Gaurav Thesis

by Gaurav Thesis

Submission date: 22-May-2026 05:13PM (UTC+0530)
Submission ID: 2967062119
File name: Thesis_Gaurav.pdf (885.2K)
Word count: 15809
Character count: 91246

IN-SILICO CHARACTERIZATION OF PA3042 FROM *Pseudomonas aeuginosa* USING AN INTEGRATED MULTI-EVIDENCE APPROACH

Thesis submitted

in partial fulfilment of the requirements for the

degree of

MASTER OF SCIENCE

in

BIOTECHNOLOGY

by

GAURAV

24/MSCBIO/01

Under the supervision of

Prof. Yasha Hasija

Head of Department

Department of Biotechnology

Department Of Biotechnology

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Shahbad Daulatpur, Main Bawana Road, Delhi – 110042, India

May, 2026

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Shahbad Daulatpur, Main Bawana Road

Delhi – 110042, India

DECLARATION

I ,Gaurav (24/MSCBIO/01) hereby certify that the work which is being presented in the thesis entitled **"In-Silico Characterization of PA3042 from *Pseudomonas aeuginosa* using an integrated multi-evidence approach"** in partial fulfillment of the requirements for the award of the Degree of Master of Science, submitted in the Department of Biotechnology, Delhi Technological University is an authentic record of my own work carried out during the period from 2025 to 2026 under the supervision of Prof. Yasha Hasija.

The matter presented in the thesis has not been submitted by me for the award of any other degree of this or any other Institute.

Candidate's Signature

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Shahbad Daulatpur, Main Bawana Road

Delhi – 110042, India

SUPERVISOR'S CERTIFICATE

Certified that Gaurav (24/MSCBIO/01) has carried out their search work presented in this thesis entitled **"In-Silico Characterizatio of PA3042 from *Pseudomonas aeuginosa* using an integrated multi-evidence approach"** for the award of Master of Science from Department of Biotechnology, Delhi Technological University, Delhi, under my supervision. The thesis embodies results of original work, and studies are carried out by the student himself and the contents of the thesis do not form the basis for the award of any other degree to the candidate or to anybody else from this or any other University/Institution.

Date:

Signature

Prof. Yasha Hasija

Head of the Department

Department of Biotechnology

Delhi Technological University

(Confirm if need to add the Project Investigator)

"IN-SILICO CHARACTERIZATION OF PA3042 FROM *Pseudomonas aeuginosa* USING AN INTEGRATED MULTI-EVIDENCE APPROACH"

"

GAURAV

(24/MSCBIO/01)

ABSTRACT

Antimicrobial resistance represents a critical global health challenge and motivates efforts to discover and characterize bacterial proteins that mediate pathogenicity and stress adaptation. Pseudomonas aeruginosa is a highly adaptable Gram-negative opportunist known for intrinsic resistance to antibiotics, capacity for persistent biofilm formation, and survival across diverse environments through coordinated regulatory networks such as quorum sensing and two-component signaling systems. Despite extensive research into these major pathways, a substantial fraction of P. aeruginosa proteins remains functionally uncharacterized. In this study, we performed an in-silico characterization of the hypothetical protein PA3042 to predict its structural and functional properties and to assess its potential role in signaling-associated processes. A comprehensive suite of bioinformatics approaches was applied, including physicochemical profiling to evaluate stability and solubility parameters, subcellular localization prediction to infer membrane association, secondary and tertiary structure modelling to generate three-dimensional models, and rigorous structural validation to assess model quality. Conserved domain searches and functional annotation methods were used to identify motifs and putative activities, while binding pocket prediction and molecular docking provided insights into possible ligand interactions. Comparative structural analysis and evolutionary conservation mapping were performed to highlight residues likely to be functionally important and to place PA3042 within a broader evolutionary context.

The collective computational evidence indicates that PA3042 is likely a membrane-associated regulatory protein that may participate in bacterial adaptation and signaling processes relevant to virulence and stress response. These findings offer preliminary structural and functional insights into PA3042 and refine our understanding of the PA3040–PA3042 operon. By generating specific hypotheses about localization, functional motifs, and potential ligand interactions, this work establishes a targeted framework for biochemical and genetic experiments to validate the predicted roles of PA3042 and to explore its contribution to P. aeruginosa biology.

ACKNOWLEDGEMENT

First of all, I'd want to extend heartfelt Sincere gratitude for my supervisor, Prof. Yasha Hasija, for their continuous support and encouragement during the duration of my research. Their expertise and constant constructive remarks have always been invaluable to me in learning and understanding new yet difficult things. They have always been an inspiration to me in pursuing my research and future goalsAdditionally, I want to thank all members of the Complex Systems and Genome Informatics laboratory for always helping and guiding me through each up and down which has come along the way and giving me an environment where I can learn and grow as a person.

I am also sincerely grateful to the faculty especially Prof. of the Department of Biotechnology, Delhi Technological University for providing me an academic environment which is coupled with theoretical as well as practical aspects of academia and providing me every necessary help and resource to carry out my work. A special thanks to Mr. Jitender Singh, Mr. C.B. Singh, Mr. Lalit, Mr. Jaspreet, and Mr. Rajesh for their technical assistance and guidance through the project.

I'd want to whole-heartedly thank you to my loved ones and friends for always being my constant support and well-wishers. Their presence and belief in me have always been a source of motivation and strength, their contribution in my life can't be put into words.

Thank you all for your valuable contributions and sacrifices which made this project a success.

Gaurav

24/MSCBIO/01

CONTENTS

1. Title i
2. Declaration ii
3. Supervisor's certificate iii
4. Abstract iv
5. Acknowledgement v
6. List of Figures vi
7. List of Tables vi
8. Abbreviations vii
9. Introduction 1
10. Literature Review 2
    - 10.1. Pathogenicity and clinical relevance of *Pseudomonas aeruginosa* 3
    - 10.2. Biofilm formation and its regulatory signaling 4
        - 10.2.1.1. Sequential stages of biofilm development 5
        - 10.2.1.2. Initial attachment 6
        - 10.2.1.3. Irreversible anchoring 7
        - 10.2.1.4. Maturation and microcolony formation 8
        - 10.2.1.5. Dispersal 9
      - 10.2.2. Regulatory pathways governing biofilm formation 10
        - 10.2.2.1. Two component system 11
        - 10.2.2.2. C-di-GMP as central regulator 12
        - 10.2.2.3. Quorum sensing 13
    - 10.3. Hypothetical proteins and their significance 14
    - 10.4. Literature background of the target protein 15
    - 10.5. Intracellular signalling architecture in bacteria 16
      - 10.5.1. Receptor proteins
        - 10.5.1.1. Histidine kinase
        - 10.5.1.2. PAS domains
        - 10.5.1.3. PilZ domain
      - 10.5.2. Second Messenger molecules
        - 10.5.2.1. ppGpp
        - 10.5.2.2. c-di-AMP

10.5.2.3. cAMP

10.5.2.4. Non-nucleotide signaling molecules

10.5.3. Effector proteins

10.5.3.1. . Transcriptional modulators

10.5.3.2. Enzymatic Effectors

11. Methodology 17

12. Results 18

13. Conclusion and future perspectives 19

14. References

LIST OF FIGURES

Figure	Description	Page no.
1	The predicted 3-D structure of the target protein	
2	Visualization of the predicted pocket 1 in the target protein by two binding site predicting algorithms	
3	Electrostatic surface potential analysis of target protein's modelled tertiary structure generated by the integrated APBS plugin	
4	Evolutionary conservation assessment of target protein produced by ConSurf	
5	Comparison of RMSD values of the best docking poses from three redocked complexes with different initialization seeds	
6	Mean binding energies from the best poses of the three separate docking seeds are compared	
7	Interacting partners of PA3042 as predicted by the STRING database	

LIST OF TABLES

Table No.	Description	Page Number
1	Summary of physiochemical properties analyzed using ProtParam tool	
2	Secondary structure analysis and relative proportions of each component	
3	Comparison of tertiary structure model generated by D-I-TASSER and AlphaFold using MolProbity analysis	
4	Effect of energy minimization on model quality based on MolProbity scores	
5	Molecular docking results showing binding energies of selected ligands at pocket 1 of the target protein	

ABBREVIATIONS

AMBER	Assisted Model Building with Energy Refinement
AMR	Anti Microbial Resistance
APBS	Adaptive Poisson-Boltzmann Solver
CASTp	Computed Atlas of Surface Topography of proteins
c-di-GMP	Cyclic di-guanosine monophosphate
DGC	Diguanylate cyclase
D-I-TASSER	Deep-learning-based Iterative Threading ASSEmbly Refinement
GRAVY	Grand Average of Hydropathy
HD	Histidine–aspartate domain
HMMER	Hidden Markov Model
HP	Hypothetical protein
IDR	Intrinsically Disordered Regions
MSA	Multiple Sequence Alignment
MUSCLE	Multiple Sequence Comparison by Log-Expectation
NCBI	National Center for Biotechnology Information
NMR	Nuclear Magnetic Resonance
PDB	Protein Data Bank
PDBQT	Protein Data Bank, Partial Charge and Atom Type
PDE	Phosphodiesterase
PSIPRED	PSI-BLAST-based Secondary Structure PREDiction
QS	Quorum Sensing
RMSD	Root Mean Square Deviation
SOPMA	Self-Optimized Prediction Method with Alignment
STRING	Search Tool for the Retrieval of Interacting Genes/Proteins
T3SS	Type III Secretion System
TCS	Two-Component Systems
TM	Template Modeling alignment
YASARA	Yet Another Scientific Artificial Reality Application

1. INTRODUCTION

One of the greatest concerns to global health in the 21st century is antimicrobial resistance (AMR), which greatly reduces the efficacy of presently available antibiotics and increases the prevalence of bacterial diseases [1]. The rise of multi-drug resistant strains has hampered treatment approaches and boosted the need to comprehend biological pathways at the molecular level which are essential for bacterial survival. *Pseudomonas aeruginosa* which is an opportunistic pathogen, responsible for causing serious nosocomial an chronic diseases is well known resistant abilities [2]. This organism shows remarkable metabolic capacity and environmental tolerance, allowing it to survive in ecological niches like water systems , soil and medical equipment [3]. Its infections mainly occur in immunocompromised people, patients with burns or with cystic fibrosis and individuals with ventilator-associated pneumonia [4]. It is one of the main causes of hospital-acquired infections which becomes very difficult to treat due to its inherent and acquired resistance against a number of antimicrobials including aminoglycosides, β-lactams, fluroquinolones [5]. Another critical factor for its persistent infections is its ability to form biofilm which are organized microbial communities enclosed in an extracellular matrix. This method of development provides a huge protection against antibiotics, immune system of the hosts, oxidative stress and environmental changes [6], [7]. Biofilm formation occurs in a coordinated manner in *P. aeruginosa* which involves reversible attachment, adhesion, microcolony formation, maturation and dispersion. During this stage, cells go through important transcriptional and metabolic changes that help in their adaptation to sessile growth conditions [8]. The biofilm's extracellular matrix is mostly made up of polysaccharides, proteins, extracellular DNA and some fats, which work together to maintain structural stability and impart antibiotic resistance [9]. Cells buried in this matrix frequently show altered states and decreased antimicrobial susceptibility, making its clearance very difficult in clinical settings. Importantly, these biofilms also allow lateral gene transfer and hence evolution of resistant phenotypes [10]. Overall, this tight relationship between biofilm formation and serious infection has made biofilm related pathways a promising target for therapeutic intervention. Highly complex signaling network control this transition from planktonic to sessile state and provide adaptability to changing environmental conditions [11]. Numerous physiological activities like motility, synthesis of virulence proteins, stress adaptation is regulated by these signaling systems. Among the various signaling pathways, cyclic-di-guanosine monophosphate (c-di-GMP) is one of the most researched secondary messenger systems [12]. Diguanylate cyclase and phosphodiesterases containing EAL or HD-GYP domain controls its levels by synthesizing and breaking down c-di-GMP and Its levels within the cell determines its state. While low concentration favors motility and planktonic growth, increased c-di-GMP levels support sessile state with biofilm formation [13]. These cascades also include a number of proteins which acts as effectors or sensors that detect the changes in c-di-GMP and modulates subsequent cellular reactions. Together, these elements support biofilm development, stress adaptation, environmental sensing and movement control [14]. Some recent literatures have even highlighted the intricacy of these signaling pathways in *P. aeruginosa*, where several interrelated routes coordinate bacterial adaptability during biofilm generation and infection [15]. However, despite substantial research on main signaling regulators, many proteins encoded are still poorly defined, and their biological functions are still unclear [16]. In this regard, Magnus Østergaard and his colleagues published a study in 2024, which offered fresh perspectives on the role of PA3040-3042 operon in *P. aeruginosa*. This study employed a transposon insertion mutagenesis and identified this operon as a

significant contributor to bacterial stress response and regulator of biofilm formation. Although the operon's functional significance was understood, but the exact molecular function of the genes in this locus are still unknown [17]. One such uncharacterized component whose structural and functional traits have not been determined yet is the hypothetical protein PA3042. Since HPs may be involved in important cellular pathways like signaling, pathogenicity and stress response, their investigation can offer important insights into bacterial physiology and possible therapeutic targets [18]. With quick developments in structural bioinformatics and computational biology, in-silico methods have become effective tools for predicting the function of unknown proteins [19]. Critical preliminary information about the function of protein and its interaction strategies can be studies with methods such as structural modelling, conserved domain analysis, molecular docking and evolutionary studies [20]. Hence, the goal of this study is to perform computational characterization of the HP PA3042 and learn more about its possible structural characteristics and functional capabilities and role in signaling pathways associated with biofilm formation. The results of this work may contribute in comprehending an unidentified protein and its wider significance in *P. aeruginosa* antibiotic resistance.

2. LITERATURE REVIEW

2.1. Pathogenicity and clinical relevance of *Pseudomonas aeruginosa*

Pseudomonas aeruginosa, a gram-negative pathogen exhibits a wide ecological range and can be isolated from soil, plants, water, humans and even hospital reservoir. It is considered as an opportunistic pathogen as it does rarely cause disease in healthy individuals but becomes extremely harmful when host's immune system is compromised or weakened [21]. Its resilience to environmental stress is clinically significant as it can easily colonize on sinks, moist surfaces and most importantly on medical equipment creating a constant source of hospital-associated exposure [22]. As mentioned, being an opportunistic pathogen, it is a great danger to patients with cancer, burns, cystic fibrosis or damaged cutaneous and mucosal barriers. In such cases, *P. aeruginosa* can cause a broad range of clinical symptoms such as pneumonia, bacteremia, urinary tract infection, wound infection and even sepsis, with significant morbidity and death [23]. Its effectiveness in causing both chronic and acute infection is due to a coordinated pathogenic strategy which involves attachment, colonization, evasion of immune system and systematic regulation [24]. For colonization, two key mediators are flagella and type IV pili as they allow adhesion to the surface and movement via twitching or swarming. These appendages create the necessary environment for invasion which involves reaching the host tissue and spreading across epithelial cells or surfaces in case of materials [25]. Apart from surface motility, P. aeruginosa also uses a wide array of secretion systems such as T1, T2, T3, T5 and T6 which together transport enzymes, toxins and several effector proteins which aids in virulence. Among these systems, T3SS is critical as it acts as a molecular syringe which directly transfers effectors into the host cells, causing cytotoxicity and invasion, similarly, T6SS also assists in toxin delivery and host interaction [26]. All of these systems are highly regulated by quorum sensing mainly via the las, rhl and pqs pathways which coordinate many genes involved in movement, pathogenicity and secretion [27]. Therefore, quorum sensing serves as a regulatory layer that coordinates surface colonization, production of toxin and other group behaviors. Secreted proteases, toxins and other compounds like siderophores then enhances the infection by causing tissue damage and evading immune cells. In this manner, the pathogen combines all of these abilities into virulence strategy and biofilm formation forming the pathogenic architecture [28], [29].

