K., Ghosh, P., Mukherjee, R., Mitter, S., Bandyopadhyay, D., 2006. In silico identification of potential therapeutic targets in the human pathogen *Helicobacter pylori*. In. Silico Biol. 6, 43–47.
Fields, F.R., Lee, S.W., Mcconnell, M.J., 2017. Using bacterial genomes and essential genes for the development of new antibiotics. Biochem. Pharmacol. 134, 74–86.
GENENGNEWS. 2005, March 5. Protein-protein interactions in drug discovery [Online]. Available: https://www.genengnews.com/magazine/18/protein-protein-interactions-in-drug-discovery/[Accessed].
Ginsburg, H., Deharo, E., 2011. A call for using natural compounds in the development of new antimalarial treatments - an introduction. Malar. J. 10 (Suppl 1), S1.
Gohari, A.R., Saeidnia, S., Mahmoodabadi, M.K., 2013. An overview on saffron, phytochemicals, and medicinal properties. Pharmacogn. Rev. 7, 61–66.
Gupta, S.K., Padmanabhan, B.R., Diene, S.M., Lopez-rojas, R., Kempf, M., Landraud, L., Rolain, J.M., 2014. Arg-annot, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. Antimicrob. Agents Chemother 58, 212–220.
Gupta, S.K., Singh, S., Gupta, M.K., Pant, K.K., Seth, P.K., 2010. Definition of potential targets in *Mycoplasma pneumoniae* through subtractive genome analysis. J. Antivir. Antiretrovir. 2, 038–041.
Güriş, D., Strebel, P.M., Bardenheier, B., Brennan, M., Tachdjian, R., Finch, E., Wharton, M., Livengood, J.R., 1999. Changing epidemiology of pertussis in the United States: increasing reported incidence among adolescents and adults, 1990-1996. Clin. Infect. Dis. 28, 1230–1237.
Haag, N.L., Velk, K.K., WU, C., 2012. Potential antibacterial targets in bacterial central metabolism. Int. J. Adv. Life Sci. 4, 21–32.
Halperin, S.A., Bortolussi, R., Langley, J.M., miller, B., Eastwood, B.J., 1997. Seven days of erythromycin estolate is as effective as fourteen days for the treatment of *Bordetella pertussis* infections. Pediatrics. 100, 65–71.
Hay, M., Thomas, D.W., Craighead, J.L., Economides, C., Rosenthal, J., 2014. Clinical development success rates for investigational drugs. Nat. Biotechnol. 32, 40–51.
Heras, B., Scanlon, M.J., Martin, J.L., 2015. Targeting virulence not viability in the search for future antibacterials. Br. J. Clin. Pharmacol. 79, 208–215.
Hoffman, C.L., Gonyar, L.A., Zacca, F., Sisti, F., Fernandez, J., Wong, T., Damron, F.H., Hewlett, E.L., 2019. *Bordetella pertussis* can be motile and express flagellum-like structures. mBio 10.
Huang, Y., Niu, B., Gao, Y., Fu, L., Li, W., 2010. CD-HIT Suite: a web server for clustering and comparing biological sequences. Bioinformatics. 26, 680–682.
Jakinovich, A., Sood, S.K., 2014. Pertussis: still a cause of death, seven decades into vaccination. Curr. Opin. Pediatr. 26, 597–604.
Jensen, L.J., Kuhn, M., Stark, M., Chaffron, S., Creevey, C., Muller, J., Doerks, T., Julien, P., Roth, A., Simonovic, M., Bork, P., Von mering, C., 2009. STRING 8–a global view on proteins and their functional interactions in 630 organisms. Nucleic. Acids. Res. 37, D412–D416.
Joseph, J., Bhaskaran, R., Kaliraj, M., Muthuswamy, M., Suresh, A., 2017. Molecular docking of phytoligands to the viral protein receptor p. monodon rab7. Bioinformation. 13, 116–121.
Khan, M.T., Mahmud, A., Iqbal, A., Hoque, S., Hasan, M., 2020. Subtractive genomics approach towards the identification of novel therapeutic targets against human *Bartonella bacilliformis*. Inform. Med. Unlocked 20, 100385.
Kim, S., Thiessen, P.A., Bolton, E.E., Chen, J., Fu, G., Gindulyte, A., Han, L., He, J., He, S., Shoemaker, B.A., Wang, J., Yu, B., Zhang, J., Bryant, S.H., 2016. Pubchem substance and compound databases. Nucleic. Acids. Res. 44, D1202–D1213.
King, B.R., GUda, C., 2007. ngLOC: an n-gram-based bayesian method for estimating the subcellular proteomes of eukaryotes. Genome Biol. 8, R68.
Kitchen, D.B., Decornez, H., Furr, J.R., Bajorath, J., 2004. Docking and scoring in virtual screening for drug discovery: methods and applications. Nat. Rev. Drug Discov. 3, 935–949.
Knox, C., Law, V., Jewison, T., Liu, P., Ly, S., Frolkis, A., Pon, A., Banco, K., Mak, C., Neveu, V., Djoumbou, Y., Eisner, R., Guo, A.C., Wishart, D.S., 2011. DrugBank 3.0: a comprehensive resource for 'omics' research on drugs. Nucleic. Acids. Res. 39, D1035–D1041.
Ko, J., Park, H., Heo, L., Seok, C., 2012. GalaxyWEB server for protein structure prediction and refinement. Nucleic. Acids. Res. 40, W294–W297.
Koparde, A.A., Doijad, R.C., Magdum, C.S., 2019. Natural Products in drug discovery. IntechOpen. https://doi.org/10.5772/intechopen.82860.
Kumar, V., Suri, S., Prasad, R., Gat, Y., Sangma, C., Jakhu, H., Sharma, M., 2019. Bioactive compounds, health benefits and utilization of Rhododendron: a comprehensive review. Agric. Food Secur. 8, 6. https://doi.org/10.1186/s40066-019-0251-3.
Kumar, A., Thotakura, P.L., Tiwary, B.K., Krishna, R., 2016. Target identification in Fusobacterium nucleatum by subtractive genomics approach and enrichment analysis of host-pathogen protein-protein interactions. BMC. Microbiol. 16, 84.
Kushwaha, S.K., Shakya, M., 2010. Protein interaction network analysis–approach for potential drug target identification in mycobacterium tuberculosis. J. Theor. Biol. 262, 284–294.
Lam, C., Octavia, S., Ricafort, L., et al., 2014. Rapid increase in pertactin-deficient *Bordetella pertussis* isolates, Australia. Emerg. Infect. Dis. 20, 626–633.
Laskowski, R.A., Rullmannn, J.A., Macarthur, M.W., Kaptein, R., Thornton, J.M., 1996. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J. Biomol. NMR 8, 477–486.
Luo, H., Lin, Y., Gao, F., Zhang, C.T., Zhang, R., 2014. DEG 10, an update of the database of essential genes that includes both protein-coding genes and noncoding genomic elements. Nucleic. Acids. Res. 42, D574–D580.
Mahmud, A., Khan, M.T., Iqbal, A., 2019. Identification of novel drug targets for humans and potential vaccine targets for cattle by subtractive genomic analysis of *Brucella abortus* strain 2308. Microb. Pathog. 137, 103731.

