Elucidating the Molecular Interaction of Zebrafish (*Danio rerio*) PGRP-2 receptor with DAP and LYS-type Peptidoglycans using *in silico* Approaches

Sunanda Paramanik

Adm. No. 05BI/15

DEPARTMENT OF BIOINFORMATICS
CENTRE FOR POST-GRADUATE STUDIES
ORISSA UNIVERSITY OF AGRICULTURE AND TECHNOLOGY
BHUBANESWAR-751003
2017

Paramanik S, M.Sc. (Bioinformatics) Thesis, 2017.
Elucidating the Molecular Interaction of Zebrafish (*Danio rerio*) PGRP-2 receptor with DAP and LYS-type Peptidoglycans using *in silico* Approaches

Elucidating the Molecular Interaction of Zebrafish (*Danio rerio*) PGRP-2 receptor with DAP and LYS-type Peptidoglycans using *in silico* Approaches

Thesis Submitted
To
Orissa University of Agriculture and Technology, Bhubaneswar
In partial fulfillment of the requirement of the degree of
M.Sc. BIOINFORMATICS

BY
Ms. SUNANDA PARAMANIK
ADM NO: 05BI/15

DEPARTMENT OF BIOINFORMATICS
CENTRE FOR POST GRADUATE STUDIES
ORISSA UNIVERSITY OF AGRICULTURE AND TECHNOLOGY
BHUBANESWAR-751003
2017

Advisor **Mr. Sukanta Kumar Pradhan**

भाकृअनुप–केन्द्रीय अंतर्स्थलीय मात्स्यिकी अनुसंधान संस्थान
आई एस ओ 9000 : 2008 प्रमाणित संगठन
(भारतीय कृषि अनुसंधान परिषद)
बैरकपुर, कोलकाता–700 120, पश्चिम बंगाल
ICAR-Central Inland Fisheries Research Institute
AN ISO 9000 : 2008 Certified Organisation
(Indian Council of Agricultural Research)
Barrackpore, Kolkata - 700 120, West Bengal

डा. बसंत कुमार दास, निदेशक
Dr. Basanta Kumar Das, Director

CERTIFICATE-I

This is to certify that dissertation entitled **"Elucidating the molecular interaction of Zebrafish (*Danio rario*) PGRP-2 receptor with DAP and LYS-type peptidoglycans using *in silico* approaches"** submitted for award for the degree of **Master of Science** in the subject of **Bioinformatics** embodies a faithful bonafied research work carried out by **Ms. Sunanda Paramanik (Adm. No. 05BI/15)** under my guidance & supervision. No part of this thesis has been submitted by her for any other degree or diploma.

I further certify that any help or information received during the course of investigation have been duly acknowledged by her.

Place: ICAR-CIFRI, Kolkata
Date:

(Basanta Kumar Das)

Phone : (033) 2592 0177 (Director), 2593 5288 (AO), 2593 3081 (SF&AO), Fax : 033 2592 0388, EPBX 2592 1190/91 Ext.
e-mail : director.cifri@icar.gov.in; director.cifri@gmail.com; Website : www.cifri.res.in
An ISO 9001:2008 Organisation

CERTIFICATE-II

This is to certify that the dissertation entitled **"Elucidating the molecular interaction of Zebrafish (*Danio rario*) PGRP-2 receptor with DAP and LYS-type peptidoglycans using *in silico* approaches"** submitted by **Ms. Sunanda Paramanik (Adm. No. 05BI/15)** to the Orissa University of Agriculture and Technology, Bhubaneswar in the partial fulfillment of the requirements for the award of the degree of **Master of Science** in the subject of **Bioinformatics** has been approved by the students advisory committee after an oral examination of the same in collaboration with external examiner.

ADVISORY COMMITTEE

1. **Dr. Basanta Kumar Das** — Chairman .....
   Director, ICAR - CIFRI

2. **Dr. Sukanta Kumar Pradhan** — Member .....
   HOD, (Bioinformatics)

3. **Mrs. Sucharita Balabantray** — Member.....
   Asst. Prof. (Bioinformatics)

4. **External Examiner** — Signature.....

DEDICATED TO

THEM FOR WHOM I EXIST

ACKNOWLEDGEMENT

While a completed dissertation bears the single name of the student, the process that leads to its completion is always accomplished in combination with the dedicated work of other people. I wish to acknowledge my appreciation to certain people.

To commence with, I pay my obeisance to GOD, the almighty to have bestowed upon me good health, courage, inspiration, zeal and the light. After GOD, I express my sincere and deepest gratitude to my supervisor I would like to thank my supervisor, **Dr. Basanta Kumar Das**, Director ICAR-CIFRI, for the patient guidance, encouragement and advice he has provided throughout my time as his student. I have been extremely lucky to have a supervisor who cared so much about my work, and who responded to my questions and queries so promptly.

At the outset, I would like to express my appreciation to my advisor **Mr. Sukanta Kumar Pradhan** for his advice during my M.Sc. schedules past two years. As my supervisor, he has constantly forced me to remain focused on achieving my goal. His observations and comments helped me to establish the overall direction f the research and to move forward with investigation in depth.

I owe my most sincere gratitude to **Mr. Ajaya Kumar Rout** who ploughed through several preliminary versions of my text, making critical suggestions and posing challenging questions.I would like to express my deepest gratitude to **Mr. Jitendra maharana**, **Dr. Buddheswar Dehury and Mr. Debashis Panda** their indispensable help while carrying out the *in-silico* experiments.I gratefully acknowledge **Mr. Hirak Jyoti Chakrobarty** for clearing my grass level basics to play with bioinformatics.

I take this opportunity to express my deep sense of gratitude and respectful regards to Mrs. Sucharita Balabantaray Assistant Professor, Dept. of Bioinformatics, O.U.A.T. Mr. Surya Narayan Rath, Assistant Professor, Dept. of Bioinformatics, O.U.A.T., Mrs Sushma Rani Martha, Assistant Professor, Dept. of Bioinformatics, O.U.A.T who taught me and gave me untiring help during my study.

I thank two of my friends Jagruti Mishra and Sheela Rani Udgata. These two friends and co-workers also provided for some much needed humor and entertainment in what could have otherwise been a somewhat stressful laboratory environment.

I would also like to thank all of the members of the ICAR-CIFRI for simultaneously encouraging, guiding, and supporting my research ideas and me.

I owe a lot to my parents, who encouraged and helped me at every stage of my personal and academic life, and longed to see this achievement come true. Above all, I owe it all to Almighty God for granting me the wisdom, health and strength to undertake this research task and enabling me to its completion.

Name of the Student	**:**	**Ms. Sunanda Paramanik**
Admission No	**:**	**05BI/15**
Title of Thesis	**:**	**Elucidating the Molecular Interaction of Zebrafish (*Danio rerio*) PGRP-2 receptor with DAP and LYS-type Peptidoglycans using *in silico* Approaches**
Thesis Submitted for	**:**	**M.Sc. Bioinformatics**
Name of the Department	**:**	**Department of Bioinformatics**
College and University	**:**	**Centre for Post Graduate Studies,** **Orissa University of Agriculture and Technology,** **Bhubaneswar-751003**
Year of submission	**:**	**2017**
Name of the Advisor	**:**	**Mr. Sukanta Kumar Pradhan, HOD** **Department of Bioinformatics,** **Centre for Post-Graduate Studies,** **Orissa University of Agriculture and Technology,** **Bhubaneswar-751003**

ABSTRACTS

Peptidoglycan recognition proteins (PGRPs) are the major constituent of innate immunity. Although PGRPs are structurally conserved through evolution but their involvement in innate immunity are different in vertebrates and invertebrates. They are highly specific for the recognition and in some cases, hydrolyze bacterial peptidoglycans (PGNs). Zebrafish PGRPs have both peptidoglycans lytic amidase activity and broad spectrum bactericidal activity, but far less is known about how these receptors recognize these microbial ligands. Such studies severely hampered by the need to know zPGRP's structural and functional configuration. So as to compensate these, 3D models of zPGRP2 was constructed and conformational and dynamic properties of zPGRP2 were studied. Docking information of microbial ligands such as Muramyl Tripeptide,Tetrapeptide,Pentapeptide-DAP(MTP-DAP,MTrP-DAP,MPP-DAP) –LYS(MTP-LYS,MTrP-LYS,MPP-LYS),Tracheal Cytotoxin(TCT) in Autodock Vina revealed β1, α2, α4, β4, and loops connecting β1−α2, α2−β2, β3−β4, and α4−α5 as the key interacting domains with the binding energy -4.5,-5.3,-5.5,-5.6,-5.3,-5.0,-5.1 for MTP-Dap, MTrP-Dap, MPP-Dap, MTP-Lys, MTrP-Lys, MPP-Lys and TCT respectively. The plasticity of the PGRP binding site revealed by these microbial ligands suggests an intrinsic capacity of the innate immune system to rapidly evolve specificities to meet new microbial challenges.

KEYWORDS

Innate immunity, Zebrafish, microbial ligands, zPGRP2, LYS, DAP

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATION

NIH	:	National Institute of Health
NLM	:	United States National Library of Medicine
NCBI	:	National Centre for Biotechnology Information
BLAST	:	Basic Local Alignment Search Tool
PDB	:	Protein Data Bank
SOPMA	:	Self Optimized Prediction Method with Alignments
DOPE	:	Discrete Optimized Protein Energy
KEGG	:	Kyoto Encyclopedia of Gene and Genomes
AA	:	Amino acid
PRRs	:	Pattern Recognition Receptors
PGRP	:	Peptidoglycan Recognition Protein
NOD	:	Nucleotide Oligomerisation Domain
PAMPs	:	Pathogen Associated Molecular Patterns
MAMPs	:	Microbe Associated Molecular Patterns
Imd	:	Immune Deficiency
PGN	:	Peptidoglycan
GlcNAc	:	N-acetylglucosamine
MurNAc	:	N-acetylmuramic acid
LTA	:	Lipoteichoic acid
LPS	:	Lipopolysaccharide

DAP	:	Diaminopimelic acid
Lys	:	Lysine
MTP	:	Muramyl Tripeptide
MTrP	:	Muramyl Tetrapeptide
MPP	:	Muramyl Pentapeptide
TCT	:	Tracheal Cytotoxin
proPO	:	Prophenoloxidase
XRD	:	X-Ray Diffraction
NMR	:	Nuclear Magnetic Resonance Spectroscopy
IUPAC	:	International Union of Pure and Applied Chemistry
GLIDE	:	Grid Based Ligand Docking with Energetics
RADAR	:	Rapid Automatic Detection and Alignment of Repeats
RMSD	:	Root Mean Square Deviation
RMSdis	:	Root Mean Square Distance
RMSF	:	Root Mean Square Fluctuation
P.E.	:	Potential Energy
EFF	:	Effective Force Field
K.E.	:	Kinetic Energy

INTRODUCTION

CHAPTER I

INTRODUCTION

The constant threat by the invasion of microorganism in hosts like vertebrates, invertebrates along with plants has evolved systems of immune defense to eliminate infective pathogens in the body. These systems categorize itself into two major type, innate and acquired immunity.

Acquired immunity uses a process of somatic cell gene rearrangement to generate an enormous repertoire of antigen receptors that are capable of fine distinctions between closely related molecules. It is concerned with the eradication of pathogens in the late stage of infection as well as the generation of immunological memory. Acquired immunity is well-characterized by specificity (Permits to distinguish among antigens, and their actions are specifically directed against the antigen that initiated the response), Diversity (Antigens are recognized by different lymphocytes), Self / non-self discrimination.

Innate immunity arose in the early multicellular organisms as the first line of defense against pathogens and has remained an essential component of defense mechanisms in all metazoans (Hoffmann and Reichhart, 2002). Innate immunity is mediated by antimicrobial peptides, phagocytes, alternative complement pathway, etc. Although all vertebrates developed acquired immunity, early vertebrates' eggs and embryos, which are laid and develop in water, must solely rely on innate immunity in defense against infections until their acquired immune system develops. It includes all aspects of the host's immune defense mechanisms relies on a limited number of receptors and secreted proteins that are encoded in their mature functional forms by the germ line genes of the host which recognize features common to many pathogens. The main functions are to differentiate self from non-self and initiate highly regulated immune cascades such as the prophenoloxidase (proPO) activating system that finally eliminate pathogens (Cerenius and Söderhäll, 2004).

Innate immunity can discriminate between pathogen and self and start a rapid defensive response through pattern recognition receptors (PRRs) like Nod-like receptors, PGRPs, CD14, Toll-like receptor-2, mannose-binding lectin, RegIIIγ C-type lectin and lysozyme (Royet and Dziarski, 2007), which detect pathogens via the recognition of pathogen/microbe associated molecular patterns (PAMPs/MAMPs), including polysaccharides, lipopolysaccharide (LPS), peptidoglycan (PGN), bacterial DNA, double stranded viral RNA

and other molecules not normally found on the surface of multicellular organisms (Janeway and Medzhitov, 2002, Medzhitov, 2007).

After recognition, the pattern recognition receptors stimulate immune responses by activating cellular reactions, including phagocytosis, and humoral reactions, including the activation of proteolytic cascades in the hemolymph and intracellular signaling pathways in immune-responsive tissues through certain downstreaming events: Firstly, PRRs are highly conserved in evolution to bind unique products of microbial metabolism not produced by the host *i.e.* pathogen-associated molecular patterns (PAMPs) that are essential for the survival of the microorganism and are therefore difficult for the microorganism to alter. Different PRRs react with specific PAMPs, show distinct expression patterns, activate specific signaling pathways, and lead to distinct anti-pathogen responses. Second, PRRs are expressed constitutively in the host and detect the pathogens regardless of their life-cycle stage. Third, PRRs are germline encoded, non-clonal, expressed on all cells of a given type, and independent of immunologic memory.

Microbial organisms express distinct cell surface molecules, such as, peptidoglycan (PGN) (thick layer) and lipoteichoic acid (LTA) on Gram-positive bacteria and lipopolysaccharide on Gram-negative bacteria that are structurally indispensable components of the cell wall. These molecules are encoded in the germ line and can distinguish invading pathogens by recognizing microbial surface molecules that are not present on the host cell surface (Janeway, 1989). PGRPs are specific for the MurNAc-pentapeptide fragment of peptidoglycans (Stowell et al., 2010, Dziarski and Gupta, 2006). The first PGRP was discovered from the hemolymph of the silkworm (*Bombyx mori*) by Ashida and coworkers in 1996 (Yoshida et al., 1996).The invertebrate and vertebrate PGRP proteins are highly conserved in their structure and recognize bacteria through their cell wall component, peptidoglycan.

All PGRPs have at least one PGRP domain, which is ~165 amino acids long and is structurally similar to bacteriophage and bacterial type 2 amidases (which catalyse the hydrolysis of amide bonds) (Royet and Dziarski, 2007, Kang et al., 1998, Dziarski and Gupta, 2006, Werner et al., 2000, Christophides et al., 2002, Liu et al., 2001); the PGRP domain constitutes most of the sequence of short PGRPs. Across all animals, the PGRP domains are approximately 42% identical and about 55% similar. Some PGRPs (for example, *Drosophila*

melanogaster PGRP-LF and mammalian PGLYRP3 and PGLYRP4) have two PGRP domains, which are homologous but not identical.

PAMPs include polysaccharides, lipopolysaccharide, peptidoglycan (PGN), bacterial DNA and double stranded viral RNA and other molecules not normally found on the surface of multicellular organisms (Ferrandon et al., 2007). PGNs are located on the surface of virtually all bacteria and, as such, constitute excellent targets for recognition by the innate immune system (Medzhitov and Janeway, 2002, Hoffmann, 2003). PGNs are polymers of alternating N-acetylglucosamine (GlcNAc) and N-acetylmuramic acid (MurNAc) in β (1-4) linkage, cross linked by short peptide stems (Heijenoort, 2001, Doyle et al., 2001). The glycan chains display little variation among different bacterial species. The cross linking peptides are composed of alternating L and D form amino acids and are similar in all Gram-negative bacteria and in Gram-positive bacilli but vary in length and amino acid composition in Gram-positive cocci. According to the residue at position 3 of the peptide stems, PGNs have been divided into two major types: L-lysine-type (Lys-type) and meso-diaminopimelic acid type (Dap-type). Dap-type PGN peptides are usually directly cross-linked, whereas Lys-type PGN peptides are interconnected by a peptide bridge that varies in length and amino acid composition in different bacteria.

Because PGN is essential to bacteria but absent from eukaryotic cells, it makes an ideal indicator for metazoan immune systems of the presence of bacteria within the organism. In addition, the PGN composition and structure can be markedly different between bacterial species, allowing the immune system to further distinguish between different types of intruders. For example, the PGN of Gram-negative and Gram-positive bacilli-type bacteria differs from the PGN of most Gram-positive bacteria by the replacement of lysine (Lys) with meso-diaminopimelic acid (DAP) at the third position in the peptide chain (Mengin-Lecreulx and Lemaitre, 2005). Although most PGRPs bind to peptidoglycan, some PGRPs also bind with high affinity to other microbial molecules, such as lipopolysaccharide (LPS) and lipoteichoic acid; these molecules usually bind outside the peptidoglycan-binding groove (Liu et al., 2000, Tydell et al., 2006, Sharma et al., 2011).

Fish are in intimate contact with their environment, which contains very high concentrations of infectious agents such as bacteria and viruses. However, under normal conditions, fish maintain a healthy state by defending themselves against these potential invaders by complex immune defense mechanisms composed of innate and adaptive immunity (Ellis, 2001).

Adaptive immunity of lower vertebrates such as teleost fish are less sophisticated than that of higher animals. In fish, the immune response is also limited in response time by the temperature constraint on fish metabolism (Ellis, 2001, Du Pasquier, 1982, Magnadóttir, 2006). Therefore, fish rely heavily on innate immunity for initial protection against infectious agents. Teleost fish occupy a key evolutionary position in the development of the immune responses in that they are the earliest class of vertebrates possessing the elements of both innate and adaptive immunity (Whyte et al., 2007). Therefore, fish are a good model to study the basic functions of various components of the innate immune response such as PGRPs. Moreover, understanding the innate defense mechanisms of teleost fish may help to develop strategies for the control of fish disease and help the long-term sustainability of fish farming (Alvarez-Pellitero, 2008).

To date, three main types of PGRPs, i.e. PGLYRP-2, PGLYRP-5, and PGLYRP-6, have been identified in teleosts including channel catfish, grass carp, large yellow croaker, rainbow trout, red drum, rockfish and zebrafish. PGLYRP-2 is homologous to mammal PGLYRP-2, while PGLYRP-5 and PGLYRP-6 are only found in teleosts. However, there is no definitive data on the function of fish PGRPs except that zebrafish PGRPs are bactericidal amidases (Li et al., 2007).These PGRPs are expressed widely in different tissues of adult fish, including in the skin, gills, liver, intestine and pancreas, indicating that they function both in the defence against bacteria that are present in the environment and during systemic infection. The high-level expression of PGRPs in the intestine also indicates a probable role of these proteins in the modulation of the fish gut microbiome.

Zebrafish Pglyrp2, Pglyrp5 and Pglyrp6 are amidases that cleave the bond between N-acetylmuramic acid and alanine in peptidoglycan. They are also bactericidal against both Gram-positive and Gram-negative bacteria. Zebrafish embryos develop in bacteria-contaminated water and rely on innate immunity for host defence. zPglyrp2 is highly expressed in eggs and in developing oocytes and Pglyrp5 is induced during early embryogenesis and is strongly expressed at 72 hours post fertilization. These proteins protect the developing embryo from infection by bacteria that are present in the surrounding water.

OBJECTIVE

- Structural characterization and validation of Peptidoglycan Recognition Protein-2 (PGRP2) of *Danio rerio* by using different Bioinformatics Tools
- To study the 3D architecture of zPGRP2 and the binding mechanism at molecular scale using molecular docking and molecular dynamics simulations Technique

REVIEW OF LITERATURE

CHAPTER II

REVIEW OF LITERATURE

Invertebrates lack an acquired immune system and their defense mechanisms against continuous threats from pathogens are based mainly on innate immune responses (Beutler, 2004, Loker et al., 2004). It has been postulated that primary immune recognition is based on structures common among invading pathogens. Janeway termed this pattern-recognition (Janeway, 1989).

In insects and mammals, a number of proteins which are pattern recognition receptors have been described. The PRRs identified in mammals include C-type lectin, proteins with leucine regions, scavenger receptors, pentraxins, lipid transferases, integrins and complement control proteins (Medzhitov and Janeway, 1997). In Anopheles and Drosophila, the PRRs identified can be divided into six gene families: peptidoglycan recognition proteins (PGRPs), thioester-containing proteins (TEPs), gram-negative binding proteins (GNBP), the multidomain scavenger receptors (SCR), C-type lectins (CTL) and galectins (GALE) (Christophides et al., 2002).

Peptidoglycan recognition proteins (PGRPs or PGLYRPs) is a family of innate immunity molecules that were first identified in insects (Yoshida et al., 1996; Kang et al., 1998) and then in mammals (Kang et al., 1998; Liu et al., 2001), other vertebrates, molluscs, and echinoderms, (Dziarski and Gupta, 2006b). PGRPs are conserved from insects to mammals, recognize bacterial peptidoglycan, and function in antibacterial immunity**.** PGRPs have been identified in mollusks, echinoderms, and vertebrates PGRPs constitute a highly diversified family of proteins present in both insects and mammals, but plants and lower metazoa, including nematodes such as *Caenorhabditis elegans*, do not have PGRPs. PGRP genes usually form clusters that suggest their origin by gene duplication. Many PGRPs are secreted, and some (such as mammalian PGLYRP1, PGLYRP3 and PGLYRP4) form disulphide-linked homodimers or heterodimers (for example, PGLYRP3– PGLYRP4). Some PGRPs (such as *D. Melanogaster* PGRP-LC and PGRP-LF) are transmembrane receptors, and some others are intracellular proteins (such as *Drosophila melanogaster* full-length PGRP-LE (Yoshida et al. 1996; Kang et al. 1998; Ochiai and Ashida 1999; Werner et al. 2000; Liu et al. 2001). The peptidoglycan-binding protein is another strong link between invertebrate and vertebrate immunity, pointing at a common origin, comparable to lysozyme, Toll, and Rel

proteins. PGRP not only is a strong link between vertebrate and invertebrate immunity but also could serve as a link between the pro-phenol oxidase cascade and humoral immunity in insects (Kang et al., 1998).

Insects usually have many PGRP genes. For example, D. melanogaster has 13 PGRP genes that are transcribed into 19 proteins (Kang et al., 1998, Werner et al., 2000), and a mosquito (*Anopheles gambiae*) has seven PGRP genes that are transcribed into nine proteins (Christophides et al., 2002). Based on the size of the mRNA transcript, these genes are usually classified into short and long forms, known as S and L, respectively (FIG. . 1). Mammals have four PGRP genes, PGLYRP1, PGLYRP2, PGLYRP3 and PGLYRP4, which were initially named PGRP-S, PGRP-L, PGRP-Iα and PGRP-Iβ (on the basis of their short (S), long (L) and intermediate (I) transcript lengths) by analogy to insect PGRPs (Liu et al., 2001).Some PGRPs, such as Drosophila PGRP-LC, are transmembrane molecules, whereas most other PGRPs have a signal peptide and are secreted, or do not have a signal peptide and therefore are either intracellular or are secreted by another mechanism.

Most PGRPs have one carboxy-terminal type 2 amidase domain (approximately 165 amino acids-long), which is evolutionary homologous to bacteriophage T7 lysozyme and bacterial type 2 amidases (Kang et al., 1998, Werner et al., 2000, Liu et al., 2001, Dziarski , 2004). It is also called a PGRP domain, because it is longer at its amino terminus than a type 2 amidase domain and contains a PGRP-specific segment not present in type 2 amidases (Kim M-S.et al., 2003). The general structural design of the PGRP domain is similar to that of bacteriophage type 2 amidases in that it has three peripheral α-helices and several central β-strands (Royet and Dziarski, 2007, Dziarski and Gupta, 2006, Kim et al., 2003, Guan et al., 2004). The front face of the molecule has a cleft that forms a peptidoglycan-binding groove, and the back of the molecule has a PGRP-specific segment that is not present in bacteriophage amidases; this region is often hydrophobic and is diverse among various PGRPs, and it might bind additional ligands or other molecules (Kim et al., 2003).

