"Selecting drug targets in contemporary Human pathogenicdiseases and screening them across a database of antimicrobial peptides for plausible therapeutic interventions"

By

PUSPITA GIRI

(Roll No. – 30017821017)

MAULANA ABUL KALAM AZAD UNIVERSITY OFTECHNOLOGY, WEST BENGAL

Under the guidance of

DR. SANKAR CHANDRA BASU

Department of Microbiology

*Asutosh College (Affilated to University Of Calcutta),*Kolkata

"Selecting drug targets in contemporary Human pathogenic diseases and screening them across a database of antimicrobial peptides for plausible therapeutic interventions"

By

PUSPITA GIRI

Dissertation submitted to the

MAULANA ABUL KALAM AZAD UNIVERSITY OF TECHNOLOGY, WEST BENGAL

In Partial fulfillment
Of there quirements for the degree of

MASTER OF SCIENCE IN BIOINFORMATICS

Under the guidance of

DR. SANKAR CHANDRA BASU

Asutosh College (Affilated to University Of Calcutta),
Department of Microbiology, Kolkata

MAULANA ABUL KALAM AZAD
UNIVERSITY OF TECHNOLOGY,
WEST BENGAL

DECLARATION BY THE STUDENT

*I'm PUSPITA GIRI student of M.Sc. (Bioinformatics) hereby declare that the project report "**Selecting drug targets in contemporary Human pathogenic diseases and screening them across a database of antimicrobial peptides for plausible therapeutic interventions**" which is submitted by me to the Department of Bioinformatics, Maulana Abul Kalam Azad University Of Technology, West Bengal, in partial fulfillment of the requirement of the degree of M.Sc.(Bioinformatics), have not been submitted in part or full to any other university or institute for the award of any degree or diploma.*

Date: 15/06/2023

Signature of the Student

Roll No. - 30017821017

ASUTOSH COLLEGE
(Estd. 1916)
92, S.P. Mukherjee Road
Kolkata – 700026
Web : www.asutoshcollege.in

Phone: 2455-4504/2486-3912
Fax : (033)2486-3006
Mail: mail @asutoshcollege.in

CERTIFICATE BY THE GUIDE

This is to certify that Ms. PUSPITA GIRI, student of MULANA ABUL KALAM AZAD UNIVERSITY OF TECHNOLOGY in West Bengal, did legitimate research on the topic of "**Selecting drug targets in contemporary Human pathogenic diseases and screening them across a database of antimicrobial peptides for plausible therapeutic interventions**" in her dissertation. From January 1, 2023through June 15, 2023. She is Department of Bioinformatics, Her Roll Number is 30017821017 and Registration Number is 213001817810017. Worked honestly under my direct supervision and guidance, in partial fulfillment of the requirements for a Master of Science degree.

Sankar Chandra Basu 15/06/2023

Signature of the Guide

Dr. Sankar Chandra Basu
Assistant Professor
Department of Microbiology
Asutosh College, Kolkata

Date: 15/06/2023

Place: Kolkata

Dr. Sankar Chandra Basu
Computational Biophysics
Assistant Professor
Asutosh College, Kolkata, India
Dept. of Microbiology
(nemo8130@gmail.com)

MAULANA ABUL KALAM AZAD UNIVERSITY OF TECHNOLOGY, WEST BENGAL

Utech

In Pursuit Of Knowledge And Excellence

MAULANA ABUL KALAM AZAD UNIVERSITY OF TECHNOLOGY, W.B

ENDORSEMENT BY THE HOD

This is to certify that the dissertation entitled ***"Selecting drug targets in contemporary Human pathogenic diseases and screening them across a database of antimicrobial peptides for plausible therapeutic interventions"*** *is a bonafide research work done by Ms. PUSPITA GIRI, Department of Bioinformatics, Roll No. - 30017821017, Reg. No. - 213001817810017 under the guidance of Dr. Sankar Chandra Basu, Assistant Professor, Dept. of Microbiology, Asutosh College (Affilated to University Of Calcutta), Kolkata.*

Soumen Kumar Pati

Date:*16/06/2023*

Signature of the HOD

Dr. Soumen Kumar Pati
Associate Professor
Department of Bioinformatics
Maulana Abul Kalam Azad University of Technology, Haringhata, West Bengal

Dr. Soumen Kumar Pati
Associate Professor
Department of Bioinformatics
Maulana Abul Kalam Azad University of Technology, W.B.
(formerly West Bengal University of Technology)
Simhat, Haringhata, Nadia - 741249

Acknowledgement

Let me begin by praising Almighty God, Most Gracious and Merciful, for givingme this opportunity and giving me the abilities I need to succeed.

In order to properly guide me through the research process, I would like to express my deepest gratitude to Dr. Sankar Chandra Basu, Assistant Professor, Department of Microbiology, Asutosh College, Kolkata. He provided me with exceptional care, guidance, patience that is beyond comprehension, and a positive outlook. Through his commitment and expertise, he has always been a source of inspiration. Without his persistent efforts to make our study and thesis work commendable and creative in every way, it would not have been feasible to finish them. I do not have words to express my thankfulness towards him for his worthy guidance, fortitude and insightful scientific expertise throughout my project. Icouldn't have asked for better mentor than him.