Resistance against antibiotics is another characteristic of P. aeruginosa which emerges from innate, acquired and adaptive mechanisms making treatment extremely challenging. Intrinsic methods include low permeability of outer membrane, modified porins, active efflux pumps and enzymes such as β-lactamases which decreases the drug exposure or degrade the antibiotic even before it starts effecting [30]. In addition to them, other β-lactamases like cephalosporinases further restricts the treatment options. Acquired resistance is another layer which mainly involves mutation or lateral gene transfer of resistance genes which modifies the drug target or inactivates antimicrobial compounds, rendering them ineffective [31]. Biofilm production also supports in resistance as a physical obstruction for the entry of compounds and promotes a metabolically transformed state which is less responsive to traditional treatments. This explains why P. aeruginosa is frequently resistant and continues to be a major nosocomial concern, mainly in patients who are immunocompromised or vulnerable [32].

2.2. Biofilm formation and its regulatory signaling

2.2.1. Sequential stages of biofilm development

The ideal way to describe biofilm formation in *P. aeruginosa* is considering it as a coordinated mechanism in which free-living cells transform into surface attached community that is structured, metabolically distinct and guarded by an extracellular polymer produced by the bacteria itself [33]. In contrast to passive coating, this method modifies gene expression, colony behavior and creates gradient of nutrients, oxygen and messenger molecules that allows growth in adverse conditions [34]. Clinically, this is significant as *P. aeruginosa* cause chronic illness that hare very difficult to treat such as airway disease in cystic fibrosis patients [35]. Also, these cells within biofilm show substantially higher tolerance to antimicrobial strategies, typically 10 to 1000-fold more resistance than planktonic cells [36]. The failure of treatment is not the only concern, tendency of recurrent infections, extensive colonization and selection of extremely tolerant sub-groups are also critical [37].

2.2.1.1. Initial attachment

The first step of biofilm production is initial reversible attachment where planktonic cells reach an appropriate surface and weakly attach via transient interactions and movement associated sensing. Although the connection is sufficient to start the biofilm formation, attachment is unstable yet and cells may separate under hostile conditions [38]. This early stage is closely associated with signaling mechanisms such as c-di-GMP signaling, which help in transforming from a motile phenotype to a surface associated phase [39].

2.2.1.2. Irreversible anchoring

Once they adhere irreversibly, cells move to stage two which involves commitment to surface and microcolony development. This transformation acts as a checkpoint, as cells not only binds to the surface rather, they proliferate, synthesize matrix proteins and form early aggregates [40]. Regulators like SadB helps capture the cells in state while enhanced c-di-GMP levels encourage the production of exopolysaccharide and formation of coherent community [41].

2.2.1.3. Maturation and microcolony formation

Now as they mature, these microcolonies grow into a more defined three-dimensional colony with greater structural diversity [42]. At this point, biofilm matrix thickens and becomes functionally specialized. Protein, polysaccharides and extracellular genetic material play a crucial role in shaping the internal framework and shielding the embedded cells. Mature biofilms also create a physiological separation where they favor subpopulations which are far less sensitive to antibiotics than other actively growing cells, making them more resilient than planktonic cells [36], [43].

2.2.1.4. Dispersion of biofilm

Dispersal, which is the final step actually a strategic developmental stage instead of just a random collapse of the community [44]. It involves coordinated return to motility, escape from the matrix and migration to new locations to seed new niches and spread infection. In order to exit from biofilm, proteins like DipA work in conjunction with BdlA via c-di-GMP signaling to ensure strict coordination as in attachment or maturation [45].

The structural and functional core is the biofilm matrix. Pel and Psl are considered as the two main exopolysaccharides that form scaffolds, mainly in non-mucoidal and early biofilms where they promote initial attachment, cohesion and durability against matrix disintegration [46]. Pel provides mechanical robustness and strengthens the matrix using charge based interactions whereas Psl is particularly important protecting nearby cells [47]. Alginate which is another polysaccharide becomes significant in mucoid colonies which frequently develop in persistent infections such as during cystic fibrosis. When overproduced, alginate increases matrix viscosity and water retention ultimately boosting resistance to phagocytosis, hostile conditions and oxidative stress, all of which contribute to stabilize chronic infection. It is not a part of every P. aeruginosa biofilm, but its buildup is a characteristic of mucoid state leading to decreased lung clearance [48], [49]. Matrix proteins provide critical structural support. An adhesin called CdrA which regulated by c-di-GMP crosslinks with Psl and enhances aggregation and helps in preserving the structure [30]. Extracellular DNA also functions as a load-bearing element that stabilize the lattice structure of matrix by interacting with polysaccharides and thus prevent antibiotic penetration [50]. Moreover, matrix synthesis is not a constitutive property, rather is controlled by signaling. So, when taken together, biofilm production can be best understood as a controlled life-cycle as opposed to a static cell aggregation.

2.2.2. Regulatory pathways governing biofilm formation

This transition from planktonic state to getting embedded in extracellular matrix is not governed by a single pathway, but rather three routes namely quorum sensing, two-component systems and c-di-GMP signaling works as integrated network, to regulate whether cells stay motile, adhere to surface, form a matrix or get dispersed [51]. These mechanisms are partially overlapping systems that alter gene expression to either cause acute infection or a biofilm associated chronic state, in response to environmental or host-induced signals [52].

2.2.2.1. Two component system

In the biofilm regulatory network TCS constitute a major input layer. It consists of sensor kinases which identifies change in the surroundings and a response regulator which transform those signals into modified transcriptional outputs [53]. The GacS/GacA system is one of the most significant example, that enhances the expression of small regulatory RNAs such as RsmY or RsmZ, which in turn sequester repressor protein RsmA causing sessile state development and biofilm formation while suppressing acute virulence [54]. This Gac/Rsm system is itself regulated by other TCS-linked modulators like RetS and LadS, which forms a balance between planktonic and biofilm stage. Another important system for associated with biofilm is BfiS/BfiR, that is necessary for irreversible attachment to the surface. It involves inhibiting RsmZ mediated growth program by activating CafA-RNase system [55]. Some other examples of TCS are BfmSR which promotes biofilm production eDNA release through phage mediated lysis and FimS/AlgR and KinB/AlgB which enhances alginate production [56]. A number of TCS pathways also regulate surface characteristics that allow attachment, for instance cup fimbrial aggregates which are regulated by Roc1, Roc2 and Rcs/Pvr. So all of these systems create an early framework which controls whether or not cells will engage in surface attachment and biofilm formation [57].

2.2.2.2. C-di-GMP as central regulator

Cyclic-di-GMP is the primary second messenger molecule responsible for transition from motile to sessile state in *P. aeruginosa* [58]. Its level inside the cell increases when cells move towards biofilm stage and its amount is maintained by reciprocal activities of diguanylate cyclase (DGCs) and phosphodiesterases (PDEs) which synthesize and breakdown c-di-GMP respectively [59]. This enzyme balance dictates whether cell continues to be motile or take surface associated state. The c-di-GMP system is really important as it is a dispersed system with stage-associated control points rather than one liner cascade [60]. It has at least five DGCs namely, WspR, RoeA, YfiN/TpbB, SadC and SiaD which are associated with transition to biofilm state whereas NicD, RdbA, GcbA and NbdA are linked with biofilm dispersal. This architecture implies c-di-GMP pathways are isolated and coordinated rather than a homogenous pool of signals [61]. Numerous receptors and downstream effector proteins sense nucleotide levels and regulate matrix production, motility and resistance. FleQ is one crucial transcription factor that responds to c-di-GMP, at lower levels it enhances the expression of flagellar genes but at higher levels it inclines towards the suppression of motility and activation of gene like psl, pel and cdr which are involved in formation of exopolysaccharides and adhesins [14]. Simultaneously, c-di-GMP also regulates BrIR which contributes to biofilm-associated resistance, suggesting that c-di-GMP influences both matrix formation and tolerance traits [62]. Another feature of this pathway is integration with more extensive regulatory networks, for instance transcriptional regulators like VqsM and AmrZ and SadC which plays a key role in GacS/GacA mediated pathway also control c-di-GMP levels and associated biofilm production [63]. Hence, c-di-GMP acts as a hub linking motility suppression, environmental sensing, matrix formation and antibiotic resistance [64].

2.2.2.3. Quorum sensing

Quorum sensing also known as density dependent signaling also plays crucial role in formation of biofilm by enabling cells to synchronize gene expression in population-dependent manner through autoinducers [65]. In *P. aeruginosa*, QS system adopts a hierarchical pattern with Las being first followed by Rhl and Pqs and these are connected to each other by multiple positive and negative feedback loops [66]. LasR which is bound to OdDHL activates rhlR, lasI and rhlI producing a positive feedback [67].

Biofilm development and maturation is closely linked to QS. Las mediated pathway regulates elastase which further modulates biofilm synthesis. Any mutations in las genes results in flat, undifferentiated biofilms [68]. On the other hand, rhlI is important for dispersal phase through formation of rhamnolipids. Moreover, elastase, LecA, pyocyanin and biofilm formation all are impacted by PQS pathway, which regulates the expression of virulence factors [69].

The fact that these pathways are interdependent is a characteristic feature of biofilm production in *P. aeruginosa*. It is specifically stated that c-di-GMP and Gac/Rsm networks are connected and this coupling in coordinating proper biofilm growth. RetS, HptB, LadS and SadC are critical points of interaction, demonstrating how TCS input signals are transferred via c-di-GMP pathway to alter phenotype [70]. Similar cross-talk also occurs between QS and c-di-GMP. Increased levels of c-di-GMP is linked to sessile state transition, as this signal changes the expression of motile genes while concurrently stimulating matrix production through regulators like AlgR and AlgU and suppressing output of RhlR [71]. At the same time VqsM connects QS to c-di-GMP by controlling HsbD and LasR, creating a regulatory loop which

involves QS influencing c-di-GMP and c-di-GMP in return, regulates biofilm associated transcriptional activity [72].

Overall, these three pathways works a interconnected model in which TCS functions as upstream sensor, c-di-GMP acts as main intracellular integrator and QS coordinates population dependent behavior, with substantial cross-regulation among them in determining whether *P. aeruginosa* should adopt a motile or biofilm phenotype [73].

2.3. Hypothetical proteins and their significance

Hypothetical proteins (HP) are gene products that are presumed to be expressed from an open reading frame but their biological functions are unclear and lack an experimental evidence for their translation [74]. These proteins show no apparent similarity with functionally defined sequences in curated databases and are only identified using computational methods. About 30 to 40% of the proteomes in bacteria is made up of uncharacterized proteins, highlighting the extent of this information gap [18].

In *P. aeruginosa*, this percentage is very striking, during the initial genome annotation about 32% of the coding region was classified as hypothetical and 14% as conserved hypothetical [75]. According to a recent report, even after two decades approximately 25% of the proteins are still considered hypothetical and 40% of genes regarded as uncharacterized. These HPs considered to be a part of "dark proteome" in bacteria, i.e., a subset whose function has not been experimentally determined [76]. However, their presence in the genome despite evolutionary selection indicates they carry out important biological functions [77]. Many functional annotation studies have linked HPs to pathogenicity, secretion systems, biofilm formation in several pathogenic microorganisms and since they are highly conserved and have no homologs in humans, they emerge as promising target for new antimicrobial treatments [78].

However, characterization of HPs is major challenge as they frequently lack identifiable homologs. This absence results in generation of weak or incomplete multiple sequence alignments. Deep MSAs are critical for determining functional relationships and therefore their absence diminishes the accuracy of these methods. [79], [80]. The frequent presence of intrinsically disordered regions (IDRs) is one key factor that complicates HPs. These IDRs are the regions that do not assume a stable conformation under physiological conditions which makes them very challenging to define using conventional experimental techniques like NMR [81]. Moreover, as these IDRs are too huge to be represented by single static model, even computational approaches struggle with them. Despite this limitation, IDRs are known to play important roles in signaling and regulatory functions, where their flexibility promotes varied interactions [82], [83]. As a result, functional analysis and modelling of these uncharacterized proteins mostly count on a combination of computational approaches and using a multi-evidence approach instead of definitive criteria [84].

2.4. Literature background of the target protein

In a recent study by Østergaard et.al., significance of an operon PA3040 – 3042 found in P. aeruginosa genome was highlighted. This experimental evidence served as a vital cornerstone for the current investigation, as they showed that a previously unidentified operon seems to be involved in biofilm formation and cell membrane stress adaptation. Despite being not fully

identified, the operon demonstrated significant phenotypic relevance under stress related conditions suggesting it might assist in bacterial survival [17].

To identify genomic regions which are required for survival in phage exposed conditions, they used transposon insertional mutagenesis method in P. aeruginosa PAO1 strain. Upon performing comparative analysis, between infected and non-infected conditions many regions demonstrated depletion patterns. Among them, PA3040-3042 operon stood out as one of the most prominent candidate. These three genes are arranged into a single transcriptional unit and insertion in PA3040 alone caused a polar effect, suppressing the expression of the downstream genes. The mutant containing insertion in this operon suggested that the locus may offer inherent advantage to the bacterium in stress or phage exposed conditions [17].

Genomic analysis revealed that this operon is a component of the *Pseudomonas* core genome and PA3042 itself exhibit strong conservation among *P. aeruginosa* strains. Annotation of the operon showed that PA3040 is homologous to a transmembrane ribosome binding protein YqjD while PA3041 displayed similarity with YqjE, a protein related to phage holin. On the other hand, the protein product of *PA3042* is yet known and has no defined molecular function assigned. However, phenotypic analysis strongly suggests that PA3042 has significant physiological role, even in the absence of an annotated function [17].

The operon is stimulated transcriptionally under stress conditions as per the expression analysis. Assays like Promoter reporter assay showed enhanced expression upon cell envelope stress and phage exposure for instance, upon treatment with D cycloserine, an antibiotic which interferes with cell wall formation, a strong activation was seen. It was also predicted that promoter region is under the control of AlgU σ factor, which is crucial regulator of alginate production and stress response in *P. aeruginosa*. All of these results points, indicate the involvement of operon in pathways related to stress adaptation and biofilm formation [17].

