Jaypee University of Information Technology
Solan (H.P.)
LEARNING RESOURCE CENTER

Acc. Num. SP04096 Call Num:

General Guidelines:

- Library books should be used with great care.
- Tearing, folding, cutting of library books or making any marks on them is not permitted and shall lead to disciplinary action.
- Any defect noticed at the time of borrowing books must be brought to the library staff immediately. Otherwise the borrower may be required to replace the book by a new copy.
- The loss of LRC book(s) must be immediately brought to the notice of the Librarian in writing.

Learning Resource Centre-JUIT

SP04096

IDENTIFICATION OF POTENTIAL DRUG BINDING SITES IN PATHOGENIC *Pseudomonas aeruginosa*

By

SABIHA -041521
SADHANA SURANA – 041529

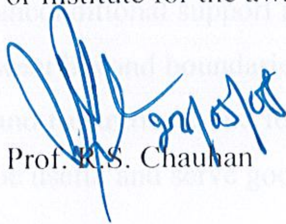
May 2008

Submitted in partial fulfillment of the Degree of Bachelors of Technology

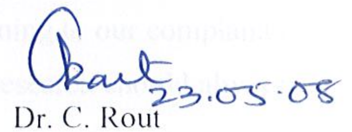
DEPARTMENT OF BIOTECHNOLOGY AND BIOINFORMATICS
JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY-WAKNAGHAT
MAY-2008

CERTIFICATE

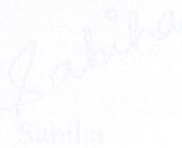
This is to certify that the work entitled, "Identification of potential drug binding sites in pathogenic *Pseudomonas* aeruginosa" submitted by Sabiha and Sadhana Surana, in partial fulfillment for the award of degree of Bachelors of Technology in Bioinformatics of Jaypee University of Information Technology has been carried out under our supervision. This work has not been submitted partially or wholly to any other University or Institute for the award of this or any other degree or diploma.

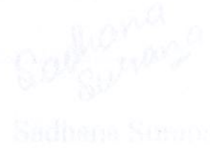


Prof. R.S. Chauhan



Dr. C. Rout





ACKNOWLEDGEMENT

We wish to express our sincere gratitude to Prof. R. S. Chauhan and Dr. C. Rout for providing us invaluable guidance and suggestions. They were always there to listen and to give advice.

They taught us how to ask questions and express our ideas and showed us different ways to approach a research problem and the need to be persistent to accomplish any goal.

We would also like to thank all the staff members of Bioinformatics Department of Jaypee University Of Information Technology, Solan, for providing us all the facilities required for the completion of this thesis.

We wish to thank all our classmates and friends for their timely suggestions and cooperation during the period of our thesis.

Last, but not least, we thank our families for educating us with aspects from sciences, for unconditional support and encouragement to pursue our interests, even when the interests went beyond boundaries of language, field and geography, for listening to our complaints and frustrations, and for believing in us, for reminding us that our research should always be useful and serve good purposes for all humankind.

Sabiha

Sadhana Surana

TABLE OF CONTENTS

CERTIFICATE	2
ACKNOWLEDGEMENT	3
LIST OF ABBREVIATIONS	6
LIST OF FIGURES	7
ABSTRACT	8
1. BACKGROUND	9
1.1 BENEFICIAL BACTERIA	9
1.2 PATHOGENIC BACTERIA	9
2. INTRODUCTION	11
2.1 *Pseudomonas* SPECIES	11
2.2 *Pseudomonas aeruginosa*	12
2.2.1 CHARACTERISTICS	12
2.2.2 PATHOGENESIS	13
2.2.3 INVASION	13
2.2.4 DISSEMINATION	14
2.2.5 TOXINOGENESIS	14
2.3 TWO STRAINS OF *Pseudomonas aeruginosa*	16
2.4 COMMON PROTEIN IDENTIFICATION IN TWO STRAINS	16
2.5 POTENTIAL TARGETS	16
3. METHODOLOGY	18
3.1 COMMON PROTEINS USING BLAST	18
3.2 FUNCTIONAL CHARACTERIZATION OF PROTEINS	18
3.3 STRUCTURE	18
3.3.1 MODELING THE STRUCTURE	18
3.4 STRUCTURE DRAWING	18
3.5 MINIMIZATION OF STRUCTURE	19
3.6 DOCKING	19
3.6.1 SCHRODINGER	19
3.6.1.1 GLIDE	19

3.6.2 FT DOCK 19

4. EXOTOXIN 22

4.1 EXOTOXIN A PRECURSOR 22

4.2 ACTION OF EXOTOXIN A 23

4.3 ACTIVATION OF EXOTOXIN A 23

5. EXOTOXIN A DOMAIN I 25

5.1 THE RECEPTOR 25

5.2 STRUCTURES 26

5.2.1 MODELING 26

5.2.2 VALIDATION 27

5.3 DOCKING 27

6. EXOTOXIN A DOMAIN III 28

6.1 ABOUT DOMAIN III 28

6.2 LIGPLOT OF NATURAL LIGANDS 29

6.3 NAD SRTUCTURE 29

6.3.1 STRUCTURE DRAWING 29

6.3.2 MINIMIZATION OF NAD STRUCTURE 30

6.4 DOCKING 30

7. RESULTS 31

8. CONCLUSION 48

BIBLIOGRAPHY 49

LIST OF ABBREVIATIONS

PA	- *Pseudomonas aeruginosa*
ADP	- Adenosine diphosphate
NAD	- Nicotinamide adenine dinucleotide
LRP	- Low density Lipoprotein Receptor-related Protein
SAVS	- Structure Analysis and Verification Server
PDB	- Protein Data Bank
NCBI	- National Center for Biotechnology Information
KEGG	- Kyoto Encyclopedia for Genes and Genomes
COG	- Cluster of Orthologus Groups
vdW	- van der Waals interaction energy
Coul	- Coulomb interaction energy
Lipo	- Lipophilic-contact plus phobic-attractive term
HBond	- Hydrogen-bonding term
Metal	- Metal-binding term (usually a reward)
Rewards	- various reward or penalty terms
RotB	- Penalty for freezing rotatable bonds
Site	- Polar interactions in the active site

LIST OF FIGURES

FIGURE: 1 Flow diagram of docking method.

FIGURE: 2 Ligplot of AMP complexed with 1DMA (Domain III) specifying the interacting residue in binding site with the ligand.

FIGURE: 3 Ligplot of NCA complexed with chain A of 1DMA (Domain III) specifying the interacting residue in binding site with the ligand.

FIGURE: 4 Ligplot of NCA complexed with chain B of 1DMA (Domain III) specifying the interacting residue in binding site with the ligand.

FIGURE: 5 Energy minimized structure of NAD through HyperChem.

FIGURE: 6 Docked structure of domain III with 16''' tautamer of NAD generated during ligand preparation showing the potential drug binding site of the protein.

ABSTRACT

Usually the drugs that are used as a cure for several diseases and disorders have side effects. The side effects might not occur only in the human body but also in surrounding environment. The drugs when excreted out into the environment harm the plants and the animals thereby creating an imbalance in the ecosystem.

The project aims to be able to give potential drug binding sites for the pathogenic *Pseudomonas aeruginosa* for which stable cure are not available. The idea is to target those proteins by drug molecules which should not produce any side effects to humans as well as beneficial microbes in the environment.

For this purpose, the proteins should be identified uniquely in each *Pseudomonas* species. Comparative genomics approaches were carried out to get drug targets *Pseudomonas aeruginosa* (protein common across all pathogenic *Pseudomonas* strains and not present in human and beneficial microbes). Exotoxin A precursor was considered as potential drug target as it is involved in protein inhibition which leads to cell death in host. It consists of three domains: Domain I is involved in cell surface receptor recognition; domain III is involved inhibition of protein synthesis; and Domain II is not significantly contributing to pathogenecity.

In this report, the binding site prediction of domain I of exotoxin A is determined by docking it to LRP receptors which is under progress as protein-protein docking takes significant amount of time. The protein-protein docking software FT-DOCK is used for docking. Similarly, the binding site of domain III is established by docking it to NAD by the use of GLIDE module of Schrodinger Suite.

1. BACKGROUND

Where there are certain microorganisms, harmful to humans, there are others which are beneficial too. Some of the beneficial and the pathogenic bacteria were identified based on the effect they cause. The bacteria included:-

1.1 Beneficial bacteria

1. *Bacillus subtilis*
2. *Escherichia coli*
3. *Streptomyces coelicolor*
4. *Pseudomonas fluorescens*
5. *Bradyrhizobium japonicum*
6. *Rhodopseudomonas palustris*
7. *Bacillus clausii*
8. *Lactobacillus acidophilus*
9. *Rhodopseudomonas palustris*
10. *Bacillus licheniformis*
11. *Bacillus halodurans*

1.2 Pathogenic bacteria

1. *Pseudomonas aeruginosa*
2. *Bacillus thuringiensis*
3. *Pseudomonas syringe*

A comparison was made between the beneficial and the pathogenic bacteria, by performing BLAST. Also for the pathogenic bacteria BLAST was performed against the human genome. This was to identify the uncommon proteins, which would only be present in the pathogenic bacteria and not in the beneficial bacteria or the human beings. The idea was to then identify drug binding sites against the disease causing microorganisms, so that only proteins present in harmful bacteria suffer the effects of drugs and no harm is caused to the beneficial bacteria.

BLAST score was obtained for all the beneficial and harmful bacteria. Proteins with low identity or no identity were thus obtained (score less than 50). These proteins were present only in the pathogenic bacteria having significant variation in their amino acid composition and three dimensional structures, against the beneficial bacteria.

After examining the harmful proteins, it was observed that the pathogenicity of *Pseudomonas aeruginosa* is quite significant. Also it was found that not many antibiotics that are available are effective against the diseases caused by *P. aeruginosa*. Thus we decided to work on *P. aeruginosa*. We would identify the potential targets (proteins) in the genome.

2. INTRODUCTION

2.1 Pseudomonas Species

Pseudomonas may be described as Gram-negative, non-spore forming, straight or slightly curved rods. They are typically motile by means of one or more polar flagella. These basic morphological characteristics, however, are common to many families of bacteria and so are of little value in the positive identification or diagnosis of a member of the genus *Pseudomonas*.