Mashiach, E., Schneidman-duhovny, D., Andrusier, N., Nussinov, R., Wolfson, H.J., 2008. FireDock: a web server for fast interaction refinement in molecular docking. Nucleic. Acids. Res. 36, W229–W232.

Meng, X.Y., Zhang, H.X., Mezei, M., Cui, M., 2011. Molecular docking: a powerful approach for structure-based drug discovery. Curr. Comput. Aided. Drug Des. 7, 146–157.

Michael, C.A., Dominey-howes, D., Labbate, M., 2014. The antimicrobial resistance crisis: causes, consequences, and management. Front. Public Health 2, 145. -145.

O'boyle, N.M., Banck, M., James, C.A., Morley, C., Vandermeersch, T., Hutchison, G.R., 2011. Open babel: an open chemical toolbox. J. Cheminform. 3, 33.

Ogata, H., Goto, S., Sato, K., Fujibuchi, W., Bono, H., Kanehisa, M., 1999. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic. Acids. Res. 27, 29–34.

Pires, D.E., Blundell, T.L., Ascher, D.B., 2015. pkCSM: Predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. J. Med. Chem. 58, 4066–4072.

Pourhajibagher, M., Bahador, A., 2016. Designing and in silico analysis of porb protein from *Chlamydia trachomatis* for developing a vaccine candidate. Drug Res. (Stuttg) 66, 479–483.

Roy, A., Kucukural, A., Zhang, Y., 2010. I-TASSER: a unified platform for automated protein structure and function prediction. Nat. Protoc. 5, 725–738.