To date, three main types of PGRPs, i.e. PGLYRP-2, PGLYRP-5, and PGLYRP-6, have been identified in teleosts including channel catfish (*Ictalurus punctatus*), grass carp (*Ctenopharyngodon idella*), large yellow croaker (*Pseudosciaena crocea*), rainbow trout (*Oncorhynchus mykiss*), red drum (*Sciaenops ocellatus*), rockfish (*Sebastes schlegeli*) and zebrafish (*Danio rerio*) (Li et al., 2007, Kim et al., 2010, Mao et al., 2010, Li et al., 2012, Jang et al., 2013, Li et al., 2013, Li et al., 2014, Yu et al., 2014 Yao et al., 2015).

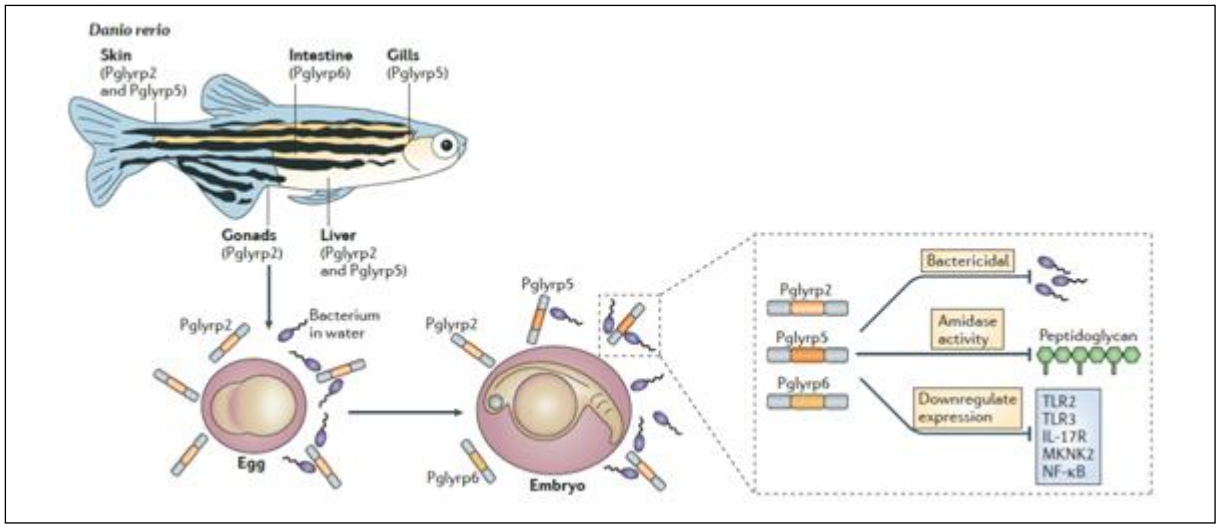

Fig. 1 Zebrafish showing PGRPs

PGRPs have been classified in two groups according to their enzymatic activity. Indeed, some PGRPs have lost their ancestral amidase activity (non-catalytic PGRPs), while others can still efficiently cleave PGN (catalytic PGRPs). Non-catalytic PGRPs have been implicated in functions as diverse as immune receptors, regulators and effectors (Royet and Dziarski, 2007). Catalytic PGRPs have been shown to down-regulate the immune response in insects, act as pro-inflammatory cytokines in mice, and have bactericidal activity in zebrafish (Zaidman-Rémy et al., 2006, Bischoff et al., 2006, Li et al., 2007, Saha et al., 2009, Saha et al., 2010). Moreover, catalytic PGRPs also participate in the establishment of symbiotic interactions in both squid and insects by preventing the activation of an immune response by their bacterial symbiont (Anselme et al., 2006, Troll et al., 2009, Wang et al., 2009, Troll et al., 2010). In addition, as a result of the well conserved key motifs, including Zn^{2+} binding sites and amidase catalytic sites, some PGRPs can also function as N-acetylmuramoyl-L-alanine amidases (NAMLAA) to cleave the lactylamide bond between muramic acid and the peptide chain in bacterial PGN (Gelius et al., 2003, Zaidman-Rémy et al., 2006, Coteur et al., 2007), and exhibit bactericidal activity like T7 lysozyme (Gelius et al., 2003, Kim et al., 2003, Mellroth et al., 2003). Furthermore, several PGRPs such as PGRPs from seastar Asterias Rubens and Drosophila are proved to serve as opsonins to induce agglutination or phagocytosis (Rämet et al., 2002, Coteur et al., 2007).

Several PGRPs have been shown to bind peptidoglycan (PG) (Yoshida et al. 1996; Werner et al. 2000; Takehana et al. 2002; Kim et al. 2003), an essential and unique cell-wall polymer found in both Gram-positive and Gram-negative bacteria. PGNs are polymers of alternating

sugars, N-acetylglucosamine (GlcNAc) and N-acetylmuramic acid (MurNAc) in β (⊢>4) linkage, crosslinked by short peptide stems (van Heijenoort, 2001, Doyle et al., 2001).

PGRPs kill bacteria by activating two-component systems

It has been hypothesized that PGRPs kill bacteria by inhibiting peptidoglycan synthesis at one of the two extracellular steps (transglycosylation and transpeptidation). This hypothesis was based on observation that bactericidal PGRPs bind peptidoglycan but do not permeabilize cytoplasmic membranes (Dziarski and Gupta, 2006, Lu et al., 2006, Wang et al., 2007) and on a structural analysis of binding of peptidoglycan fragments to PGRPs, which was thought to prevent formation of cross-links between peptidoglycan's peptide stems in the growing cell wall (Cho et al., 2007). Peptidoglycan recognition proteins (PGRPs) kill bacteria by utilizing bacterial two-component stress response systems, such as CssR–CssS in *Bacillus subtilis* and CpxA–CpxR in *Escherichia* coli (Kashyap et al., 2011). These two-component systems consist of a transmembrane sensor and a cytoplasmic regulator; they detect extracytoplasmic misfolded and aggregated bacterial proteins, which are generated under stress, exported from the cell and degraded by proteases (Hyyryläinen et al., 2001, Kohanski et al., 2008).

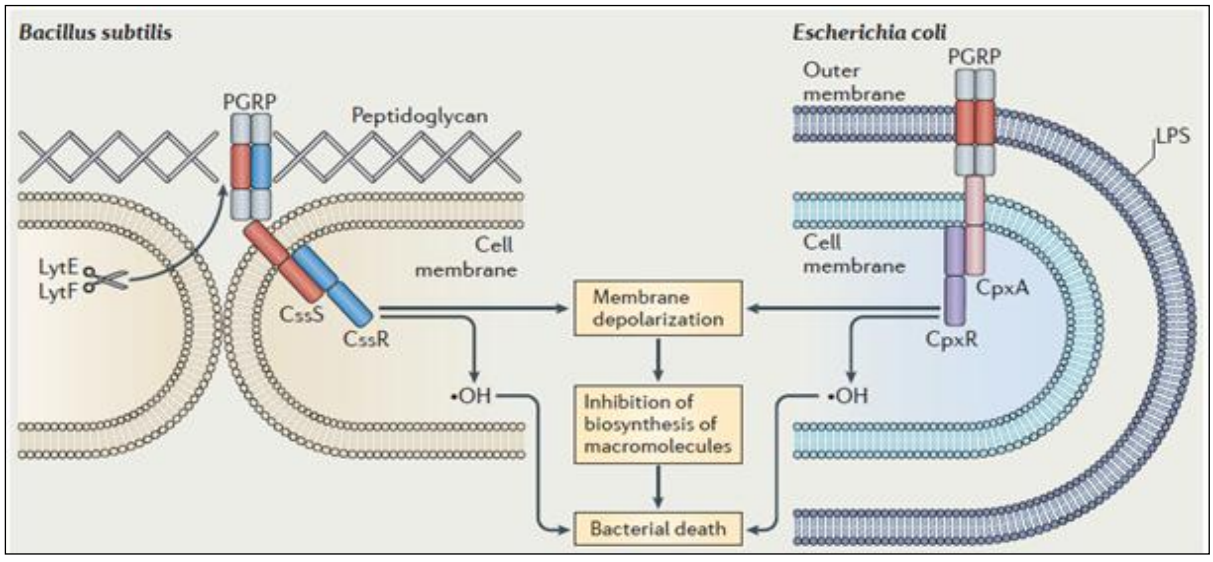

Fig. 2 PGRPs kill bacteria by two component system: CssR-CssR in *Bacillus subtilis* and CpxA-CpxR in *Escherichia coli*

Recently, it has been reported that mammalian PGRPs can prevent aberrant inflammatory responses by modulating the composition of the intestinal bacterial flora, a function in accordance with the strong expression of PGRPs along the digestive tract (Saha et al.,

2010).In view of the significant roles of PGRPs in the innate immune defense, study of PGRPs has also been performed in mollusc animals recently, though compared with that in arthropods and mammals, it is still in its infancy. Two short PGRPs, CfPGRP-S1 and AiPGRP, have been identified from scallop *Chlamys farreri* and *Argopecten irradians* respectively, and both of them can be induced by the stimulation of PGN (Yang et al., 2010, Rämet et al., 2002). Zebrafish Pglyrp5 and Pglyrp6 might also affect multiple intracellular pathways. Inhibition of Pglyrp5 expression in the developing embryo using small interfering RNA modified the expression of genes involved in several pathways, including immune and inflammatory responses, signalling pathways, transcription and metabolism (Chang et al., 2008, Chang et al., 2009).

Insect PGRPs activate the Toll or IMD signal transduction pathways or induce proteolytic cascades that lead to the protection of insects from infections by generating antimicrobial products and inducing phagocytosis (Charroux et al., 2009). By contrast, mammalian PGRPs do not act through host signaling pathways but are directly bactericidal (Cho et al., 2005, Dziarski and Gupta, 2006). Almost all PGRPs have two closely spaced conserved cysteines in the middle of the PGRP domain that form a disulfide bond, which is needed for the activity of PGRPs. A mutation in one of these cysteines in Drosophila PGRP-SA (Cys80Tyr) abolishes the ability of PGRP-SA to activate the Toll pathway and to induce a protective response against Gram-positive bacteria (Michel et al., 2001), whereas a mutation in one of these cysteines in human PGLYRP Cys419Ala) abolishes its amidase activity (Wang et al., 2003).The Toll signaling pathway responds mainly to Gram-positive bacteria or fungal infections, which lead to the proteolytic processing of the cytokine-like polypeptide Spätzle. Binding of the cleaved Spätzle to the transmembrane receptor Toll activates an intracellular signaling cascade that results in the degradation of the IκB-like protein Cactus and the nuclear localization of the NF-κB–like proteins Dif and Dorsal, which induce the transcription of several antimicrobial peptide (AMP) genes, such as Drosomycin (Lemaitre et al. 1996, 1997; Meng et al. 1999; Rutschmann et al. 2000b; Tauszig-Delamasure et al. 2002; Weber et al. 2003). By contrast, the immune deficiency (Imd) pathway mediates defense reactions against primarily Gramnegative bacteria through different signaling components and regulates the cleavage and activation of another NF-κB– related nuclear factor, Relish, which activates a different set of antimicrobial peptide genes, including Diptericin (Lemaitre et al. 1995; Hedengren et al. 1999; Leulier et al. 2000; Rutschmann et al. 2000a; Vidal et al. 2001). Recently, Leulier and colleagues have shown that the Toll pathway is activated

primarily by lysine-type PG found in most Gram-positive bacteria but responds weakly to DAP-type PG from Gram negative bacteria (Leulier et al., 2003).

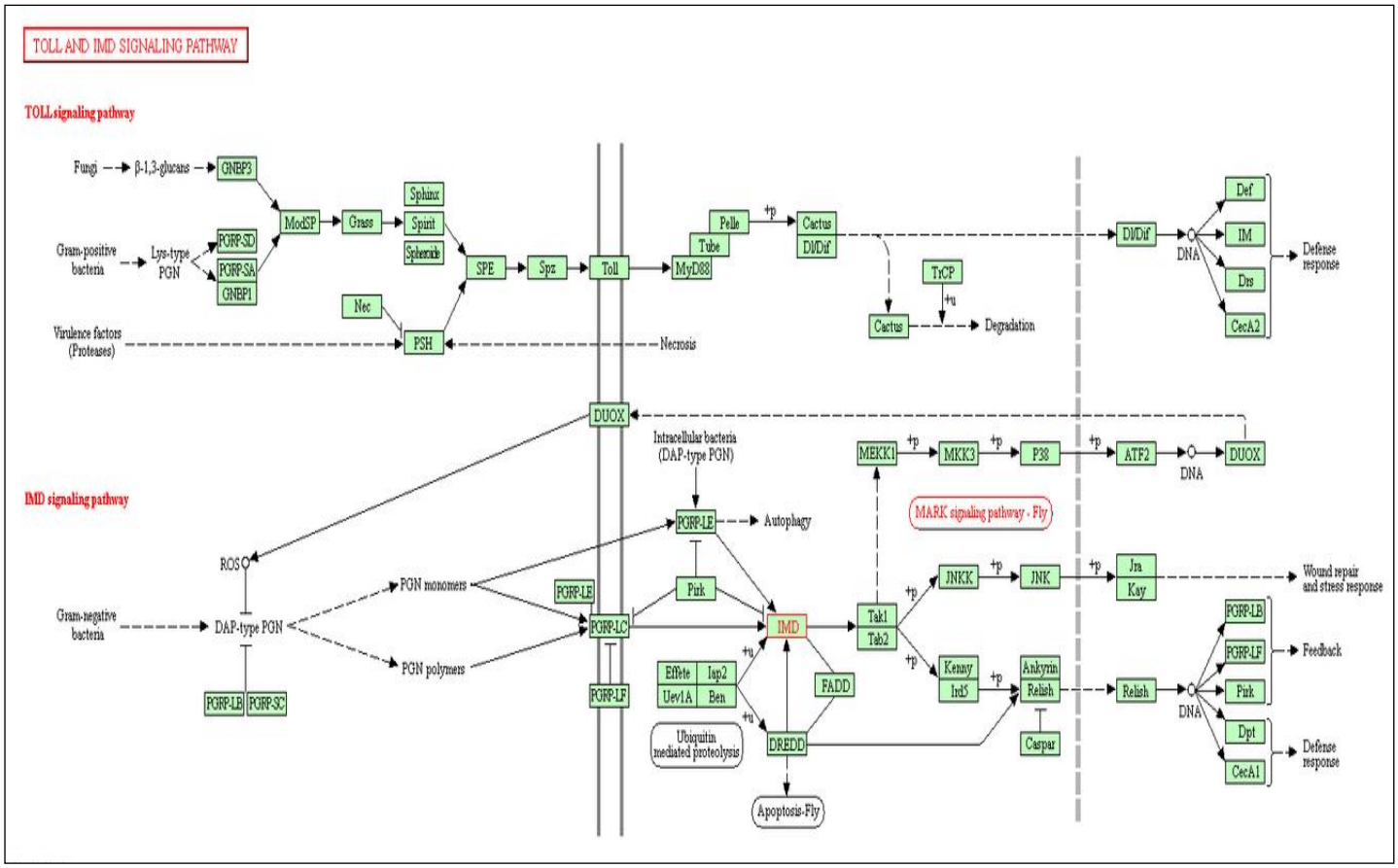

Fig. 3 Toll and IMD pathway in *Drosophila melanogaster*

MATERIALS AND METHODS

CHAPTER-III

MATERIALS AND METHODS

3.1 Materials

3.1.1 Online Databases

3.1.1.1 NCBI (https://www.ncbi.nlm.nih.gov/)

National Center for Biotechnology Information (NCBI) was established on November 4, 1988, as a division of the National Library of Medicine (NLM) at the National Institutes of Health (NIH) by the late senator Claude Pepper. NCBI's mission is to develop new information technologies to aid in the understanding of fundamental molecular and genetic processes that control health and disease with the major databases include GenBank for DNA sequences and PubMed, a bibliographic database for the biomedical literature. Other databases include the NCBI Epigenomics database. All these databases are available online which is accessed through Entrez search engine.

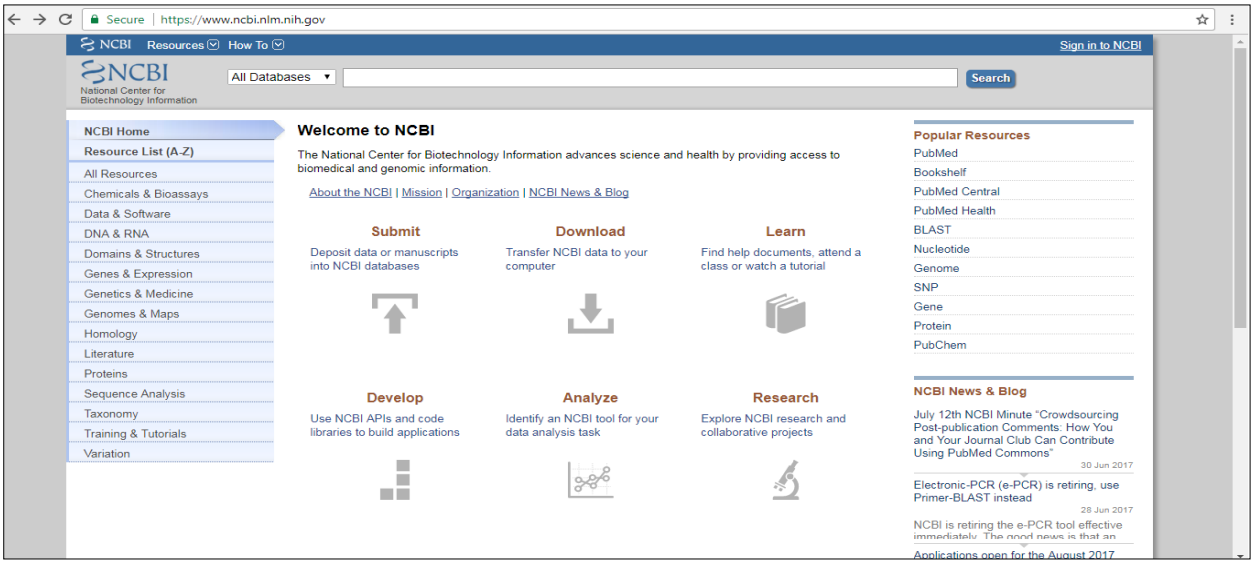

Fig. 4 NCBI homepage

3.1.1.2 PubMed (https://www.ncbi.nlm.nih.gov/pubmed/)

PubMed is a free resource that provides access to MEDLINE, the National Library of Medicine database of citations and abstracts in the fields of medicine, nursing, dentistry, veterinary medicine, health care systems, and preclinical sciences. Relevant information on

PGRP2 and its inhibitors were manually curated from public domain like PubMed, PMC and other bibliographic databases.

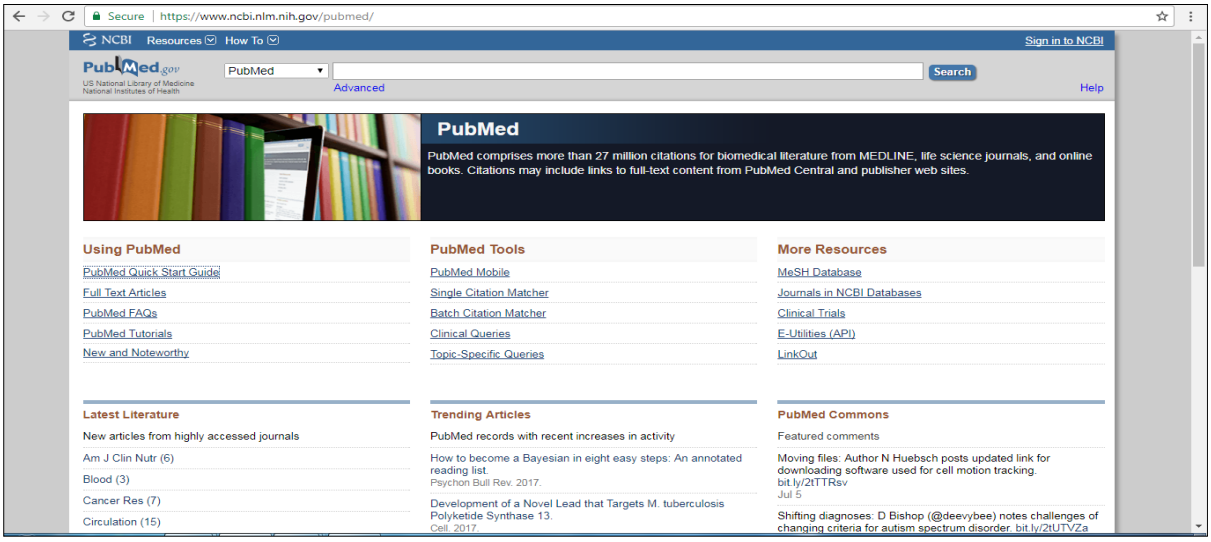

Fig. 5 PubMed Homepage

3.1.1.3 Protein Data Bank (PDB) (https://www.rcsb.org/pdb/home/)

The Protein Data Bank (PDB) archive is the single worldwide repository of information about the 3D structures of large biological molecules, including proteins and nucleic acids. The data typically obtained by X-ray crystallography or NMR spectroscopy and submitted by biologists and biochemists from around the world and it can be accessed for free. The PDB was established in 1971 at Brookhaven National Laboratory under the leadership of Walter Hamilton and originally contained seven structures.

3.1.1.4 UniProt (http://www.uniprot.org/)

The Universal Protein Resource (UniProt) is a comprehensive resource for protein sequence and annotation data. The UniProt databases are the UniProt Knowledgebase (UniProtKB), the UniProt Reference Clusters (UniRef), and the UniProt Archive (UniParc). UniProt is collaboration between the European Bioinformatics Institute (EMBL-EBI), the Swiss Institute of Bioinformatics (SIB) and the Protein Information Resource (PIR). The UniProt Knowledgebase (UniProtKB) is the central hub for the collection of functional information on proteins, with accurate, consistent and rich annotation which consists of two sections: a section containing manually-annotated records with information extracted from literature and

curator-evaluated computational analysis, and a section with computationally analyzed records that await full manual annotation.

3.1.2 *In-Silico* Tools

3.1.2.1 SMART (http://smart.embl.de/)

SMART (Simple Modular Architecture Research Tool) is a web resource providing simple identification and extensive annotation of protein domains and the exploration of protein domain architectures. In the current version, SMART contains manually curated models for more than 1200 protein domains. The underlying protein databases were synchronized with UniProt, Ensembl and STRING, bringing the total number of annotated domains and other protein features above 100 million. SMART's Genomic mode, which annotates proteins from completely sequenced genomes was greatly expanded and now includes 2031 species.

Fig. 6 SMART Homepage

3.1.2.2 Pfam (http://pfam.xfam.org/)

The Pfam database is a large collection of protein domain families. Each family is represented by multiple sequence alignments and a hidden Markov model (HMMs). Each Pfam family, often referred to as a Pfam-A entry, consists of a curated seed alignment containing a small set of representative members of the family, profile hidden Markov models (profile HMMs) built from the seed alignment, and an automatically generated full alignment, which contains all detectable protein sequences belonging to the family, as defined by profile HMM searches of primary sequence databases.

3.1.2.3 CDD (https://www.ncbi.nlm.nih.gov/cdd/)

The Conserved Domain Database is a resource for the annotation of functional units in proteins. Its collection of domain models includes a set curated by NCBI, which utilizes 3D structure to provide insights into sequence/structure/function relationships.

3.1.2.4 PRALINE (http://www.ibi.vu.nl/programs/pralinewww/)

PRALINE is a multiple sequence alignment program with many options to optimise the information for each of the input sequences; e.g. homology-extended alignment, predicted secondary structure and/or transmembrane structure information and iteration capabilities.