Generally, common to all constituent species of the genus *Pseudomonas* are certain physiological properties such as chemoorganotrophic nutrition, aerobic metabolism, absence of fermentation, absence of photosynthesis, inability to fix nitrogen, and capacity for growth at the expense of a large variety of organic substrates.

Pseudomonas and certain other pseudomonads include species pathogenic for humans, domestic animals, and cultivated plants. *Pseudomonas* species, as well as species included in the newly-created genera *Burkholderia* and *Ralstonia* (ex-*Pseudomonas*) are among the most important bacteria that are pathogens of plants. They cause economically significant crop disease and crop loss world-wide.

The Genus *Pseudomonas*

1. Gram-negative
2. Rod-shaped, 0.5-0.8 um x 1-3 um
3. Strictly aerobic; the only anaerobic activities may be denitrification and arginine degradation to ornithine
4. Motile by polar flagella; some strains also produce lateral flagella
5. Oxidative, chemoorganotrophic metabolism
6. Catalase-positive
7. Usually oxidase-positive
8. No organic growth factors are required
9. Diffusible and/or insoluble pigments may be produced
10. GC content of the DNA: 58-68 mol%

2.2 *Pseudomonas aeruginosa*

Pseudomonas aeruginosa is a Gram-negative, aerobic rod belonging to the bacterial family Pseudomonadaceae. These bacteria are common inhabitants of soil and water.

It is an opportunistic pathogen. These bacteria are common inhabitants of soil and water. They occur regularly on the surfaces of plants and occasionally on the surfaces of animals. *Pseudomonas aeruginosa* is an opportunistic pathogen, meaning that it exploits some break in the host defenses to initiate an infection. It is primarily a nosocomial pathogen.

It causes:-

1. Urinary tract infections
2. Respiratory system infections
3. Dermatitis
4. Soft tissue infections
5. Bacteremia
6. Bone and joint infections
7. Gastrointestinal infections

A variety of systemic infections, particularly in patients with severe burns and in cancer and AIDS patients who are immuno suppressed.

Pseudomonas aeruginosa infection is a serious problem in patients hospitalized with cancer, cystic fibrosis, and burns. The case fatality rate in these patients is 50 percent.

2.2.1 *Characteristics*

Pseudomonas aeruginosa is a

1. Gram-negative rod measuring 0.5 to 0.8 μm by 1.5 to 3.0 μm.
2. Almost all strains are motile by means of a single polar flagellum.
3. The bacterium is ubiquitous in soil and water, and on surfaces in contact with soil or water.
4. Its metabolism is respiratory and never fermentative, but it will grow in the absence of O2 if NO3 is available as a respiratory electron acceptor.
5. Its optimum temperature for growth is 37 degrees, and it is able to grow at temperatures as high as 42 degrees.

6. It is tolerant to a wide variety of physical conditions, including temperature. It is resistant to high concentrations of salts and dyes, weak antiseptics, and many commonly used antibiotics.

2.2.2 *Pathogenesis*

For an opportunistic pathogen such as *Pseudomonas aeruginosa*, the disease process begins with some alteration or circumvention of normal host defenses.

The pathogenesis of *Pseudomonas* infections is multifactorial, as suggested by the number and wide array of virulence determinants possessed by the bacterium.

Multiple and diverse determinants of virulence are expected in the wide range of diseases caused, which include septicemia, urinary tract infections, pneumonia, chronic lung infections, endocarditis, dermatitis, and osteochondritis.

Most *Pseudomonas* infections are both invasive and toxinogenic. The ultimate *Pseudomonas* infection may be seen as composed of three distinct stages:

1. Bacterial attachment and colonization
2. Local invasion
3. Disseminated systemic disease.

However, the disease process may stop at any stage. Particular bacterial determinants of virulence mediate each of these stages and are ultimately responsible for the characteristic syndromes that accompany the disease.

2.2.3 *Invasion*

The ability of *Pseudomonas aeruginosa* to invade tissues depends upon production of extracellular enzymes and toxins that break down physical barriers and damage host cells, as well as resistance to phagocytosis and the host immune defenses. As mentioned above, the bacterial capsule or slime layer effectively protect cells from opsonization by antibodies; complement deposition, and phagocyte engulfment.

Virulence Determinants of Pathogenic *Pseudomonas aeruginosa*

Adhesins	Fimbriae (N-methyl-phenylalanine pili) Polysaccharide capsule (glycocalyx) Alginate slime (biofilm)
Invasions	Elastase Alkaline protease Hemolysins (phospholipase and lecithinase) Cytotoxin (leukocidin)
Motility/Chemotaxis	Flagella Exoenzyme S Exotoxin A
Defense against immune responses	Capsules Slime layers Protease enzymes

2.2.4 ***Dissemination***

Blood stream invasion and dissemination of *Pseudomonas* from local sites of infection is probably mediated by the cell-associated and extra cellular products responsible for the localized disease, although it is not entirely clear how the bacterium produces systemic illness. It is that *Pseudomonas* Exotoxin A exerts some pathologic activity during the dissemination stage.

2.2.5 ***Toxinogenesis***

P. aeruginosa produces two extracellular protein toxins, Exoenzyme S and Exotoxin A. Exoenzyme S is probably an exotoxin. It has the characteristic subunit structure of the A-component of a bacterial toxin, and it has ADP-ribosylating activity (for a variety of eukaryotic proteins) characteristic of exotoxins.

Exoenzyme S is produced by bacteria growing in burned tissue and may be detected in the blood before the bacteria are. Exotoxin A has exactly the same mechanism of action as the diphtheria toxin; it causes the ADP ribosylation of eukaryotic elongation factor 2. It is partially-identical to diphtheria toxin, but it is antigenically-distinct. It utilizes a different receptor on host cells, but otherwise it enters cells in the same manner as the diphtheria toxin and it has the exact enzymatic mechanism. The production of Exotoxin

A in is regulated by exogenous iron, but the details of the regulatory process are distinctly different in *C. diphtheriae* and *P. aeruginosa*.

Exotoxin A appears to mediate both local and systemic disease processes caused by *Pseudomonas aeruginosa*. It has necrotizing activity at the site of bacterial colonization and is thereby thought to contribute to the colonization process. Toxinogenic strains cause a more virulent form of pneumonia than nontoxinogenic strains. In terms of its systemic role in virulence, purified Exotoxin A is highly lethal for animals including primates. Indirect evidence involving the role of exotoxin A in disease is seen in the increased chance of survival in patients with *Pseudomonas* septicemia that is correlated with the titer of anti-exotoxin A antibodies in the serum

Diseases Caused by *Pseudomonas aeruginosa*

1. Endocarditic - *Pseudomonas aeruginosa* infects heart valves of IV drug users and prosthetic heart valves. The organism establishes itself on the endocardium by direct invasion from the blood stream.

2. Respiratory infections - Respiratory infections caused by *Pseudomonas aeruginosa* occur almost exclusively in individuals with a compromised lower respiratory tract or a compromised systemic defense mechanism.

3. Bacteremia and Septicemia - *Pseudomonas aeruginosa* causes bacteremia primarily in immunocom promised patients. Predisposing conditions include hematologic malignancies, immunodeficiency relating to AIDS, neutropenia, diabetes mellitus, and severe burns.

4. Central Nervous System infections - *Pseudomonas aeruginosa* causes meningitis and brain abscesses. The organism invades the CNS from a contiguous structure such as the inner ear or paranasal sinus.

5. Ear infections including external otitis - *Pseudomonas aeruginosa* is the predominant bacterial pathogen in some cases of external otitis including "swimmer's ear". The bacterium is infrequently found in the normal ear, but often inhabits the external auditory canal in association with injury, maceration, inflammation, or simply wet and humid conditions.

6. Eye infections - *Pseudomonas aeruginosa* can cause devastating infections in the human eye. *Pseudomonas* can colonize the ocular epithelium by means of a fimbrial

attachment to sialic acid receptors. If the defenses of the environment are compromised in any way the bacterium can proliferate rapidly and, through the production of enzymes such as elastase, alkaline protease and exotoxin A.

7. Bone and joint infections - *Pseudomonas* infections of bones and joints result from direct inoculation of the bacteria or the hematogenous spread of the bacteria from other primary sites of infection.

8. Urinary tract infections - Urinary tract infections (UTI) caused by *Pseudomonas aeruginosa* are usually hospital-acquired and related to urinary tract catheterization, instrumentation or surgery.

9. Gastrointestinal infections - *Pseudomonas aeruginosa* can produce disease in any part of the gastrointestinal tract from the oropharynx to the rectum. As in other forms of *Pseudomonas* disease, those involving the GI tract occur primarily in immunocompromised individuals.

2.3 Two strains of Pseudomonas aeruginosa were considered

The following 2 strains were considered

1. *P. aeruginosa* 2516(NP_002516) – PAO1
2. *P. aeruginosa* 8463(NP_008463) - UCBPP-PA14

2.4 Common protein identification in two strains

We saw that both the strains have different proteins present in them. 72 proteins of PAO1 and 99 proteins of UCBPP-PA14 were identified. We then found out the common proteins between both the strains, which came out to be 21 in number (Table 3).

2.5 Potential Targets

Identified 6 potential drug targets, proteins which are involved in secretion pathway (Table 4) and as we know that in *P. aeruginosa*, toxins are responsible for causing disease. So have chosen secretary pathway protein as our target for which we will be designing lead molecules. Having the accession IDs of the proteins, we obtained the structures for these proteins from the Protein Data Bank (PDB). Now we are working on the 6th table (Table 5) i.e. Exotoxin A precursor (NAD-dependent ADP

ribosyltransferase) as we find it the most potential because responsible for the, catalysis of the transfer of the ADP ribosyl moiety of oxidized NAD onto elongation factor 2 (EF-2) thus arresting protein synthesis. Out of all the proteins that were common amongst the two strains, it was found that the toxin producing protein Exotoxin A precursor was present. As has been mentioned earlier, exotoxin is the major virulence factor involved in pseudomonas aeruginosa infection causing ability in most of the cases. So we chose exotoxin A precursor to be the target protein, whose action on the host cell is required to be controlled.