Sarkar, M., Maganti, L., Ghoshal, N., Dutta, C., 2012. In silico quest for putative drug targets in Helicobacter pylori HPAG1: molecular modeling of candidate enzymes from lipopolysaccharide biosynthesis pathway. J. Mol. Model. 18, 1855–1866.

Schmidtke, A.J., Boney, K.O., Martin, S.W., et al., 2012. Population diversity among *Bordetella pertussis* isolates, United States, 1935–2009. Emerg. Infect. Dis. 18, 1248–1255.

Schneidman-duhovny, D., Inbar, Y., Nussinov, R., Wolfson, H.J., 2005. Patchdock and symmdock: servers for rigid and symmetric docking. Nucleic. Acids. Res. 33, W363–W367.

Shin, H.K., Kang, Y.-M., No, K.T., 2016. Predicting ADME Properties of chemicals. ed.. In: Leszczynski, J. (Ed.), Handbook of Computational Chemistry. Springer Netherlands, Dordrecht.

Storsaeter, J., Hallander, H.O., Gustafsson, L., Olin, P., 1998. Levels of anti-pertussis antibodies related to protection after household exposure to *Bordetella pertussis*. Vaccine 16, 1907–1916.

Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-cepas, J., Simonovic, M., Doncheva, N.T., Morris, J.H., Bork, P., Jensen, L.J., Mering, C.V., 2019. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic. Acids. Res. 47, D607–d613.

Theofiles, A.G., Cunningham, S.A., Chia, N., Jeraldo, P.R., Quest, D.J., Mandrekar, J.N., Patel, R., 2014. Pertussis outbreak, southeastern Minnesota, 2012. Mayo Clin. Proc. 89, 1378–1388.

Torres, R.S.L.A., Moraes, P.A., Oliveira, K., 2015. *Bordetella pertussis* Antibiotic Resistance in Southern Brazil: A 3-Year Surveillance Study. 28O Congresso Brasileiro de Microbiologia, Santa Catarina, Brasil.

Totsika, M., 2017. Disarming pathogens: benefits and challenges of antimicrobials that target bacterial virulence instead of growth and viability. Future Med. Chem. 9, 267–269.

Wang, Q., He, J., Wu, D., Wang, J., Yan, J., Li, H., 2015. Interaction of α-cyperone with human serum albumin: determination of the binding site by using discovery studio and via spectroscopic methods. J. Lumin. 164.

Warfel, J.M., Zimmerman, L.I., Merkel, T.J., 2014. Acellular pertussis vaccines protect against disease but fail to prevent infection and transmission in a nonhuman primate model. Proc. Natl. Acad. Sci. u S. a 111, 787–792.

Winter, K., Harriman, K., Zipprich, J., Schechter, R., Talarico, J., Watt, J., Chavez, G., 2012. California pertussis epidemic, 2010. J. Pediatr. 161, 1091–1096.

Wishart, D.S., Feunang, Y.D., Guo, A.C., Lo, E.J., Marcu, A., Grant, J.R., Sajed, T., Johnson, D., Li, C., Sayeeda, Z., Assempour, N., Iynkkaran, I., Liu, Y., Maciejewski, A., Gale, N., Wilson, A., Chin, L., Cummings, R., Le, D., Pon, A., Knox, C., Wilson, M., 2018. Drugbank 5.0: a major update to the drugbank database for 2018. Nucleic. Acids. Res. 46, D1074–d1082.

Yang, Y., Yao, K., Ma, X., Shi, W., Yuan, L., Yang, Y., 2015. Variation in *Bordetella pertussis* susceptibility to erythromycin and virulence-related genotype changes in China (1970-2014). PLoS. One 10 (9), e0138941.

Yeung, K.H.T., Duclos, P., Nelson, E.A.S., Hutubessy, R.C.W., 2017. An update of the global burden of pertussis in children younger than 5 years: a modelling study. Lancet Infect. Dis. 17, 974–980.

Yu, C.S., Chen, Y.C., Lu, C.H., Hwang, J.K., 2006. Prediction of protein subcellular localization. Proteins. 64, 643–651.

Yu, N.Y., Wagner, J.R., Laird, M.R., Melli, G., Rey, S., Lo, R., Dao, P., Sahinalp, S.C., Ester, M., Foster, L.J., Brinkman, F.S., 2010. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. Bioinformatics. 26, 1608–1615.