The physiological significance of the operon was further supported by characterization performed in clinical isolate PAC3. While deletion and increased expression in PAO1 had little impact on the phage sensitivity or growth, mutation in PAC3 strain showed a number of unique characteristics. Growth was hindered, surface biofilm was greatly reduced and pyocyanin production was increased when the operon was knocked out. Together these results indicate that PA3040-3042 operon, specifically PA3042 gene is significant but remain underexplored component of P. aeruginosa genome. Moreover, its role in cell envelope stress and biofilm production makes it a desirable target for more structural and functional research [17].

2.5. Intracellular signalling architecture in bacteria

Signal transduction pathways are used by bacteria to sense changes in their surroundings and initiate an essential physiological reaction. Three components make up their architecture, a receptor that detects the signal, the stimulus which is transmitted by conformational change, cellular response and the signal is terminated to revert to the initial state [85]. Two component systems, where sensor histidine kinases detect outside signals and change the phosphorylation state of the respective response regulator that carry out the downstream output, serve as an example of this concept [86]. Chemosensory routes expand this via specific adaptor proteins, methylating domains and several other systems to allow adaptive sensing across a broad range [87]. P. aeruginosa which has a very complex signaling network among bacterial systems, integrates several interconnected layers that regulate virulence and biofilm formation [88].

Several unusual receptors, such as cytosolic receptors lacking transmembrane domains, chemoreceptors with carboxyl-terminal binding motifs and receptors that operate without any ligand binding domains at all are some of the examples which highlights this complexity [89]. These non-canonical designs require experimental methods that extend beyond homology-based identification and pose challenge to classical sequence based annotation.

2.5.1. Receptor proteins

Receptors are characterized by their ability to recognize a particular signal, such small chemical molecules, a protein or even physical stimulus and subsequently transmit this information to a downstream protein through conformational change [90]. This mechanistic definition is independent of primary sequence annotation. When a protein lacks identifiable pattern, receptor activity must be deduced from structural characteristics, expected ligand-binding sites and genomic context rather than from domains only [91].

2.5.1.1. Histidine kinase

The histidine kinase family which is the predominant model for transmembrane perception is the most extensively researched class of bacterial receptors [92]. The membrane bound histidine kinases has two membrane spanning helices on either of its cytoplasmic ligand-binding motifs [93]. Molecules binding to this motif induces a conformational change which is then transferred across the membrane to the autokinase domain to modulate ATP-binding and phosphotransfer to the regulator proteins [94]. *P. aeruginosa* produces 64 sensor histidine kinases and yet activating signals for only a small number of them have been experimentally determined which highlights the gap between genomic understanding and functional information [95], [96].

2.5.1.2. PAS domains

One of the most adaptable sensor modules in bacterial signaling is PAS domain [97]. They bind to a remarkable range of ligands including haem, 4Fe-4S clusters, flavin mononucleotides and other signaling compounds that can be found in both membrane bound and cytosolic conditions [98], [99]. This ligand diversity sets an important precedent that the presence of PAS domains alone cannot be used to predict the activating signal and single type of fold can allow the binding of chemically diverse signals [100].

2.5.1.3. PilZ domain

Another very well-characterized domain is the PilZ domain which specifically binds to c-di-GMP [101]. In PilZ, the nucleotide binding pocket is formed by the conserved RXXXR and [D/N]hSXXG motifs which allows interaction with c-di-GMP [102]. Eight such PilZ domains are encoded by P. aeruginosa and majority of unknown functions. PilZ domains can be found in a variety of structures such as protein like MapZ, cellulose synthases like BcsA, or connected to alginate production by Alg44 where its production is allosterically regulated by c-di-GMP [103].

Despite the abundance of known domain families, a large subset of proteins acts as receptors without possessing identifiable domains [104]. The flagellar main regulator FleQ that reacts to c-di-GMP levels lacks any GGDEF, EAL, PilZ or any other recognized motif, is the most notable example in *P. aeruginosa.* Similarly, a transcriptional regulator Vfr that binds to cAMP does not have any classical cyclic nucleotide binding domain. Hence, domain based annotation

systems often fails to predict activities of such proteins and requires integration of structure prediction, ligand docking and several other methods to understand the functional significance [105].

2.5.2. Second Messenger molecules

C-di-GMP which is the most thoroughly studied nucleotide second messenger controls the shift between motile to sessile phenotype in a number of phyla [106]. Its synthesis is catalysed by DCGs containing GGDEF domains and phosphodiesterases perform its degradation. Nearly 41 predicted c-di-GMP metabolizing enzymes are encoded by P. aeruginosa genome and each of them contains at least one sensory domain allowing it to combine a variety of environmental signals into one nucleotide output [107]. This extensive enzymatic collection explains why a novel c-di-GMP receptor would be logically plausible and since all the 41 enzymes are dispersed throughout the sub-cellular components, c-di-GMP acts through spatially and temporally separate pools, signaling network is dense enough that other unidentified receptors are likely to exist [108].

2.5.2.1. ppGpp

Another important second messenger is ppGpp which forms a complex signaling system, that quickly builds up when there is a depletion of amino acids and initiates a complete transition from growth to survival mode [109]. In *P. aeruginosa* Nucleoside diphosphate kinase (Ndk) which is positively regulated by AlgQ-AlgR2 pathway, synthesizes ppGpp and deletion of algQ results in cell death at late exponential phase, highlighting the crucial role of these molecules in stationary-phase survival [110]. Additionally, the strict response links metabolic stress to behavioral adaptation by controlling the expression of chemosensory pathway genes [111].

2.5.2.2. c-di-AMP

In gram-positive bacteria, c-di-AMP which is a more recently identified second messenger plays crucial role in cell wall homeostasis and potassium transport [112]. It is generated by diadenylate cyclase and binds to a number of receptor proteins such as PstA, KtrA phosphate transporters and sensor kinases like KdpD [113]. C-di-AMP is an example of growing range of cyclic dinucleotide signals in prokaryotes, although being less explored [114].

2.5.2.3. cAMP

Unlike other messenger molecules, the signaling chemistry of cAMP is very different [115]. The adenylate cyclases such as CyaB and some CyA are mainly responsible for its synthesis in *P. aeruginosa*. cAMP forms a complex with Vfr which is a transcriptional activator and virulence regulator and modulates the expression of T3SS, exotoxin A, type IV pili and Las quorum sensing pathway [116]. C-di-GMP and cAMP are inversely correlated as increased c-di-GMP suppresses the production of acute virulence factor by lowering cAMP levels [117].

2.5.2.4. Non-nucleotide signaling molecules

Apart from nucleotides, bacteria also use some non-nucleotide signaling molecules. Diffusible signal factors (DSFs), like cis-2-dodecenoic acid, represent a fatty acid signaling family that affects biofilm dispersion, Similarly N-acyl homoserine lactones (AHLs) are classical quorum sensing molecules synthesized by LasI and RhlI respectively [118], [119]. Additionally, lipid based signaling molecules like phosphatidylethanolamine (PE) and Phosphatidylcholine (PC)

have been linked to interkingdom communication and host-pathogen interaction, thus expanding the landscape of signal transduction pathway in bacteria beyond nucleotide mediated routes [120].

2.5.3. Effector proteins

Molecules which sense the concentration of second messenger and carry out a functional response are referred as effector proteins. They can be divided into two main groups, the ones which use direct enzymatic activity and the ones that act as transcriptional regulators [121].

2.5.3.1. Transcriptional modulators

One of the most well studied transcriptional regulator is FleQ, often called as the master regulator of flagellar synthesis. It senses c-di-GMP levels and modulates the expression of pel operon producing Pel polysaccharide, which is a crucial component of biofilm matrix [122]. C-di-GMP occupies AAA+ ATPase domain in FleQ which is located very close to the but separate from the ATP binding domain and thus, couples the ATPase activity with transcriptional output [123]. At decreased c-di-GMP levels, FleQ suppresses the expression of Pel while at higher c-di-GMP concentrations pel operon becomes active and the repression is released. This dual-input method which places it at the center of motility to biofilm switch, makes it one of the cleanest example transcription factor (TF) that converts second messenger levels into phenotypic change [124].

2.5.3.2. Enzymatic Effectors

Enzymatic effectors use catalytic activity to carry out messages instead of binding to DNA. In bacterial systems the distinction between effector function and second messenger metabolism is blurred. An enzyme that breaks down a second messenger in response to an upstream signal is considered an effector for that signal, as its catalytic activity results in the output [105]. This idea is demonstrated by phosphodiesterases and diguanylate cyclases. Several PDEs and DGCs are multi-domain containing proteins that combine catalytic activity with sensory domains to directly link signal detection to nucleotide turnover [125]. For instance, in *E. coli*, DosC/DosP system uses haem-based oxygen sensors with variable affinities to couple DGC and PDE. This equilibrium between synthesis and degradation changes with oxygen levels and the resulting c-di-GMP signal regulates the RNA processing ability of PNPase which is a c-di-GMP receptor as well. In this case, the enzymes are both metabolizers and oxygen signal's effectors [126].

Apart from hydrolases and cyclases, another class of effector molecules exist which are GTPases. They employ a typical GTP/GDP switch in which when attached to GTP, they are active and hydrolysis to GDP due to their inherent GTPase activity results in their inactivation. Since the rate of hydrolysis controls that time for which, the protein will remain active and consequently the amplitude of downstream response, the GTPase activity itself becomes a signal processing event [127]. This concept is widely present in eukaryotic signal transduction and is now being recognized in bacterial systems. The GTPase activity of proteins like Era in P. aeruginosa has been connected to ribosome assembly and cell cycle progression, the nucleotide attached state of these proteins dictates how they interact with the downstream effector proteins [128]. Hence, GTPase activity adds another level regulatory complexity by offering a signal processing mechanism that is complementary to the common phosphorylation systems.

3. METHODOLOGY

3.1. Sequence retrieval

A hypothetical protein from *P. aeruginosa* (UniProt ID Q9HZG6) was selected from UniProt Knowledgebase (UniProtKB) [129]. Its 107 amino acid long protein sequence was retrieved in FASTA format for structural analysis of the protein.

3.2. Physiochemical analysis

The physiochemical properties play a crucial role in determining the structure and functional capabilities of a protein. The ProtParam tool of the ExPasy server was employed for our HP. This tool analyzes several properties including isoelectric point (pI), molecular weight of the protein, instability score and many more [130].

3.3. Sub-cellular localization

Different proteins in bacteria gets localized to different locations within cell which is often related to its function. It is also a key criterion for identifying vaccine and drug targets. $\text{SOSUI}_{\text{GramN}}$ was utilized for predicting sub-cellular location which was further correlated with localization predicted by D-I-TASSER tool [131].

3.4. Determination of secondary structures

2° structure prediction was performed to identify the arrangement of amino acid residues into structures like alpha-helices, random coils and beta-sheets. SOPMA tool was utilized to perform this task subsequently its result was validated and using an additional tool PSIPRED [132], [133].

3.5. Predicting the 3-D structure of HP

Tertiary structure which is dictated by the unique amino acid sequence of the protein determines its biological activity as it allows various groups and domains to come together. Consequently, structural insights are essential not only for determining molecular mechanisms but also for performing functional annotation of uncharacterized proteins. The 3-D structure of the HP was produced using AlphaFold 2 and D-I-TASSER servers [134], [135]. AF2 employees a deep neural network-based method and can either use templates or operate in an ab-initio manner, while D-I-TASSER is hybrid tool which combines deep learning algorithm with conventional iterative threading approach. For the subsequent analysis of the protein, the model generated by D-I-TASSER was used.

3.6. Energy Minimization

Energy optimization of the model produced by D-I-TASSER was performed using YASARA software. This method improves the accuracy of the structure predicted [136].

3.7. Assessing quality of structure

After performing the energy minimization, determining the quality of the model is an imperative task to identify the feasibility and physical stability of the protein. MolProbity server was used for performing this task [137].

3.8. Active site and binding pocket determination

Proteins perform their biological activity via specific active sites and binding pockets present on them. In order to identify them, two tools CASTp and DoGSiteScorer server were utilized. Both of them finds out pockets, curved regions on the proteins but with different algorithms. CASTp employes an analytic geometry based algorithm while DoGSiteScorer uses grid-based method. Lastly, the results from both the tools were evaluated visually using software PyMol

3.9. Molecular function prediction

The COFACTOR server was used to predict the probable molecular function of PA3042, which determines this by threading the submitted structure against a library of proteins and finding homologs based on similar folds, sequence conservation [138]. The best ranked predictions are assessed based on C-score, which is a confidence metric with values ranging from 0 to 1. Hits with C-score greater than 0.5 are considered significant. All predicted molecular function gene ontology terms were then analyzed along with other evidence obtained for PA3042.

3.10. Structural homology analysis

Since PA3042 does not show any sequence homology, it becomes critical to identify structural homologs for which Dali server was utilized, which performs pairwise structural comparison between the entered structure and PDB entries using a distance matrix method [139]. The predicted structure of PA3042 was submitted to Dali which returned the results ranked by Z-score with Z-score greater than 0.2 regarded as statistically significant. The top hit was selected for additional structural comparison. The TM-align server was used to align the Dali hit with predicted structure and evaluate the degree of similarity [140]. This server performs sequence independent superimposition of the two structures and returns TM-score, to assess similarity. TM-score above 0.5 indicate similar fold. The resulting alignment was then visualized using PyMOL to evaluate equivalence between the two structures [141].

3.11. Electrostatic characterization of pocket 1

The electrostatic surface potential was determined using Adaptive Poisson-Boltzmann Solver (APBS), which is an in-built plugin in PyMOL [142]. Before calculation, hydrogen atoms and partial charges are assigned to the protein using another complementary plugin in PDB2PQR which uses AMBER force field at physiological pH. The resulting electrostatic potential is then mapped on the protein surface and visualized with color, that vary from deep blue to deep red indicating positive and negative potential respectively. The electrostatic properties of pocket 1 was investigated to determine the possible ligand preferences.

3.12. Identification of evolutionary conserved residues

The ConSurf server was used to examine the evolutionary conservation of residues in the query protein. For which, homologous sequences were identified using jackhmmer search from HMMER suite against the UniProtKB database, with three rounds of iterative search [143]. Then using MUSCLE algorithm, a multiple sequence alignment (MSA) was constructed using the top 50 hits [144]. Finally, the generated MSA along with the query protein was submitted to ConSurf, which returned per-residue conservation score on a scale of 1 to 9, with 1 being most variable and 9 being highly conserved [145]. The obtained conservation scores were then mapped, to identify evolutionary conserved residues in pocket 1.