3. METHODOLOGY

3.1 Common Proteins using BLAST

Both the strains were BLASTed against all the other beneficial microorganisms, to identify their unique proteins (Table 2). In order to obtain the common proteins between both the strains, each protein from one strain was BLASTed against all the proteins of the other strain. For a BLAST score of 98% similarity was considered between the proteins.

3.2 Functional Characterization of Proteins

After identification of common proteins, these proteins were classified into classes, based on their function in the cell. Several resources such as NCBI, Swissprot, KEGG and COG were used

3.3 Structure

PDB was searched for the protein structure of the Exotoxin A and its receptor LRP. While structure for Exotoxin A was available, there was no structure for LRP.

3.3.1 Modeling the structure

LRP structure was modeled using modeler. For this a template was obtained after performing PDB BLAST. The structures were then validated using SAVS. Low results showed that with this template modeler did not give optimum structures. Hence Threader was used. It is used for fold recognition. It aligns the given sequence to the database of sequences so as to identify folds and generates a score. It gave a better template and again modeler was used to generate structure.

3.4 Structure drawing

ISIS draw is used to draw the structure of NAD. It provides various options to draw any molecule easily and efficiently. This is the artificial ligand to be docked in the protein.

3.5 Minimization of structure

The structure is minimized for optimization of energy using Hyper Chem by geometry optimization and molecular dynamics, using both semi-empirical quantum mechanics and molecular mechanics methods.

3.6 Docking

3.6.1 Schrodinger

Schrödinger was used for the docking of Domain III with NAD to study the binding site and the interaction of ligand with domain III. For this purpose Glide was used.

3.6.1.1 Glide

Glide is designed to assist you in high-throughput screening of potential ligands based on binding mode and affinity for a given receptor molecule. You can compare ligand scores with those of other test ligands, or compare ligand geometries with those of a reference ligand. Additionally, you can use Glide to generate one or more plausible binding modes for a newly designed ligand. Once you locate favorable structures or bonding conformations with Glide, you can use Liaison™ or QSite™ to obtain binding energies for ligand-receptor pairs.

Impact is a molecular mechanics and dynamics program that provides the molecular mechanics component of Glide calculations. Some of the basic capabilities of Impact are also available from Maestro. These capabilities include molecular mechanics (MM) energy minimization, molecular dynamics (MD) simulations, hybrid Monte Carlo (HMC) simulations, and addition of explicit water solvent to a structure.

Protein Preparation is also required for Glide calculations. It can be performed for most protein and protein-ligand complex PDB structures using the Protein Preparation panel in Maestro. Command-line utilities complete the protein preparation facility.

3.6.2 FT Dock

The receptor and the ligand were docked using software called FT Dock. It first of all preprocesses both the ligand and the receptor molecule. Then performs the global scan of translational and rotational space for possible positions interactions in two molecules.

limited by surface complementarity and an electrostatic filter. Then an empirical scoring of the possible complexes is generated using residue level pair potentials. If any biological information is available then the possible complexes are screened on its basis. Finally step performs energy minimization and removal of steric clashes on the side-chains of interface.

Learning Resource Center
Acc. No. SP04096
JUIT, Waknaghat, Solan (H.P)

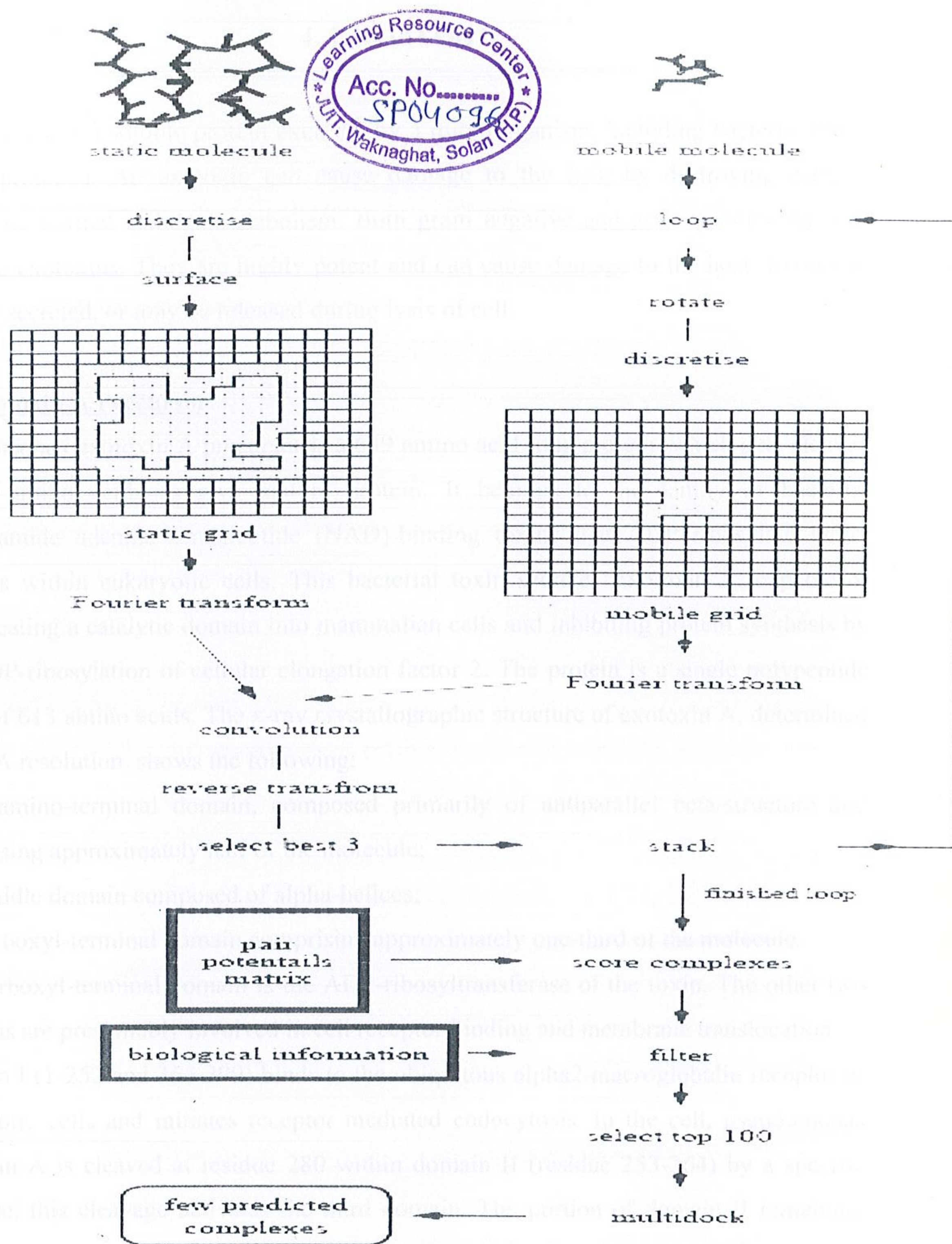


Figure: 1 Flow diagram of docking method used in FT Dock

4. EXOTOXIN

An exotoxin is a soluble protein excreted by a microorganism, including bacteria, fungi, algae, protozoa. An exotoxin can cause damage to the host by destroying cells or disrupting normal cellular metabolism. Both gram negative and gram positive bacteria produce exotoxins. They are highly potent and can cause damage to the host. Exotoxins may be secreted, or may be released during lysis of cell.

4.1 Exotoxin A Precursor

Pseudomonas exotoxin A precursor is a 639 amino acid long protein, which gets cleaved by 25 amino acids to give mature protein. It belongs to the family of bacterial nicotinamide adenine dinucleotide (NAD)-binding toxins that ADP ribosylate target proteins within eukaryotic cells. This bacterial toxin secretes Exotoxin A, capable of translocating a catalytic domain into mammalian cells and inhibiting protein synthesis by the ADP-ribosylation of cellular elongation factor 2. The protein is a single polypeptide chain of 613 amino acids. The x-ray crystallographic structure of exotoxin A, determined to 3.0-A resolution, shows the following:

1. an amino-terminal domain, composed primarily of antiparallel beta-structure and comprising approximately half of the molecule;

2. a middle domain composed of alpha-helices;

3. a carboxyl-terminal domain comprising approximately one-third of the molecule.

The carboxyl-terminal domain is the ADP-ribosyltransferase of the toxin. The other two domains are presumably involved in cell receptor binding and membrane translocation. Domain I (1-252 and 365-399) binds to the ubiquitous alpha2-macroglobulin receptor of eukaryotic cells and initiates receptor mediated endocytosis. In the cell, pseudomonas exotoxin A is cleaved at residue 280 within domain II (residue 253-364) by a specific protease, this cleavage activates the third domain. The portion of domain II remaining with the C terminal fragment appears to translocate the fragment through intracellular membranes into the cytosol, where domain III (residue 400-613) acts by transferring the ADP ribose from NAD to a modified histidine in elongation factor-2(EF-2). This irreversible covalent modification inhibits protein synthesis and leads to cell death.

4.2 <u>Action of Exotoxin A</u>

Pseudomonas Exotoxin A is synthesized as a single chain bacterial protein composed of three structural domains. The N-terminal domain mediates binding to LRP: specifically, Lys-57 and possibly other nearby residues mediate binding. Toxicity for cells is reduced by at least 100-fold when Lys-57 is changed to glutamic acid or when most or this entire domain is deleted. The middle domain of PE has two functions:

1. It contains sequences necessary for translocation to the cell cytosol
2. It serves as a substrate for cell-mediated cleavage.

The C-terminal domain has the ADP-ribosylating activity and contains a putative endoplasmic reticulum (ER) retention sequence.

Functionally DT is quite similar to PE. However, the binding and ADP-ribosylating domains of DT are located in the opposite orientation to PE, i.e. the binding domain of DT is at the C terminus while the ADP-ribosylating activity is at the N terminus. Like PE, the middle domain of DT contains sequences that mediate translocation to the cell cytosol. While the exact intracellular location for the translocation of either toxin has not been clearly defined, existing data suggests that the A chain of DT reaches the cytosol from an acidic endosomal compartment while the corresponding PE fragment needs to reach the ER to facilitate its translocation. Thus it appears that these functionally similar toxins use different intracellular pathways to reach the same cytosolic location.