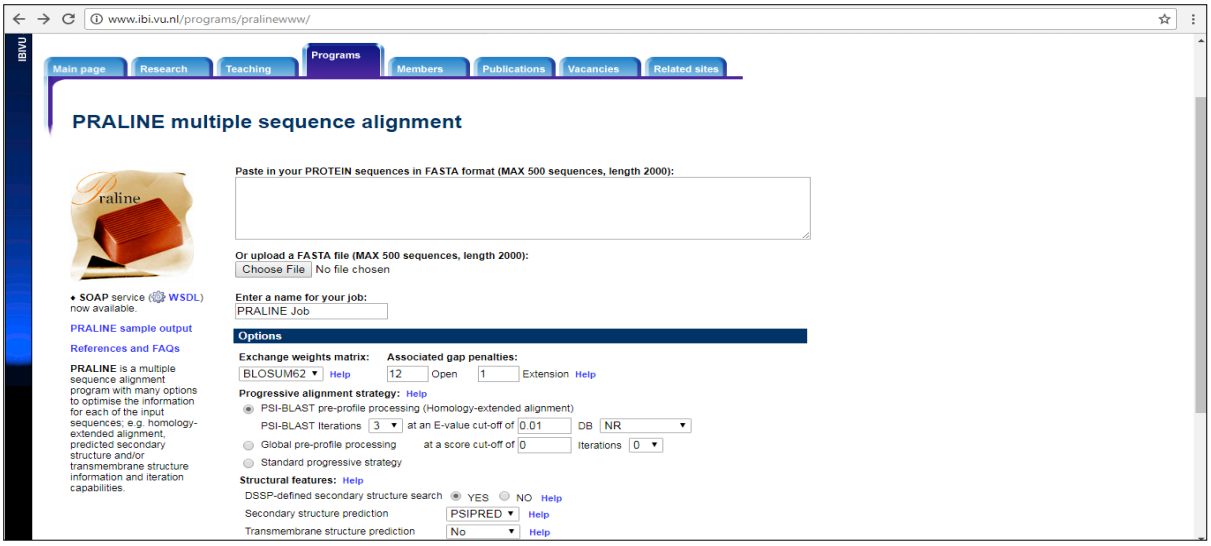

Fig. 7 PRALINE Homepage

3.1.2.5 RADAR (http://www.ebi.ac.uk/Tools/pfa/radar/)

RADAR stands for Rapid Automatic Detection and Alignment of Repeats in protein sequences. RADAR identifies gapped approximate repeats and complex repeat architectures involving many different types of repeats. The algorithm takes in a single query sequence and returns an explicit multiple alignment of repeats found in the sequence, an automatic algorithm, RADAR, for segmenting a query sequence into repeats.

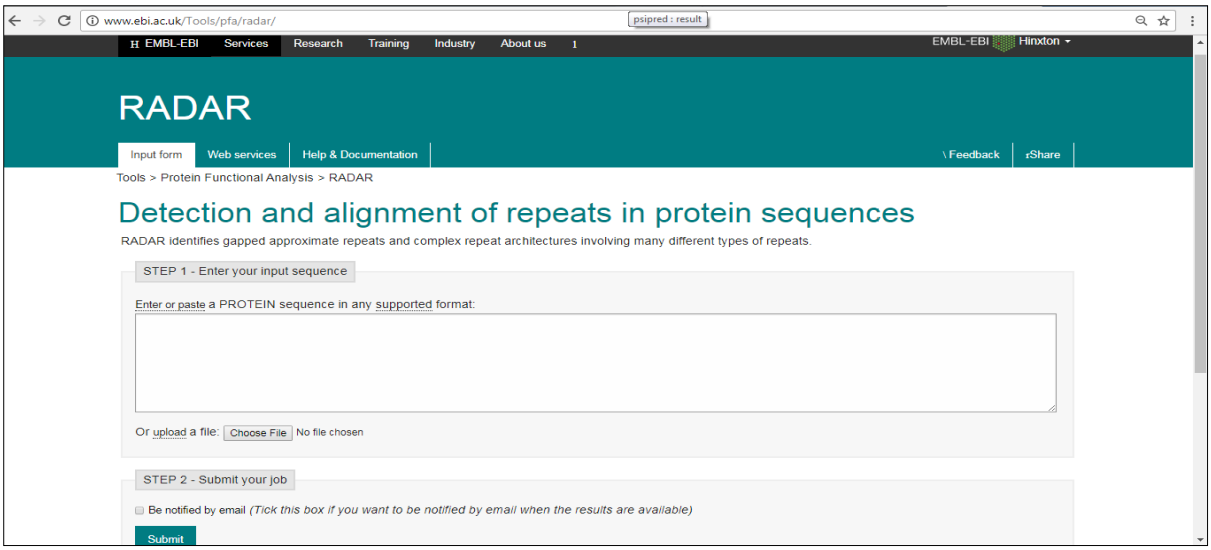

Fig. 8 RADAR Homepage

3.1.2.6 PROTPARAM (http://www.expasy.org/tools/protparam.html)

ProtParam computes various physicochemical properties that can be deduced from a protein sequence either from Swiss-Prot or TrEMBL or for a user entered protein sequence. The computed parameters include the molecular weight, theoretical pI, amino acid composition, atomic composition, extinction coefficient, estimated half-life, instability index, aliphatic index and grand average hydropathicity (GRAVY).

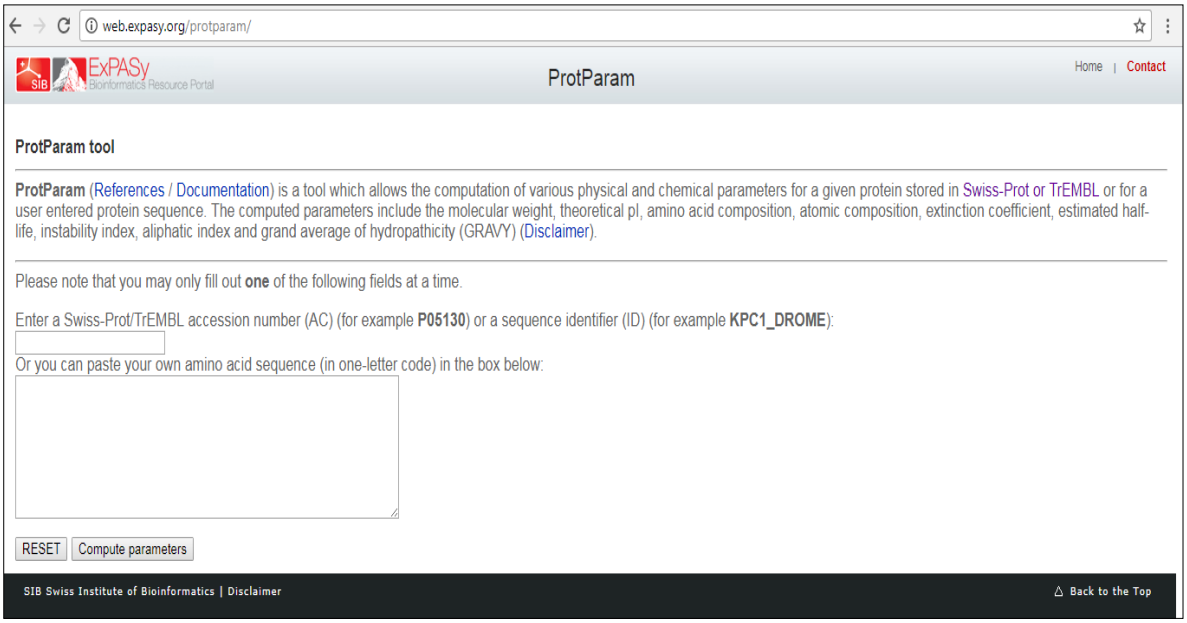

Fig. 9 PROTPARAM Homepage

3.1.2.7 PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred/)

PSIPRED is a simple and accurate secondary structure prediction method, incorporating two feed-forward neural networks which perform an analysis on output obtained from PSI-BLAST (Position Specific Iterated - BLAST). Using a very stringent cross validation method to evaluate the method's performance, PSIPRED 3.2 achieves an average Q_3 score of 81.6%.in the server PSIPRED incorporated with other servers like DISOPRED, DOMPRED, MEMSAT-SVM etc.

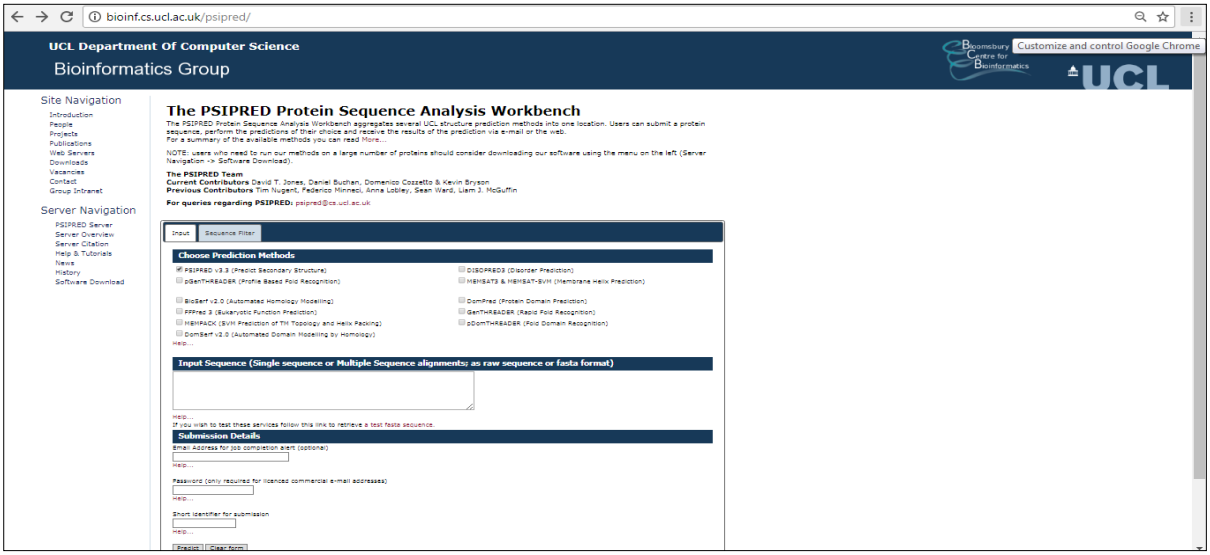

Fig. 10 PSIPRED Homepage

3.1.2.8 PDBsum (http://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/)

The PDBsum is a pictorial database that provides an at-a-glance overview of the contents of each 3D structure deposited in the Protein Data Bank (PDB).It shows the molecule(s) that make up the structure (ie protein chains, DNA, ligands and metal ions) and schematic diagrams of their interactions. Extensive use is made of the freely available RasMol molecular graphics program to view the molecules and their interactions in 3D.

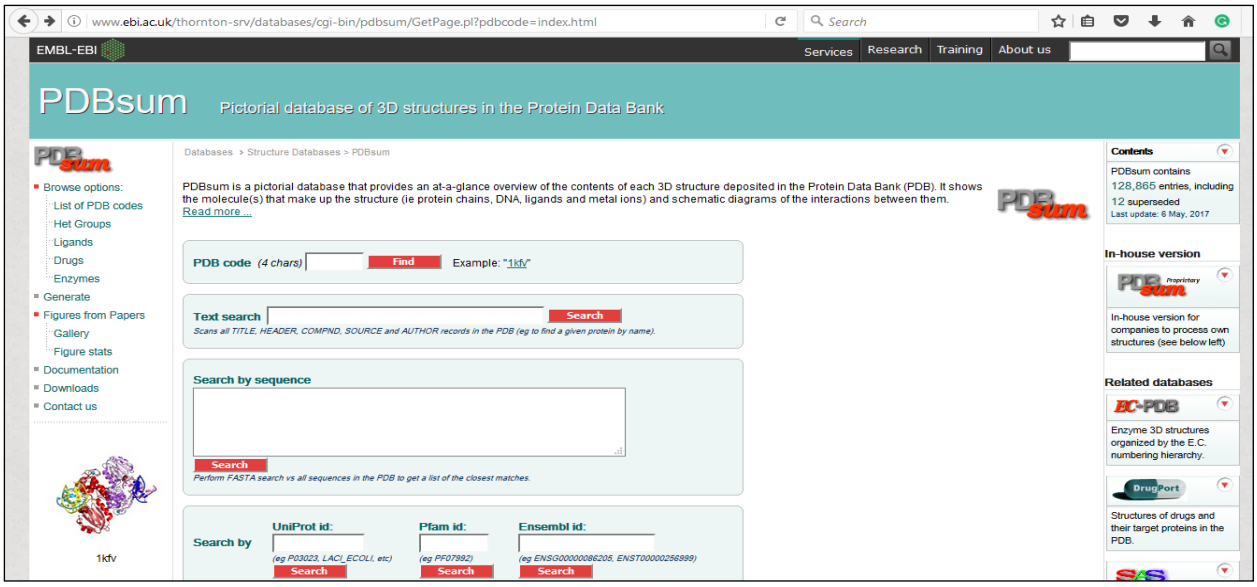

Fig. 11 PDBsum homepage

3.1.2.9 DiANNA1.1 (http://clavius.bc.edu/~clotelab/DiANNA/)

DiANNA (DiAminoacid Neural Network Application) is a web server that provides two services:

i. cysteine classification prediction:

DiANNA 1.1 determines the cysteine species (free cysteine, half-cystine or ligand-bound) by using a support vector machine (SVM) with degree 2 polynomial kernel for the spectrum representation. Additionally, if a cysteine is predicted to be ligand-bound, then the most likely of the four most common ligands (iron, zinc, cadmium, carbon) is proposed.

ii. Disulfide connectivity prediction:

DiANNA 1.1 determines the disulfide connectivity is predicted using a state-of-the-art method involving a novel architecture neural network. By disulfide connectivity, we mean, for example, in the case of four half-cystines, to determine that (1,2) and (3,4) are the disulfide bonds, or that (1,3) and (2,4) are the disulfide bonds, etc.

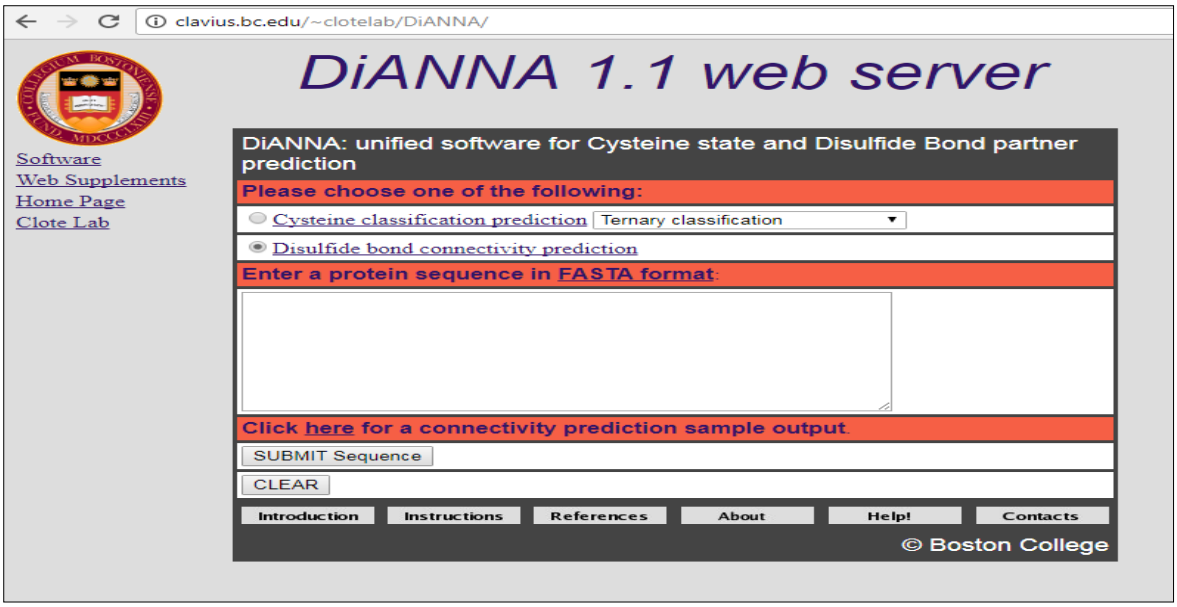

Fig. 12 DiANNA1.1 Homepage

3.1.2.11 ProSA (https://prosa.services.came.sbg.ac.at/prosa.php)

The ProSA program (Protein Structure Analysis) is an established tool which has a large user base and is frequently employed in the refinement and validation of experimental protein structures and in structure prediction and modeling. ProSA program which exploits the advantages of interactive web-based applications for the display of scores and energy plots that highlight potential problems spotted in protein structures. The service specifically addresses the needs encountered in the validation of protein structures obtained from X-ray analysis, NMR spectroscopy and theoretical calculations.

3.1.2.12 ProQ (http://www.sbc.su.se/~bjornw/ProQ/ProQ.html)

ProQ is a neural network based predictor that based on a number of structural features predicts the quality of a protein model. ProQ is optimized to find correct models in contrast to other methods which are optimized to find native structures. Two quality measures are predicted LGscore and Maxsub. LGscore is -log of a P-value and MaxSub ranges from 0-1, were 0 is insignificant and 1 very significant. Different ranges of quality:

Correct	Good	Very good
LGscore > 1.5	LGscore > 3	LGscore > 5
MaxSub > 0.1	MaxSub > 0.5	MaxSub > 0.8

3.1.2.13 PROCHECK (http://services.mbi.ucla.edu/PROCHECK/)

The PROCHECK provides a detailed check on the stereochemistry of a protein structure. It analyses the Ramachandran plot quality, peptide bond planarity non-bonded interaction, main chain hydrogen bond energy, Cα chiralities and overall G factor.

3.1.2.14 Verify3D (http://services.mbi.ucla.edu/Verify_3D/)

The three-dimensional (3D) profile of a protein structure is a table computed from the atomic coordinates of the structure that can be used to score the compatibility of the 3D structure model with any amino acid sequence. The correctness of a protein model can be verified by its 3D profile, regardless of whether the model has been derived by X-ray, nuclear magnetic resonance (NMR), or computational procedures. For this reason, 3D profiles are useful in the evaluation of undetermined protein models, based on low-resolution electron-density maps, on NMR spectra with inadequate distance constraints, or on computational procedures.

3.1.2.15 ERRAT (http://services.mbi.ucla.edu/ERRAT/)

ERRAT is a program for verifying protein structures determined by crystallography. Error values are plotted as a function of the position of a sliding 9-residue window. The error function is based on the statistics of non-bonded atom-atom interactions in the reported structure (compared to a database of reliable high-resolution structures).

3.1.3 Software

3.1.3.1 MODELLER 9.18 (https://salilab.org/modeller/)

MODELLER is used for homology or comparative modeling of protein three-dimensional structures. The user provides an alignment of a sequence to be modeled with known related structures and MODELLER automatically calculates a model containing all non-hydrogen atoms. MODELLER implements comparative protein structure modeling by satisfaction of spatial restraints, and can perform many additional tasks, including de novo modeling of loops in protein structures, optimization of various models of protein structure with respect to a flexibly defined objective function, multiple alignment of protein sequences and/or structures, clustering, searching of sequence databases, comparison of protein structures, etc.

3.1.3.2 PYMOL(https://www.pymol.org/)

PyMOL is a stand-alone molecular visualization program that is very popular with protein crystallographers because of the high quality of its rendering, its speed and versatility. A

large percentage of the Fig. ures in journal publications reporting new macromolecular structures are created using PyMOL. Effective use of PyMOL requires knowledge with highly abbreviated menus, and/or a command scripting language and also creation of Warren DeLano.

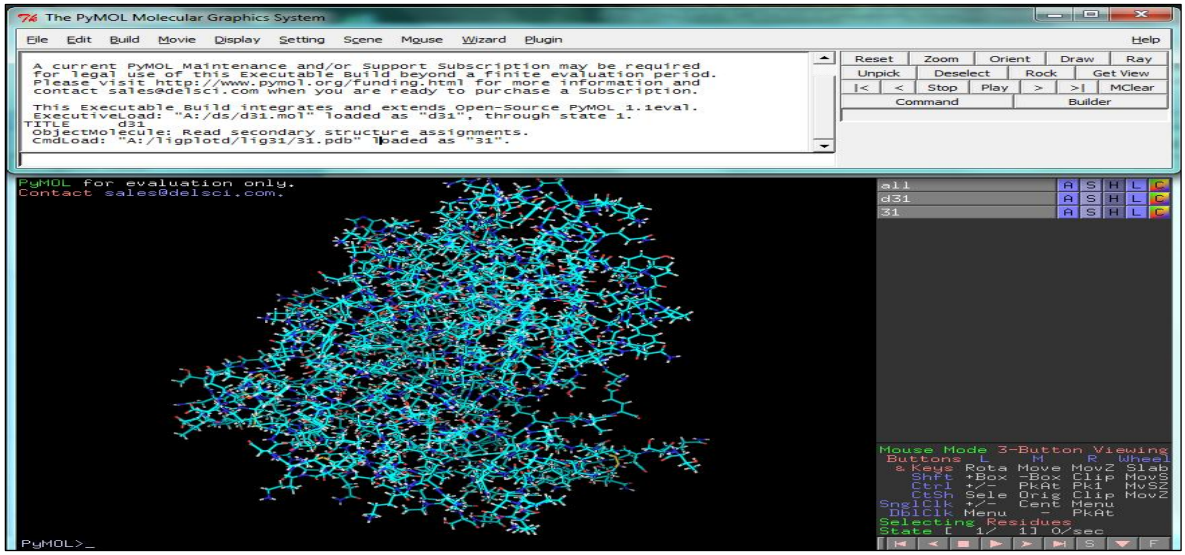

Fig. 13 Pymol viewer

3.1.3.3 LIGPLOT(http://www.ebi.ac.uk/thornton-srv/software/LIGPLOT/)

Ligplot, Automatically generates schematic diagrams of protein-ligand interactions for a given PDB file. These diagrams illustrate the pattern of interactions between the two molecules and are particularly useful for comparing different structures or for studying the interactions between different ligands and the same protein.The program are completely general and will work for any ligand. Indeed, it has also been used for segments of proteins to show, for example, interactions between a helix and the residues in its vicinity.

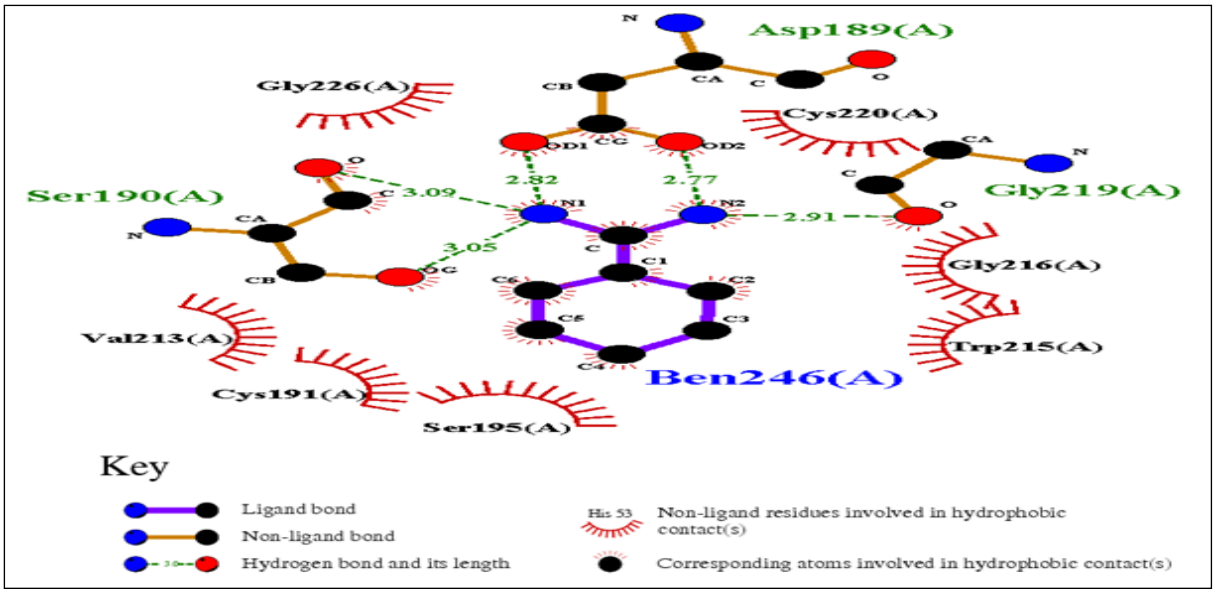

Fig. 14 Ligplot

3.1.3.4 ChemSketch (http://www.acdlabs.com/resources/freeware/chemsketch/)

ChemSketch is a chemical structure drawing program developed by ACD/Labs. Among other features, ChemSketch has the ability to:

- Draw and view structures in 2D or render in 3D to view from any angle
- Draw reactions and reaction schemes and calculate reactant quantities
- Generate structures from InChI and SMILES strings
- Generate IUPAC systematic names for molecules of up to 50 atoms and 3 ring structures
- Predict logP for individual structures
- Search for structures in the built-in dictionary of over 165,000 systematic, trivial, and trade names

ChemSketch uses many standard file formats for the import and export of drawings. The full list of available file formats can be found, under the link "standard file formats." The program allows the user to draw chemical structures including organics, organometallics, polymers, and Markush structures. Users can download a freeware version of the software on the ACD/Labs website. The full version of the software is also available for purchase.