3.13. Molecular docking

A thorough literature review was carried out to identify signaling molecules using good, updated reviews on c-di-GMP signaling, quorum sensing and other pathways involved in biofilm production in P. aeruginosa. 14 potential ligands were gathered which included nucleotides, cyclic nucleotides, variants of homoserine lactones and non-nucleotide molecules. Apart from them two control ligands were also used which were citrate and glutamate. The three – dimensional structures of all ligands were obtained from PubChem database and prepared by combining non-polar hydrogens, assigning Gasteiger charges and defining rotatable bonds using AutoDock Tools 1.5.7 and saved in PDBQT format [146]. Similarly, receptor was also prepared and converted into PDBQT format. Now, molecular docking was performed using AutoDock Vina with the search space oriented towards pocket 1 [147]. The following grid parameters were used center x = 58.448, center y = 82.540, center z = 74.920 and grid box dimensions as size x = 28, size y = 30, size z = 34. An exhaustiveness of 32 was selected for each docking run. The top-ranked pose of each ligand was then analyzed using PyMOL.

3.14. Reproducibility assessment of docking results

The top two ligands underwent re-docking in order to evaluate the consistency of docking results. Each ligand was docked three times using same docking parameters but distinct random seeds. Convergence was then assessed by extracting the best pose from each seed and pairwise RMSD values were calculated between the three poses. Additionally, the binding free energies from all the three seeds were used to determine the mean binding energy ± standard deviation to quantitatively measure docking stability.

3.15. Protein interaction network analysis

The STRING database, which combines data from co-expression, genomic neighborhood, literature mining and experimental interaction was used to investigate the interacting partners of PA3042 [148]. The protein's sequence was submitted and the generated interaction network was analyzed for high confidence relationship.

In order to verify the chromosomal proximity of the PA3042 with an adjacent gene PA3043 as suggested by the STRING network, NCBI gene page for PA3042 was utilized, validating the identity and annotation of dgt2.

3.16. Analysis of Phylogenetic co-occurrence

A phylogenetic co-occurrence study was performed to determine if genomic relationship between PA30432 and PA3043 is conserved. To find organisms which encode dgt2, phmmer search was performed against UniProtKB database and the resulting taxonomic distribution was cross-referenced with PA3042 result. Overlapping taxa were identified and examined, which would provide an independent support for a functional link between the two proteins.

4. RESULTS

4.1. Physiochemical characterization

Physiochemical analysis of PA3042 showed a predicted molecular weight of 11,739 Da. The isoelectric point (pI) which is estimated to be 12 indicates a substantial basic character due to the dominance of positively charged residues in the sequence. The GRAVY score for the protein is -0.164, suggesting a hydrophilic nature and high aliphatic index of 114 implies that PA3042 is thermostable and adopts a rigid, well-folded structure that may sustain stable interactions.

Table I: Summary of physiochemical properties analyzed using ProtParam tool.

Parameter	Value
Molecular weight	11,739.0 Da
Theoretical pI	12
Aliphatic index	114.0
GRAVY score	-0.164

4.2. Sub-cellular localization

Sub-cellular localization was predicted by two separate tools and both consistently predict the protein to be membrane associated. This prediction was better presented by $SOSUI_{GramN}$ which localized the protein on inner membrane. The agreement between two separate prediction tools highlights the dependability of this localization prediction.

4.3. Secondary structure composition

Secondary structures which were predicted using SOPMA server reveals that PA3042 is mainly composed of alpha helices, which make up 53.27% of all residues. Other structures include, random coils, beta turns and extended strand which are presented in the Table II. The high alpha helical percentage is consistent with well-organized structure.

Table II: Secondary structure analysis and relative proportions of each component.

Structure	Percentage (%)
α-helix	53.27
Extended strand	11.21
β-turn	14.20
Random coils	21.50

4.5. Predicted 3-D structure

The tertiary structure of PA3042 was determined independently using two tools and both the models were subjected to stereochemical quality evaluation which are summarized in the Table III.

Table III: Comparison of tertiary structure model generated by D-I-TASSER and AlphaFold using MolProbity analysis.

Parameter	D-I-TASSER	Alpha Fold 2
Ramachandran favoured (%)	93.33	87.62
Clashscore, all atoms	1.16	34.82
MolProbity score	1.33	3.49
Bad angles (%)	0.18	2.69
Poor rotamers (%)	1.23	13.58

The D-I-TASSER model showed noticeably superior stereochemical quality in every evaluated parameter. However, there were fewer problematic angles (0.18%) and poor rotamers (1.23%) in it. The AF model on the other hand, generated model with very high clash score of 34.82 and MolProbity score of 3.49 indicating significant abnormalities in the structure. Hence, D-I-TASSER model was chosen for further investigations.

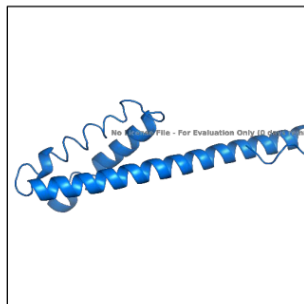


Fig. I: The predicted 3-D structure of the target protein is shown in the figure: The tertiary structure depicts the spatial orientation of the amino acids within protein and it was produced using the deep learning assisted I-TASSER suite. The resulting model offers structural information about the arrangement of secondary structures and possible functional regions that could support biological function.

4.6. Structural refinement and energy minimization

Following the energy minimization from -64,244.1 Kj/mol to -71,935.3 kj/mol, the overall structure improved significantly, suggesting a change from physically stressed conformation to a more stable state. The energy minimized structure was reassessed and the following results were obtained (Table IV).

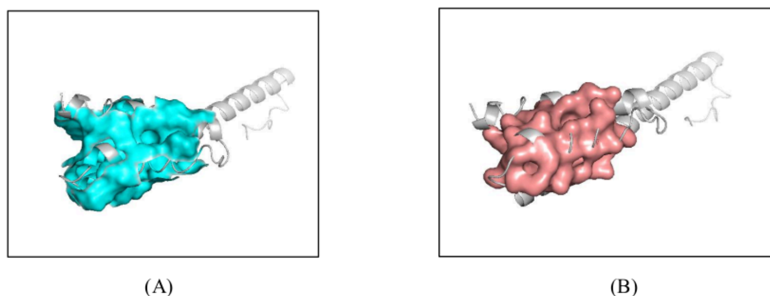
Table IV: Effect of energy minimization on model quality based on MolProbity scores.

Parameter	Pre-YASARA	Post-YASARA
Ramachandran favoured (%)	93.33	98.33
Clashscore, all atoms	1.16	0.58
MolProbity score	1.33	1.13
Bad angles (%)	0.18	0.00
Poor rotamers (%)	1.23	0.00

All the stereochemical metrics showed great improved, clashscore decreased from 1.16 to 0.58 and MolProbity score was enhanced to 1.13. Poor rotamers and bad angles were completely removed, indicating that energy minimization has immensely improved geometric configuration apart from stabilizing the structure.

4.7. Predicted ligand-binding sites

Two independent pocket identification tools were used in parallel and both of them converged on a particular dominant binding cavity at a constant location, which is referred as pocket 1. The volume of pocket 1 was determined by CASTp was 923.086 ų. This number is considerably in the range of functional ligand-binding regions in proteins. DoGSiteScorer also assigned a druggability score of 0.82, which is much higher than widely accepted druggability threshold of 0.5. Additionally, the spatial and methodological concordance between two entirely different pocket identification tools, provides strong evidence for the genuineness of pocket 1.



(A) (B)

Fig. II: Visualisation of the predicted pocket 1 in the target protein by two binding site predicting algorithms: (A) Pocket 1 anticipated by CASTp highlighting surface accessible cavity. (B) DoGSiteScorer's prediction of pocket 1 demonstrating the ligand-binding region. The comparative examination of two tools shows that predicted binding site position is consistent and justifies the use of pocket 1 for further analysis.

4.8. Putative molecular function

The two functional annotations obtained by submitting the predicted 3-D structure of PA3042 at the COFACTOR server were signaling receptor activity and phosphodiesterase activity with C-scores of 0.6 and 0.4 respectively. Since PA3042 lacks any identifiable domains and these predictions were solely based on structural comparison with the COFACTOR's database consisting of large number of characterized proteins, a C-score of 0.6 represents a significant signal to generate hypothesis. This prediction of signaling receptor activity served as a computational premise for examining PA3042 as a possible component of signal transduction pathway. COFACTOR's second prediction of phosphodiesterase activity was not completely disregarded but the notion of a single polypeptide independently containing both the sensing and catalytic domain was considered structurally contrasting due to the small size of PA3042.

4.9. Structural homology analysis

Structural analysis of PA3042 using Dali server revealed an important hit to a "GAF-domain containing protein of DUF484 family" from P. aeruginosa, with a Z-score of 0.5. To further measure the structural similarity, both the structures were superimposed and TM-score of 0.564 was obtained. This suggests that both the proteins share a structural fold, but does not imply that PA3042 has a GAF domain. Instead, the helical architecture of pocket 1 acquires a GAF-like fold, especially in regard to the arrangement of 2°structural elements. Moreover, GAF domains are well-established cyclic nucleotide and other nucleotide binding modules and this structural correspondence might suggest a potential role of PA3042 in signaling activity.

4.9. Electrostatic characterization of pocket 1

Visualization of the electrostatic map generated using APBS reveals that pocket 1 consists of a deep positively charged groove. A mixture of polar and hydrophobic residues lines the pocket, forming a chemically diverse binding region. The strong electropositive character of the pocket 1 groove, indicates that it has a physiochemical preference of negatively charged ligands. Thus, this charge-based profiling offers a separate physiochemical justification for signaling molecules as potential P3042 binding partners.

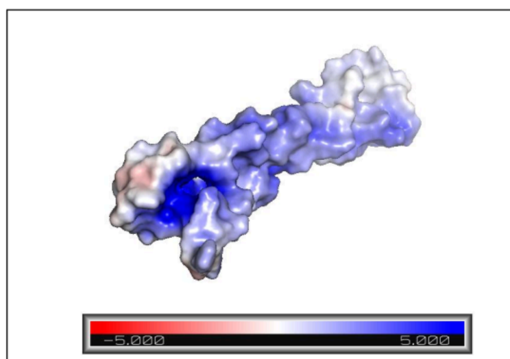


Figure III: Electrostatic surface potential analysis of target protein's modelled tertiary structure generated by the integrated APBS plugin: The electric potential dispersion is visualized as color on the molecular surface, with red color indicating a negative charge and the areas with positive potential appear in blue. The analysis reveals possible ligand attaching regions, intermolecular interaction sites and functionally significant surface exposed residues which may contribute to protein stability and molecular recognition apart from highlighting the charge distribution.

4.10. Conserved residue mapping

The HMMER search resulted in several significant hits with E-values in the range of e-22 to e-8 from variety of Pseudomonas species, indicating that PA3042 has identifiable homologs within genus. The top 50 hits along with the query sequence was used to generate an MSA and

subsequently used for ConSurf analysis which revealed that 14 residues in pocket 1 had conservation score in the range of 7 to 9, indicating evolutionary conserved regions. The presence of clustered conserved residues specifically within the binding pocket indicates that this area likely has a functional role.

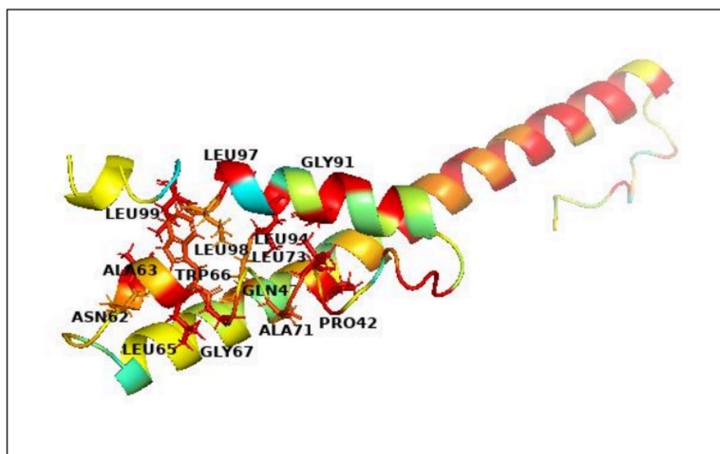


Figure IV: Evolutionary conservation assessment of target protein produced by ConSurf: A cyan-red-yellow gradient was used to color the protein's tertiary structure based on residue conservation scores, cyan represents variable residues, yellow stands for moderately conserved amino acids and red highlights highly conserved residues. In Pocket 1, conserved amino acids with ConSurf score between 7 to 9 are shown as sticks and labelled to highlight their distribution within the anticipated pocket. Presence of highly conserved residues may suggest a possible functional role.

4.11. Molecular docking results of selected ligands

Using AutoDock Vina, a list of 14 compounds involved in biofilm-associated signaling in P. aeruginosa were docked within pocket 1 and the binding affinity results for all ligands are presented in the Table V.

Table V: Molecular docking results showing binding energies of selected ligands at pocket 1 of the target protein.

Ligand	Binding Affinity (kcal/mol)
c-di-GMP	−8.8
pGpG	−8.0
ATP	−7.4
cGMP	−7.0
cAMP	−6.9
GTP	−6.7
PQS	−6.7
AMP	−6.5

3-oxo-C12HSL	−6.2
HHQ	−6.0
C4-HSL	−5.4
Citrate	−5.2
Phosphatidylethanolamine	−5.1
Glutamate	−4.3

The docking result showed a very distinct order of binding affinities throughout the ligands. The best binding ligands were c-di-GMP with a score of -8.8 kcal/mol and pGpG with -8.0 kcal/mol and both are guanosine-based nucleotides. They were followed by other nucleotides namely, ATP (-7.4), cGMP (-7.0), cAMP (-6.9), GTP (-6.7) and AMP (-6.5), Other non-nucleotide signaling molecules and controls occupied lowest positions. This consistent separation between non-nucleotide and nucleotide-class ligands shows that observed differences in binding affinities are due to actual pocket geometry and physiochemical complementarity rather than non-specific interactions. Noatbly, the top three positions were occupied by guanosine based molecules including c-di-GMP, pGpG and GTP, which suggests that pocket 1 has a preference for them over other nucleotide molecules.

4.12. Assessment of docking poses

In order to validate the docking results, they were repeated using the same parameter but three separate random seeds. For c-di-GMP remarkable reproducibility was observed as all the docking runs consistently predicted its binding the same site within pocket 1. The pairwise RMSD calculation of the top binding pose for all three independent seeds gave a value of 0.001 Å, demonstrating a nearly identical ligand binding orientation throughout the docking runs. Infact the binding energies for all the three seeds were consistent at -8.8 kcal/mol indicating extremely stable and repeatable binding interaction.

Docking of pGpG, on the other hand revealed far more fluctuation than c-di-GMP between the separate docking runs. Its pairwise RMSD comparison between the best pose across the three docking seeds ranged from 0.064 to 2.104 Å. However, despite the variation all the RMSD values were below the recognized threshold of 3 Å, demonstrating acceptable reproducibility. Consistent with this result, the binding energies also showed a small fluctuation, with mean binding energy being -8.07 kcal/mol and a standard deviation of ± 0.11.