4.3 <u>Activation of Exotoxin A</u>

Furin plays an important role in activating the exotoxin A, by proteolysing. This is important for the toxin so as to express the full toxic activity. Proteolysis is also important to release catalytic domain from other domains having receptor binding and translocation functions. The protein encoded by Furin gene belongs to the subtilisin-like proprotein convertase family. The members of this family are proprotein convertases that process latent precursor proteins into their biologically active products. This encoded protein is a calcium-dependent serine endoprotease that can efficiently cleave precursor proteins at their paired basic amino acid processing sites. Furin is enriched in the Golgi apparatus, where it functions to cleave other proteins into their mature/active forms. Furin cleaves proteins just downstream of a basic amino acid target sequence (canonically,

Arg-X-(Arg/Lys) -Arg'). In addition to processing cellular precursor proteins, furin is also utilized by a number of pathogens. Pseudomonas exotoxin must be processed by furin during their initial entry into host cells.

5. EXOTOXIN A DOMAIN I

Exotoxin A domain I is an amino-terminal domain, composed primarily of antiparallel beta-structure and comprising approximately half of the molecule. It is composed of two parts:-

1. Domain Ia
2. Domain Ib

The domain Ia ranges from amino acid 1 to 252. The structure is then protruded by the domain II which ranges upto amino acid 364. Then comes the domain Ib from amino acid 365 to 399.

The domain Ia is involved in receptor recognition and receptor binding. The receptor for binding of domain Ia is low density lipoprotein receptor-related protein (LRP). The function of domain Ib is unknown and it can entirely be deleted without loss of toxin activity.

The N-terminal domain mediates binding to LRP; specifically Lys-57 and possibly other nearby residues. Toxicity for cells is reduced at least by 100 fold when Lys-57 is changed to glutamic acid or when most of the domain is deleted.

5.1 <u>The receptor – Lipoprotein Receptor related Protein</u>

The receptor for binding of Domain Ia is the LRP also known as the alpha-2-macroglobulin receptor. It is one of the largest membrane-associated proteins characterized to date and is located both on surface and intracellular membranes of many eukaryotic cell types. LRP is biosynthesized as a single chain ,4525-amino acid precursor. After synthesis, possibly in the Golgi, single chain LRP is cleaved that is subsequently processed into a hetradimer composed of C-terminal 601 residues light chain(β) and an N-terminal 3923 residues heavy chain(α).

After synthesis, single chain LRP is cleaved to give a

i) 515-kDa heavy chain

II) 85-kDa light chain.

The light chain contains both a membrane-spanning region and cytoplasmic tail. The heavy chain which remains membrane-associated via non-covalent interactions with the light chain contains the ligand-binding sites.

It comprises

i) a small cytoplasmic domain

ii) a single transmembrane helix and

iii) clusters of two types of small cystine rich repeats on extra cellular side.

1. Epidermal Growth Factor-like repeats of about 50 residues.
2. Complement like repeat (C8 and C9) of 42 residues.

Each of these types of repeat contains 6 cystines that are involved in 3 intradomain disulfide bonds.

LRP contains a total of 31 such repeats organized into 4 clusters of 2,8,7,11 repeats (CI, CII, CIII, CIV). Cluster II is a major locus for protein ligand binding. LRP binds to 30 ligands. The ability of LRP to bind to such a diverse array of ligands is related to presence of many slightly different complement like and EGF like repeats. Most of the ligands of LRP bind to Cluster II but Exotoxin A domain Ia is an exception. It binds to the Cluster IV.

5.2 Structures

Protein Data Bank was searched for the availability of structure of both Exotoxin A domain Ia and the receptor LRP. Although hit was obtained for the Exotoxin A, no results were obtained for LRP.

The Exotoxin A structure ID was 1IKQ, which is Pseudomonas Aeruginosa Exotoxin A, wild type. The structure determination is done using X-Ray diffraction method with expression system *E. Coli.* It is classified as a transferase.

5.2.1 *Modeling receptor structure*

The structure for LRP was modeled using software called modeller. For the purpose PDB blast was performed to obtain the template. The highest hit was obtained with 1N7D, which is the structure of Extra cellular domain of the LDL receptor. This was used as the template for modeling the LRP structure. Modeller gives files. After the structure was

odeled, it was validated using SAVS. ProCheck was performed, which gave the amachandran plot. Scores were obtained for the residues in allowed, additionally llowed, generously allowed and disallowed regions.

or the structure to be considered potential, the score for allowed and additionally llowed residues must be greater than 95%. It was obtained that the score was not more an 93% for any of the model generated. Thus this model could not be considered.

n order to come up with a better template a program called threader was implemented. hreading is an approach to fold recognition in the absence of an evolutionary elationship. It uses a detailed 3-D representation of protein structure. It aligns the given equence to the database of sequences so as to identify folds and generates a score. The etter the score, the better is the matching. This gave 1 K32A3 as the potential template.

he structure was again modeled using 1K32. The validation resulted to a better score of 7.6%.

.2.2 *Validation*.

he obtained structures were validated using SAVS server. Procheck was performed, which gave the Ramachandran Plot with the scores listed in Table 6 and Table 7 for 1n7d nd 1k32 respectively. Where the criteria of selection of a structure is that the percentage core of residues in most favored and additionally allowed regions should be more than 5%. Since none of the above structures in Table 6 fulfilled the criteria, hence none was elected. The structure 3 of Table 7 has the highest score and hence was selected as nodeling template.

.3 Docking

After having obtained the structure for LRP and already having the structure for Exotoxin a, the two were docked. For this not only the domain Ia but the entire Exotoxin A PDB was taken. Docking is important to be carried out as there is hardly any information about he binding site or the interacting residues barring Lys 57, mutating which leads to eduction of toxicity, which indicates that it might be involved in binding along with nearby residues. For docking purpose, software called FT Dock was used.

6. EXOTOXIN A DOMAIN III

6.1 About domain III

Exotoxin domain III length is 213 amino acids ranging from 400-613 amino acids it has NAD binding site from 465-481 amino acid having two chains that is chain A and chain B complexed with AMP and Nicotinamide which acts by transferring the ADP ribose from NAD to a modified histidine in elongation factor-2 (EF-2). It comprise of both alpha and beta sheets.

Further validation of Domain III i.e. Exotoxin A precursor (NAD-dependent ADP-ribosyltransferase) BLAST-p was performed with entire human genome so as to check if there is any similar protein in the humans. A single hit was found.

The similarity is only with Hypothetical protein and that is very low so we concluded that Domain III is a good target because the designed lead molecule can't bind with any of human proteins.

For domain III there are 9 structures in PDB database and they are:-

List of available structure of Domain III in Protein Data Bank

PDB ID	Title	Resolution
1AER	Domain III of *Pseudomonas aeruginosa* Exotoxin complexed with Beta-TAD	2.30
1DMA	Domain III of *Pseudomonas aeruginosa* Exotoxin complexed with Nicotinamide and AMP	2.50
1IKP	Pseudomonas Aeruginosa Exotoxin A, P201Q, W281A mutant	1.45
1IKQ	Pseudomonas Aeruginosa Exotoxin A, wild type	1.62
1XK9	Pseudomanas exotoxin A in complex with the PJ34 inhibitor	2.10
1ZM2	Structure of ADP-ribosylated eEF2 in complex with catalytic fragment of ETA	3.07
1ZM3	Structure of the eEF2-ETA-bTAD complex	2.90
1ZM4	Structure of the eEF2-ETA-bTAD complex	2.90

1ZM9	Structure of eEF2-ETA in complex with PJ34	2.80

So out of 9 structures from PDB we considered first two:

a) 1AER - Title-Domain III of *Pseudomonas aeruginosa* Exotoxin complexed with Beta-TAD. Resolution – 2.30Å

b) 1DMR – Title-Domain III of *Pseudomonas aeruginosa* Exotoxin complexed with Nicotinamide and AMP. Resolution – 2.50Å

Then 1DMA protein from PDB was selected for final studies of domain III because it was complexed with natural ligands that are AMP and nicotinamide.

VMD (Visual molecular dynamics) was used to study the dynamically protein and the natural ligand interactions.

6.2 <u>Ligplot of natural ligands</u>

Domains III have two natural ligand complexed with it that is AMP and nicotinamide in the two chains of A and B. AMP is complexed in chain B only but nicotinamide is complexed with both the chains A and B.

Ligplot is performed to get the information about the interacting residue with the natural ligands.

a) AMP (ADENOSINE MONOPHOSPHATE) complexed with 1DMA (domain III)

b) NCA complexed with Chain A of 1DMA (domain III)

c) NCA complexed with Chain B 1DMA (domain III)

Ligplot shows the interacting residue of the protein with the ligands. The interacting residues are important for docking any other ligand and give an idea of binding cavity of the protein.

6.3 <u>NAD structure</u>

NAD (Nicotinamide) structure is needed to study the potential drug binding site of the protein 1DMA of domain III by docking NAD in the active site of Domain III.

6.3.1 <u>Structure of NAD</u>

ISIS draw is used to draw the structure of NAD.

6.3.2 *Minimization of NAD structure*

The structure drawn in ISIS draw is minimized for optimization of energy of NAD molecule. It was done using Hyper Chem. On RMS gradient of 0.001 kcal/(Å mol) and maximum cycles 10,000.

Minimized structure is further used for the docking purpose in the protein-ligand dock.

6.4 Docking

After having the minimized structure of NAD. docking was performed with the protein that is 1DMA obtained from PDB to fit the ligand into the the protein and study the interaction of the potential binding site. For docking purpose Schrödinger is used.

Many applications of Schrödinger were used for ligand energy minimization and preparation as Impact was used for ligand energy minimization, for protein preparation Protein Preparation Wizard is used and for docking ligand into protein Glide is used.

In ligand preparation through Impact of NAD three are 29 tautamers generation of the molecule and out of that 16th tautamer having lowest glide core -8.56 is the best tautamer to be used for further studies.