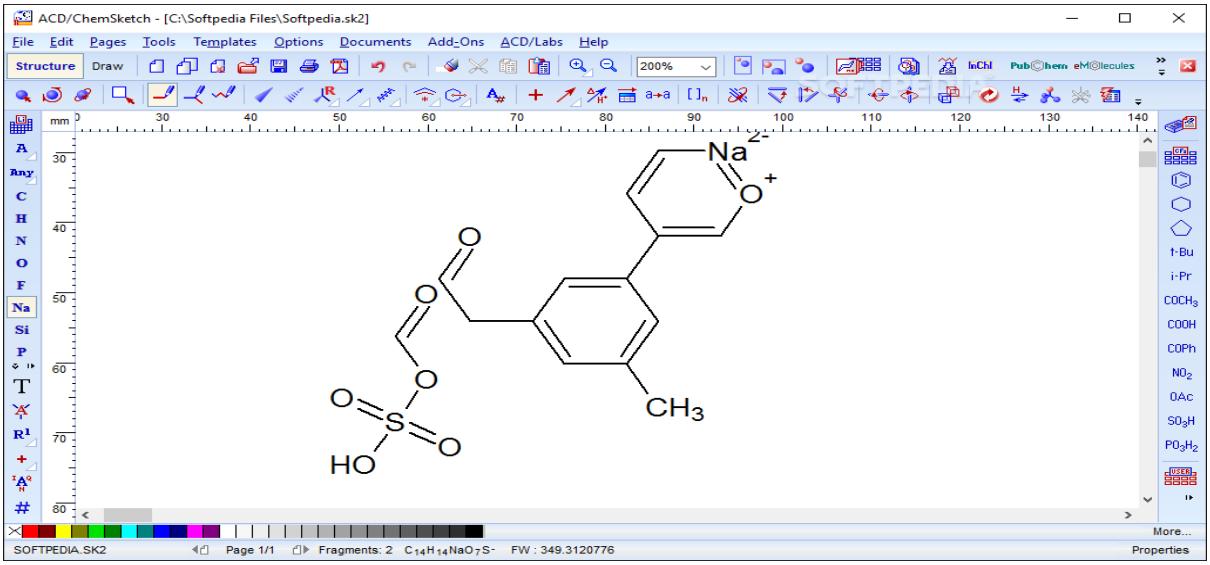

Fig. 15 ChemSketch Homepage

3.1.3.5 AutoDock (http://autodock.scripps.edu/)

AutoDock is a suite of automated docking tools. It is designed to predict how small molecules, such as substrates or drug candidates, bind to a receptor of known 3D structure. Current distributions of AutoDock consist of two generations of software: AutoDock 4 and AutoDock Vina. AutoDock 4 actually consists of two main programs: *autogrid* pre-calculates these grids; *autodock* performs the docking of the ligand to a set of grids describing the target protein. In addition to using them for docking, the atomic affinity grids can be visualized. This can help, for example, to guide organic synthetic chemists design better binders (Morris et al., 2008). Auto Dock Vina does not require choosing atom types and pre-calculating grid maps for them. Instead, it calculates the grids internally, for the atom types that are needed, and it does this virtually instantly.

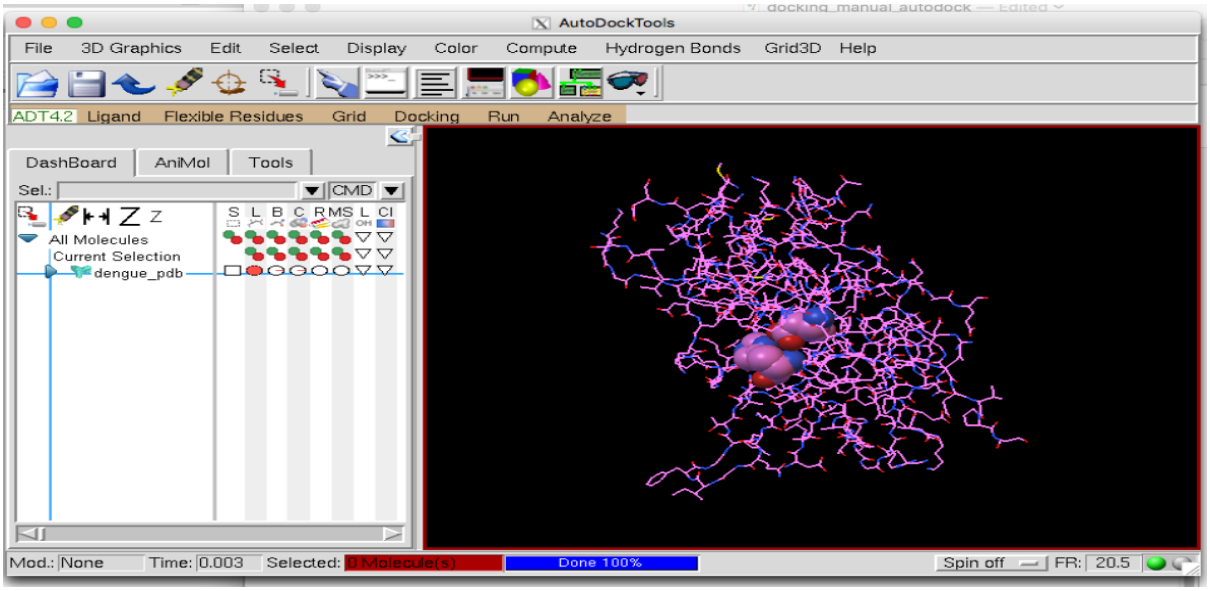

Fig. 16 AutoDock Homepage

3.1.3.6 AUTODOCK VINA (http://vina.scripps.edu/)

Autodock Vina, a turnkey computational docking program based on a simple scoring function and rapid gradient-optimization conformational search.It is a new generation of docking software from the Molecular Graphics Lab. It achieves significant improvements in the average accuracy of the binding mode predictions. AutoDock Vina is an open-source program used for drug discovery, molecular docking and virtual screening, offering multi-core capability, high performance and enhanced accuracy and ease of use. Auto Dock Vina significantly improves the average accuracy of the binding mode predictions compared to Auto Dock 4, judging by our tests on the training set used in AutoDock 4 development. Auto Dock Vina has been tested against a virtual screening benchmark called the Directory of Useful Decoys by the Watowich group, and was found to be "a strong competitor against the other programs and at the top of the pack in many cases". It should be noted that all six of the other docking programs, to which it was compared, are distributed commercially.

3.1.3.6 GROMACS (http://www.gromacs.org/)

GROMACS (GROningen MAchine for Chemical Simulations) is a molecular dynamics package primarily designed for simulations of proteins, lipids and nucleic acids. Generally, the dynamics behaviour of a macromolecular structure is being studied using this software. It was originally developed in the Biophysical Chemistry department of University of Groningen, and is now maintained by contributors in universities and research centres across the world. GROMACS is one of the fastest and most popular software packages available,

and can run on CPUs as well as GPUs. It is free, open source released under the GNU General Public License (Van Der Spoel et al., 2005; Pronk et al., 2013).

3.1.3.7 Grace (http://plasma-gate.weizmann.ac.il/Grace/)

Grace is a free WYSIWYG 2D graph plotting tool, for Unix-like operating systems. The package name stands for "GRaphing Advanced Computation and Exploration of data." Grace uses the X Window System and Motif for its GUI. In 1996, Linux Journal described Xmgr (an early name for Grace) as one of the two most prominent graphing packages for Linux. Here, Grace is used to visualizing the graphs produced by GROMACS.

3.2 METHODOLOGY

3.2.1 Data retrieval and identification of functional domains

The amino acid sequence of PGRP2 (Zebrafish) was retrieved from NCBI protein database and UniProt database with Accession No: NP_001038631.1 and Q1W1Y3 respectively. Domain architecture of zPGRP2 was analysed by SMART (Schultz et al., 1998, Letunic et al. 2012), CDD (Marchler et al., 2016) followed by Pfam (Sammut et al. 2008; Finn et al. 2009) database search which results two overlapping domains PGRP and Ami-2. To detect repeats in protein sequences an automatic algorithm, RADAR (Heger and Holm, 2000) used. RADAR is sufficiently fast to scan databases while producing an explicit multiple alignment of each repeat type found. For the construction of Multiple Sequence Alignment, a total of 3 PGRP sequences from Zebrafish PGRP2, PGRP5, PGRP6 (Accession No: Q1W1Y3,Q1W1Y2,Q1W1Y1 respectively) were retrieved from the UniProt databases were aligned using PRALINE online MSA tool (Simossis et al., 2005).

3.2.2 Protein Sequence Characterization

3.2.2.1 Sequence Based Primary Analysis

The computation of various physical and chemical parameters such as theoretical isoelectric point (PI), molecular weight, total number of positive and negative residues, extinction coefficient (Gill et al., 1989), instability index, aliphatic index and grand average hydropathicity (GRAVY) for the PGRP2 was done ExPASy - ProtParam tool .The presence and location of signal peptide cleavage sites and the glycolysation sites in amino acid sequences was predicted by SignalP 4.1 Server and NetNGlyc 1.0 Server(Gupta et al., 2002).

3.2.2.2 Secondary Structure Analysis

The secondary structure of protein was predicted through PSIPRED (Buchan et al., 2013) and SOPMA (Geourjon and Deleage, 1995) based on the query (primary sequence of a protein). DiANNA 1.1 web server (Ferrè and Clote, 2005) was further used to predict disulfide bonds in the protein.

3.2.2.3 3D model building of zPGRP2

To infer the structure and function of PGRP-2, the PGRP domain region was considered for model building. Comparative modeling of proteins is often considered as a method of choice when a clear relationship of homology between the target and template is found. It is used to generate reliable 3D models of proteins from its primary sequence. Till date there is no crystal structure zPGRP2 has present in the Protein Databank (PDB). Hence, the comparative modelling of zPGRP2 has been carried out to predict protein 3D structure.

Post-secondary structure prediction, the standard modelling procedure was followed. Primarily, the suitable template for PGRP was searched against PDB database using BLASTp search engine (Altschul et al., 1990). The significant template was considered for 3D model building using MODELLER 9.18 (Webb and Sali, 2014). The suitable template was considered on the basis of high sequence identity, query coverage and E-value. A total of 50 rough models of PGRP-2 were obtained from MODELLER 9.18 and the best model was chosen based on discrete optimized protein energy (DOPE) score. The evaluated model of PGRP2 Protein was then considered for further studies.

3.2.3 Model Validation

The obtained crude structure of PGRP-2 was refined using GalaxyRefine (Heo et al., 2013) web server. ProSA-web (Wiederstein et al., 2007) was used to evaluate the energy potential. The model quality of PGRP-2 was calculated using ProQ (Wallner et al., 2003) web server. The analysis of bond lengths and angles were carried out in MolProbity (Chen et al., 2010) web server. The quality of the final energy-optimized model was evaluated with respect to its stereo-chemical geometry and energy. The PROCHECK (Laskowski et al., 1993) provides an idea the stereo-chemical quality of the protein. The ERRAT (Colovos et al., 1993) is a program for verifying protein structure determined by crystallography. It checks the overall quality factor of the protein and was used to check the statistics of non-bonded interactions between different atom types. Verify3D (Eisenberg et al., 1991) was used to access the

compatibility of the atomic models with its own amino acid sequence. A high verify3D profile score indicates the high quality of protein model.

3.2.4 Molecular dynamics of zPGRP2

MD simulations of the zPGRP2 model was carried out using Amber-99sb ILDN force field (Lindorff et al., 2010) in GROMACS v4.6 package (Pronk et al., 2013). All the systems were solvated with TIP3P water models (Jorgensen et al., 1983) in cubic boxes with minimum distance of 15 Å from the protein surface and box edge. Periodic boundary condition was applied to the simulation system. To neutralize the system, a physiological ionic strength (0.15 M NaCl) of counter ions was added. To remove steric conflicts between atoms and to avoid high energy interactions, the electro-neutralized zPGRP2-simualtion system was subjected to energy minimization using the steepest descent algorithm with a tolerance of 1000 kJ mol−1nm−1. The particle mesh Ewald (PME) summation method was used to treat the long-range electrostatic interactions. The final MD was performed for 50ns.

3.2.5 Analysis of MD trajectories

The trajectory analysis was carried out using modules of GROMACS and visual molecular dynamics (VMD 1.9.1) (Humphrey et al., 1996). The backbone root mean square deviation (RMSD), root mean square fluctuation (RMSF) of Cα atoms and radius of gyration (Rg) of proteins (apo) was calculated using gmx rmsd, gmx rmsf and gmx gyrate tool, respectively.

3.2.6 Ligand optimization and Docking

The ligand structures of Muramyl pentapeptide (MPP), Muramyl tetrapeptide (MTrP), Muramyl tripeptide (MTP), Tracheal cytotoxin (TCT) were obtained from the PDB databank with PDB IDs: 2APH, 4KNL 1TWQ and 2F2L respectively. The lysine residue at 3rd position was carefully replaced with Dap followed by ligand optimization and energy minimization in PRODRG2 server (SchuÈttelkopf and Van Aalten, 2004). The final structures of MPP-Dap, MTrP-Dap, MTP-Dap, MPP-Lys, MTrP-Lys and MTP-Lys were cross-checked with the synthetic PGN compounds. The 3D coordinates of these ligands along with TCT were considered for docking simulation studies with zPGRP2.

In the present work, AutoDock Vina (henceforth referred to as ADVINA), a freely available program for academic and commercial applications, was used. ADVINA provides some degree of protein flexibility by allowing predefined side-chain residues to be flexible during docking. ADVINA 1.0.2 (Trott and Olson, 2010) was used to perform docking using all the

flexible and rigid zPGRP-2 conformations. A docking grid with a size of 60×60×60 was used. For each PDB structure, the center co-ordinate was obtained from the central atom of the co-crystallized inhibitor. The grid size was selected in order to encompass the co-crystallized inhibitors (15Å) on each direction. ADVINA presents the docking scores as free energy of binding (DG). The resultant protein-ligands interactions were plotted with pymol and ligplot (2D plot) to show their H-bond, pi-bond and electrostatic interactions.

RESULTS AND DISCUSSIONS

CHAPTER IV

RESULTS AND DISCUSSION

4.1 Sequence retrieval, functional domain prediction, and multiple sequence alignment

Danio rerio Peptidoglycan Recognition protein (PGRP2) sequence (458 amino acids) was retrieved from NCBI protein database (NP_001038631.1). The domain prediction of PGRP2 by SMART and Pfam revealed two putative domains that are overlapping with each other viz., PGRP (285-431) and Ami-2(enzyme domain (297-438). PGRP is a ligand binding domain Conserved domain (CD) search of zPGRP2 revealed that it belongs to super family cl02712 which is homologous to Bacteriophage T3 lysozyme and some bacterial homologues. The detailed domain boundaries are visualized in Fig. 20.

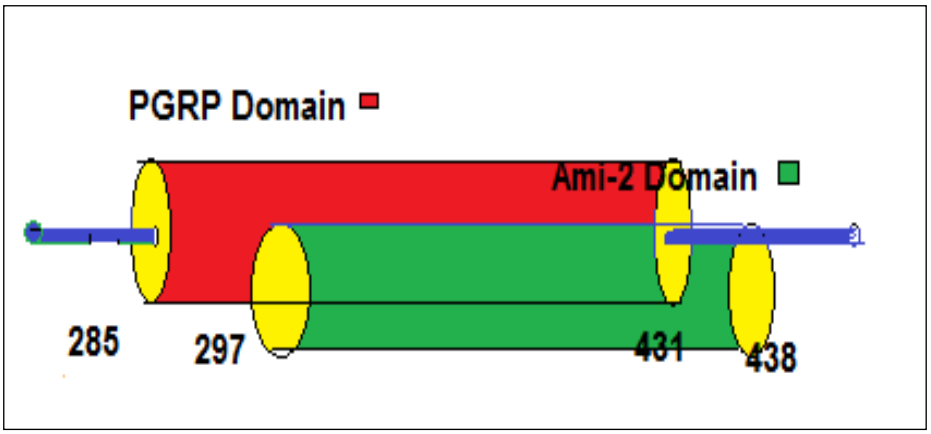

Fig. 17 Domain organization of PGRP and Ami-2 Domain

While performing the multiple sequence alignment through PRALINE server, PGRP2, PGRP5, PGRP6 suggests the homology between the sequences and their effective potentials in zebra fish. The high conserved region in the three domains depicts their strong evolutionary relationships and the hydrophobicity suggests that pgrp5 is more hydrophobic than the other two PGRPs in Zebrafish.

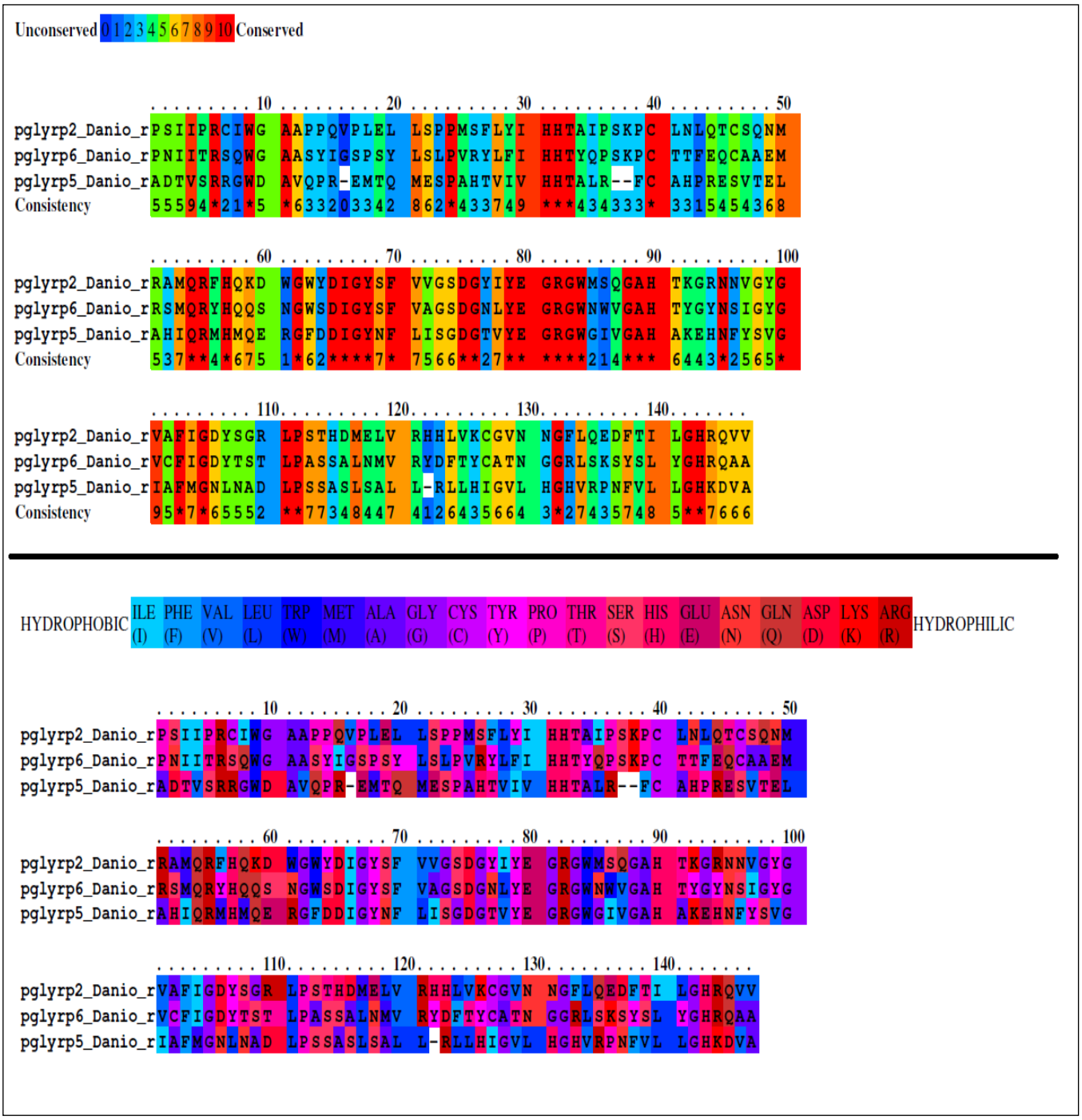

Fig. 18 Multiple sequence alignment of PGRP2, PGRP5, PGRP6 sequences in Zebrafish in PRALINE program. The consistencies of the amino acid conservation percentage in a column are shown in different colors with 0 as completely unconserved and 10 as conserved. (B) The prediction of hydrophobicity from the amino acid sequences for the aligned PGRP sequences Zebrafish.

4.2 Primary sequence analysis and secondary structure prediction

After the domain prediction, the Physico-chemical property of PGRP was computed using ProtParam tool of ExPasy proteomic server; the results are presented in table-1. The molecular weight of PGRP2 (147 aa long) is 53212.8 KDa. Isoelectric point (pI) is the pH at which the surface of protein is covered with charge but net charge of protein is zero. Computational results showed that PGRP2 is basic in nature (pI >7.0) .The computed isoelectric point (8.55) will be useful for developing buffer system for purification by isoelectric focusing method. The aliphatic index (AI) which is defined as the relative volume of a protein occupied by aliphatic side chains is regarded as a positive factor for the increase of thermal stability of globular proteins (Ikai AJ, 1980). Aliphatic index of PGRP2 is 76.87. High aliphatic index of PGRP2 indicates that of *Danio rerio* PGRP2 protein may be stable for a wide range of temperature. The instability index less than40 indicates stability of protein. Hydropathicity Analysis of zPGRP2 revealed that the protein is hydrophilic in nature.