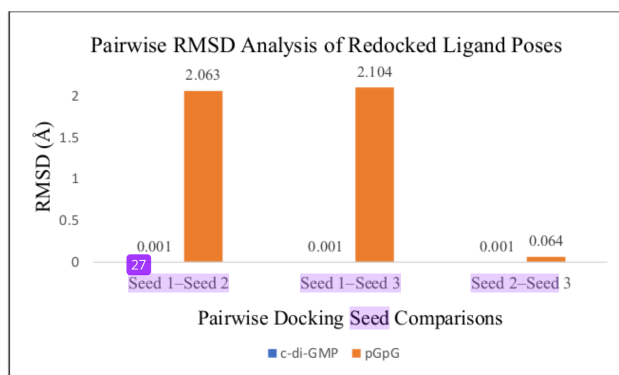


Figure V: Comparison of RMSD values of the best docking poses from three redocked complexes with different initialization seeds: The RMSD values are illustrated as bar graphs and indicates the structural stability and reproducibility of the various docking runs. Low RMSD values show that the poses are highly similar and represents dependable conformation of the ligand within the particular binding site.

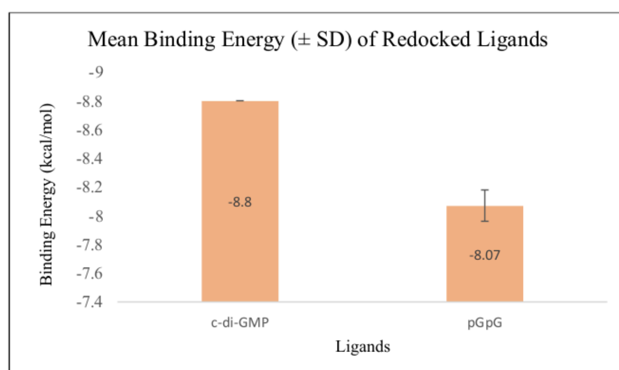


Figure VI: Mean binding energies from the best poses of the three separate docking seeds are compared: The bar graph displays the mean energy and their corresponding standard deviation are represented as error bars which give an assessment of consistency and reproducibility of interaction between protein and ligand and show variability between docking runs.

4.13. STRING analysis and Phylogenetic co-occurrence

The STRING database which was used to analyze interacting partners of PA3042, revealed strong associations with proteins encoded by primarily neighboring genomic genes and some co-expressing proteins like PA1323, OsmC, PA1324. Among the adjacent genes, PA3040 and

PA3042 are essentially the part of the same operon as PA3042 and their roles do not directly correspond to signaling model presented here, but PA3043 which encodes for HD domain family, dgt2 (dGTP triphosphohydrolase) protein was noteworthy. This close genomic vicinity suggests a potential functional interaction between the two proteins. Similar genomic organizations are frequently found in bacterial systems involving nucleotide signaling, in which nucleotide sensing or regulatory protein is found next to enzymatic proteins which degrade them, allowing for localized and tightly controlled signaling responses. Using phmmer based homology search on PA3043, phylogenetic co-occurrence was identified to further explore evolutionary conservation. Numerous *Pseudomonas* species and many other genera have been found to express homologs of PA3043 and this widespread occurrence and conservation of both proteins suggests that the link between PA3042 like protein and a nucleotide metabolizing enzyme may indicate towards a conserved functional module in the bacteria, however since this is an in-silico prediction, it would require further experimental validation to confirm this relationship.

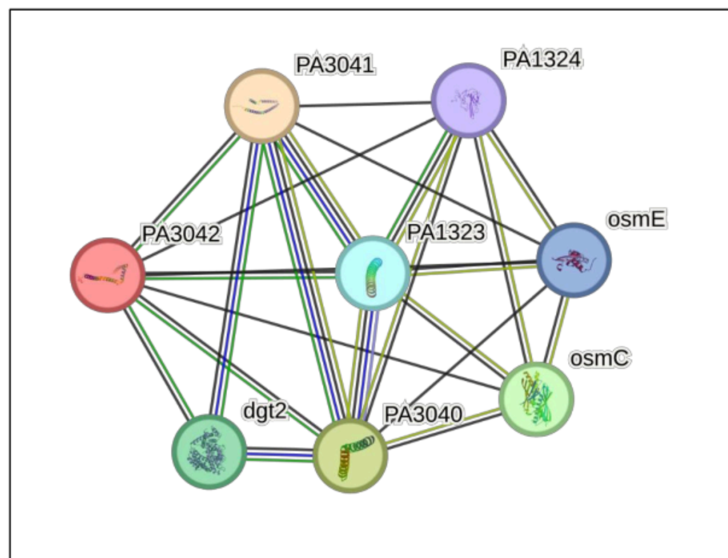


Figure VII: Interacting partners of PA3042 as predicted by the STRING database: The probable functional association partners are shown in the interaction map and only those interaction with confidence score of higher than 0.65 were included in the network. Proteins are represented by nodes and the connecting edges show the expected functional relationship based on computational predictions, so-expression, genomic context and experimental data.

5. CONCLUSION AND FUTURE PERSPECTIVE

The present study employed multiple computational and bioinformatics approaches to perform an in-silico characterization of the hypothetical protein PA3042 from Pseudomonas aeruginosa. Sequence analysis, physicochemical profiling, subcellular localization prediction, secondary and tertiary structure modeling, binding pocket identification, and molecular docking studies were collectively used to obtain preliminary insights into the structural and functional properties of the protein. The findings suggest that PA3042 may possess membrane-associated characteristics and could potentially participate in cellular processes related to environmental adaptation, regulatory activity, or signaling-associated mechanisms. Structural modeling predicted a stable three-dimensional conformation containing identifiable ligand-binding regions, while docking analysis demonstrated favorable interactions with selected ligands. In addition, functional annotation indicated similarities with proteins involved in regulatory or interaction-based cellular processes; however, these observations remain predictive and require further validation.

Overall, the computational evidence generated in this study provides a preliminary functional framework for PA3042 and contributes toward the characterization of previously unannotated proteins in Pseudomonas aeruginosa. Nevertheless, the biological role of the protein cannot be conclusively established through computational methods alone, and the present findings should therefore be considered exploratory in nature. Future investigations may further examine the possible involvement of PA3042 in nucleotide-associated regulatory pathways, membrane-linked cellular processes, and adaptive bacterial responses under varying environmental conditions. Additional studies focusing on interaction patterns, ligand specificity, and evolutionary conservation may help clarify whether PA3042 contributes to signaling or regulatory networks within Pseudomonas aeruginosa, thereby improving the broader understanding of bacterial adaptation and pathogenicity

REFERENCES

[1] B. Aslam *et al.*, "AMR and Sustainable Development Goals: at a crossroads," *Global. Health*, vol. 20, no. 1, p. 73, Oct. 2024, doi: 10.1186/s12992-024-01046-8.

[2] M. Kolář, "Bacterial Infections, Antimicrobial Resistance and Antibiotic Therapy," *Life*, vol. 12, no. 4, p. 468, Mar. 2022, doi: 10.3390/life12040468.

[3] W. Wu, Y. Jin, F. Bai, and S. Jin, "Pseudomonas aeruginosa," in *Molecular Medical Microbiology*, Elsevier, 2015, pp. 753–767. doi: 10.1016/B978-0-12-397169-2.00041-X.

[4] M. F. Moradali, S. Ghods, and B. H. A. Rehm, "Pseudomonas aeruginosa Lifestyle: A Paradigm for Adaptation, Survival, and Persistence," *Front. Cell. Infect. Microbiol.*, vol. 7, Feb. 2017, doi: 10.3389/fcimb.2017.00039.

[5] N. Sathe, P. Beech, L. Croft, C. Suphioglu, A. Kapat, and E. Athan, "Pseudomonas aeruginosa: Infections and novel approaches to treatment 'Knowing the enemy' the threat of Pseudomonas aeruginosa and exploring novel approaches to treatment," *Infect. Med.*, vol. 2, no. 3, pp. 178–194, Sep. 2023, doi: 10.1016/j.imj.2023.05.003.

[6] O. Ciofu, C. Moser, P. Ø. Jensen, and N. Høiby, "Tolerance and resistance of microbial biofilms," *Nat. Rev. Microbiol.*, vol. 20, no. 10, pp. 621–635, Oct. 2022, doi: 10.1038/s41579-022-00682-4.

[7] R. Yin, J. Cheng, J. Wang, P. Li, and J. Lin, "Treatment of Pseudomonas aeruginosa infectious biofilms: Challenges and strategies," *Front. Microbiol.*, vol. 13, Aug. 2022, doi: 10.3389/fmicb.2022.955286.

[8] S.-K. Kim and J.-H. Lee, "Biofilm dispersion in Pseudomonas aeruginosa," *J. Microbiol.*, vol. 54, no. 2, pp. 71–85, Feb. 2016, doi: 10.1007/s12275-016-5528-7.

[9] P. S. Stewart, M. J. Franklin, K. S. Williamson, J. P. Folsom, L. Boegli, and G. A. James, "Contribution of Stress Responses to Antibiotic Tolerance in Pseudomonas aeruginosa Biofilms," *Antimicrob. Agents Chemother.*, vol. 59, no. 7, pp. 3838–3847, Jul. 2015, doi: 10.1128/AAC.00433-15.

[10] J. Walker and G. Moore, "Pseudomonas aeruginosa in hospital water systems: biofilms, guidelines, and practicalities," *J. Hosp. Infect.*, vol. 89, no. 4, pp. 324–327, Apr. 2015, doi: 10.1016/j.jhin.2014.11.019.

[11] K. Vadakkan *et al.*, "Critical review on plant-derived quorum sensing signaling inhibitors in pseudomonas aeruginosa," *Bioorg. Chem.*, vol. 151, p. 107649, Oct. 2024, doi: 10.1016/j.bioorg.2024.107649.

[12] S. Park and K. Sauer, "Controlling Biofilm Development Through Cyclic di-GMP Signaling," 2022, pp. 69–94. doi: 10.1007/978-3-031-08491-1_3.

[13] Z. Liang *et al.*, "Transcription of the Alginate Operon in Pseudomonas aeruginosa Is Regulated by c-di-GMP," *Microbiol. Spectr.*, vol. 10, no. 4, Aug. 2022, doi: 10.1128/spectrum.00675-22.

[14] C. L. Hall and V. T. Lee, "Cyclic-di-GMP regulation of virulence in bacterial pathogens," *WIREs RNA*, vol. 9, no. 1, Jan. 2018, doi: 10.1002/wrna.1454.

[15] F. Khan, D. T. N. Pham, S. F. Oloketuyi, and Y.-M. Kim, "Regulation and controlling

the motility properties of Pseudomonas aeruginosa," *Appl. Microbiol. Biotechnol.*, vol. 104, no. 1, pp. 33–49, Jan. 2020, doi: 10.1007/s00253-019-10201-w.

[16] S. Caprari, V. Brandi, A. Pasquadibisceglie, and F. Polticelli, "Uncovering the structure and function of Pseudomonas aeruginosa periplasmic proteins by an in silico approach," *J. Biomol. Struct. Dyn.*, vol. 38, no. 15, pp. 4508–4520, Oct. 2020, doi: 10.1080/07391102.2019.1683468.

[17] M. Z. Østergaard, F. D. Nielsen, M. H. Meinfeldt, and C. L. Kirkpatrick, " The uncharacterized PA3040-3042 operon is part of the cell envelope stress response and a tobramycin resistance determinant in a clinical isolate of Pseudomonas aeruginosa ," *Microbiol. Spectr.*, vol. 12, no. 8, 2024, doi: 10.1128/spectrum.03875-23.

[18] M. Naveed *et al.*, "The Virulent Hypothetical Proteins: The Potential Drug Target Involved in Bacterial Pathogenesis," *Mini-Reviews Med. Chem.*, vol. 22, no. 20, pp. 2608–2623, Nov. 2022, doi: 10.2174/1389557522666220413102107.

[19] B. Kuhlman and P. Bradley, "Advances in protein structure prediction and design," *Nat. Rev. Mol. Cell Biol.*, vol. 20, no. 11, pp. 681–697, Nov. 2019, doi: 10.1038/s41580-019-0163-x.

[20] A. Shehu, D. Barbará, and K. Molloy, "A Survey of Computational Methods for Protein Function Prediction," in *Big Data Analytics in Genomics*, Cham: Springer International Publishing, 2016, pp. 225–298. doi: 10.1007/978-3-319-41279-5_7.

[21] C. Liao, X. Huang, Q. Wang, D. Yao, and W. Lu, "Virulence Factors of Pseudomonas Aeruginosa and Antivirulence Strategies to Combat Its Drug Resistance," *Front. Cell. Infect. Microbiol.*, vol. 12, no. July, pp. 1–17, 2022, doi: 10.3389/fcimb.2022.926758.

[22] X. Zhang *et al.*, "A comprehensive review of the pathogenic mechanisms of Pseudomonas aeruginosa: synergistic effects of virulence factors, quorum sensing, and biofilm formation," *Front. Microbiol.*, vol. 16, no. July, pp. 1–14, 2025, doi: 10.3389/fmicb.2025.1619626.

[23] D. Reynolds and M. Kollef, "The Epidemiology and Pathogenesis and Treatment of Pseudomonas aeruginosa Infections: An Update," *Drugs*, vol. 81, no. 18, pp. 2117–2131, Dec. 2021, doi: 10.1007/s40265-021-01635-6.

[24] M. Garcia-Clemente *et al.*, "Impact of Pseudomonas aeruginosa Infection on Patients with Chronic Inflammatory Airway Diseases," *J. Clin. Med.*, vol. 9, no. 12, p. 3800, Nov. 2020, doi: 10.3390/jcm9123800.

[25] S. Finch, M. J. McDonnell, H. Abo-Leyah, S. Aliberti, and J. D. Chalmers, "A Comprehensive Analysis of the Impact of Pseudomonas aeruginosa Colonization on Prognosis in Adult Bronchiectasis," *Ann. Am. Thorac. Soc.*, vol. 12, no. 11, pp. 1602–1611, Nov. 2015, doi: 10.1513/AnnalsATS.201506-333OC.

[26] M. Buhl, S. Peter, and M. Willmann, "Prevalence and risk factors associated with colonization and infection of extensively drug-resistant Pseudomonas aeruginosa : a systematic review," *Expert Rev. Anti. Infect. Ther.*, vol. 13, no. 9, pp. 1159–1170, Sep. 2015, doi: 10.1586/14787210.2015.1064310.

[27] J. Fong *et al.*, "Combination Therapy Strategy of Quorum Quenching Enzyme and Quorum Sensing Inhibitor in Suppressing Multiple Quorum Sensing Pathways of P. aeruginosa," *Sci. Rep.*, vol. 8, no. 1, p. 1155, Jan. 2018, doi: 10.1038/s41598-018-19504-

w.