7. RESULTS

Table 1

The unique proteins in UCBPP-PA14 (Strain of *P.aeruginosa*)

Metabolic process:

1	116053747	putative peptidyl-prolyl isomerase [Pseudomonas aeruginosa UCBPP-PA14]
2	116054097	putative acetyltransferase [Pseudomonas aeruginosa UCBPP-PA14]
3	116054148	putative chemotaxis protein [Pseudomonas aeruginosa UCBPP-PA14]
4	116052248	probable phenazine biosynthesis protein [Pseudomonas aeruginosa UCBPP-PA14]
5	116052247	probable phenazine biosynthesis protein [Pseudomonas aeruginosa UCBPP-PA14]
7	116052047	possible O-methyltransferase [Pseudomonas aeruginosa UCBPP-PA14]
9	116052046	possible hydrolase [Pseudomonas aeruginosa UCBPP-PA14]
10	116051902	putative integrase protein [Pseudomonas aeruginosa UCBPP-PA14]
11	116051710	putative oxidase [Pseudomonas aeruginosa UCBPP-PA14]
12	116050958	putative oxidase [Pseudomonas aeruginosa UCBPP-PA14]
13	116050938	putative cytochrome b561 [Pseudomonas aeruginosa UCBPP-PA14]
14	116050794	putative sialidase [Pseudomonas aeruginosa UCBPP-PA14]
15	116050262	putative enzyme of the cupin superfamily [Pseudomonas aeruginosa UCBPP-PA14]
16	116050201	pyoverdine biosynthesis protein PvcA [Pseudomonas aeruginosa UCBPP-PA14]
17	116049854	probable phenazine biosynthesis protein [Pseudomonas aeruginosa UCBPP-PA14]
18	116049853	probable phenazine biosynthesis protein [Pseudomonas aeruginosa UCBPP-PA14]
19	116049621	putative enzyme [Pseudomonas aeruginosa UCBPP-PA14]
20	116049516	putative ATP-NAD kinase [Pseudomonas aeruginosa UCBPP-PA14]
21	116049161	putative clavaminic acid synthetase [Pseudomonas aeruginosa UCBPP-PA14]
22	116052382	putative flavoprotein [Pseudomonas aeruginosa UCBPP-PA14]
23	116052603	Pyocin S5 [Pseudomonas aeruginosa UCBPP-PA14]
24	116052604	Colicin immunity protein [Pseudomonas aeruginosa UCBPP-PA14]
25	116052697	type 4 fimbrial biogenesis protein PilX [Pseudomonas aeruginosa UCBPP-PA14]
26	116052942	putative peroxiredoxin [Pseudomonas aeruginosa UCBPP-PA14]

Signaling

27	116053746	phospholipase C, PlcB [Pseudomonas aeruginosa UCBPP-PA14]
28	116053754	putative histidine phosphotransfer domain [Pseudomonas aeruginosa UCBPP-PA14]
29	116053845	putative plasmid stabilisation system protein [Pseudomonas aeruginosa UCBPP-PA14]
30	116054235	putative c-type cytochrome [Pseudomonas aeruginosa UCBPP-PA14]
31	116054243	putative c-type cytochrome precursor [Pseudomonas aeruginosa UCBPP-PA14]
32	116052076	possible ABC transporter component [Pseudomonas aeruginosa UCBPP-PA14]
33	116052075	possible ABC transporter permease component [Pseudomonas aeruginosa UCBPP-PA14]
34	116054245	nitrite reductase precursor [Pseudomonas aeruginosa UCBPP-PA14]
35	116051805	putative TRAP-type C4-dicarboxylate transport protein [Pseudomonas aeruginosa UCBPP-PA14]
36	116051782	possible oriT-binding protein, TraK [Pseudomonas aeruginosa UCBPP-PA14]
37	116051781	oriT-binding protein, TraJ [Pseudomonas aeruginosa UCBPP-PA14]
38	116051349	putative acyl carrier protein [Pseudomonas aeruginosa UCBPP-PA14]
39	116051169	putative 7,8-dihydro-8-oxoguanine-triphosphatase [Pseudomonas aeruginosa UCBPP-PA14]
40	116054461	adenylate cyclase ExoY [Pseudomonas aeruginosa UCBPP-PA14]
41	116049974	putative ring-cleaving dioxygenase [Pseudomonas aeruginosa UCBPP-PA14]
42	116048928	Quinolone signal response protein [Pseudomonas aeruginosa UCBPP-PA14]
43	116048766	phospholipase accessory protein PlcR precursor [Pseudomonas aeruginosa UCBPP-PA14]

Secretary pathway

44	116053766	Exoenzyme T [Pseudomonas aeruginosa UCBPP-PA14]
45	116053806	putative secretion protein [Pseudomonas aeruginosa UCBPP-PA14]
46	116051100	secretion protein XcpP [Pseudomonas aeruginosa UCBPP-PA14]

47 116051096 general secretion pathway outer membrane protein H precursor [Pseudomonas aeruginosa UCBPP-PA14]

48 116051092 general secretion pathway protein L [Pseudomonas aeruginosa UCBPP-PA14]

49 116051091 general secretion pathway protein M [Pseudomonas aeruginosa UCBPP-PA14]

50 116050660 Type II secretion system protein [Pseudomonas aeruginosa UCBPP-PA14]

51 116050659 type II secretion system protein [Pseudomonas aeruginosa UCBPP-PA14]

52 116050658 putative type II secretion system protein [Pseudomonas aeruginosa UCBPP-PA14]

53 116048925 PqsB [Pseudomonas aeruginosa UCBPP-PA14]

54 116048915 SpcU [Pseudomonas aeruginosa UCBPP-PA14]

55 116048914 ExoU [Pseudomonas aeruginosa UCBPP-PA14]

Transcriptional regulator:

56 116053770 putative transcriptional regulator [Pseudomonas aeruginosa UCBPP-PA14]

57 116051789 putative replication protein, RepC [Pseudomonas aeruginosa UCBPP-PA14]

58 116050517 PA-I galactophilic lectin [Pseudomonas aeruginosa UCBPP-PA14]

59 116049299 putative transcriptional regulator [Pseudomonas aeruginosa UCBPP-PA14]

60 116049235 putative transcriptional regulator [Pseudomonas aeruginosa UCBPP-PA14]

61 116048750 probable transcriptional regulator [Pseudomonas aeruginosa UCBPP-PA14]

62 116048614 Transcriptional regulator ToxR [Pseudomonas aeruginosa UCBPP-PA14]

63 116049662 Exoenzyme S synthesis protein B [Pseudomonas aeruginosa UCBPP-PA14]

64 116049656 type III secretion protein PcrV [Pseudomonas aeruginosa UCBPP-PA14]

65 116049650 putative type III secretion protein [Pseudomonas aeruginosa UCBPP-PA14]

66 116049649 putative protein in type III secretion [Pseudomonas aeruginosa UCBPP-PA14]

67 116049648 Type III secretion outer membrane protein PopN precursor [Pseudomonas aeruginosa UCBPP-PA14]

68 116049080 exotoxin A precursor [Pseudomonas aeruginosa UCBPP-PA14]

69 116052556 type IV pilin structural subunit [Pseudomonas aeruginosa UCBPP-PA14]

70 116052608 type IV B pilus protein [Pseudomonas aeruginosa UCBPP-PA14]

Cell wall protein

No.	GI	Protein
71	116053843	putative hemolysin [Pseudomonas aeruginosa UCBPP-PA14]
72	116054148	putative chemotaxis protein [Pseudomonas aeruginosa UCBPP-PA14]
73	116054238	heme d1 biosynthesis protein NirH [Pseudomonas aeruginosa UCBPP-PA14]
74	116054239	heme d1 biosynthesis protein NirG [Pseudomonas aeruginosa UCBPP-PA14]
75	116054240	heme d1 biosynthesis protein NirL [Pseudomonas aeruginosa UCBPP-PA14]
76	116054241	heme d1 biosynthesis protein NirD [Pseudomonas aeruginosa UCBPP-PA14]
77	116054389	putative integral membrane protein [Pseudomonas aeruginosa UCBPP-PA14]
78	116051974	possible membrane protein [Pseudomonas aeruginosa UCBPP-PA14]
79	116050071	fimbrial subunit CupA4 [Pseudomonas aeruginosa UCBPP-PA14]
80	116049133	coat protein B of bacteriophage Pf1 [Pseudomonas aeruginosa UCBPP-PA14]
81	116049130	helix destabilizing protein of bacteriophage Pf1 [Pseudomonas aeruginosa UCBPP-PA14]
82	116052648	fimbrial subunit [Pseudomonas aeruginosa UCBPP-PA14]
83	116052697	type 4 fimbrial biogenesis protein PilX [Pseudomonas aeruginosa UCBPP-PA14]

Protein synthesis & structure:

84 116051871 probable chaperone [Pseudomonas aeruginosa UCBPP-PA14]

Transcriptional repressor:

No.	GI	Protein
85	116051787	Mercuric resistence transcriptional repressor protein MerD [Pseudomonas aeruginosa UCBPP-PA14]
86	116054400	Mercuric transport protein MerT [Pseudomonas aeruginosa UCBPP-PA14]
87	116050818	putative RNA polymerase sigma factor [Pseudomonas aeruginosa UCBPP-PA14]

Translocation protein:

88 116051376 Fucose-binding lectin PA-IIL [Pseudomonas aeruginosa UCBPP-PA14]
89 116051001 Na+-translocating NADH:ubiquinone oxidoreductase subunit Nrq1 [Pseudomonas aeruginosa UCBPP-PA14]
90 116051000 Na+-translocating NADH:ubiquinone oxidoreductase subunit Nrq3 [Pseudomonas aeruginosa UCBPP-PA14]
91 116049659 Translocator outer membrane protein PopD precursor [Pseudomonas aeruginosa UCBPP-PA14]

Regulator:

92 116049658 Translocator protein PopB [Pseudomonas aeruginosa UCBPP-PA14]
93 116049376 regulatory protein RsaL [Pseudomonas aeruginosa UCBPP-PA14]
94 116053356 putative phosphate transport regulator [Pseudomonas aeruginosa UCBPP-PA14]