TABLE 1

Characterization of the physico-chemical properties of peptidoglycan recognition protein of *Danio rerio*

Characteristic Properties	**Values**
Number of amino acids	**147 aa**
Molecular weight	**53212.8 KDa**
Theoretical Pi	**8.55**
Total number of positively charged residues (Arg + Lys)	**12**
Total number of negatively charged residues (Asp + Glu)	**10**
The N-terminal of the sequence considered is	**P (Pro)**
Instability index	**36.66**
Aliphatic index	**76.87**
Grand average of hydropathicity (GRAVY)	**-0.248**

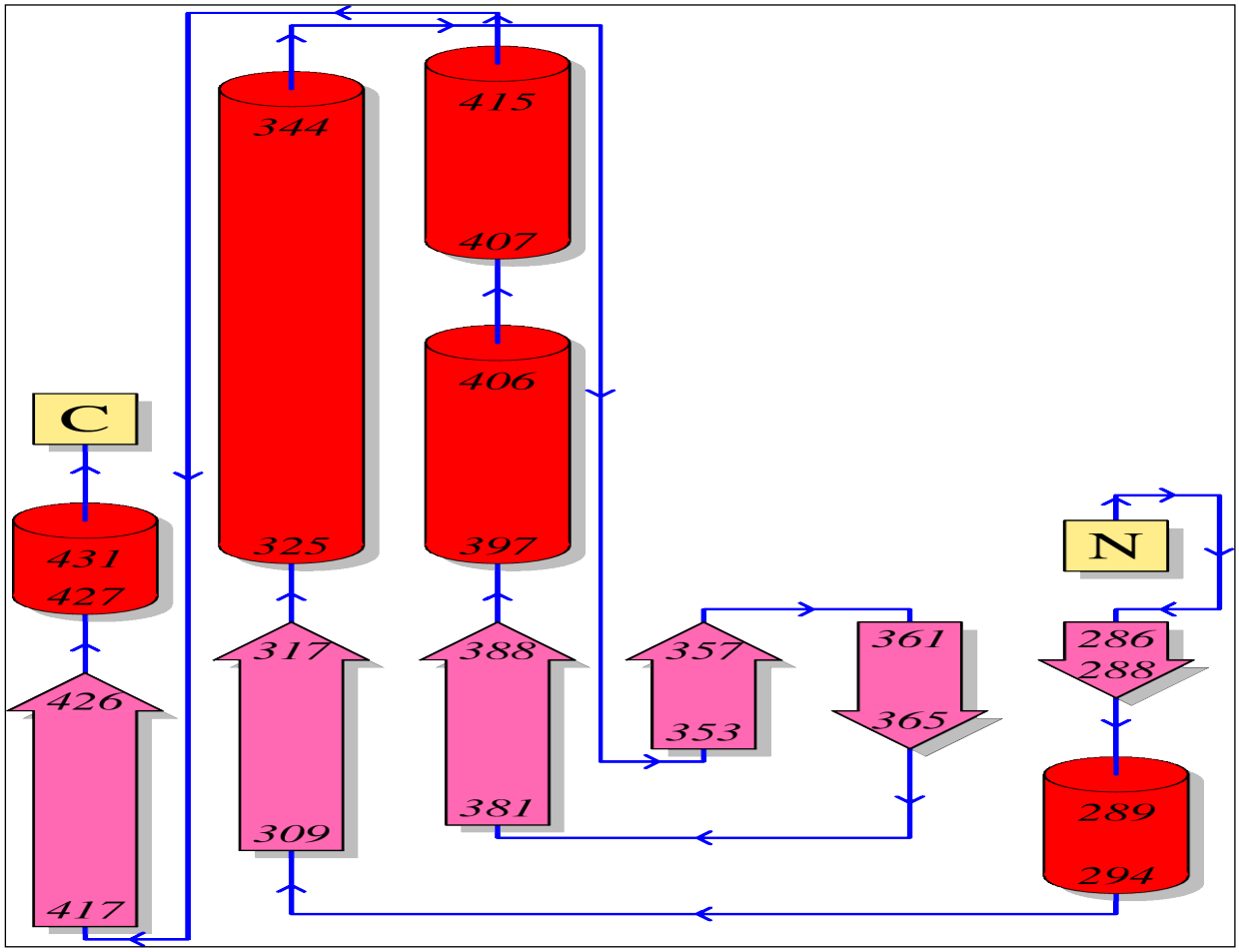

Fig. 19 Topology analysis of PGRP domain of zPGRP2.

Analysis of the sequence with SignalP4.1 tool did not predict any signal peptide in PGRP2 protein. The N-glycosylation sites prediction using NetNGlyc 1.0 server proposed six glycosylation sites, out of which two were highly potential (N79 and N203), two moderate (N51 and N203) , and two (N62and N208) was predicted with low potential. These potential glycosylation sites suggested their role in PGN ligand recognition and bactericidal activity (Guan et al., 2004, Reiser et al., 2004). Disulfide bond prediction at DiANNA showed the presence of one potential bond between C40-C46 residues.

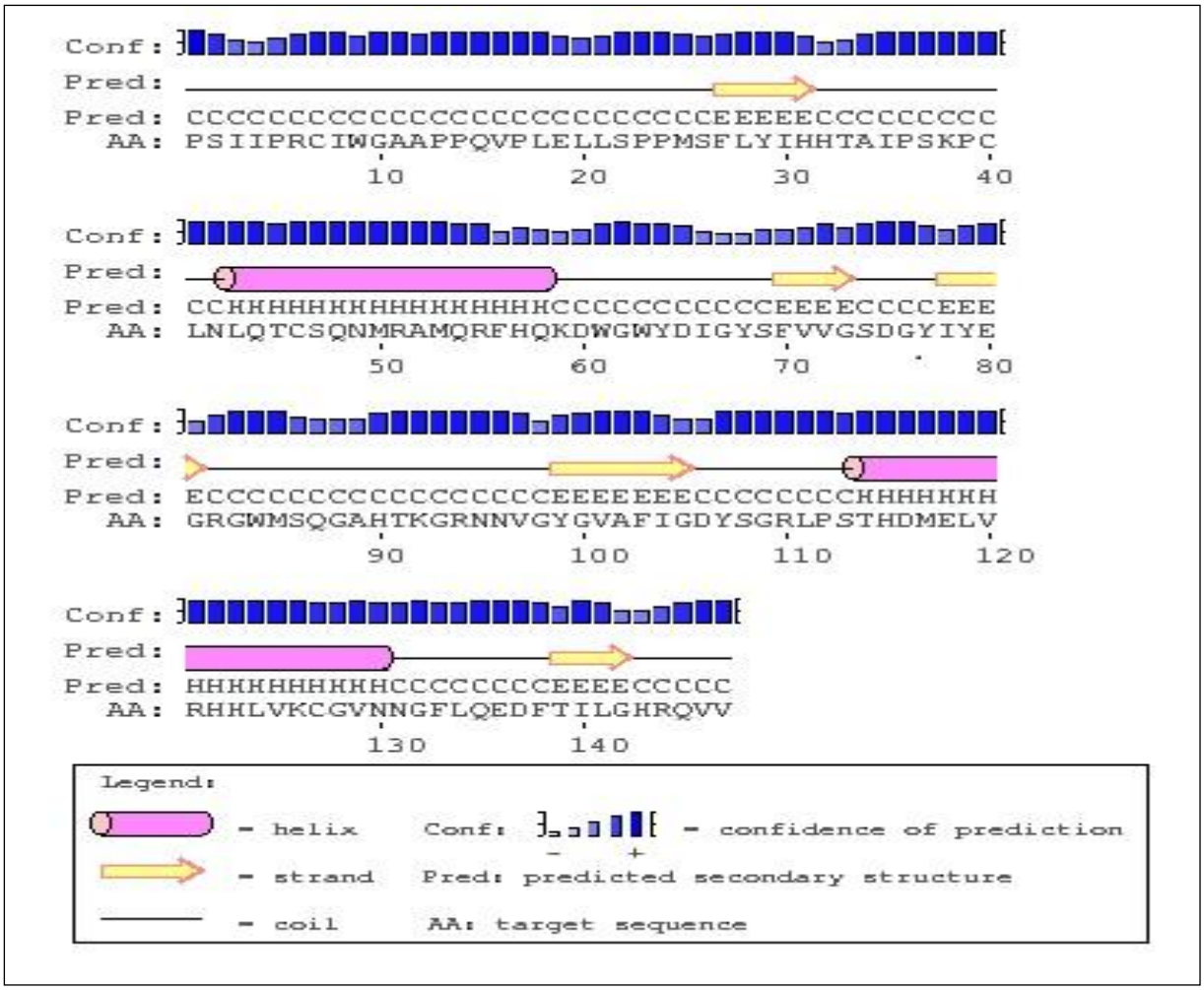

Fig. 20 The Fig. ure indicates predicted secondary structure of zPGRP2. The legends provided in bottom line of the Fig. s indicate the sign of secondary structure elements

The SOPMA showed that the protein contains 14.97% alpha helix, with random coil is 42.86% and beta turn is 14.97%. The extended strand is 27.21%. It also shows that the window width is 17, with a similarity threshold of 8 and number of states. The abundance of coiled region indicates higher conservation.

TABLE 2

Secondary structure analysis through SOPMA of zPGRP2

Secondary Structure	Percentage (%)
Alpha helix (Hh)	14.97
Beta turn (Tt)	14.97
Extended strand (Ee)	27.21
Random coil (Cc)	42.86
310 Helix	0.00
pi helix	0.00
Bend region	0.00

4.3 Comparative modeling Analysis of zPGRP2 protein

Template search for zPGRP2 in PDB using Protein BLAST showed the crystal structure of Chain A, Drosophila Peptidoglycan Recognition Protein (PGRP-SA) (PDB ID: 1SXR_A) shared the highest homology (identity-37%and similarity-59%).considering the above homology, a target template alignment was made and thereafter 50 models built over Modeller9.18, among which best model (model 18) identified by analyzing the lowest DOPE score. The model structure was further refined and energy minimized by GalaxyRefine and presented before different validation servers to optimize its potential.

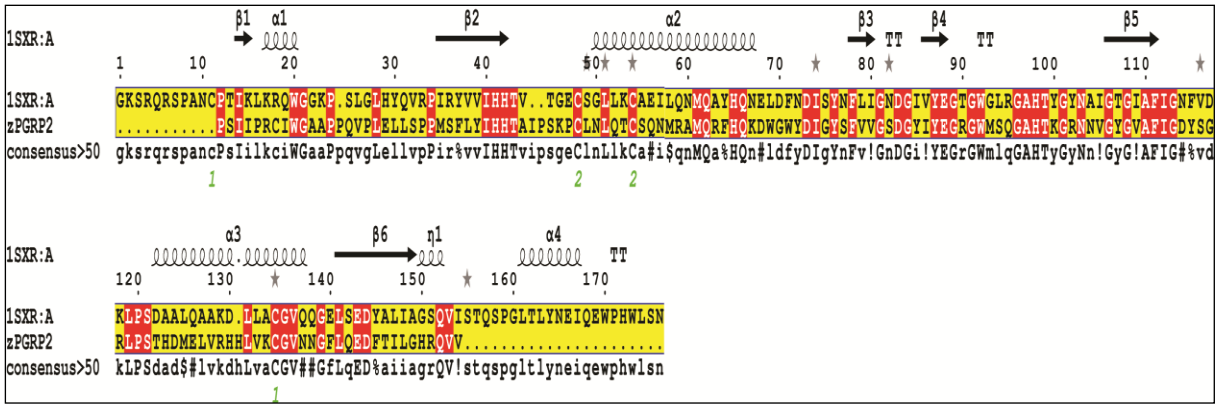

Fig. 21 Target-Template alignment of zPGRP2-1SXR_A

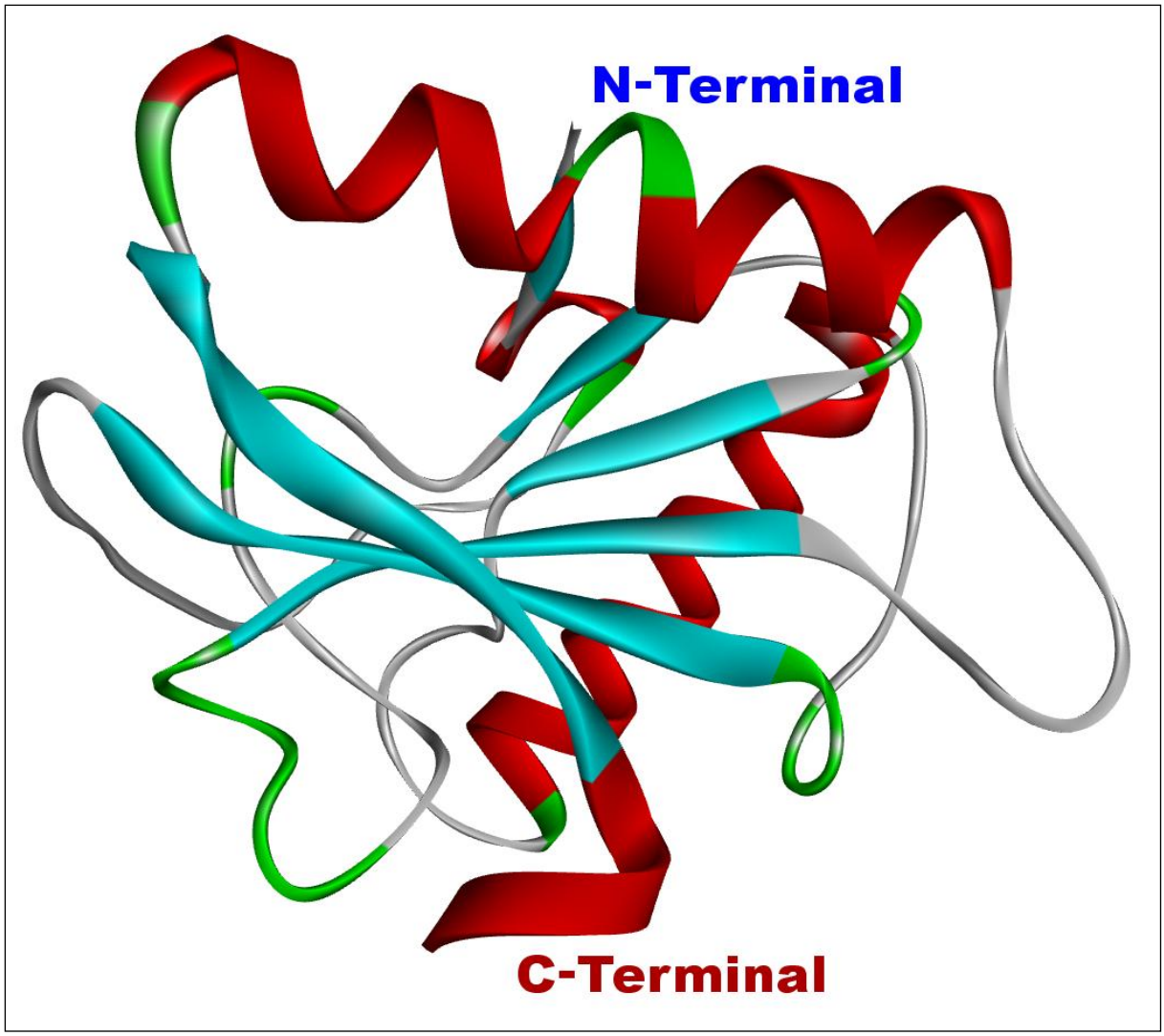

Fig. 22 Three Dimensional structure of zPGRP2

4.4 Molecular dynamics simulation analysis of homology model

To ensure the stability of modeled zPGRP2, MD simulation of the model was performed over a simulation time period of 50 ns. It was well equilibrated during this 50 ns MD simulation as revealed from the potential energy, temperature, and volume analysis. The RMSD of the protein backbone atoms are plotted as a function of time to check the stability of the system throughout the simulation time. The RMSD analysis of the backbone atoms showed an initial upraise till ~10 ns, and thereafter achieved a steady plateau with an average RMSD value of ~ 3.7 Å. The radius of gyration (Rg) analysis gyration radii of 1.52 Å revealed good compactness for the modeled structure. To understand the residual fluctuation of the model, the root mean square fluctuations (RMSF) of the Cα atoms were calculated RMSF analysis of each residue with respect to simulation time period revealed that the maximum residues

shown as the flipping regions with RMSF values (>2Å). The C-terminal loop region in zPGRP2 comprised of 25 residues exhibited the maximum fluctuation (>4Å) as compared to N-terminal region. During the simulation MD simulation highlighting that the model is stable in nature which is stable for further studies.

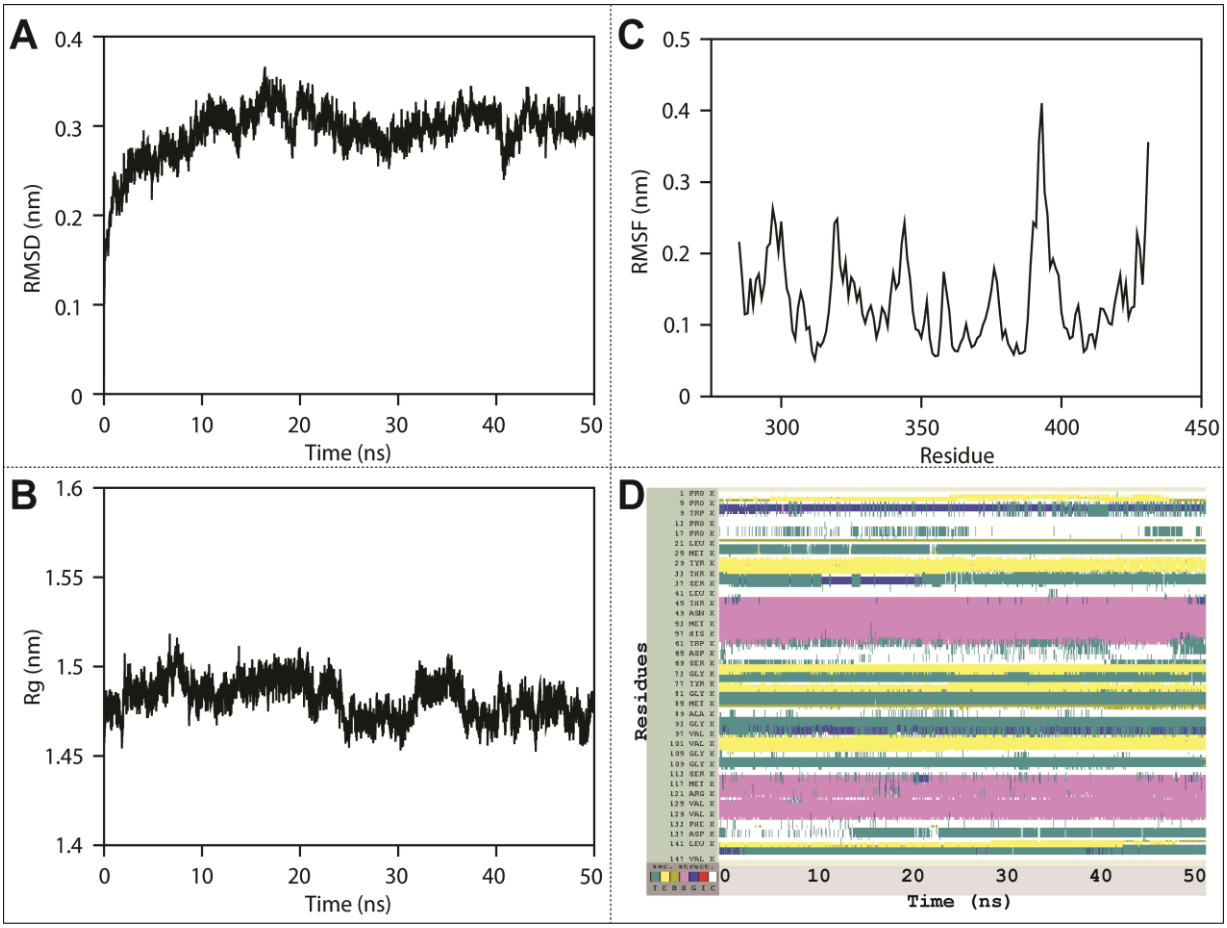

Fig. 23 Stability parameters for Zebrafish PGRP2 homology model during 50 ns MD simulation in Gromacs 4.5.5. (A) RMSD of C-α atoms, (B) Radius of gyration of C-α atoms, (C) RMSF of different amino acid residues and (D) Helix numbers

4.5 Structure Quality Assessment and Evaluation

The simulated zPGRP2 model comprised of 4 α-helices and 5 β-sheets (parallel and 1 anti-parallel).These β-sheets located at the inner surfaces surrounded by the helices. The accuracy of dihedral angles (Φ/Ψ) of zPGRP2 model was measured using Ramachandran plot in Procheck. The refined model showed good percent of residues in most favored regions (112 amino acids, 93.3%), additional allowed regions (6 amino acids, 5.0%) and disallowed regions (2 amino acids. 1.7%).

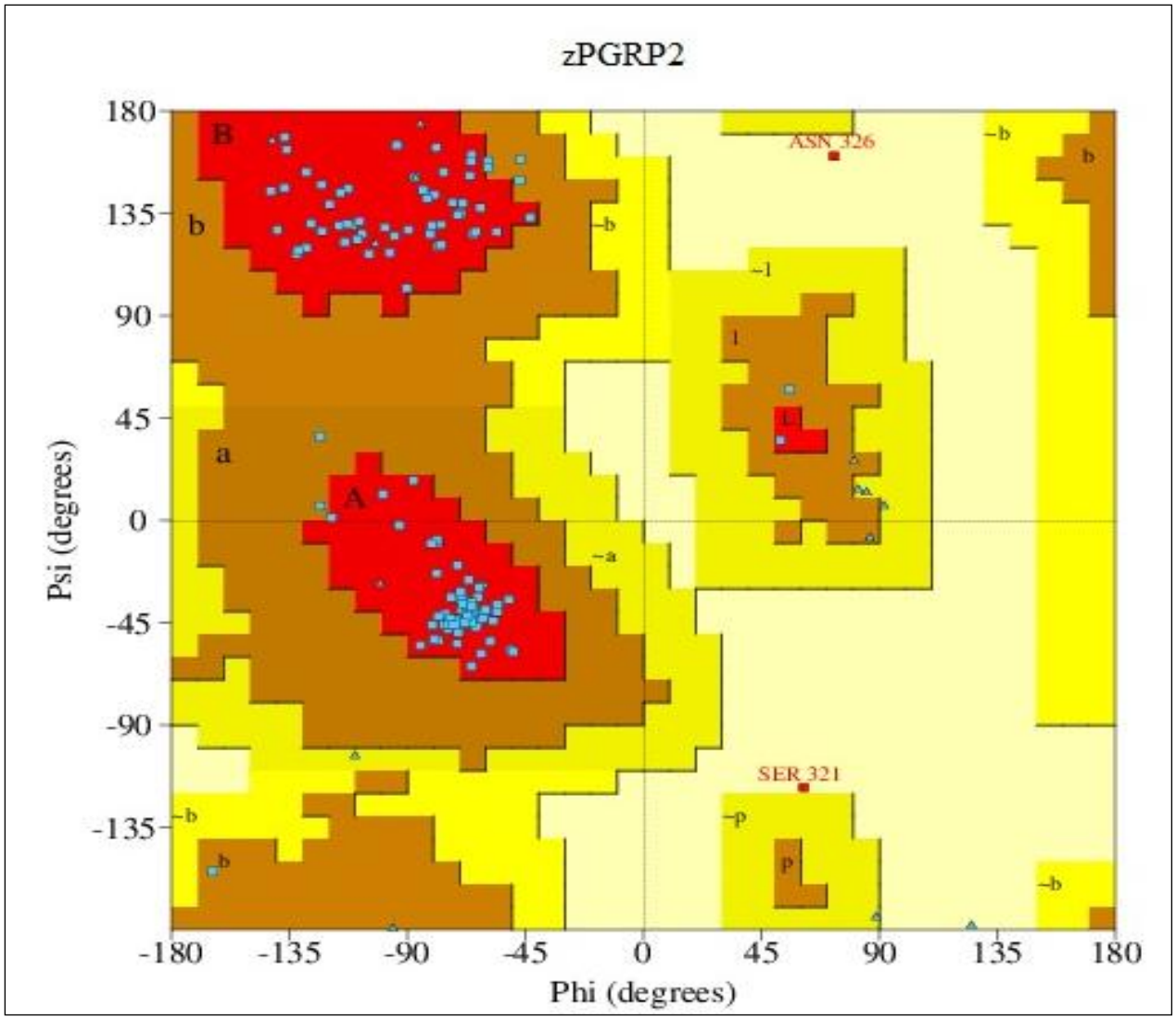

Fig. 24 Ramachandran plot analysis of zPGRP2. Here, red region indicates favored region, yellow for allowed & light yellow shows generously allowed region & white for disallowed region. The phi and psi angles determine torsion angles.