[28] K. Vadakkan, A. K. Ngangbam, K. Sathishkumar, N. P. Rumjit, and M. K. Cheruvathur, “A review of chemical signaling pathways in the quorum sensing circuit of Pseudomonas aeruginosa,” *Int. J. Biol. Macromol.*, vol. 254, p. 127861, Jan. 2024, doi: 10.1016/j.ijbiomac.2023.127861.

[29] M. W. Azam and A. U. Khan, “Updates on the pathogenicity status of Pseudomonas aeruginosa,” *Drug Discov. Today*, vol. 24, no. 1, pp. 350–359, Jan. 2019, doi: 10.1016/j.drudis.2018.07.003.

[30] F. F. Tuon, L. R. Dantas, P. H. Suss, and V. S. Tasca Ribeiro, “Pathogenesis of the Pseudomonas aeruginosa Biofilm: A Review,” *Pathogens*, vol. 11, no. 3, 2022, doi: 10.3390/pathogens11030300.

[31] N. Seder, W. A. Rayyan, L. Al-Fawares, M. Hilmi, and A. Bakar, “Pseudomonas aeruginosa Virulence Factors and Antivirulence mechanisms to Combat Drug Resistance; A Systematic Review,” *Sapporo Med. J.*, vol. 56, no. 12, pp. 1–23, 2022.

[32] S. Qin *et al.*, “Pseudomonas aeruginosa: pathogenesis, virulence factors, antibiotic resistance, interaction with host, technology advances and emerging therapeutics,” *Signal Transduct. Target. Ther.*, vol. 7, no. 1, p. 199, Jun. 2022, doi: 10.1038/s41392-022-01056-1.

[33] P. Banerjee, P. K. Sahoo, Sheenu, A. Adhikary, R. Ruhal, and D. Jain, “Molecular and structural facets of c-di-GMP signalling associated with biofilm formation in Pseudomonas aeruginosa,” *Mol. Aspects Med.*, vol. 81, no. February, p. 101001, 2021, doi: 10.1016/j.mam.2021.101001.

[34] M. T. T. Thi, D. Wibowo, and B. H. A. Rehm, “Pseudomonas aeruginosa Biofilms,” *Int. J. Mol. Sci.*, vol. 21, no. 22, p. 8671, Nov. 2020, doi: 10.3390/ijms21228671.

[35] N. M. Maurice, B. Bedi, and R. T. Sadikot, “Pseudomonas aeruginosa Biofilms: Host Response and Clinical Implications in Lung Infections,” *Am. J. Respir. Cell Mol. Biol.*, vol. 58, no. 4, pp. 428–439, Apr. 2018, doi: 10.1165/rcmb.2017-0321TR.

[36] T. Rasamiravaka, Q. Labtani, P. Duez, and M. El Jaziri, “The Formation of Biofilms by Pseudomonas aeruginosa : A Review of the Natural and Synthetic Compounds Interfering with Control Mechanisms,” *Biomed Res. Int.*, vol. 2015, pp. 1–17, 2015, doi: 10.1155/2015/759348.

[37] M. del M. Cendra and E. Torrents, “Pseudomonas aeruginosa biofilms and their partners in crime,” *Biotechnol. Adv.*, vol. 49, p. 107734, Jul. 2021, doi: 10.1016/j.biotechadv.2021.107734.

[38] A. Vetrivel *et al.*, “Pseudomonas aeruginosa Biofilm Formation and Its Control,” *Biologics*, vol. 1, no. 3, pp. 312–336, 2021, doi: 10.3390/biologics1030019.

[39] C.-Y. Chang, “Surface Sensing for Biofilm Formation in Pseudomonas aeruginosa,” *Front. Microbiol.*, vol. 8, Jan. 2018, doi: 10.3389/fmicb.2017.02671.

[40] S. Skariyachan, V. S. Sridhar, S. Packirisamy, S. T. Kumargowda, and S. B. Challapilli, “Recent perspectives on the molecular basis of biofilm formation by Pseudomonas aeruginosa and approaches for treatment and biofilm dispersal,” *Folia Microbiol. (Praha).*, vol. 63, no. 4, pp. 413–432, Jul. 2018, doi: 10.1007/s12223-018-0585-4.

[41] M. Fernández-Billón, A. E. Llambías-Cabot, E. Jordana-Lluch, A. Oliver, and M. D. Macià, “Mechanisms of antibiotic resistance in Pseudomonas aeruginosa biofilms,” *Biofilm*, vol. 5, p. 100129, Dec. 2023, doi: 10.1016/j.bioflm.2023.100129.

[42] D. Shokri, M. R. Khorasgani, M. Mohkam, S. M. Fatemi, Y. Ghasemi, and A. Taheri-Kafrani, “The Inhibition Effect of Lactobacilli Against Growth and Biofilm Formation of Pseudomonas aeruginosa,” *Probiotics Antimicrob. Proteins*, vol. 10, no. 1, pp. 34–42, Mar. 2018, doi: 10.1007/s12602-017-9267-9.

[43] B. A. Omran, B. S. Tseng, and K.-H. Baek, “Nanocomposites against Pseudomonas aeruginosa biofilms: Recent advances, challenges, and future prospects,” *Microbiol. Res.*, vol. 282, p. 127656, May 2024, doi: 10.1016/j.micres.2024.127656.

[44] Z. Chegini, A. Khoshbayan, M. Taati Moghadam, I. Farahani, P. Jazireian, and A. Shariati, “Bacteriophage therapy against Pseudomonas aeruginosa biofilms: a review,” *Ann. Clin. Microbiol. Antimicrob.*, vol. 19, no. 1, p. 45, Dec. 2020, doi: 10.1186/s12941-020-00389-5.

[45] D.-G. Ha and G. A. O’Toole, “c-di-GMP and its Effects on Biofilm Formation and Dispersion: a Pseudomonas Aeruginosa Review,” in *Microbial Biofilms*, Washington, DC, USA: ASM Press, 2015, pp. 301–317. doi: 10.1128/9781555817466.ch15.

[46] S. N. Stoner, J. J. Baty, and J. A. Scoffield, “Pseudomonas aeruginosa polysaccharide Psl supports airway microbial community development,” *ISME J.*, vol. 16, no. 7, pp. 1730–1739, Jul. 2022, doi: 10.1038/s41396-022-01221-y.

[47] M. Rybtke, P. Ø. Jensen, C. H. Nielsen, and T. Tolker-Nielsen, “The Extracellular Polysaccharide Matrix of Pseudomonas aeruginosa Biofilms Is a Determinant of Polymorphonuclear Leukocyte Responses,” *Infect. Immun.*, vol. 89, no. 1, Dec. 2020, doi: 10.1128/IAI.00631-20.

[48] L. Z. Ma, D. Wang, Y. Liu, Z. Zhang, and D. J. Wozniak, “Regulation of Biofilm Exopolysaccharide Biosynthesis and Degradation in Pseudomonas aeruginosa,” *Annu. Rev. Microbiol.*, vol. 76, no. 1, pp. 413–433, Sep. 2022, doi: 10.1146/annurev-micro-041320-111355.

[49] O. J. Hills, J. Smith, A. J. Scott, D. A. Devine, and H. F. Chappell, “Cation complexation by mucoid Pseudomonas aeruginosa extracellular polysaccharide,” *PLoS One*, vol. 16, no. 9, p. e0257026, Sep. 2021, doi: 10.1371/journal.pone.0257026.

[50] B. S. Tseng *et al.*, “A Biofilm Matrix-Associated Protease Inhibitor Protects Pseudomonas aeruginosa from Proteolytic Attack,” *MBio*, vol. 9, no. 2, May 2018, doi: 10.1128/mBio.00543-18.

[51] F. G. Vital-Lopez, J. Reifman, and A. Wallqvist, “Biofilm Formation Mechanisms of Pseudomonas aeruginosa Predicted via Genome-Scale Kinetic Models of Bacterial Metabolism,” *PLOS Comput. Biol.*, vol. 11, no. 10, p. e1004452, Oct. 2015, doi: 10.1371/journal.pcbi.1004452.

[52] T. de Sousa *et al.*, “Genomic and Metabolic Characteristics of the Pathogenicity in Pseudomonas aeruginosa,” *Int. J. Mol. Sci.*, vol. 22, no. 23, p. 12892, Nov. 2021, doi: 10.3390/ijms222312892.

[53] V. I. Francis, E. C. Stevenson, and S. L. Porter, “Two-component systems required for virulence in Pseudomonas aeruginosa,” *FEMS Microbiol. Lett.*, vol. 364, no. 11, Jun.

2017, doi: 10.1093/femsle/fnx104.

[54] V. R. Flores-Vega *et al.*, "Contribution of two-component regulatory systems to the acute-to-chronic infection transition of Pseudomonas aeruginosa in cystic fibrosis," *J. Bacteriol.*, vol. 208, no. 3, Mar. 2026, doi: 10.1128/jb.00471-25.

[55] A. Y. Bhagirath *et al.*, "Two Component Regulatory Systems and Antibiotic Resistance in Gram-Negative Pathogens," *Int. J. Mol. Sci.*, vol. 20, no. 7, p. 1781, Apr. 2019, doi: 10.3390/ijms20071781.

[56] A. R. Tierney and P. N. Rather, "Roles of Two-Component Regulatory Systems in Antibiotic Resistance," *Future Microbiol.*, vol. 14, no. 6, pp. 533–552, Apr. 2019, doi: 10.2217/fmb-2019-0002.

[57] M. Valentini, D. Gonzalez, D. A. Mavridou, and A. Filloux, "Lifestyle transitions and adaptive pathogenesis of Pseudomonas aeruginosa," *Curr. Opin. Microbiol.*, vol. 41, pp. 15–20, Feb. 2018, doi: 10.1016/j.mib.2017.11.006.

[58] J. De Smet *et al.*, "Bacteriophage-mediated interference of the c-di-GMP signalling pathway in Pseudomonas aeruginosa," *Microb. Biotechnol.*, vol. 14, no. 3, pp. 967–978, May 2021, doi: 10.1111/1751-7915.13728.

[59] S. J. Cole and V. T. Lee, "Cyclic Di-GMP Signaling Contributes to Pseudomonas aeruginosa-Mediated Catheter-Associated Urinary Tract Infection," *J. Bacteriol.*, vol. 198, no. 1, pp. 91–97, Jan. 2016, doi: 10.1128/JB.00410-15.

[60] C. Kennelly, P. Tran, and A. Prindle, "Environmental purines decrease Pseudomonas aeruginosa biofilm formation by disrupting c-di-GMP metabolism," *Cell Rep.*, vol. 43, no. 5, p. 114154, May 2024, doi: 10.1016/j.celrep.2024.114154.

[61] G. Guła, A. Dorotkiewicz-Jach, K. Korzekwa, M. A. Valvano, and Z. Drulis-Kawa, "Complex Signaling Networks Controlling Dynamic Molecular Changes in Pseudomonas aeruginosa Biofilm," *Curr. Med. Chem.*, vol. 26, no. 11, pp. 1979–1993, 2018, doi: 10.2174/0929867325666180912110151.

[62] F. Wang, Q. He, J. Yin, S. Xu, W. Hu, and L. Gu, "BrlR from Pseudomonas aeruginosa is a receptor for both cyclic di-GMP and pyocyanin," *Nat. Commun.*, vol. 9, no. 1, p. 2563, Jul. 2018, doi: 10.1038/s41467-018-05004-y.

[63] J. B. Andersen *et al.*, "Induction of Native c-di-GMP Phosphodiesterases Leads to Dispersal of Pseudomonas aeruginosa Biofilms," *Antimicrob. Agents Chemother.*, vol. 65, no. 4, Mar. 2021, doi: 10.1128/AAC.02431-20.

[64] Q. Wei *et al.*, "Diguanylate Cyclases and Phosphodiesterases Required for Basal-Level c-di-GMP in Pseudomonas aeruginosa as Revealed by Systematic Phylogenetic and Transcriptomic Analyses," *Appl. Environ. Microbiol.*, vol. 85, no. 21, Nov. 2019, doi: 10.1128/AEM.01194-19.

[65] A. Kariminik, M. Baseri-Salehi, and B. Kheirkhah, "Pseudomonas aeruginosa quorum sensing modulates immune responses: An updated review article," *Immunol. Lett.*, vol. 190, pp. 1–6, Oct. 2017, doi: 10.1016/j.imlet.2017.07.002.

[66] J. H. Hammond, E. F. Dolben, T. J. Smith, S. Bhuju, and D. A. Hogan, "Links between Anr and Quorum Sensing in Pseudomonas aeruginosa Biofilms," *J. Bacteriol.*, vol. 197, no. 17, pp. 2810–2820, Sep. 2015, doi: 10.1128/JB.00182-15.

[67] J. Luo *et al.*, “Baicalein attenuates the quorum sensing-controlled virulence factors of Pseudomonas aeruginosa and relieves the inflammatory response in P. aeruginosa-infected macrophages by downregulating the MAPK and NFκB signal-transduction pathways,” *Drug Des. Devel. Ther.*, p. 183, Jan. 2016, doi: 10.2147/DDDT.S97221.

[68] J. Chadha, K. Harjai, and S. Chhibber, “Revisiting the virulence hallmarks of <scp> *Pseudomonas aeruginosa* </scp> : a chronicle through the perspective of quorum sensing,” *Environ. Microbiol.*, vol. 24, no. 6, pp. 2630–2656, Jun. 2022, doi: 10.1111/1462-2920.15784.

[69] Q. Li, S. Mao, H. Wang, and X. Ye, “The Molecular Architecture of Pseudomonas aeruginosa Quorum-Sensing Inhibitors,” *Mar. Drugs*, vol. 20, no. 8, p. 488, Jul. 2022, doi: 10.3390/md20080488.

[70] H. Almblad *et al.*, “The Cyclic AMP-Vfr Signaling Pathway in Pseudomonas aeruginosa Is Inhibited by Cyclic Di-GMP,” *J. Bacteriol.*, vol. 197, no. 13, pp. 2190–2200, Jul. 2015, doi: 10.1128/JB.00193-15.

[71] J. Wang, Y. Wang, H. Lou, and W. Wang, “AlgU controls environmental stress adaptation, biofilm formation, motility, pyochelin synthesis and antagonism potential in Pseudomonas protegens SN15-2,” *Microbiol. Res.*, vol. 272, p. 127396, Jul. 2023, doi: 10.1016/j.micres.2023.127396.

[72] S. Lin Chua *et al.*, “Reduced Intracellular c-di-GMP Content Increases Expression of Quorum Sensing-Regulated Genes in Pseudomonas aeruginosa,” *Front. Cell. Infect. Microbiol.*, vol. 7, Oct. 2017, doi: 10.3389/fcimb.2017.00451.

[73] M. A. Matilla, D. Martín-Mora, J. A. Gavira, and T. Krell, “Pseudomonas aeruginosa as a Model To Study Chemosensory Pathway Signaling,” *Microbiol. Mol. Biol. Rev.*, vol. 85, no. 1, Feb. 2021, doi: 10.1128/MMBR.00151-20.