Lowmatch_nohits in human:

1 116054086 putative phosphate acyltransferase [Pseudomonas aeruginosa UCBPP-PA14]
2 116052217 Pvds-regulated endoprotease, lysyl class [Pseudomonas aeruginosa UCBPP-PA14]
3 116054401 mercuric resistance protein MerE [Pseudomonas aeruginosa UCBPP-PA14]
4 116050230 Putative protease [Pseudomonas aeruginosa UCBPP-PA14]
5 116050202 pyoverdine biosynthesis protein PvcB [Pseudomonas aeruginosa UCBPP-PA14]
6 116050189 Possible acetyltransferase [Pseudomonas aeruginosa UCBPP-PA14]
7 116052372 Putative acyltransferase [Pseudomonas aeruginosa UCBPP-PA14]

Table 2 The unique proteins In PAO1 (Strain of *P.aeruginosa*)

Metabolic process:

1 15595614 Probable chemotaxis protein [Pseudomonas aeruginosa PAO1]
2 15595706 Probable c-type cytochrome [Pseudomonas aeruginosa PAO1]
3 15595714 Probable c-type cytochrome precursor [Pseudomonas aeruginosa PAO1]
4 15595716 nitrite reductase precursor [Pseudomonas aeruginosa PAO1]
5 15596181 colicin immunity protein [Pseudomonas aeruginosa PAO1]
6 15596182 pyocin S5 [Pseudomonas aeruginosa PAO1]
7 15597068 LasA protease precursor [Pseudomonas aeruginosa PAO1]
8 15597096 Probable phenazine biosynthesis protein [Pseudomonas aeruginosa PAO1]

9 15597097 Probable phenazine biosynthesis protein [Pseudomonas aeruginosa PAO1]

10 15597220 Probable ring-cleaving dioxygenase [Pseudomonas aeruginosa PAO1]

11 15597450 pyoverdine biosynthesis protein PvcA [Pseudomonas aeruginosa PAO1]

12 15597766 PA-I galactophilic lectin [Pseudomonas aeruginosa PAO1]

13 15597890 Probable thioredoxin [Pseudomonas aeruginosa PAO1]

14 155981[illegible] Na(+)-translocating NADH-quinone reductase subunit C [Pseudomonas aeruginosa PAO1]

15 15598195 Na(+)-translocating NADH-quinone reductase subunit A [Pseudomonas aeruginosa PAO1]

16 15599037 Probable chaperone [Pseudomonas aeruginosa PAO1]

17 15599405 Probable phenazine biosynthesis protein [Pseudomonas aeruginosa PAO1]

18 15599406 Probable phenazine biosynthesis protein [Pseudomonas aeruginosa PAO1]

Cell signaling

19 15595224 phospholipase C, PlcB [Pseudomonas aeruginosa PAO1]

20 15596194 beta-keto-acyl-acyl-carrier protein synthase-like protein [Pseudomonas aeruginosa PAO1]

21 15596197 Quinolone signal response protein [Pseudomonas aeruginosa PAO1]

22 15598530 Probable acyl carrier protein [Pseudomonas aeruginosa PAO1]

Secretary pathway

23 15595242 exoenzyme T [Pseudomonas aeruginosa PAO1]

24 15596345 Exotoxin A precursor [Pseudomonas aeruginosa PAO1]

25 15596628 regulatory protein RsaL [Pseudomonas aeruginosa PAO1]

26 15596903 type III secretion protein PcrV [Pseudomonas aeruginosa PAO1]

27 15596907 ExsC, exoenzyme S synthesis protein C precursor. [Pseudomonas aeruginosa PAO1]

28 15596908 ExsE [Pseudomonas aeruginosa PAO1]

29 15596911 ExsD [Pseudomonas aeruginosa PAO1]

30 15596914 type III export protein PscD [Pseudomonas aeruginosa PAO1]

31 15596916 type III export protein PscF [Pseudomonas aeruginosa PAO1]

32 15596917 type III export protein PscG [Pseudomonas aeruginosa PAO1]

33 15596918 type III export protein PscH [Pseudomonas aeruginosa PAO1]

34 15596919 type III export protein PscI [Pseudomonas aeruginosa PAO1]

35 15598291 general secretion pathway protein M [Pseudomonas aeruginosa PAO1]

36 15598292 general secretion pathway protein L [Pseudomonas aeruginosa PAO1]

37 15598296 General secretion pathway outer membrane protein H precursor [Pseudomonas aeruginosa PAO1]

38 15598300 secretion protein XcpP [Pseudomonas aeruginosa PAO1]

39 15598349 O-antigen translocase [Pseudomonas aeruginosa PAO1]

40 15598350 B-band O-antigen polymerase [Pseudomonas aeruginosa PAO1]

41 15598356 O-antigen chain length regulator [Pseudomonas aeruginosa PAO1]

Transcriptional regulator:

42 15595246 Probable transcriptional regulator [Pseudomonas aeruginosa PAO1]

43 15595450 Probable transcriptional regulator [Pseudomonas aeruginosa PAO1]

44 15595614 Probable chemotaxis protein [Pseudomonas aeruginosa PAO1]

45 15595710 Probable transcriptional regulator [Pseudomonas aeruginosa PAO1]

46 15595712 Probable transcriptional regulator [Pseudomonas aeruginosa PAO1]

47 15595904 transcriptional regulator ToxR [Pseudomonas aeruginosa PAO1]

48 15596025 Probable transcriptional regulator [Pseudomonas aeruginosa PAO1]

49 15596480 Probable transcriptional regulator [Pseudomonas aeruginosa PAO1]

50 15596544 Probable transcriptional regulator [Pseudomonas aeruginosa PAO1]

51 15597387 adenylate cyclase ExoY [Pseudomonas aeruginosa PAO1]

52 15597868 Probable type II secretion system protein [Pseudomonas aeruginosa PAO1]

53 15597869 Probable type II secretion system protein [Pseudomonas aeruginosa PAO1]

54 15597870 Probable type II secretion system protein [Pseudomonas aeruginosa PAO1]

55 15599036 exoenzyme S [Pseudomonas aeruginosa PAO1]

Cell wall protein

56 15595711 heme d1 biosynthesis protein NirL [Pseudomonas aeruginosa PAO1]

57 15595920 coat protein B of bacteriophage Pf1) [Pseudomonas aeruginosa PAO1]

58 [illegible] [illegible]pase accessory protein PilcK precursor [Pseudomonas aeruginosa PAO1]

59 15597327 Fimbrial subunit CupA4 [Pseudomonas aeruginosa PAO1]

60 15599745 type 4 fimbrial biogenesis protein FimT [Pseudomonas aeruginosa PAO1]

61 15599746 type 4 fimbrial biogenesis protein FimU [Pseudomonas aeruginosa PAO1]

62 15599749 type 4 fimbrial biogenesis protein PilX [Pseudomonas aeruginosa PAO1]

63 15599751 type 4 fimbrial biogenesis protein PilY2 [Pseudomonas aeruginosa PAO1]

64 15599752 type 4 fimbrial biogenesis protein PilE [Pseudomonas aeruginosa PAO1]

Translocation protein:

65 15596905 translocator protein PopB [Pseudomonas aeruginosa PAO1]

66 15596906 Translocator outer membrane protein PopD precursor [Pseudomonas aeruginosa PAO1]

67 15598349 O-antigen translocase [Pseudomonas aeruginosa PAO1]

Regulator:

68 15596628 regulatory protein RsaL [Pseudomonas aeruginosa PAO1]

69 15596904 regulatory protein PcrH [Pseudomonas aeruginosa PAO1]

Low match nohits in human:

1 15596921 type III export protein PscK [Pseudomonas aeruginosa PAO1]

2 15597451 pyoverdine biosynthesis protein PvcB [Pseudomonas aeruginosa PAO1]

3 15599370 Pvds-regulated endoprotease, lysyl class [Pseudomonas aeruginosa PAO1]

Table 3 Common proteins of both the strains of *P.aeruginosa*

Involved in Metabolism

Gi	Name of the protein	Swissprot	KEGG	NCBI
15595614	Chemotactic transduction protein chpE		Probable chemotaxis protein	
15596181	Colicin-Ia immunity protein	This protein is able to protect a cell, which harbors the plasmid ColIa-CA53 encoding colicin Ia, against colicin Ia.		
15596182	Colicin-Ib	This colicin is a channel-forming colicin. This class of transmembrane toxins depolarize the cytoplasmic membrane, leading to dissipation of cellular energy. Colicins are polypeptide toxins produced by and active against, Escherichia coli and closely related bacteria.		This colicin is a channel-forming colicin. This class of transmembrane toxins depolarize the cytoplasmic membrane, leading to dissipation of cellular energy. Colicins are polypeptide toxins produced by and active against,

		yopE.		Escherichia and closely related bacteria.
15597096	Phenazine biosynthesis protein phzA 2	Involved in the biosynthesis of the antibiotic, phenazine, a nitrogen-containing heterocyclic molecule having important roles in virulence, competition and biological control.	Probable phenazine biosynthesis protein	Involved in the biosynthesis of the antibiotic, phenazine, a nitrogen-containing heterocyclic molecule having important roles in virulence, competition and biological control.
15597450	Spore wall maturation protein DIT1	Involved in spore wall maturation.Catalyzes two step reaction that leads to the LL-dityrosine containing precursor of spore wall	Dit1p	Involved in spore wall maturation. Catalyzes a two step reaction that leads to the LL-dityrosine containing precursor of the spore wall
15597766	PA-I galactophilic lectin	D-galactose specific lectin. Binds in decreasing order of affinity: melibiose, methyl-alpha-D-galactoside, D-galactose, methyl-beta-D-galactoside, N-acetyl-D-galactosamine. Similar to plant lectins in its selective (carbohydrate-specific) hemagglutinating activity.	PA-I galactophilic lectin	D-galactose specific lectin. Binds in decreasing order of affinity: melibiose, methyl-alpha-D-galactoside, D-galactose, methyl-beta-D-galactoside, N-acetyl-D-galactosamine. Similar to plant lectins in its selective (carbohydrate-specific) hemagglutinating activity.
15599037	YopE regulator	Positive regulator of		

		yopE.		