The packing quality of each residue as assessed by the Verify-3D program represents the profile obtained with respect to the residues. Compatibility of the model residues with their environment is assessed by a score function. Residues with a score over 0.2 should be considered reliable. Score for our model zPGRP2 maximally lied above 0.2 which corresponds to acceptable side chain environment. The statistics of non-bonded interactions between different atom types calculated by ERRAT also revealed a good validation score (95.652), and suggesting the reliability of the initial model.

Table 3

Structure Validation of zPGRP2through online server

Servers	Parameters	Validation scores
Procheck (Ramachandran plot)	Most favoured regions (%) Additional allowed regions (%) Generously allowed regions (%) Disallowed regions (%)	93.3 5.0 0.0 1.7
Verify 3D	Averaged 3D-1D score >= 0.2(%)	89.08
Errat	Overall quality (%)	95.652
ProSA	Z score	-5.01
ProQ	LG score Max Sub	3.068 0.258
Prove	Z score mean	0.528

Energy profile of the proposed model and the z-score value was obtained using ProSA program which calculates the interaction energy per residue using a distance-based pair potential. ProSA revealed a z-score value of the model which lies between -5.0 to -9.0 (negative value imply model accuracy) as depicted in Fig. . 25. ProQ analysis indicated that the quality of zPGRP2 was extremely good.

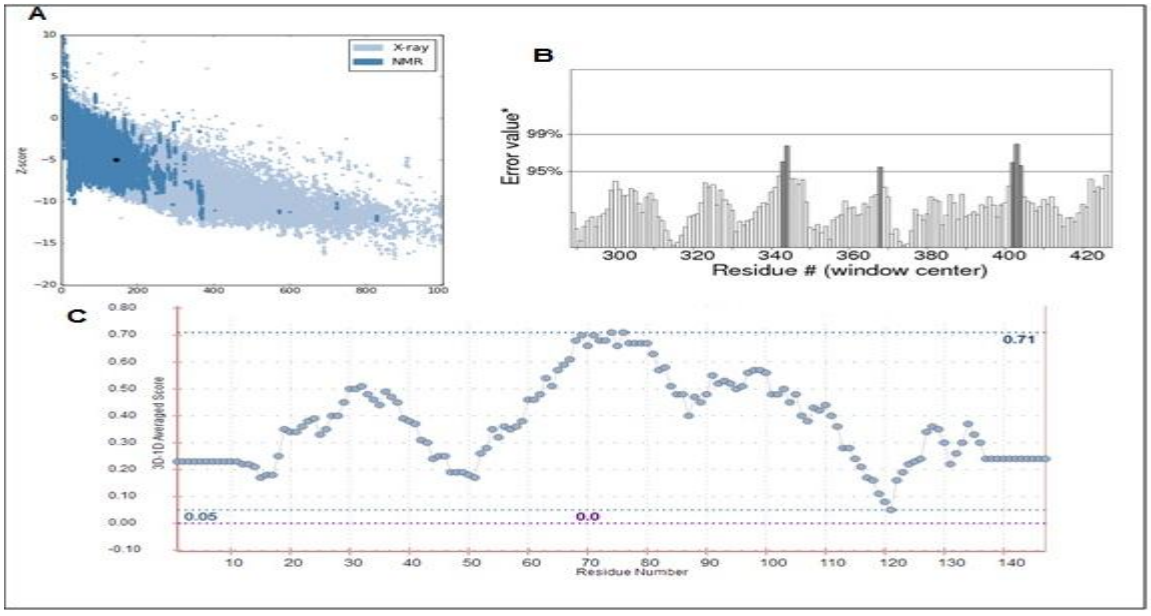

Fig. 25 Validation results showing A. ProSA, B. ERRAT, C. Verify 3D Analysis

4.6 Molecular Docking of protein-ligand complexes

The zPGRP2 consisted of two potential N-glycosylation sites, among these the N79 and N127 were situated on the loop surrounded by the L-shaped cleft. Earlier studies in different PGRP receptors highlighted this region as the potential binding region of PGN compounds. The large molecular size of muramyl-peptides with more than ~ 25 rotatable bonds may not be very accurate because of algorithm limitations. The docking results in AutoDock Vina generated good binding energy for all constructed PGN compounds, and the binding energy obtained in both programs was given in Table . The Dap-type PGN-fragments showed that, the acetylmuramic acid interacts with the C-terminal and loop regions. The polar side group containing Trp63 residue in zPGRP2 was observed to form potential hydrogen bonds with the acetylmuramic acid atoms of all Dap-type PGNs. The His31, Ala89, His143 and Asp60 were also found to be actively engaged in Dap-type ligand interactions.

The binding sites of C-terminal peptide stems in Dap-type ligands were varied significantly with different orientations. The ligand binding pockets and corresponding 2D atomic interaction plots between zPGRP2 and Dap-type ligands were shown in Fig. .26, 28, 30. The Lys-type PGN fragments rendered an approximately same binding orientation for all muramyl peptides. In MTP-Lys and MTrP-Lys, the L-lysine actively bonded with the Arg144 of zPGRP2. The acetylmuramic acid was also observed to interact with the C-terminal loop and helix regions as like as Dap-type PGNs. The 3D and 2D diagrams for Lys-type PGN-fragment docking interactions were illustrated in Fig. 27, 29, 31.

The docking simulation presented both type ligands shared an approximately same interacting sites for the acetylmuramic acid atoms in zPGRP2; however, varied at the peptide stem and 3rd place comprising L-Lys or Dap binding sites. The key binding residues predicted were well compatible with the previous *in silico* and experimental studies Drosophila. The His31 and Arg144 that were actively engaged in acetylmuramic acid binding may be highly responsible for the PGN-fragment orientation as revealed from previous studies (Mellroth et al., 2003,Guan et al., 2005), and upon mutation may affect the stability of the complex. This indicated, in Zebrafish PGRP2, the His31 and Arg144 residues may be highly essential to mediate the bactericidal activity and PGRP2-sense innate immune signaling.

TABLE 4

The binding energy table of PGRP2-LIGAND interaction

Ligands	Binding energy
Muramyl Pentapeptide-DAP (MPP-Dap)	-5.5
Muramyl Pentapeptide-LYS (MPP-Lys)	-5.0
Muramyl Tripeptide-DAP (MTP-Dap)	-4.5
Muramyl Tripeptide-Lys (MTP-Lys)	-5.6
Muramyl Tetrapeptide-DAP (MTr-Dap)	-5.2
Muramyl Tetrapeptide-LYS (MTr-Lys)	-5.3
Tracheal Cytotoxin (TCT)	-5.1