[74] A. T. Vincent, “Bacterial hypothetical proteins may be of functional interest,” *Front. Bacteriol.*, vol. 3, Mar. 2024, doi: 10.3389/fbrio.2024.1334712.

[75] J. Prava, G. Pranavathiyani, and A. Pan, “Functional assignment for essential hypothetical proteins of Staphylococcus aureus N315,” *Int. J. Biol. Macromol.*, vol. 108, pp. 765–774, 2018, doi: 10.1016/j.ijbiomac.2017.10.169.

[76] N. Hanif *et al.*, “In Silico Characterization of Hypothetical Protein AZJ53_10480 in Streptococcus pneumoniae,” *Biosci. Rev.*, vol. 6, no. 4, pp. 1–12, 2024, doi: 10.32350/bsr.64.01.

[77] J. Ijaq *et al.*, “A model to predict the function of hypothetical proteins through a nine-point classification scoring schema,” *BMC Bioinformatics*, vol. 20, no. 1, p. 14, Dec. 2019, doi: 10.1186/s12859-018-2554-y.

[78] A. Bin Hafeez, S. Sappati, R. Krzemieniecki, R. Worobo, and P. Szweda, “In Silico Functional Annotation and Structural Characterization of Hypothetical Proteins in Bacillus paralicheniformis and Bacillus subtilis Isolated from Honey,” *ACS Omega*, vol. 10, no. 9, pp. 8993–9006, Mar. 2025, doi: 10.1021/acsomega.4c07105.

[79] M. A. Rahman, U. H. Heme, and M. A. K. Parvez, “In silico functional annotation of hypothetical proteins from the Bacillus paralicheniformis strain Bac84 reveals proteins with biotechnological potentials and adaptational functions to extreme environments,”

PLoS One, vol. 17, no. 10, p. e0276085, Oct. 2022, doi: 10.1371/journal.pone.0276085.

[80] B. Huang *et al.*, "Protein Structure Prediction: Challenges, Advances, and the Shift of Research Paradigms," *Genomics. Proteomics Bioinformatics*, vol. 21, no. 5, pp. 913–925, Oct. 2023, doi: 10.1016/j.gpb.2022.11.014.

[81] R. Chen, X. Li, Y. Yang, X. Song, C. Wang, and D. Qiao, "Prediction of protein-protein interaction sites in intrinsically disordered proteins," *Front. Mol. Biosci.*, vol. 9, Sep. 2022, doi: 10.3389/fmolb.2022.985022.

[82] S. Bhattacharya and X. Lin, "Recent Advances in Computational Protocols Addressing Intrinsically Disordered Proteins," *Biomolecules*, vol. 9, no. 4, p. 146, Apr. 2019, doi: 10.3390/biom9040146.

[83] L. X. Peterson, A. Roy, C. Christoffer, G. Terashi, and D. Kihara, "Modeling disordered protein interactions from biophysical principles," *PLOS Comput. Biol.*, vol. 13, no. 4, p. e1005485, Apr. 2017, doi: 10.1371/journal.pcbi.1005485.

[84] S. Sivashankari and P. Shanmughavel, "Functional annotation of hypothetical proteins – A review," *Bioinformation*, vol. 1, no. 8, pp. 335–338, Dec. 2006, doi: 10.6026/97320630001335.

[85] S. Gottesman, "Trouble is coming: Signaling pathways that regulate general stress responses in bacteria," *J. Biol. Chem.*, vol. 294, no. 31, pp. 11685–11700, Aug. 2019, doi: 10.1074/jbc.REV119.005593.

[86] S. Tiwari *et al.*, "Two-Component Signal Transduction Systems of Pathogenic Bacteria As Targets for Antimicrobial Therapy: An Overview," *Front. Microbiol.*, vol. 8, Oct. 2017, doi: 10.3389/fmicb.2017.01878.

[87] K. He and C. E. Bauer, "Chemosensory signaling systems that control bacterial survival," *Trends Microbiol.*, vol. 22, no. 7, pp. 389–398, Jul. 2014, doi: 10.1016/j.tim.2014.04.004.

[88] G. Sharma, S. Rao, A. Bansal, S. Dang, S. Gupta, and R. Gabrani, "Pseudomonas aeruginosa biofilm: Potential therapeutic targets," *Biologicals*, vol. 42, no. 1, pp. 1–7, Jan. 2014, doi: 10.1016/j.biologicals.2013.11.001.

[89] D. Flores-Percino *et al.*, "Mechanisms of the Quorum Sensing Systems of Pseudomonas aeruginosa : Host and Bacteria," *Curr. Med. Chem.*, vol. 31, no. 35, pp. 5755–5767, Oct. 2024, doi: 10.2174/0929867331666230821110440.

[90] M. Y. Galperin, "What bacteria want," *Environ. Microbiol.*, vol. 20, no. 12, pp. 4221–4229, Dec. 2018, doi: 10.1111/1462-2920.14398.

[91] Y. Kawaharada *et al.*, "Receptor-mediated exopolysaccharide perception controls bacterial infection," *Nature*, vol. 523, no. 7560, pp. 308–312, Jul. 2015, doi: 10.1038/nature14611.

[92] A. E. Bem, N. Velikova, M. T. Pellicer, P. van Baarlen, A. Marina, and J. M. Wells, "Bacterial Histidine Kinases as Novel Antibacterial Drug Targets," *ACS Chem. Biol.*, vol. 10, no. 1, pp. 213–224, Jan. 2015, doi: 10.1021/cb5007135.

[93] S. Borland, C. Prigent-Combaret, and F. Wisniewski-Dyé, "Bacterial hybrid histidine kinases in plant–bacteria interactions," *Microbiology*, vol. 162, no. 10, pp. 1715–1734, Oct. 2016, doi: 10.1099/mic.0.000370.

[94] E. Ishii and Y. Eguchi, "Diversity in Sensing and Signaling of Bacterial Sensor Histidine Kinases," *Biomolecules*, vol. 11, no. 10, p. 1524, Oct. 2021, doi: 10.3390/biom11101524.

[95] J. W. Willett and S. Crosson, "Atypical modes of bacterial histidine kinase signaling," *Mol. Microbiol.*, vol. 103, no. 2, pp. 197–202, Jan. 2017, doi: 10.1111/mmi.13525.

[96] A. Ali-Ahmad *et al.*, "Structural and functional insights into the periplasmic detector domain of the GacS histidine kinase controlling biofilm formation in Pseudomonas aeruginosa," *Sci. Rep.*, vol. 7, no. 1, p. 11262, Sep. 2017, doi: 10.1038/s41598-017-11361-3.

[97] E. C. Stuffle, M. S. Johnson, and K. J. Watts, "PAS domains in bacterial signal transduction," *Curr. Opin. Microbiol.*, vol. 61, pp. 8–15, Jun. 2021, doi: 10.1016/j.mib.2021.01.004.

[98] T. H. Mann and L. Shapiro, "Integration of cell cycle signals by multi-PAS domain kinases," *Proc. Natl. Acad. Sci.*, vol. 115, no. 30, Jul. 2018, doi: 10.1073/pnas.1808543115.

[99] M. A. Matilla, F. Velando, D. Martín-Mora, E. Monteagudo-Cascales, and T. Krell, "A catalogue of signal molecules that interact with sensor kinases, chemoreceptors and transcriptional regulators," *FEMS Microbiol. Rev.*, vol. 46, no. 1, pp. 1–31, 2022, doi: 10.1093/femsre/fuab043.

[100] J. Xing, V. M. Gumerov, and I. B. Zhulin, "Origin and functional diversification of PAS domain, a ubiquitous intracellular sensor," *Sci. Adv.*, vol. 9, no. 35, Sep. 2023, doi: 10.1126/sciadv.adi4517.

[101] M. Y. Galperin and S.-H. Chou, "Structural Conservation and Diversity of PilZ-Related Domains," *J. Bacteriol.*, vol. 202, no. 4, Jan. 2020, doi: 10.1128/JB.00664-19.

[102] Q. W. Cheang, L. Xin, R. Y. F. Chea, and Z.-X. Liang, "Emerging paradigms for PilZ domain-mediated C-di-GMP signaling," *Biochem. Soc. Trans.*, vol. 47, no. 1, pp. 381–388, Feb. 2019, doi: 10.1042/BST20180543.

[103] X.-F. Yan *et al.*, "Structural analyses unravel the molecular mechanism of cyclic di-GMP regulation of bacterial chemotaxis via a PilZ adaptor protein," *J. Biol. Chem.*, vol. 293, no. 1, pp. 100–111, Jan. 2018, doi: 10.1074/jbc.M117.815704.

[104] M. W. Orr, M. Y. Galperin, and V. T. Lee, "Sustained sensing as an emerging principle in second messenger signaling systems," *Curr. Opin. Microbiol.*, vol. 34, pp. 119–126, Dec. 2016, doi: 10.1016/j.mib.2016.08.010.

[105] H. Mak and T. L. M. Thurston, "Interesting Biochemistries in the Structure and Function of Bacterial Effectors," *Front. Cell. Infect. Microbiol.*, vol. 11, Feb. 2021, doi: 10.3389/fcimb.2021.608860.

[106] R. Hengge, "High-specificity local and global c-di-GMP signaling," *Trends Microbiol.*, vol. 29, no. 11, pp. 993–1003, Nov. 2021, doi: 10.1016/j.tim.2021.02.003.

[107] M. Valentini and A. Filloux, "Multiple Roles of c-di-GMP Signaling in Bacterial Pathogenesis," *Annu. Rev. Microbiol.*, vol. 73, no. 1, pp. 387–406, Sep. 2019, doi: 10.1146/annurev-micro-020518-115555.

[108] U. Jenal, A. Reinders, and C. Lori, "Cyclic di-GMP: second messenger extraordinaire,"

Nat. Rev. Microbiol., vol. 15, no. 5, pp. 271–284, May 2017, doi: 10.1038/nrmicro.2016.190.

[109] V. Hauryliuk, G. C. Atkinson, K. S. Murakami, T. Tenson, and K. Gerdes, "Recent functional insights into the role of (p)ppGpp in bacterial physiology," *Nat. Rev. Microbiol.*, vol. 13, no. 5, pp. 298–309, May 2015, doi: 10.1038/nrmicro3448.

[110] W. Steinchen and G. Bange, "The magic dance of the alarmones (p)ppGpp," *Mol. Microbiol.*, vol. 101, no. 4, pp. 531–544, Aug. 2016, doi: 10.1111/mmi.13412.

[111] O. Pacios *et al.*, "(p)ppGpp and Its Role in Bacterial Persistence: New Challenges," *Antimicrob. Agents Chemother.*, vol. 64, no. 10, Sep. 2020, doi: 10.1128/AAC.01283-20.

[112] W. Yin *et al.*, "A decade of research on the second messenger c-di-AMP," *FEMS Microbiol. Rev.*, vol. 44, no. 6, pp. 701–724, Nov. 2020, doi: 10.1093/femsre/fuaa019.

[113] T. N. Huynh and J. J. Woodward, "Too much of a good thing: regulated depletion of c-di-AMP in the bacterial cytoplasm," *Curr. Opin. Microbiol.*, vol. 30, pp. 22–29, Apr. 2016, doi: 10.1016/j.mib.2015.12.007.

[114] C. Opoku-Temeng, J. Zhou, Y. Zheng, J. Su, and H. O. Sintim, "Cyclic dinucleotide (c-di-GMP, c-di-AMP, and cGAMP) signalings have come of age to be inhibited by small molecules," *Chem. Commun.*, vol. 52, no. 60, pp. 9327–9342, 2016, doi: 10.1039/C6CC03439J.

[115] A. Bock, R. Irannejad, and J. D. Scott, "cAMP signaling: a remarkably regional affair," *Trends Biochem. Sci.*, vol. 49, no. 4, pp. 305–317, Apr. 2024, doi: 10.1016/j.tibs.2024.01.004.

[116] A. Berry *et al.*, "cAMP and Vfr Control Exolysin Expression and Cytotoxicity of Pseudomonas aeruginosa Taxonomic Outliers," *J. Bacteriol.*, vol. 200, no. 12, Jun. 2018, doi: 10.1128/JB.00135-18.

[117] K. Lefkimmiatis and M. Zaccolo, "cAMP signaling in subcellular compartments," *Pharmacol. Ther.*, vol. 143, no. 3, pp. 295–304, Sep. 2014, doi: 10.1016/j.pharmthera.2014.03.008.

[118] K. Papenfort and B. L. Bassler, "Quorum sensing signal–response systems in Gram-negative bacteria," *Nat. Rev. Microbiol.*, vol. 14, no. 9, pp. 576–588, Sep. 2016, doi: 10.1038/nrmicro.2016.89.

[119] R. P. Ryan, S. An, J. H. Allan, Y. McCarthy, and J. M. Dow, "The DSF Family of Cell–Cell Signals: An Expanding Class of Bacterial Virulence Regulators," *PLOS Pathog.*, vol. 11, no. 7, p. e1004986, Jul. 2015, doi: 10.1371/journal.ppat.1004986.

[120] D. Patel and S. N. Witt, "Ethanolamine and Phosphatidylethanolamine: Partners in Health and Disease," *Oxid. Med. Cell. Longev.*, vol. 2017, no. 1, Jan. 2017, doi: 10.1155/2017/4829180.

[121] L. E. Reddick and N. M. Alto, "Bacteria Fighting Back: How Pathogens Target and Subvert the Host Innate Immune System," *Mol. Cell*, vol. 54, no. 2, pp. 321–328, Apr. 2014, doi: 10.1016/j.molcel.2014.03.010.

[122] B. Y. Matsuyama, P. V. Krasteva, C. Baraquet, C. S. Harwood, H. Sondermann, and M. V. A. S. Navarro, "Mechanistic insights into c-di-GMP–dependent control of the biofilm

regulator FleQ from Pseudomonas aeruginosa," *Proc. Natl. Acad. Sci.*, vol. 113, no. 2, Jan. 2016, doi: 10.1073/pnas.1523148113.

[123] E. Blanco-Romero *et al.*, "Genome-wide analysis of the FleQ direct regulon in Pseudomonas fluorescens F113 and Pseudomonas putida KT2440," *Sci. Rep.*, vol. 8, no. 1, p. 13145, Sep. 2018, doi: 10.1038/s41598-018-31371-z.

[124] V. I. Oladosu, S. Park, and K. Sauer, "Flip the switch: the role of FleQ in modulating the transition between the free-living and sessile mode of growth in Pseudomonas aeruginosa," *J. Bacteriol.*, vol. 206, no. 3, Mar. 2024, doi: 10.1128/jb.00365-23.

[125] E. H. Junkermeier and R. Hengge, "Local signaling enhances output specificity of bacterial c-di-GMP signaling networks," *microLife*, vol. 4, Jan. 2023, doi: 10.1093/femsml/uqad026.

[126] "Escherichia coli DosC and DosP: a role of c-di-GMP in compartmentalized sensing by degradosomes," 2019, pp. 53–67. doi: 10.1016/bs.ampbs.2019.05.002.