Involved in Signaling

15595224	Apical membrane antigen 1 [Precursor]	Involved in parasite invasion of erythrocytes		Involved in parasite invasion of erythrocytes
15596197	Uncharacterized protein PA1000		hypothetical protein	

Involved in Secretary pathway

15595242	ADP-ribosyltransferase toxin aexT (Exoenzyme T).	Directly involved in the toxicity for RTG-2 (rainbow trout gonad) fish cells.		Characterization of an ADP-ribosyltransferase toxin (AexT) from Aeromonas salmonicida subsp. Salmonicida
15598291	General secretion pathway protein M.	Involved in a general secretion pathway (GSP) for the export of proteins.	General secretion pathway protein M	Protein secretion in gram-negative bacteria: transport across the outer membrane involves common mechanisms in different bacteria
15598292	General secretion pathway protein L.	Involved in a general secretion pathway (GSP) for the export of proteins.	General secretion pathway protein L	Protein secretion in gram-negative bacteria: transport across the outer membrane involves common mechanisms in different bacteria
15598296	General secretion pathway protein H precursor	Involved in a general secretion pathway (GSP) for the export of proteins. Required for the	General secretion pathway protein H	Protein secretion in Pseudomonas aeruginosa: characterization of seven xcp genes and

		translocation of a variety of enzymes across the outer membrane.		processing of secretory apparatus components by prepilin peptidase
15598300	General secretion pathway protein N	Involved in a general secretion pathway (GSP) for the export of proteins. Required for the translocation of a variety of enzymes across the outer membrane.	Secretion protein XcpP	Xcp-mediated protein secretion in Pseudomonas aeruginosa identification of two additional genes and evidence for regulation
15596345	Exotoxin A precursor (NAD-dependent ADP-ribosyltransferase).	This toxin is a NAD-dependent ADP-ribosyltransferase. It catalyzes the transfer of the ADP ribosyl moiety of oxidized NAD onto elongation factor 2 (EF-2) thus arresting protein synthesis.	exotoxin A precursor	Active site of Pseudomonas aeruginosa exotoxin A. Glutamic acid 553 is photolabeled by NAD and shows functional homology with glutamic acid 148 of diphtheria toxin

Involved in Receptor

15596628	Peptidyl-prolyl cis-trans isomerase-like 4	PPIases accelerate the folding of proteins. It catalyzes the cis-trans isomerization of proline imidic peptide bonds in oligopeptides (By similarity).		Peptidyl-prolyl cis-trans isomerase-like 4 (PPIase) (Rotamase).

Involved in Cell wall synthesis

11559571	Protein nirL		heme d1 biosynthesis protein NirL	Gene cluster for dissimilatory nitrite reductase (nir) from Pseudomonas aeruginosa: sequencing and identification of a locus for heme d1 biosynthesis
15595920	Coat protein B [Precursor]	Coat protein B is the major coat protein of the virion.		DNA sequence of the filamentous bacteriophage Pf1
15597327	NAD-dependent malic enzyme 65 kDa isoform, mitochondrial [Precursor]			Cloning and analysis of the C4 photosynthetic NAD-dependent malic enzyme of amaranth mitochondria
15599749	DNA-directed RNA polymerase subunit beta	DNA-dependent RNA polymerase catalyzes the transcription of DNA into RNA using the four ribonucleoside triphosphates as substrates.		

Table 4
Secretary pathway proteins involved in toxicity of *P.aeruginosa*

S.No.	Gene ID	Name of protein
1.	15595242	ADP-ribosyltransferase toxin aexT(Exoenzyme T).
2.	15598291	General secretion pathway protein M.

	15598292	General secretion pathway protein L.
	15598296	General secretion pathway protein H precursor.
	15598300	General secretion pathway protein N.
	15596345	Exotoxin A precursor (NAD-dependent ADP-ribosyltransferase).

able 5

alidation for structure using 1n7d as template using SAVS (procheck)

ructure	Residues in most favoured regions	Residues in additional allowed regions	Residues in generously allowed regions	Residues in disallowed regions
ructure1	231 (60.0%)	127 (33.0%)	16 (4.2%)	11 (2.9%)
ructure 2	229 (59.5%)	123 (31.9%)	25 (6.5%)	8 (2.1%)
ructure 3	234 (60.8%)	129(33.5%)	14(3.6%)	8 (2.1%)
ructure 4	241(62.6%)	122(31.7%)	17(4.4%)	5 (1.3%)
ructure 5	240 (62.3%)	116 (30.1%)	20 (5.2%)	9 (2.3%)

able 6

alidation for structure using 1k32 as template using SAVS (ProCheck)

tructure	Residues in most favoured regions	Residues in additional allowed regions	Residues in generously allowed regions	Residues in disallowed regions
tructure1	289 (75.1%)	84 (21.8%)	7 (1.8%)	5 (1.3%)
tructure 2	305 (79.2%)	63 (16.4%)	13 (3.4%)	4 (1.0%)
tructure 3	305 (79.2%)	71(18.4%)	5 (1.3%)	4 (1.0%)
tructure 4	290(75.3%)	83(21.6%)	10(2.6%)	2 (0.5%)
tructure 5	297 (77.1%)	72 (18.7%)	12 (3.1%)	4 (1.0%)

able: 7

Blast-p result of *P.aeruginosa* with entire human genome having 1 hit i.e.

Hypothetical protein LOC652806 [*Homo sapiens*]

Score = 37.4 bits (85)

Expect = 0.008

Identities = 41/145 (28%),

Positives = 59/145 (40%)

Gaps = 10/145 (6%)

Table: 8

Glide Score (Schrödinger results) of 1DMA (Domain III) docked with ligand NAD.

S. No.	RANK	LIG #	G Score
1	NAD	16	-8.56
2	NAD	3	-7.83
3	NAD	9	-6.82
4	NAD	17	-6.35
5	NAD	12	-6.29
6	NAD	21	-6.19
7	NAD	8	-6.14
8	NAD	6	-6.13
9	NAD	15	-6.04
10	NAD	10	-5.89
11	NAD	5	-5.87
12	NAD	22	-5.77
13	NAD	13	-5.76
14	NAD	25	-5.70
15	NAD	26	-5.68
16	NAD	7	-5.66
17	NAD	11	-5.70
18	NAD	28	-5.68
19	NAD	24	-5.34

20	NAD	2	-5.29
21	NAD	4	-5.24
22	NAD	30	-5.18
23	NAD	29	-5.16
24	NAD	19	-5.07
25	NAD	20	-4.89
26	NAD	14	-4.75
27	NAD	27	-4.29
28	NAD	18	-4.24
29	NAD	23	-4.23

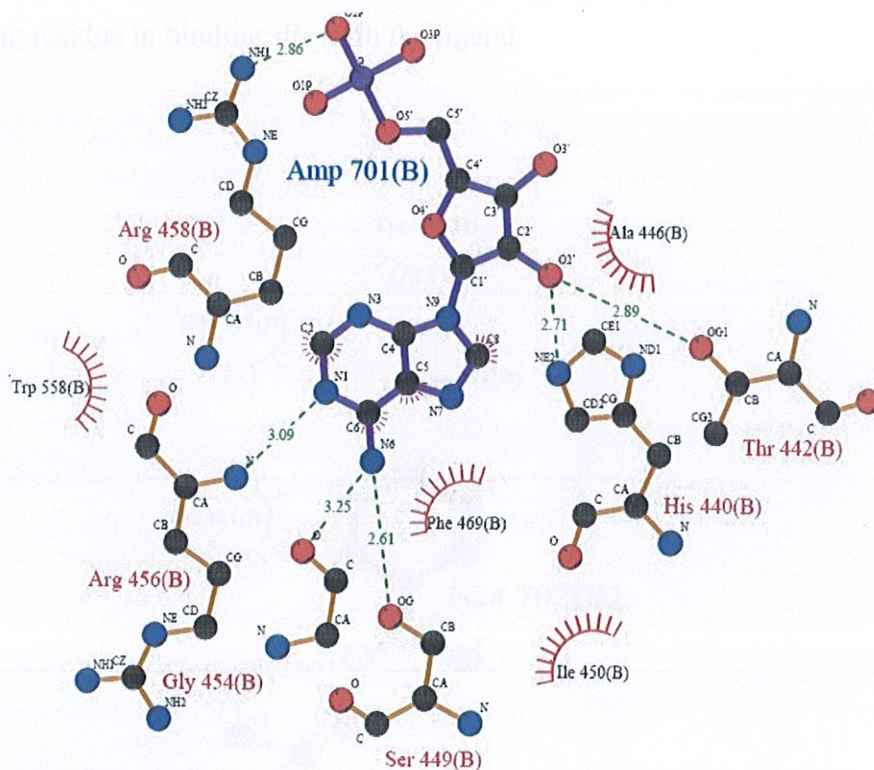


Fig: 2 Ligplot of AMP complexed with 1DMA (Domain III) specifying the interacting residue in binding site with the ligand

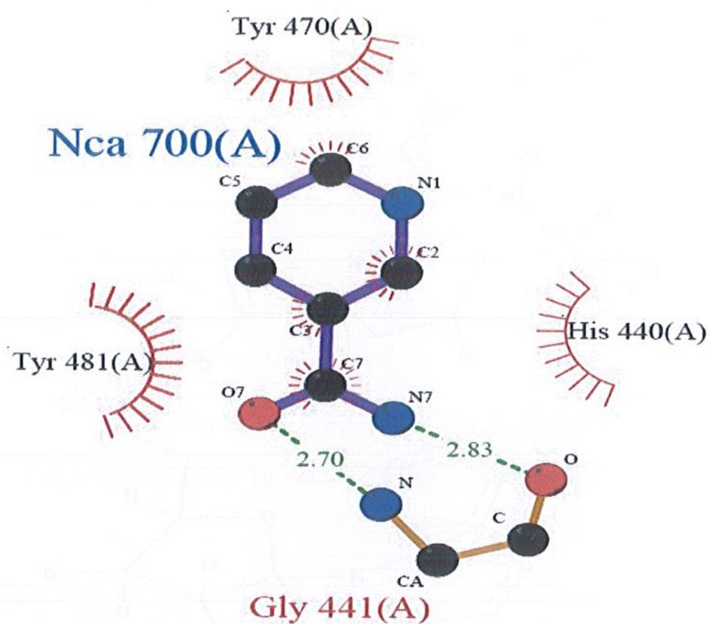


Fig: 3 Ligpolt of NCA complexed with chain A of 1DMA (Domain III) specifying the interacting residue in binding site with the ligand.