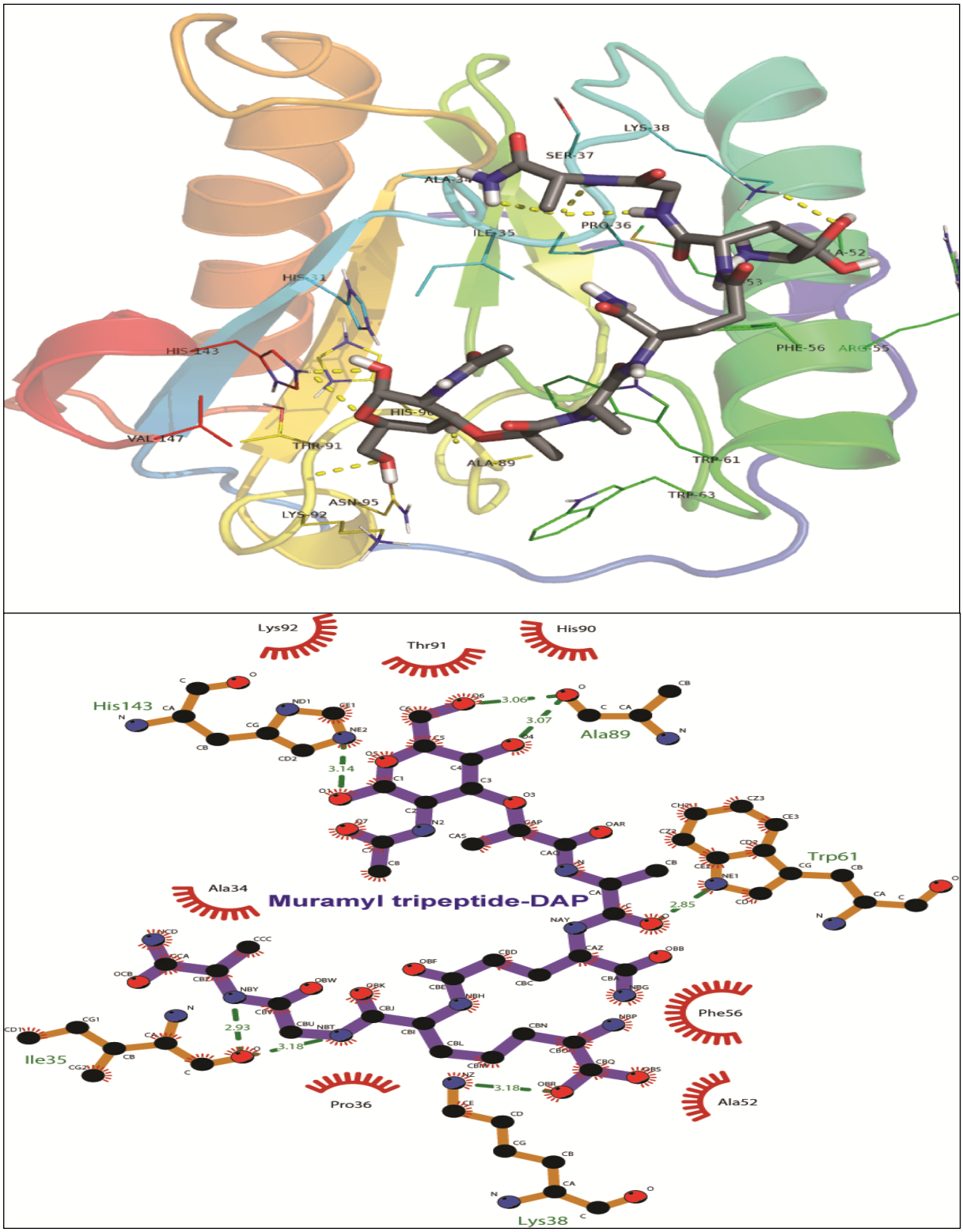

Fig. 26 Molecular interaction of zPGRP2 and Muramyl Tripeptide-DAP

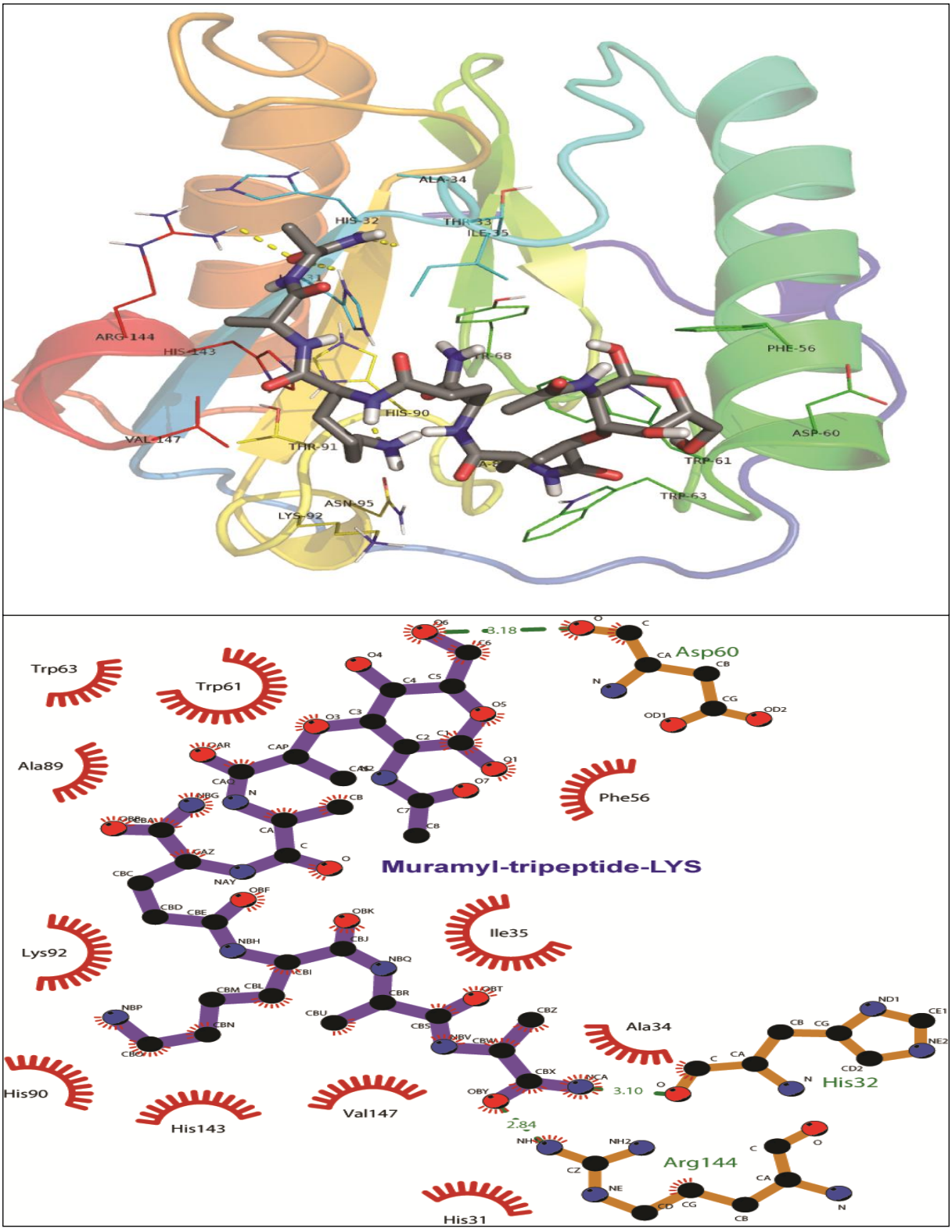

Fig. 27 Molecular interaction of zPGRP2 and Muramyl Tripeptide-LYS

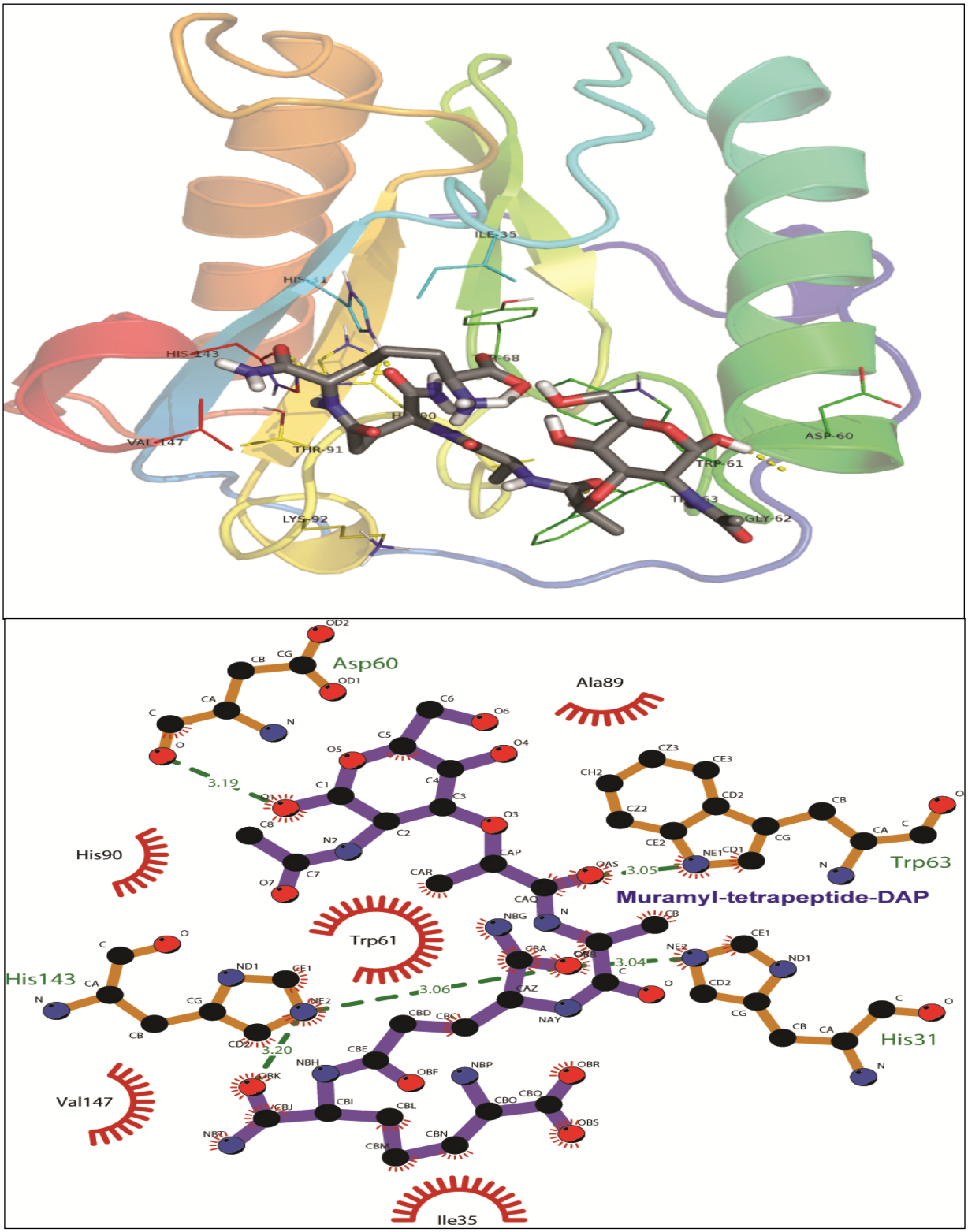

Fig. 28 Molecular interaction of zPGRP2 and Muramyl Tetrapeptide-DAP

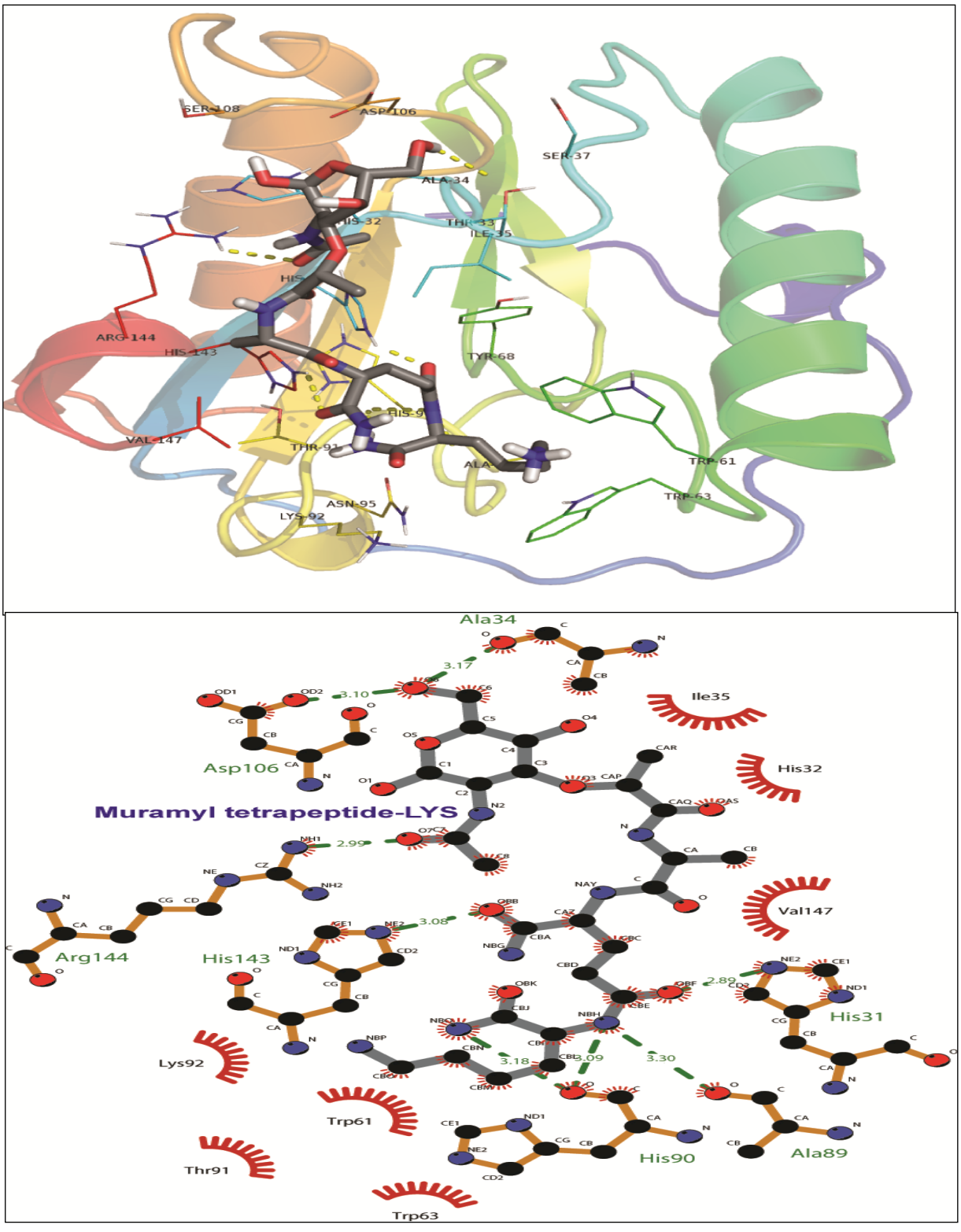

Fig. 29 Molecular interaction of zPGRP2 and Muramyl Tetrapeptide-LYS

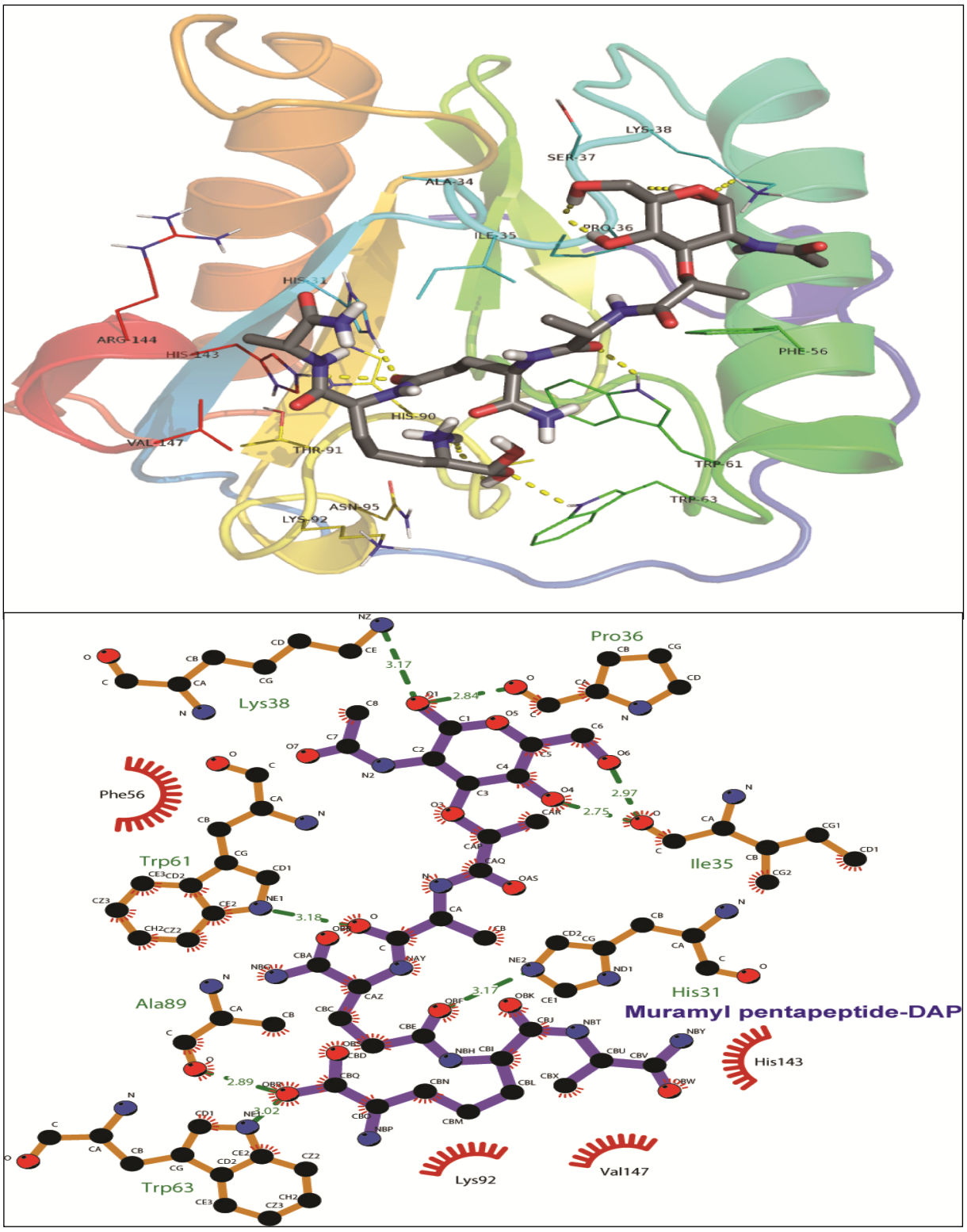

Fig. 30 Molecular interaction of zPGRP2 and Muramyl Pentapeptide-DAP

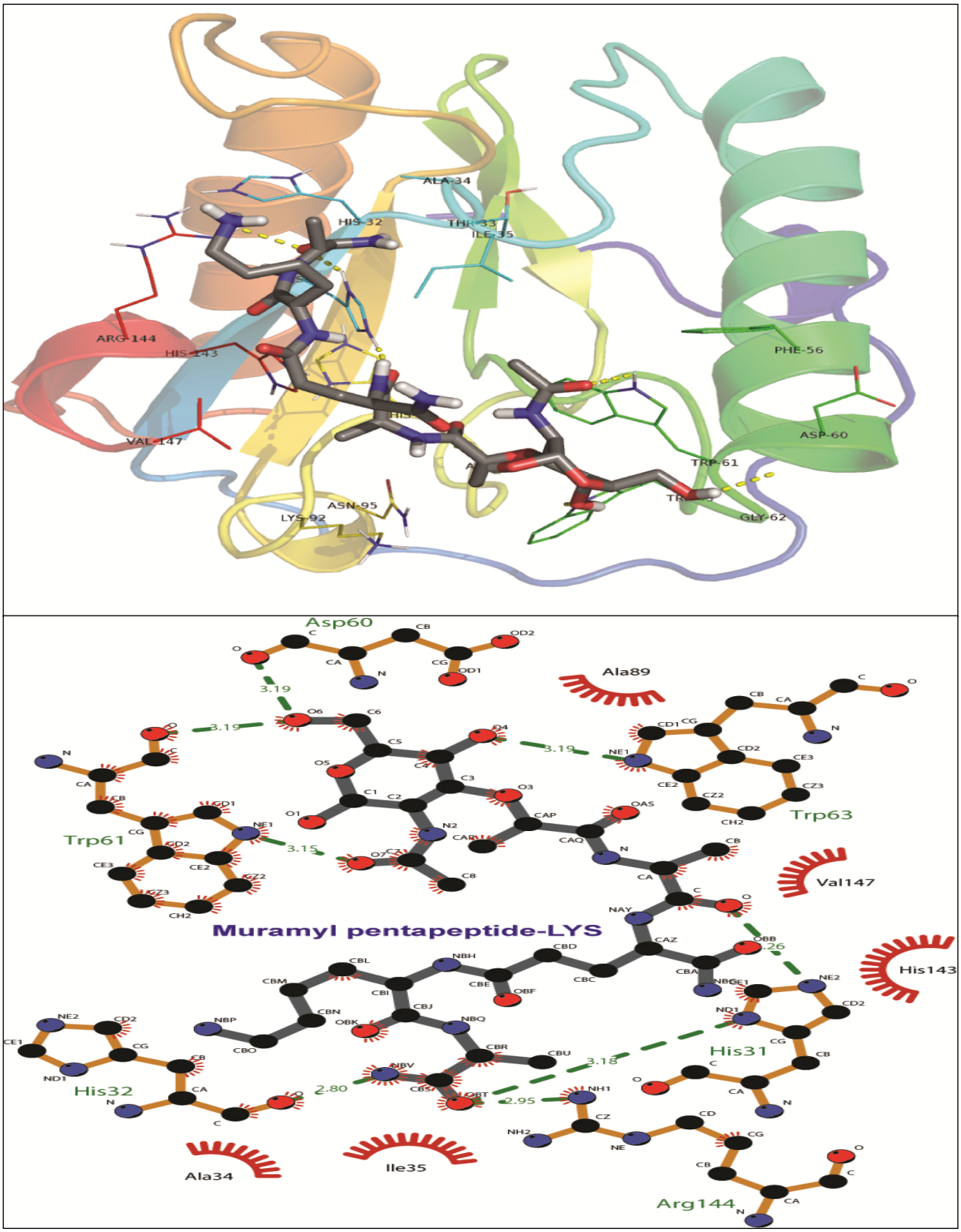

Fig. 31 Molecular interaction of zPGRP2 and Muramyl Pentapeptide-LYS

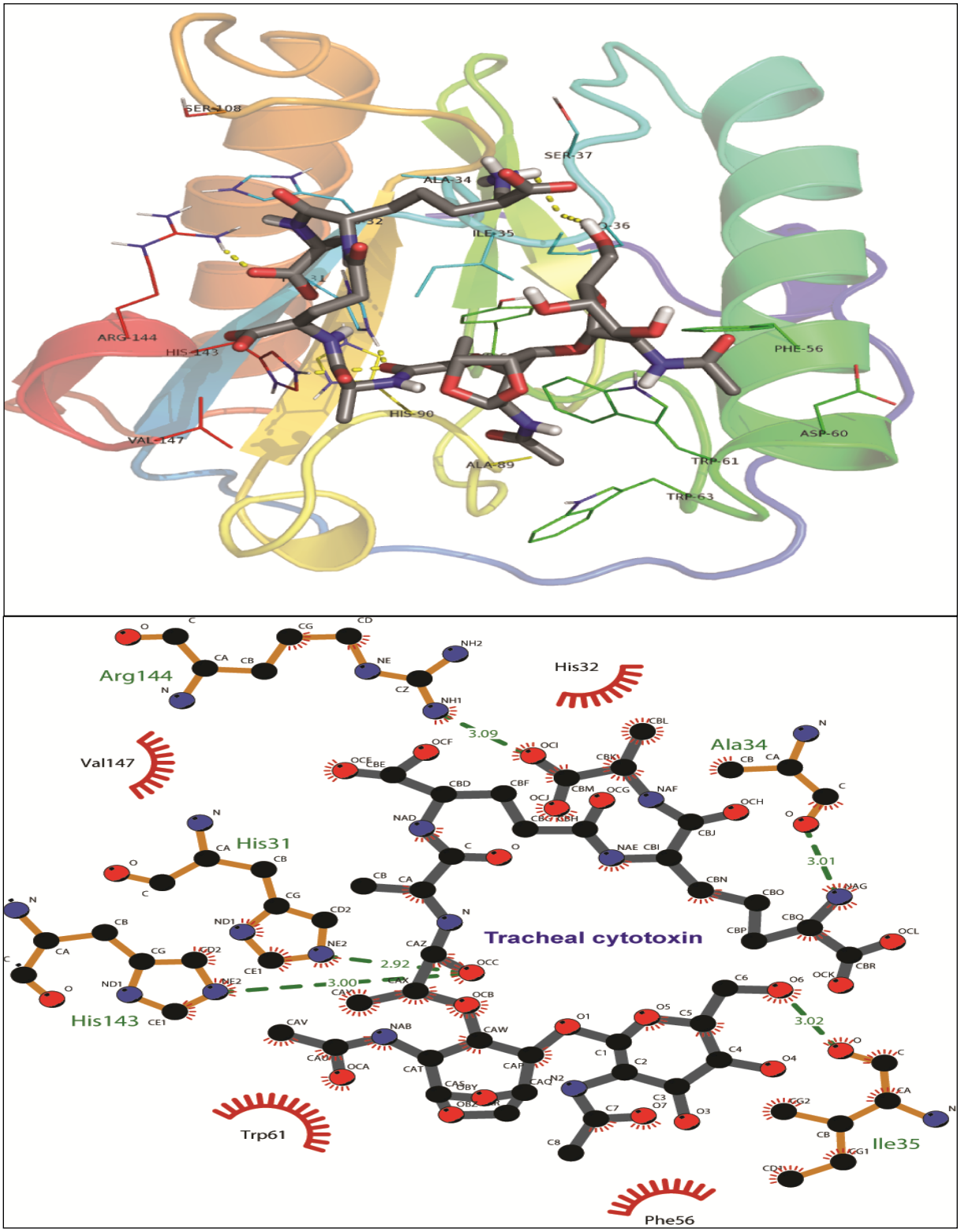

Fig. 32 Molecular interaction of zPGRP2 and Tracheal Cytotoxin

TABLE 5

Intermolecular interaction of ligands with PGRP2

Complexes	Name	Distance	category	Types	From	From chemistry	To	To chemistry
Complex I	:LYS38:HZ1 - :DRG1:OBR	2.50237	H-Bond	Conventional H-Bond	:LYS38:HZ1	H-Donor	:DRG1:OBR	H-Acceptor
	:TRP61:HE1 - :DRG1:O	2.06086	H-Bond	Conventional H-Bond	:TRP61:HE1	H-Donor	:DRG1:O	H-Acceptor
	:DRG1:HBT - :ILE35:O	2.32253	H-Bond	Conventional H-Bond	:DRG1:HBT	H-Donor	:ILE35:O	H-Acceptor
	:DRG1:HBY - :ILE35:O	2.18353	H-Bond	Conventional H-Bond	:DRG1:HBY	H-Donor	:ILE35:O	H-Acceptor
	:DRG1:HCD - :ILE35:O	2.87023	H-Bond	Conventional H-Bond	:DRG1:HCD	H-Donor	:ILE35:O	H-Acceptor
	:DRG1:H42 - :ALA89:O	2.08039	H-Bond	Conventional H-Bond	:DRG1:H42	H-Donor	:ALA89:O	H-Acceptor
	:DRG1:H9H - :ALA34:O	2.62392	H-Bond	Conventional H-Bond	:DRG1:H9H	H-Donor	:ALA34:O	H-Acceptor
	:DRG1:NBP - :PHE56	3.75365	Electrostatic	Pi-Cation	:DRG1:NBP	Positive	:PHE56	Pi-Orbitals
	:DRG1:H - :TRP61	2.94461	H-Bond	Pi-Donor H-Bond	:DRG1:H	H-Donor	:TRP61	Pi-Orbitals
	:DRG1:CAS - :TRP61	3.84334	Hydrophobic	Pi-Sigma	:DRG1:CAS	C-H	:TRP61	Pi-Orbitals
	:DRG1:CAS - :TRP61	3.80945	Hydrophobic	Pi-Sigma	:DRG1:CAS	C-H	:TRP61	Pi-Orbitals
Complex II	:HIS31:HD1 - :DRG1:OBY	2.78748	H-Bond	Conventional H-Bond	:HIS31:HD1	H-Donor	:DRG1:OBY	H-Acceptor
	:ARG144:HH11 - :DRG1:OBY	2.55792	H-Bond	Conventional H-Bond	:ARG144:HH11	H-Donor	:DRG1:OBY	H-Acceptor
	:ARG144:HH12 - :DRG1:OBY	2.50476	H-Bond	Conventional H-Bond	:ARG144:HH12	H-Donor	:DRG1:OBY	H-Acceptor
	:DRG1:HCA - :HIS32:O	2.213	H-Bond	Conventional H-Bond	:DRG1:HCA	H-Donor	:HIS32:O	H-Acceptor
	:DRG1:H9G - :ALA89:O	2.93689	H-Bond	Conventional H-Bond	:DRG1:H9G	H-Donor	:ALA89:O	H-Acceptor
	:DRG1:H9G - :HIS90:O	2.33626	H-Bond	Conventional H-Bond	:DRG1:H9G	H-Donor	:HIS90:O	H-Acceptor
	:LYS92:CE - :DRG1:O	3.78334	H-Bond	Carbon H-Bond	:LYS92:CE	H-Donor	:DRG1:O	H-Acceptor
Complex III	:HIS31:HE2 - :DRG1:OBB	2.04453	H-Bond	Conventional H-Bond	:HIS31:HE2	H-Donor	:DRG1:OBB	H-Acceptor
	:TRP63:HE1 - :DRG1:OAS	2.39016	H-Bond	Conventional H-Bond	:TRP63:HE1	H-Donor	:DRG1:OAS	H-Acceptor
	:HIS143:HE2 - :DRG1:OBB	2.56613	H-Bond	Conventional H-Bond	:HIS143:HE2	H-Donor	:DRG1:OBB	H-Acceptor
	:HIS143:HE2 - :DRG1:OBK	2.51779	H-Bond	Conventional H-Bond	:HIS143:HE2	H-Donor	:DRG1:OBK	H-Acceptor
	:DRG1:H12 - :ASP60:O	2.31725	H-Bond	Conventional H-Bond	:DRG1:H12	H-Donor	:ASP60:O	H-Acceptor
	:TRP61:CD1 - :DRG1:O1	3.34121	H-Bond	Carbon H-Bond	:TRP61:CD1	H-Donor	:DRG1:O1	H-Acceptor
	:DRG1:CA - :TRP61	3.96026	Hydrophobic	Pi-Sigma	:DRG1:CA	C-H	:TRP61	Pi-Orbitals

Complexes	Name	Distance	category	Types	From	From chemistry	To	To chemistry
Complex IV	:HIS31:HD1 - :DRG1:O7	3.04424	H-Bond	Conventional H-Bond	:HIS31:HD1	H-Donor	:DRG1:O7	H-Acceptor
	:HIS31:HE2 - :DRG1:OBF	2.05129	H-Bond	Conventional H-Bond	:HIS31:HE2	H-Donor	:DRG1:OBF	H-Acceptor
	:HIS143:HE2 - :DRG1:OBB	2.37745	H-Bond	Conventional H-Bond	:HIS143:HE2	H-Donor	:DRG1:OBB	H-Acceptor
	:ARG144:HH11 - :DRG1:O7	2.58917	H-Bond	Conventional H-Bond	:ARG144:HH11	H-Donor	:DRG1:O7	H-Acceptor
	:ARG144:HH12 - :DRG1:O7	2.73139	H-Bond	Conventional H-Bond	:ARG144:HH12	H-Donor	:DRG1:O7	H-Acceptor
	:DRG1:HBH - :ALA89:O	2.67351	H-Bond	Conventional H-Bond	:DRG1:HBH	H-Donor	:ALA89:O	H-Acceptor
	:DRG1:HBH - :HIS90:O	2.4197	H-Bond	Conventional H-Bond	:DRG1:HBH	H-Donor	:HIS90:O	H-Acceptor
	:DRG1:HBQ - :HIS90:O	2.18215	H-Bond	Conventional H-Bond	:DRG1:HBQ	H-Donor	:HIS90:O	H-Acceptor
	:DRG1:H63 - :ALA34:O	2.38789	H-Bond	Conventional H-Bond	:DRG1:H63	H-Donor	:ALA34:O	H-Acceptor
	:LYS92:CE - :DRG1:OBK	3.44209	H-Bond	Carbon H-Bond	:LYS92:CE	H-Donor	:DRG1:OBK	H-Acceptor
	:DRG1:NBP - :TRP61	4.9481	Electrostatic	Pi-Cation	:DRG1:NBP	Positive	:TRP61	Pi-Orbitals
Complex V	:HIS31:HE2 - :DRG1:OBF	2.15349	H-Bond	Conventional H-Bond	:HIS31:HE2	H-Donor	:DRG1:OBF	H-Acceptor
	:LYS38:HZ2 - :DRG1:O1	2.68662	H-Bond	Conventional H-Bond	:LYS38:HZ2	H-Donor	:DRG1:O1	H-Acceptor
	:TRP61:HE1 - :DRG1:O	2.32317	H-Bond	Conventional H-Bond	:TRP61:HE1	H-Donor	:DRG1:O	H-Acceptor
	:TRP63:HE1 - :DRG1:OBR	2.61974	H-Bond	Conventional H-Bond	:TRP63:HE1	H-Donor	:DRG1:OBR	H-Acceptor
	:HIS143:HE2 - :DRG1:OBF	2.85957	H-Bond	Conventional H-Bond	:HIS143:HE2	H-Donor	:DRG1:OBF	H-Acceptor
	:DRG1:HBP - :HIS90:O	2.8937	H-Bond	Conventional H-Bond	:DRG1:HBP	H-Donor	:HIS90:O	H-Acceptor
	:DRG1:HBR - :ALA89:O	2.47166	H-Bond	Conventional H-Bond	:DRG1:HBR	H-Donor	:ALA89:O	H-Acceptor
	:DRG1:H12 - :PRO36:O	2.40142	H-Bond	Conventional H-Bond	:DRG1:H12	H-Donor	:PRO36:O	H-Acceptor
	:DRG1:H42 - :ILE35:O	1.76727	H-Bond	Conventional H-Bond	:DRG1:H42	H-Donor	:ILE35:O	H-Acceptor
	:DRG1:H63 - :ILE35:O	2.00587	H-Bond	Conventional H-Bond	:DRG1:H63	H-Donor	:ILE35:O	H-Acceptor
	:DRG1:H9D - :ALA89:O	2.10009	H-Bond	Conventional H-Bond	:DRG1:H9D	H-Donor	:ALA89:O	H-Acceptor
	:DRG1:HBG - :TRP61	3.01836	H-Bond	Pi-Donor H-Bond	:DRG1:HBG	H-Donor	:TRP61	Pi-Orbitals
Complex VI	:HIS31:HD1 - :DRG1:OBT	2.49813	H-Bond	Conventional H-Bond	:HIS31:HD1	H-Donor	:DRG1:OBT	H-Acceptor
	:HIS31:HE2 - :DRG1:O	2.26356	H-Bond	Conventional H-Bond	:HIS31:HE2	H-Donor	:DRG1:O	H-Acceptor
	:HIS32:HN - :DRG1:OBT	2.82914	H-Bond	Conventional H-Bond	:HIS32:HN	H-Donor	:DRG1:OBT	H-Acceptor
	:TRP61:HE1 - :DRG1:O7	2.63723	H-Bond	Conventional H-Bond	:TRP61:HE1	H-Donor	:DRG1:O7	H-Acceptor
	:ARG144:HH12 - :DRG1:OBT	2.3686	H-Bond	Conventional H-Bond	:ARG144:HH12	H-Donor	:DRG1:OBT	H-Acceptor
	:DRG1:OAS - :ALA89:O	3.08048	H-Bond	Conventional H-Bond	:DRG1:OAS	H-Donor	:ALA89:O	H-Acceptor
	:DRG1:HBV - :HIS32:O	1.85963	H-Bond	Conventional H-Bond	:DRG1:HBV	H-Donor	:HIS32:O	H-Acceptor
	:DRG1:H63 - :ASP60:O	2.85887	H-Bond	Conventional H-Bond	:DRG1:H63	H-Donor	:ASP60:O	H-Acceptor
	:HIS32:CD2 - :DRG1:OBT	3.69135	H-Bond	Carbon H-Bond	:HIS32:CD2	H-Donor	:DRG1:OBT	H-Acceptor
	:HIS143:CD2 - :DRG1:OBK	3.57746	H-Bond	Carbon H-Bond	:HIS143:CD2	H-Donor	:DRG1:OBK	H-Acceptor

Complexes	Name	Distance	category	Types	From	From chemistry	To	To chemistry
Complex VII	:HIS31:NE2 - :UNK1:OCJ	4.12377	Electrostatic	Attractive Charge	:HIS31:NE2	Positive	:UNK1:OCJ	Negative
	:HIS143:NE2 - :UNK1:OCJ	4.23915	Electrostatic	Attractive Charge	:HIS143:NE2	Positive	:UNK1:OCJ	Negative
	:ARG144:NH1 - :UNK1:OCJ	4.4003	Electrostatic	Attractive Charge	:ARG144:NH1	Positive	:UNK1:OCJ	Negative
	:HIS31:HE2 - :UNK1:OCC	1.94549	H-Bond	Conventional H-Bond	:HIS31:HE2	H-Donor	:UNK1:OCC	H-Acceptor
	:TRP61:HE1 - :UNK1:O1	2.94059	H-Bond	Conventional H-Bond	:TRP61:HE1	H-Donor	:UNK1:O1	H-Acceptor
	:HIS143:HE2 - :UNK1:OCC	2.50436	H-Bond	Conventional H-Bond	:HIS143:HE2	H-Donor	:UNK1:OCC	H-Acceptor
	:ARG144:HH11 - :UNK1:OCI	2.10934	H-Bond	Conventional H-Bond	:ARG144:HH11	H-Donor	:UNK1:OCI	H-Acceptor
	:UNK1:HAG - :ALA34:O	2.19129	H-Bond	Conventional H-Bond	:UNK1:HAG	H-Donor	:ALA34:O	H-Acceptor
	:UNK1:H63 - :ILE35:O	2.70401	H-Bond	Conventional H-Bond	:UNK1:H63	H-Donor	:ILE35:O	H-Acceptor
	:UNK1:H9G - :ILE35:O	2.66423	H-Bond	Conventional H-Bond	:UNK1:H9G	H-Donor	:ILE35:O	H-Acceptor
	:HIS31:CE1 - :UNK1:OCJ	3.35248	H-Bond	Carbon H-Bond	:HIS31:CE1	H-Donor	:UNK1:OCJ	H-Acceptor
	:UNK1:NAB - :TRP61	4.26617	Electrostatic	Pi-Cation	:UNK1:NAB	Positive	:TRP61	Pi-Orbitals
	:UNK1:NAB - :TRP61	4.16813	Electrostatic	Pi-Cation	:UNK1:NAB	Positive	:TRP61	Pi-Orbitals
	:UNK1:CAV - :TRP61	3.87525	Hydrophobic	Pi-Sigma	:UNK1:CAV	C-H	:TRP61	Pi-Orbitals
	:UNK1:CAV - :TRP61	3.80177	Hydrophobic	Pi-Sigma	:UNK1:CAV	C-H	:TRP61	Pi-Orbitals

CONCLUSION

CHAPTER V

CONCLUSION

Several studies on PGRP-sense innate signalling has been reported in human and *Drosophila*. The binding site of Gram positive and negative bacteria cell-wall PGN fragments in PGRP is very important to understand these mechanisms. In this study, we reported the key binding domains and amino acid residues that recognize the PGN structures of both Gram positive and negative bacteria. Results showed the Dap-type PGN has a greater binding affinity to PGRP2 receptor than Lys-type in Zebrafish. The conserved amino acid residues sensing the acetylmuramic acid and D-Ala in both Dap- and Lys type PGN-fragments warrants in vivo investigation to understand the hydrolyzing and bactericidal activity of PGRP2 in Zebrafish and other fish species. The Dap- and Lys-type PGN key binding domains suggests the need of further investigation for the therapeutic application of the PGRP2 related disorders and innate immune signalling in teleosts and other higher eukaryotes.