[127] R. Hengge, A. Gründling, U. Jenal, R. Ryan, and F. Yildiz, "Bacterial Signal Transduction by Cyclic Di-GMP and Other Nucleotide Second Messengers," *J. Bacteriol.*, vol. 198, no. 1, pp. 15–26, Jan. 2016, doi: 10.1128/JB.00331-15.

[128] S. Spanò, X. Gao, S. Hannemann, M. Lara-Tejero, and J. E. Galán, "A Bacterial Pathogen Targets a Host Rab-Family GTPase Defense Pathway with a GAP," *Cell Host Microbe*, vol. 19, no. 2, pp. 216–226, Feb. 2016, doi: 10.1016/j.chom.2016.01.004.

[129] A. Bateman *et al.*, "UniProt: the Universal Protein Knowledgebase in 2025," *Nucleic Acids Res.*, vol. 53, no. D1, pp. D609–D617, Jan. 2025, doi: 10.1093/nar/gkae1010.

[130] E. Gasteiger *et al.*, "The Proteomics Protocols Handbook," *Proteomics Protoc. Handb.*, pp. 571–608, 2005, doi: 10.1385/1592598900.

[131] K. Imai *et al.*, "SOSUI-GramN: high performance prediction for sub-cellular localization of proteins in Gram-negative bacteria," *Bioinformation*, vol. 2, no. 9, pp. 417–421, 2008, doi: 10.6026/97320630002417.

[132] C. Geourjon and G. Deléage, "SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments," *Bioinformatics*, vol. 11, no. 6, pp. 681–684, 1995, doi: 10.1093/bioinformatics/11.6.681.

[133] D. T. Jones, "Protein secondary structure prediction based on position-specific scoring matrices 1 1Edited by G. Von Heijne," *J. Mol. Biol.*, vol. 292, no. 2, pp. 195–202, Sep. 1999, doi: 10.1006/jmbi.1999.3091.

[134] J. Jumper *et al.*, "Highly accurate protein structure prediction with AlphaFold," *Nature*, vol. 596, no. 7873, pp. 583–589, Aug. 2021, doi: 10.1038/s41586-021-03819-2.

[135] W. Zheng, C. Zhang, Y. Li, R. Pearce, E. W. Bell, and Y. Zhang, "Folding non-homologous proteins by coupling deep-learning contact maps with I-TASSER assembly simulations," *Cell Reports Methods*, vol. 1, no. 3, p. 100014, Jul. 2021, doi: 10.1016/j.crmeth.2021.100014.

[136] E. Krieger, G. Koraimann, and G. Vriend, "Increasing the precision of comparative models with YASARA NOVA—a self-parameterizing force field," *Proteins Struct. Funct. Bioinforma.*, vol. 47, no. 3, pp. 393–402, May 2002, doi: 10.1002/prot.10104.

[137] C. J. Williams *et al.*, “MolProbity: More and better reference data for improved all-atom structure validation,” *Protein Sci.*, vol. 27, no. 1, pp. 293–315, Jan. 2018, doi: 10.1002/pro.3330.

[138] A. Roy, J. Yang, and Y. Zhang, “COFACTOR: an accurate comparative algorithm for structure-based protein function annotation,” *Nucleic Acids Res.*, vol. 40, no. W1, pp. W471–W477, Jul. 2012, doi: 10.1093/nar/gks372.

[139] L. Holm, “DALI and the persistence of protein shape,” *Protein Sci.*, vol. 29, no. 1, pp. 128–140, Jan. 2020, doi: 10.1002/pro.3749.

[140] Y. Zhang, “TM-align: a protein structure alignment algorithm based on the TM-score,” *Nucleic Acids Res.*, vol. 33, no. 7, pp. 2302–2309, Apr. 2005, doi: 10.1093/nar/gki524.

[141] L. Schrödinger, “The PyMOL Molecular Graphics System,” 2026. [Online]. Available: https://www.pymol.org

[142] E. Jurrus *et al.*, “Improvements to the <scp>APBS</scp> biomolecular solvation software suite,” *Protein Sci.*, vol. 27, no. 1, pp. 112–128, Jan. 2018, doi: 10.1002/pro.3280.

[143] S. R. Eddy, “Accelerated Profile HMM Searches,” *PLoS Comput. Biol.*, vol. 7, no. 10, p. e1002195, Oct. 2011, doi: 10.1371/journal.pcbi.1002195.

[144] R. C. Edgar, “MUSCLE: multiple sequence alignment with high accuracy and high throughput,” *Nucleic Acids Res.*, vol. 32, no. 5, pp. 1792–1797, Mar. 2004, doi: 10.1093/nar/gkh340.

[145] H. Ashkenazy *et al.*, “ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules,” *Nucleic Acids Res.*, vol. 44, no. W1, pp. W344–W350, Jul. 2016, doi: 10.1093/nar/gkw408.

[146] G. M. Morris *et al.*, “AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility,” *J. Comput. Chem.*, vol. 30, no. 16, pp. 2785–2791, Dec. 2009, doi: 10.1002/jcc.21256.

[147] O. Trott and A. J. Olson, “AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading,” *J. Comput. Chem.*, vol. 31, no. 2, pp. 455–461, Jan. 2010, doi: 10.1002/jcc.21334.

[148] D. Szklarczyk *et al.*, “The STRING database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets,” *Nucleic Acids Res.*, vol. 49, no. D1, pp. D605–D612, Jan. 2021, doi: 10.1093/nar/gkaa1074.

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, New Delhi, 110042

PLAGIARISM VERIFICATION

Title of the Thesis **"The p53/SOX2 axis in modulating the cell fate in neurons: A paradigm of neurodegeneration and brain tumours"** Total Pages **66** Name of the Scholar **Gaurav (24/MSCBIO/01).**

Supervisor

Prof. Yasha Hasija

Department of Biotechnology

This is to report that the above thesis was scanned for similarity detection. Process and outcome is given below:

Software used: **Turnitin,** Similarity Index: **4%**, Total Word Count: **19,259**

Date: ____________

Candidate's Signature **Signature of Supervisor**

Plagiarism report

Software used: Turnitin

Gaurav Thesis

ORIGINALITY REPORT

10%	9%	7%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

#	Source	Type	%
1	Submitted to Delhi Technological University	Student Paper	1%
2	www.mdpi.com	Internet Source	1%
3	dtu.ac.in	Internet Source	1%
4	dspace.dtu.ac.in:8080	Internet Source	<1%
5	"Microbial Cyclic Di-Nucleotide Signaling", Springer Science and Business Media LLC, 2020	Publication	<1%
6	new.esp.org	Internet Source	<1%
7	docsplayer.org	Internet Source	<1%
8	www.tandfonline.com	Internet Source	<1%

9	Anjaneyulu Musini, Narasaiah Kolliputi, Vinod Kumar Yata. "Staphylococcus aureus - Pathogenesis, Drug Resistance, and Innovative Therapeutics", CRC Press, 2026 Publication	<1%
10	www.researchgate.net Internet Source	<1%
11	spiral.imperial.ac.uk Internet Source	<1%
12	academic.oup.com Internet Source	<1%
13	idr.l1.nitk.ac.in Internet Source	<1%
14	tudr.thapar.edu:8080 Internet Source	<1%
15	www.frontiersin.org Internet Source	<1%
16	www.biorxiv.org Internet Source	<1%
17	www.esp.org Internet Source	<1%
18	Jingyuan Fu, Yuning Zhang, Shiyu Lin, Wei Zhang et al. "Strategies for Interfering With Bacterial Early Stage Biofilms", Frontiers in Microbiology, 2021	<1%

Publication

19	Erez Mills. "The bacterial second messenger c-di-GMP: mechanisms of signalling : c-di-GMP signalling", Cellular Microbiology, 08/2011 Publication	<1 %
20	Submitted to Rajiv Gandhi National University of Law Student Paper	<1 %
21	ir.juit.ac.in:8080 Internet Source	<1 %
22	Submitted to University of Reading Student Paper	<1 %
23	Wei Zheng, Yang Li, Chengxin Zhang, Xiaogen Zhou, Robin Pearce, Eric W. Bell, Xiaoqiang Huang, Yang Zhang. " Protein structure prediction using deep learning distance and hydrogen-bonding restraints in ", Proteins: Structure, Function, and Bioinformatics, 2021 Publication	<1 %
24	C. Albert-Weissenberger, T. Sahr, O. Sismeiro, J. Hacker, K. Heuner, C. Buchrieser. "Control of Flagellar Gene Regulation in Legionella pneumophila and Its Relation to Growth Phase", Journal of Bacteriology, 2009 Publication	<1 %

25	Carly A. Shanahan, Barbara L. Gaffney, Roger A. Jones, Scott A. Strobel. "Differential Analogue Binding by Two Classes of c-di-GMP Riboswitches", Journal of the American Chemical Society, 2011 Publication	<1 %
26	bmcbioinformatics.biomedcentral.com Internet Source	<1 %
27	www.wiaawi.org Internet Source	<1 %
28	Shuxun Liu, Xujie Feng, Hangjia Zhang, Ping Li, Baoru Yang, Qing Gu. "Decoding bacterial communication: Intracellular signal transduction, quorum sensing, and cross-kingdom interactions", Microbiological Research, 2025 Publication	<1 %
29	file.scirp.org Internet Source	<1 %
30	institut-agro-dijon.hal.science Internet Source	<1 %
31	jgeb.springeropen.com Internet Source	<1 %
32	Abhishek Samuel Krishnamoorthy, Bhavya S. G., Rajkuberan Chandrasekaran, Selvakumari Palani et al. "Computational analysis of LasR	<1 %

inhibitors: A novel approach to disrupt quorum sensing in Pseudomonas aeruginosa", In Silico Research in Biomedicine, 2026
Publication

33 Jiri Neuzil. "Affinity of vitamin E analogues for the ubiquinone complex II site correlates with their toxicity to cancer cells", Molecular Nutrition & Food Research, 10/2011 — <1 %
Publication

34 Mutaz Mohammed Abdallah, Ruaa Abdalla Ibrahim Suliman, Yousra Tagelsir Ahmed, Mawada Yahia. "In silico analysis of hypothetical proteins in Pseudomonas aeruginosa PAC1: Structural and functional insights", Journal of Genetic Engineering and Biotechnology, 2025 — <1 %
Publication

35 Ramiro Vílchez, André Lemme, Britta Ballhausen, Verena Thiel et al. "Streptococcus mutans Inhibits Candida albicans Hyphal Formation by the Fatty Acid Signaling Molecule trans-2-Decenoic Acid (SDSF)", ChemBioChem, 2010 — <1 %
Publication

36 Weihan Gu, Yufan Chen, Congcong Xie, Zhongqiao Chen, Huagui Gao, Yaping Zhu, Lian-hui Zhang, Lisheng Liao. " The cyclic di- — <1 %

GMP receptor YcgR links the second messenger with the putrescine quorum sensing system in modulation of motility ", mBio, 2025

Publication

---

37 Zhuo Wang, Li Song, Xiaozhen Liu, Xihui Shen, Xin Li. "Bacterial second messenger c-di-GMP: emerging functions in stress resistance", Microbiological Research, 2023 — <1 %

Publication

---

38 blog.csdn.net — <1 %

Internet Source

---

39 docusalut.com — <1 %

Internet Source

---

40 rke.abertay.ac.uk — <1 %

Internet Source

---

41 www.hznu.edu.cn — <1 %

Internet Source

---

42 Guercio, Danielle Rose. "Characterizing the Role of the Histidine Kinase NahK in Pseudomonas Aeruginosa.", State University of New York at Stony Brook — <1 %

Publication

---

43 Hanne Hendrix, Annabel Itterbeek, Hannelore Longin, Lize Delanghe et al. "PlzR regulates type IV pili assembly in Pseudomonas — <1 %

aeruginosa via PilZ binding", Nature Communications, 2024
Publication

---

44 Justin L. Burns, Parth B. Jariwala, Shannon Rivera, Benjamin M. Fontaine, Laura Briggs, Emily E. Weinert. " Oxygen-Dependent Globin Coupled Sensor Signaling Modulates Motility and Virulence of the Plant Pathogen ", ACS Chemical Biology, 2017 <1%
Publication

---

45 Qishun Feng, Jianuan Zhou, Lianhui Zhang, Yang Fu, Liang Yang. " Insights into the molecular basis of c-di-GMP signalling in ", Critical Reviews in Microbiology, 2022 <1%
Publication

---

46 Sazzad Hossain, Anisa Jannat. "In Silico Characterization of a Hypothetical Protein from Pseudomonas aeruginosa LESB58: A Structural and Functional Perspective", In Silico Research in Biomedicine, 2026 <1%
Publication

---

47 Veronika Zdarska, Milan Kolar, Patrik Mlynarcik. "Occurrence of beta-lactamases in bacteria", Infection, Genetics and Evolution, 2024 <1%
Publication

48	Viola Camilla Scoffone, Giulia Barbieri, Samuele Irudal, Gabriele Trespidi, Silvia Buroni. "New Antimicrobial Strategies to Treat Multi-Drug Resistant Infections Caused by Gram-Negatives in Cystic Fibrosis", Antibiotics, 2024 Publication	<1%
49	biblio.ugent.be Internet Source	<1%
50	kth.diva-portal.org Internet Source	<1%
51	mafiadoc.com Internet Source	<1%
52	mdpi-res.com Internet Source	<1%
53	Kristin N. Kovach, Derek Fleming, Marilyn J. Wells, Kendra P. Rumbaugh, Vernita Diane Gordon. " Specific Disruption of Established Biofilms Using Polymer-Attacking Enzymes ", Langmuir, 2020 Publication	<1%
54	Lian Zhou, Lian-Hui Zhang, Miguel Cámara, Ya-Wen He. "The DSF Family of Quorum Sensing Signals: Diversity, Biosynthesis, and Turnover", Trends in Microbiology, 2017 Publication	<1%

55 Suchetana Bhattacharyya, Sayanti Karmakar, Debarupa Hajra, Santanu Paul. "Integrated computational biology and network pharmacology predict anti-cancer potential of the novel pteridophyte, Leucostegia immersa C. Presl, against non-small cell lung cancer", In Silico Research in Biomedicine, 2026 — <1%
Publication

56 archives.univ-biskra.dz — <1%
Internet Source

57 bioinfo.nchc.org.tw — <1%
Internet Source

58 dspace.iiti.ac.in — <1%
Internet Source

59 hal.science — <1%
Internet Source

60 intor.torlakinstitut.com — <1%
Internet Source

61 japsonline.com — <1%
Internet Source

62 journals.sagepub.com — <1%
Internet Source

63 pmc.ncbi.nlm.nih.gov — <1%
Internet Source

prism.ucalgary.ca

64	Internet Source	<1%
65	www.research.unipd.it Internet Source	<1%
66	www.scirp.org Internet Source	<1%
67	www2.mdpi.com Internet Source	<1%

Exclude quotes	On	Exclude matches	< 8 words
Exclude bibliography	On		