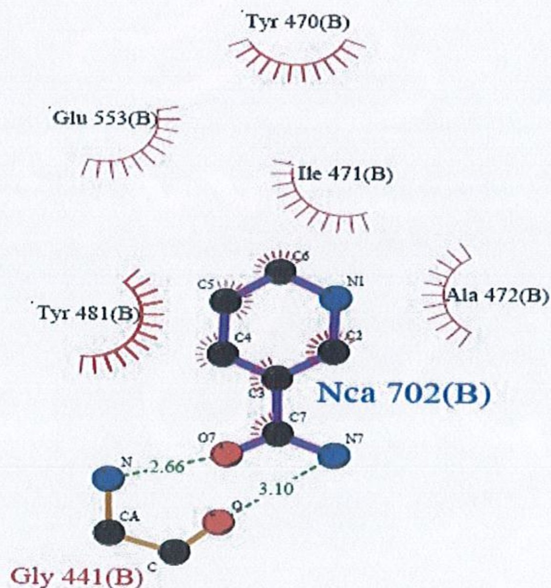


Fig: 4 Ligplot of NCA complexed with chain B of 1DMA (Domain III) specifying the interacting residue in binding site with the ligand.

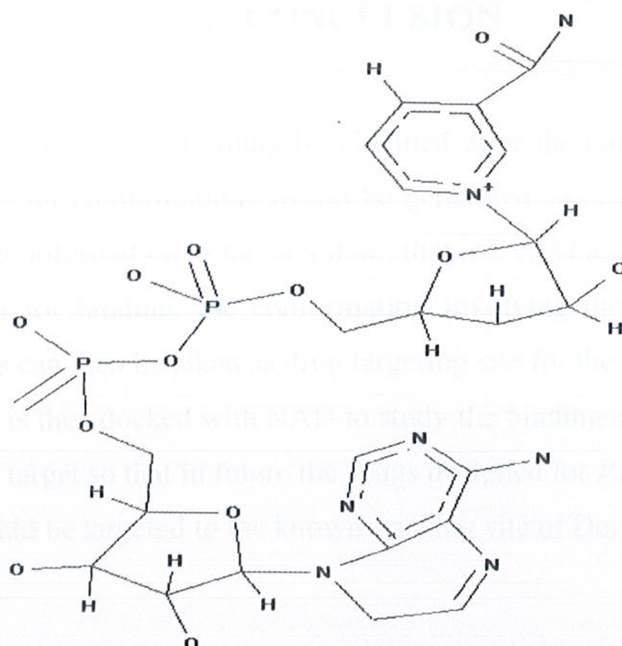


Fig 5: Energy minimized structure of NAD through Hyper Chem

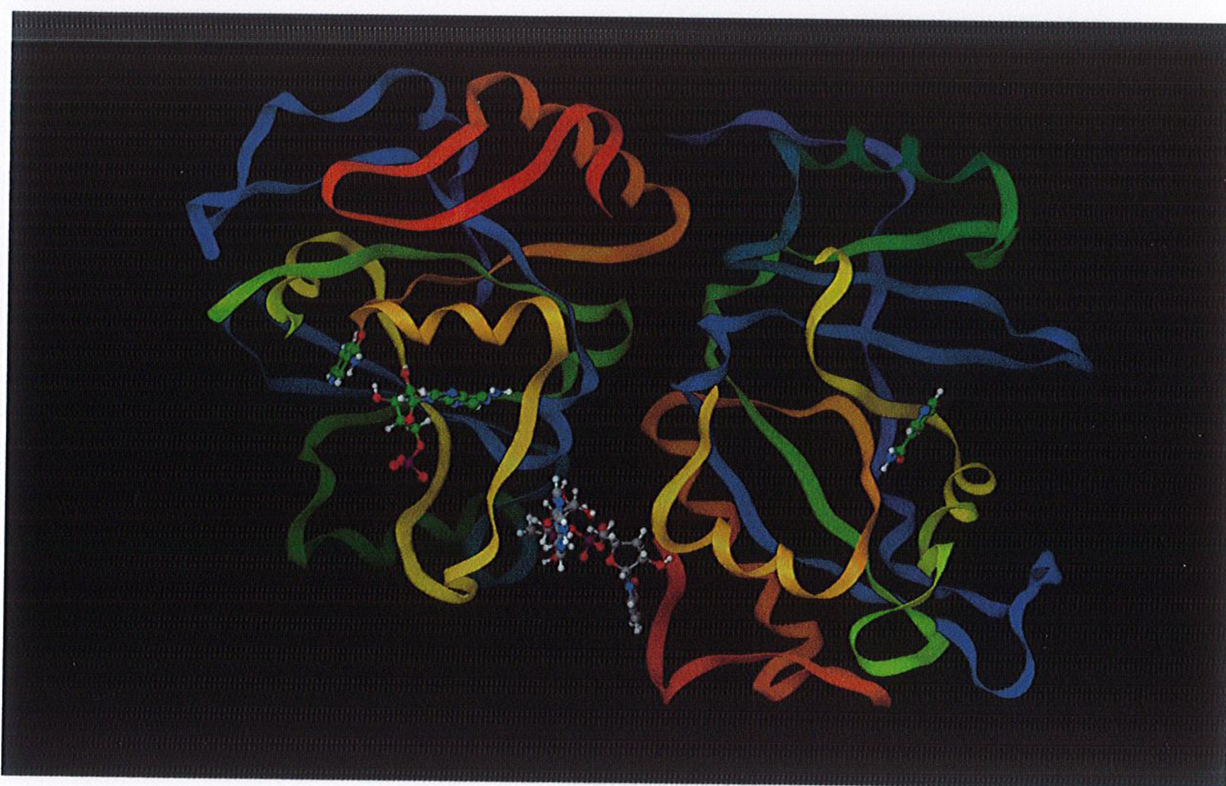


Fig: 6 Docked structure of domain III with 16th tautamer of NAD generated during ligand preparation showing the potential drug binding site of the protein.

8. CONCLUSION

The binding site for domain I would be obtained after the completion of run of FT Dock. Since several conformations would be generated so after the filtration on the basis of available information in the literature, that is Lys 57 and neighboring residues being important for binding, the conformation involving these residues would be considered. This can then be taken as drug targeting site for the future work.

The Domain III is then docked with NAD to study the binding site on the domain for the further drug target so that in future the drugs designed for *Pseudomonas aeruginosa* should be targeted to the known docking site of Domain III of exotoxin A.

BIBLIOGRAPHY

1) www.rcsb.org

2) www.expasy.org

3) www.ncbi.nlm.nih.org

4) www.genome.jp/kegg/

5) Li M., Dyda F.,Benhatrt I., Pastant I. & David R. Davies. *Crystal structure of the catalytic domain of Pseudomonas exotoxin A complexed with a nicotinamide adenine dinucleotide analog: Implications for the activation process and for ADP ribosylation.* Biochemistry Vol. 93, pp. 6902-6906, July 1996.

6) Li M., Dyda F.,Benhatrt I., Pastant I. & David R. Davies. *The crystal structure of Pseudomonas aeruginosa exotoxin domain III with nicotinamide and AMP: Conformational differences with the intact exotoxin.* Biochemistry Vol. 92, pp. 9308-9312, September 1995.

7) Kenneth, *Todar's Online Textbook of Bacteriology, Pseudomonas aeruginosa.* Todar. University of Wisconsin-Madison Department of Bacteriology 2004.

8) Glide Quick Start user manual copyright 2006 .

9) Topics in infectious Diseases Newsletter, August 2001.

10) Gidon Moont, Graham R. Smith and Michael J.E. *3D-DOCK incorporating FTDOCK (version 2.0) R.P. Score and Multidock.* Sternberg 2001.

11) Kounnas M., Morris R., Thompson M., FitzGerald D. *The alpha2-Macroglobulin Receptor/Low Density Lipoprotein Receptor-related Protein Binds and Internalizes Pseudomonas Exotoxin A.* THE JOURNAL OF BIOLOGICAL CHEMISTRY *Vol 267,,pp.12420-12423,June 1992.*

12) *Yates. S, Merril A A Catalytic Loop within Pseudomonas aeruginosa Exotoxin A Modulates Its Transferase Activity.* THE JOURNAL OF BIOLOGICAL CHEMISTRY. Vol 76, pp 35029-35036, 2001.

13) Wedekind J., Trame C., Dorywalska M., Koehl P., Raschke T., McKee M, FitzGerald D., Collier R. and McKay D. *Refined Crystallographic Structure of Pseudomonas aeruginosa Exotoxin A and its implications for the Molecular Mechanisms of toxicity.* Journal of Mol. Biol. Vol.314,pp- 823-837 2001.

14) Neilsen P., Ellgaard L., Etzerodt M., Thogersen H.and Poulsen F.*The solution structure of the N-terminal domain of alpha2-macroglobulin receptor associated protein* Proc. Natl. Acad. Sci. USA Vol 94, pp 7521-7525, July 1997

15) Horn I., van den Berg B., Meijden P., Pannekoek H., Zonneveld A. *Molecular Analysis of Ligand Binding to the Second Cluster of Complement-type Repeats of the Low Density Lipoprotein Receptor-related Protein* THE JOURNAL OF BIOLOGICAL CHEMISTRY. Vol 272, pp 13608-13613, 1997

16) Nielsen M., Nykjaer A., Warshawsky I., Scwartz A.and Gliemann J. *Analysis of Ligand Binding to the alpha2-Macroglobulin Receptor/Low Density Lipoprotein Receptor-related Protein.* THE JOURNAL OF BIOLOGICAL CHEMISTRY. Vol 270, pp 23713-23719, 1995.