REFERENCE

REFERENCE

1. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990; 215: 403-10.
2. Alvarez-Pellitero P. Fish immunity and parasite infections: from innate immunity to immunoprophylactic prospects. Vet Immunol Immunopathol. 2008; 126: 171-98.
3. Anselme C, Vallier A, Balmand S, Fauvarque MO, Heddi A. Host PGRP gene expression and bacterial release in endosymbiosis of the weevil Sitophilus zeamais. Appl Environ Microbiol. 2006; 72: 6766-72.
4. Beutler B. Innate immunity: an overview. Mol Immunol. 2004; 40:845-59.
5. Bischoff V, Vignal C, Duvic B, Boneca IG, Hoffmann JA, Royet J. Downregulation of the Drosophila immune response by peptidoglycan-recognition proteins SC1 and SC2. PLoS Pathog. 2006; 2: e14.
6. Buchan DW, Minneci F, Nugent TC, Bryson K, Jones DT. Scalable web services for the PSIPRED Protein Analysis Workbench. Nucleic Acids Res. 2013 ; 41: W349–W357.
7. Cerenius L, Söderhäll K. The prophenoloxidase-activating system in invertebrates. Immunol Rev. 2004; 198: 116-26.
8. Charroux B, Rival T, Narbonne-Reveau K, Royet J. Bacterial detection by Drosophila peptidoglycan recognition proteins. Microbes Infect. 2009; 11: 631-6.
9. Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr. 2010; 66: 12-21.
10. Cho JH, Fraser IP, Fukase K, Kusumoto S, Fujimoto Y, Stahl GL, Ezekowitz RA. Human peptidoglycan recognition protein S is an effector of neutrophil-mediated innate immunity. Blood. 2005; 106: 2551–2558.
11. Cho S, Wang Q, Swaminathan CP, Hesek D, Lee M, Boons GJ, Mobashery S, Mariuzza RA. Structural insights into the bactericidal mechanism of human peptidoglycan recognition proteins. Proc Natl Acad Sci U S A. 2007; 104: 8761-6.

12. Christophides GK, Zdobnov E, Barillas-Mury C, Birney E, Blandin S, Blass C, Brey PT, Collins FH, Danielli A, Dimopoulos G, Hetru C. Immunity-related genes and gene families in Anopheles gambiae. Science. 2002; 298: 159-65.
13. Colovos C, Yeates TO. ERRAT: an empirical atom-based method for validating protein structures. Protein Sci. 1993; 2: 1511-9.
14. Colovos C, Yeates TO. Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci. 1993; 2: 1511-9.
15. Coteur G, Mellroth P, De Lefortery C, Gillan D, Dubois P, Communi D, Steiner H. Peptidoglycan recognition proteins with amidase activity in early deuterostomes (Echinodermata). Dev Comp Immunol. 2007; 31: 790-804.
16. Doyle RJ, Dziarski R. The Bacterial Cell-7: Peptidoglycan.
17. Du Pasquier L. Antibody diversity in lower vertebrates—why is it so restricted? Nature. 1982; 296: 311-3.
18. Dziarski R, Gupta D. Mammalian PGRPs: novel antibacterial proteins. Cell Microbiol. 2006; 8: 1059-69.
19. Dziarski R, Gupta D. The peptidoglycan recognition proteins (PGRPs). Genome Biol. 2006; 7: 232.
20. Dziarski R. Peptidoglycan recognition proteins (PGRPs). Mol Immunol. 2004; 40: 877-86.
21. Eisenberg D, Lüthy R, Bowie JU. [20] VERIFY3D: Assessment of protein models with three-dimensional profiles. Methods Enzymol. 1997; 277: 396-404.
22. Ellis AE. Innate host defense mechanisms of fish against viruses and bacteria. Dev Comp Immunol; 25: 827-39.
23. Ferrandon D, Imler JL, Hetru C, Hoffmann JA. The Drosophila systemic immune response: sensing and signalling during bacterial and fungal infections. Nat Rev Immunol. 2007; 7: 862-74.
24. Ferrè F, Clote P. DiANNA: a web server for disulfide connectivity prediction. Nucleic Acids Res. 2005 ; 33: W230-2.
25. Finn RD, Mistry J, Tate J, Coggill P, Heger A, Pollington JE, Gavin OL, Gunasekaran P, Ceric G, Forslund K, Holm L. The Pfam protein families database. Nucleic acids research. 2009 Nov 17;38(suppl_1):D211-22.
26. Gelius E, Persson C, Karlsson J, Steiner H. A mammalian peptidoglycan recognition protein with N-acetylmuramoyl-L-alanine amidase activity. Biochem Biophys Res Commun. 2003; 306: 988-94.

27. Geourjon C, Deleage G. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Comput Appl Biosci. 1995; 11: 681-4.
28. Guan R, Roychowdhury A, Ember B, Kumar S, Boons GJ, Mariuzza RA. Structural basis for peptidoglycan binding by peptidoglycan recognition proteins. Proc Natl Acad Sci U S A. 2004; 101: 17168–17173.
29. Guan R, Wang Q, Sundberg EJ, Mariuzza RA. Crystal structure of human peptidoglycan recognition protein S (PGRP-S) at 1.70 Å resolution. J Mol Biol. 2005; 347: 683-91.
30. Gupta R, Brunak S. Prediction of glycosylation across the human proteome and the correlation to protein function. Pac Symp Biocomput. 2002:310-22.
31. Hedengren M, Dushay MS, Ando I, Ekengren S, Wihlborg M, Hultmark D. Relish, a central factor in the control of humoral but not cellular immunity in Drosophila. Mol Cell. 1999; 4: 827-37.
32. Heger A, Holm L. Rapid automatic detection and alignment of repeats in protein sequences. Proteins. 2000; 41: 224-37.
33. Heijenoort JV. Formation of the glycan chains in the synthesis of bacterial peptidoglycan. Glycobiology. 2001; 11: 25R-36R.
34. Heo L, Park H, Seok C. GalaxyRefine: protein structure refinement driven by side-chain repacking. Nucleic Acids Res. 2013; 41: W384-8.
35. Hoffmann JA, Reichhart JM. Drosophila innate immunity: an evolutionary perspective. Nat Immunol. 2002; 3: 121-6.
36. Hoffmann JA. The immune response of Drosophila. Nature. 2003; 426: 33-8.
37. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. J Mol Graph. 1996; 14: 33-38, 27-28.
38. Janeway CA. Approaching the asymptote? Evolution and revolution in immunology. Cold Spring Harb Symp Quant Biol 1989. 54: 1-13.
39. Janeway Jr CA, Medzhitov R. Innate immune recognition. Annu Rev Immunol. 2002; 20: 197-216.
40. Jang JH, Kim H, Cho JH. Rainbow trout peptidoglycan recognition protein has an anti-inflammatory function in liver cells. Fish Shellfish Immunol. 2013; 35: 1838-47.
41. Kang D, Liu G, Lundström A, Gelius E, Steiner H. A peptidoglycan recognition protein in innate immunity conserved from insects to humans. Proc Natl Acad Sci U S A. 1998 ; 95: 10078-82.

42. Kim MS, Byun M, Oh BH. Crystal structure of peptidoglycan recognition protein LB from Drosophila melanogaster. Nat Immunol. 2003 ; 4: 787-93.
43. Kim MY, Jang JH, Lee JW, Cho JH. Molecular cloning and characterization of peptidoglycan recognition proteins from the rockfish, Sebastes schlegeli. Fish Shellfish Immunol. 2010; 28: 632-9.
44. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Cryst.1993. 26, 283-291.
45. Lemaitre B, Kromer-Metzger E, Michaut L, Nicolas E, Meister M, Georgel P, Reichhart JM, Hoffmann JA. A recessive mutation, immune deficiency (imd), defines two distinct control pathways in the Drosophila host defense. Proc Natl Acad Sci U S A. 1995; 92: 9465-9.
46. Lemaitre B, Nicolas E, Michaut L, Reichhart JM, Hoffmann JA. The dorsoventral regulatory gene cassette spätzle/Toll/cactus controls the potent antifungal response in Drosophila adults. Cell. 1996; 86: 973-83.
47. Leulier F, Parquet C, Pili-Floury S, Ryu JH, Caroff M, Lee WJ, Mengin-Lecreulx D, Lemaitre B. The Drosophila immune system detects bacteria through specific peptidoglycan recognition. Nat Immunol. 2003; 4: 478-84.
48. Leulier F, Rodriguez A, Khush RS, Abrams JM, Lemaitre B. The Drosophila caspase Dredd is required to resist Gram-negative bacterial infection. EMBO Rep. 2000; 1: 353-8.
49. Li JH, Chang MX, Xue NN, Nie P. Functional characterization of a short peptidoglycan recognition protein, PGRP5 in grass carp Ctenopharyngodon idella. Fish Shellfish Immunol. 2013; 35: 221-30.
50. Li JH, Yu ZL, Xue NN, Zou PF, Hu JY, Nie P, Chang MX. Molecular cloning and functional characterization of peptidoglycan recognition protein 6 in grass carp Ctenopharyngodon idella. Dev Comp Immunol. 2014; 42: 244-55.
51. Li MF, Zhang M, Wang CL, Sun L. A peptidoglycan recognition protein from Sciaenops ocellatus is a zinc amidase and a bactericide with a substrate range limited to Gram-positive bacteria. Fish Shellfish Immunol. 2012; 32: 322-30.
52. Li X, Wang S, Qi J, Echtenkamp SF, Chatterjee R, Wang M, Boons GJ, Dziarski R, Gupta D. Zebrafish peptidoglycan recognition proteins are bactericidal amidases essential for defense against bacterial infections. Immunity. 2007; 27: 518-29.

53. Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO, Shaw DE. Improved side-chain torsion potentials for the Amber ff99SB protein force field. Proteins. 2010; 78:1950-8.
54. Liu C, Gelius E, Liu G, Steiner H, Dziarski R. Mammalian peptidoglycan recognition protein binds peptidoglycan with high affinity, is expressed in neutrophils, and inhibits bacterial growth. J Biol Chem. 2000; 275: 24490-9.
55. Liu C, Xu Z, Gupta D, Dziarski R. Peptidoglycan recognition proteins a novel family of four human innate immunity pattern recognition molecules. J Biol Chem. 2001; 276: 34686-94.
56. Loker ES, Adema CM, Zhang SM, Kepler TB. Invertebrate immune systems–not homogeneous, not simple, not well understood. Immunol Rev. 2004;198:10-24.
57. Magnadóttir B. Innate immunity of fish (overview). Fish Shellfish Immunol. 2006; 20: 137-51.
58. Mao Y, Wang J, Zhang Z, Ding S, Su Y. Cloning, mRNA expression, and recombinant expression of peptidoglycan recognition protein II gene from large yellow croaker (Pseudosciaena crocea). Mol Biol Rep. 2010; 37: 3897-908.
59. Medzhitov R, Janeway CA. Decoding the patterns of self and nonself by the innate immune system. Science. 2002 Apr 12;296(5566):298-300.
60. Medzhitov R, Janeway CA. Innate immunity: the virtues of a nonclonal system of recognition. Cell. 1997; 91: 295-8.
61. Medzhitov R. Recognition of microorganisms and activation of the immune response. Nature. 2007 Oct 18; 449(7164):819-26.
62. Mellroth P, Karlsson J, Steiner H. A scavenger function for a DrosophilaPeptidoglycan recognition protein. J Biol Chem. 2003 Feb 28;278(9):7059-64.
63. Meng X, Khanuja BS, Ip YT. Toll receptor-mediated Drosophila immune response requires Dif, an NF-κB factor. Genes Dev. 1999; 13: 792-7.
64. Mengin-Lecreulx D, Lemaitre B. Structure and metabolism of peptidoglycan and molecular requirements allowing its detection by the Drosophila innate immune system. J Endotoxin Res. 2005; 11:105-11.
65. Michel T, Reichhart JM, Hoffmann JA, Royet J. Drosophila Toll is activated by Gram-positive bacteria through a circulating peptidoglycan recognition protein. Nature. 2001 Dec 13;414(6865):756-9.

66. Pronk S, Pall S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, et al. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. Bioinformatics. 2013; 29: 845–54.
67. Rämet M, Manfruelli P, Pearson A, Mathey-Prevot B, Ezekowitz RA. Functional genomic analysis of phagocytosis and identification of a Drosophila receptor for E. coli Nature. 2002; 416: 644-8.
68. Reiser JB, Teyton L, Wilson IA. Crystal structure of the Drosophila peptidoglycan recognition protein (PGRP)-SA at 1.56 Å resolution. J Mol Biol. 2004; 340: 909-17.
69. Royet J, Dziarski R. Peptidoglycan recognition proteins: pleiotropic sensors and effectors of antimicrobial defences. Nature Reviews Microbiology. 2007 Apr 1;5(4):264-77.
70. Rutschmann S, Jung AC, Hetru C, Reichhart JM, Hoffmann JA, Ferrandon D. The Rel protein DIF mediates the antifungal but not the antibacterial host defense in Drosophila. Immunity. 2000 May; 12: 569-80.
71. Rutschmann S, Jung AC, Zhou R, Silverman N, Hoffmann JA, Ferrandon D. Role of Drosophila IKKγ in a Toll-independent antibacterial immune response. Nat Immunol. 2000;1: 342-7.
72. Saha S, Jing X, Park SY, Wang S, Li X, Gupta D, Dziarski R. Peptidoglycan recognition proteins protect mice from experimental colitis by promoting normal gut flora and preventing induction of interferon-γ. Cell Host Microbe. 2010; 8: 147-62.
73. Saha S, Qi J, Wang S, Wang M, Li X, Kim YG, Núñez G, Gupta D, Dziarski R. PGLYRP-2 and Nod2 are both required for peptidoglycan-induced arthritis and local inflammation. Cell Host Microbe. 2009; 5: 137-50.
74. Sahoo BR, Dubey PK, Goyal S, Bhoi GK, Lenka SK, Maharana J, Pradhan SK, Kataria RS. Exploration of the binding modes of buffalo PGRP1 receptor complexed with meso-diaminopimelic acid and lysine-type peptidoglycans by molecular dynamics simulation and free energy calculation. Chem Biol Interact. 2014; 220: 255-68.
75. SchuÈttelkopf AW, Van Aalten DM. PRODRG: a tool for high-throughput crystallography of protein–ligand complexes. Acta Crystallogr D Biol Crystallogr. 2004; 60: 1355-63.
76. Schultz J, Milpetz F, Bork P, Ponting CP. SMART, a simple modular architecture research tool: identification of signaling domains. Proc Natl Acad Sci U S A. 1998; 95: 5857-64.

77. Sharma P, Dube D, Singh A, Mishra B, Singh N, Sinha M, Dey S, Kaur P, Mitra DK, Sharma S, Singh TP. Structural basis of recognition of pathogen-associated molecular patterns and inhibition of proinflammatory cytokines by camel peptidoglycan recognition protein. J Biol Chem. 2011; 286: 16208-17.
78. Simossis VA, Heringa J. PRALINE: a multiple sequence alignment toolbox that integrates homology-extended and secondary structure information. Nucleic Acids Res. 2005; 33: W289–W294.
79. Stowell SR, Arthur CM, Dias-Baruffi M, Rodrigues LC, Gourdine JP, Heimburg-Molinaro J, Ju T, Molinaro RJ, Rivera-Marrero C, Xia B, Smith DF. Innate immune lectins kill bacteria expressing blood group antigen. Nat Med. 2010; 16: 295-301.
80. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol. 2013; 30: 2725-9.
81. Tauszig-Delamasure S, Bilak H, Capovilla M, Hoffmann JA, Imler JL. Drosophila MyD88 is required for the response to fungal and Gram-positive bacterial infections. Nat Immunol. 2002; 3: 91-7.
82. Troll JV, Adin DM, Wier AM, Paquette N, Silverman N, Goldman WE, Stadermann FJ, Stabb EV, McFall-Ngai MJ. Peptidoglycan induces loss of a nuclear peptidoglycan recognition protein during host tissue development in a beneficial animal-bacterial symbiosis. Cell Microbiol. 2009; 11: 1114-27.
83. Troll JV, Bent EH, Pacquette N, Wier AM, Goldman WE, Silverman N, McFall-Ngai MJ. Taming the symbiont for coexistence: a host PGRP neutralizes a bacterial symbiont toxin. Environ Microbiol. 2010; 12: 2190-203.
84. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem. 2010; 31: 455-61.
85. Tydell CC, Yuan J, Tran P, Selsted ME. Bovine peptidoglycan recognition protein-S: antimicrobial activity, localization, secretion, and binding properties. J Immunol. 2006; 176: 1154-62.
86. Vidal S, Khush RS, Leulier F, Tzou P, Nakamura M, Lemaitre B. Mutations in the Drosophila dTAK1 gene reveal a conserved function for MAPKKKs in the control of rel/NF-κB-dependent innate immune responses. Genes Dev. 2001; 15: 1900-12.
87. Wallner B, Fang H, Elofsson A. Automatic consensus-based fold recognition using Pcons, ProQ, and Pmodeller Proteins. 2003; 53: 534-41.

88. Wang J, Wu Y, Yang G, Aksoy S. Interactions between mutualist Wigglesworthia and tsetse peptidoglycan recognition protein (PGRP-LB) influence trypanosome transmission. Proc Natl Acad Sci U S A. 2009; 106: 12133-8.
89. Wang M, Liu LH, Wang S, Li X, Lu X, Gupta D, Dziarski R. Human peptidoglycan recognition proteins require zinc to kill both gram-positive and gram-negative bacteria and are synergistic with antibacterial peptides. J Immunol. 2007; 178: 3116-25.
90. Wang ZM, Li X, Cocklin RR, Wang M, Wang M, Fukase K, Inamura S, Kusumoto S, Gupta D, Dziarski R. Human peptidoglycan recognition protein-L is an N-acetylmuramoyl-L-alanine amidase. J Biol Chem. 2003; 278: 49044-52.
91. Webb B, Sali A. Protein structure modeling with MODELLER. Curr Protoc Bioinformatics. 2014; 47: 1-32.
92. Weber AN, Tauszig-Delamasure S, Hoffmann JA, Lelièvre E, Gascan H, Ray KP, Morse MA, Imler JL, Gay NJ. Binding of the Drosophila cytokine Spätzle to Toll is direct and establishes signaling. Nat Immunol. 2003; 4: 794-800.
93. Werner T, Liu G, Kang D, Ekengren S, Steiner H, Hultmark D. A family of peptidoglycan recognition proteins in the fruit fly Drosophila melanogaster. Proc Natl Acad Sci U S A. 2000; 97: 13772-7.
94. Whyte SK. The innate immune response of finfish–a review of current knowledge. Fish Shellfish Immunol. 2007; 23: 1127-51.
95. Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res. 2007; 35: W407-10.
96. Yao J, Mu W, Liu S, Zhang J, Wen H, Liu Z. Identification, phylogeny and expression analysis of suppressors of cytokine signaling in channel catfish. Mol Immunol. 2015; 64: 276-84.
97. Yoshida H, Kinoshita K, Ashida M. Purification of a peptidoglycan recognition protein from hemolymph of the silkworm, Bombyx mori. J Biol Chem. 1996; 271: 13854-60.
98. Yu ZL, Li JH, Xue NN, Nie P, Chang MX. Expression and functional characterization of PGRP6 splice variants in grass carp Ctenopharyngodon idella. Dev Comp Immunol. 2014; 47: 264-74.
99. Zaidman-Rémy A, Hervé M, Poidevin M, Pili-Floury S, Kim MS, Blanot D, Oh BH, Ueda R, Mengin-Lecreulx D, Lemaitre B. The Drosophila amidase PGRP-LB modulates the immune response to bacterial infection. Immunity. 2006; 24: 463-73.

CURRICULUM VITAE

Ms. SUNANDA PARAMANIK
Department of Bioinformatics
Orissa University of Agriculture and Technology (O.U.A.T)
Bhubaneswar-751003
Email: sunanda050@gmail.com
Contact no: +91-8984756144

CAREER OBJECTIVE

Seeking to build on strong research skills gained through studies and practical lab experience while working in a creative and dynamic environment with competitive educational venture. Ambition is to eventually have broad-based academic career including teaching, research and management responsibilities.

EDUCATION

Exam Passed	College/School	University/Board	Year	% Obtained
M. Sc (Bioinformatics)	Centre for Post Graduate Studies	O.U.A.T Bhubaneswar	Currently Pursuing	7.58 (OGPA)
B. Sc (Biotechnology)	Trident Academy of Creative Technology, Bhubaneswar	UTKAL UNIVERSITY, Bhubaneswar	2015	65.5
12th Standard (10+2 equivalent)	Jawahar Navodaya Vidyalaya, Mundali, Cuttack	CBSE, Odisha	2012	71.2
10th Standard	Jawahar Navodaya Vidyalaya, Mundali, Cuttack	CBSE, Odisha	2010	91.2

RESEARCH INTERESTS

My principal research interests lie in the field of structural bioinformatics (working with software and tools), genetics & genomics (Comparative, Functional, Structural), proteomics (Structure of the Protein, Protein-Protein Interaction), gene expression and regulation as well as a bit of transcriptomics studies, My future research plans are to build on the foundations of my PhD to further investigate on various omics with computational strategies. r

DISSERTATION

Four Months project work for partial fulfillment of M.sc Degree in ICAR-Central Inland Fisheries Research Institute, Barrackpore, Kolkata-700120, India, March-June 2017 under the Supervision of Dr. Basanta Kumar Das, Director, ICAR-CIFRI.

BINFORMATICS SKILLS AND DATABASE

OPERATING SYSTEMS: MS-DOS, Windows, Red Hat Linux9
DATABASE AND RDBMS: MySQL, PHP
PROGRAMMING LANGUAGE: C, JAVA, VB.NET, Perl, HTML
BIOINFORMATICS DATABASE: NCBI, GenBank, SwissProt, EMBL, DDBJ, PDB, UniProt, PDBsum

BIOINFORMATICS TOOLS: BLAST, FASTA, EMBOSS, CLUSTAL, PYMOL, MEGA, MODELLER, AUTODOCK, MAFFT, VMD, CHEMSKETCH, Discovery Studio

WORKSHOPS/CONFERENCES/SYMPOSIUM ATTENDED

- Presented in poster presentation on "29th All India Congress of Zoology and International Symposium on culture based fisheries inland open waters and satellite Symposium on Fish Immunology" organized By: Zoological Society of India, Inland Fisheries society of India and ICAR-Central Inland Fisheries Research Institute, 9-11 June, 2017 held at ICAR-CIFRI, Barrackpore, Kolkata on the Session of Molecular Biology and Bioinformatics
- Helped in demonstration of Bioinformatics tools in the International Workshop on "Bioinformatics in Fisheries and Aquaculture." 19-21 June, 2017 Organized by ICAR-Central Inland Fisheries Research Institute, Barrackpore, Kolkata
- Participated as delegate in 19th Workshop on "Computational and Structural Bioinformatics", 09-10 Jan, 2017 Organized by Bioinformatics Centre, Mahatma Gandhi Institute of Medical Science, Sevagram
- Participated in National Seminar on "Microbial Technology: Prospects and Application" 25-26 Dec, 2015 Organized by Orissa University of Agriculture &Technology (O.U.A.T), Bhubaneswar-751003

PERSONAL SKILLS

- Good Communication skills with an analytical bent of mind to grasp new concepts easily.
- Outstanding loyalty and commitment to the people and ability to work hard and smart.
- Ability to use sound judgment in decision-making.
- Determined, Punctual, positive attitude.
- Singing

PERSONAL DETAILS

Name: Sunanda Paramanik
Father's Name: Mr. Premananda Paramanik
Date of Birth: 10/05/1995
Category: OBC
Sex: Female
Marital Status: Unmarried
Nationality: Indian
Address: At-Sahapur, P.O.-Kumuda Jaypur, Via-Kuanpal, Dist-Cuttack, Odisha -754204

Languages known: English, Hindi, Odia

REFERENCES

Dr. Basanta Kumar Das, Ph.D, PDF (U.K.)

Director

ICAR-Central Inland Fisheries Research Institute

Indian Council of Agriculture and Research

Ministry of Agriculture & Farmers Welfare, Govt. of India

Barrackpore-700120, Kolkata, West Bengal, India

Email: director.cifri@gmail.com

Mob: +918420229567

Fax: +91-3325920388

Mr. Sukanta Kumar Pradhan

Head, Department of Bioinformatics

Coordinator (DBT- Bioinformatics Infrastructure Facility Centre)

Orissa University of Agriculture and Technology

Bhubaneswar-751003, Odisha, India

Email: ksukantapradhan@gmail.com

Tel/Fax: +91-6742565760

Mob: +91-9437442622

DECLARATION

I confirm that the above information has been furnished by me and is true to the best of my knowledge and belief.

Date: 26.07.2017 Sunanda Paramanik
Place: Kolkata **Signature**