I also thankful to our H.O.D. Dr. Soumen Kumar Pati (Associate Professor, Dept. of Bioinformatics) for his continuous support throughout the project. He deserves my gratitude for his timely support and advice during my study. His methodical technical counsel, unwavering commitment, and flawless execution throughout the project piqued my attention and gave me the courage to persevere in my efforts to complete my research. Which is a big asset for us.

I also thankful to our course co-ordinator Dr. Hridoy Ranjan Bairagya (Assistant Professor, Dept. of Bioinformatics) for conduct our course in a healthy manner and to provide all our information on time. He deserves my thanks for his quick assistance and suggestions during my studies.

And also, thankful to Dr. Chittabrata Mal (Assistant Professor, Dept. of Bioinformatics), for his motivation and valuable feedback regarding the project work. I am truly appreciative and convey my heartfelt gratitude to the non- teaching staff, Samar Naskar and Ritu Maity from the Department of Bioinformatics, for assisting us in solving all obstacles and doing our research. I am profoundly thankful to all my M.Sc. and M.Tech. classmates and friends at MAKAUT, Haringhata. They have helped me to focus towards this work.

My best friends Rijwana Parvin and Sucharita Biswas deserve a special note of gratitude and appreciation for their unwavering support, inspiration, and connection to me. They helped me a lot when I needed it.

And finally, There are no words to express how grateful I am to my parents, Manindranath Giri and Jyotsna Giri, and grandparents Narayan Giri and Ganga Giri and my extended family for their unwavering support in completing this report. I am grateful to my parents and family, whose precious blessings and selfless sacrifices have exceeded whatever accomplishment I have attained in my life. I pray to almighty that I stand up to their expectation.

*Thank you **Almighty**, for providing me with the best opportunity to thank everyone who has supported and mentored me throughout my life. I completely submit to him and thank him for the benefits that have been showered upon me.*

Regards,
PUSPITA GIRI
Roll No.- 30017821017
Dept. of Bioinformatics
Maulana Abul Kalam Azad University Of Technology, West Bengal

TABLE OF CONTENTS

Sl. No.	*Particulars*	*Page No.*
1.	*List of Figures*	*9, 10*
2.	*List of Tables*	*11*
3.	*Abstract*	*12*
4.	*Introduction*	*13 – 17*
5.	*Methods and Materials*	*18 - 25*
6.	*Results and Discussion*	*26 - 48*
7.	*Conclusion*	*49*
8.	*References*	*50 - 53*

List of Figures

Figure No.	*Figures Content*	*Page No.*
1.	**Filovirus taxonomy and Ebola virus transmission**	**15**
2.	Taxonomical classification of Ebola viruses	**21**
3.	A schematic representation of EBOV viralprotein 40 (VP40)	**22**
4.	A schematic representation of EBOV viralprotein 35 (VP35)	**23**
5.	A schematic representation of EBOV Lprotein	**23**
6.	A schematic representation of EBOVnucleoprotein (NP)	**24**
7.	Different structures for different FunctionsOf Ebola virus VP40	**25**
8.	A chain of 7JZJ PDB ID	**26**
9.	ClusPro Docking Server	**26**
10.	EnCPdock Server	**27**
11.	PyMOL	**27**
12.	ClusPro docking model	**32**
13.	Displayed Docking model by PyMOL	**33**
14.	PyMOL makes it evident that the receptor7JZJ_A and ligand 2G9P_A are docked.	**40**

Figure No.	*Figures Content*	*Page No.*
15.	PyMOL makes it evident that the receptor 7JZJ_A and ligand 2NA3_A are docked	**41**
16.	PyMOL makes it evident that the receptor7JZJ_A and ligand 2K6O_A are docked	**42**
17.	PyMOL makes it evident that the receptor7JZJ_A and ligand 2N8D_A are docked	**43**
18.	PyMOL makes it evident that the receptor7JZJ_A and ligand 2G9P_A are docked	44
19.	Scores & Plot of Top ranked model (model2) byEnCPdock	45
20.	Complementarity Plot of Top ranked model(model2) by EnCPdock	46
21.	Top ranked model's (Model2)dynamicvisualization by EnCPdock	47
22.	Contact map of Top ranked model (model2) byEnCPdock	48
23.	Features trends of top ranked model (model2) by EnCPdock	49

List of Tables

Tables No.	Tables Content	Page No.
1.	The Ebola virus's target proteins	28
2.	Anti-microbial peptide filtered	29
3.	Reranked by EnCPdock	35, 36
4.	Table of the top ranked model	38

Abstract

The spread of numerous infectious diseases caused by multi-drug resistant bacteria calls for new approaches in drug discovery. In the coming years, advances in experimental techniques and computational methods in biology together should encourage the development of new drugs. These advancements include a better understanding of the metabolism, regulation of diseases – which increases the knowledge- base from genomic studies of disease-born microorganisms, leading to more representative disease models. Anti-infective medications serve a key role in the modern world, helping to reduce the global mortality rates significantly, brought about by infectious diseases. Anti-Microbial peptides (AMPs) exhibit antimicrobial activity against a variety of pathogenic viruses, bacteria, fungi; and, are multi-functional effectors of the innate immune system on mucosal surfaces. In this time-constrained master's thesis, a critical therapeutic target amid a plethora of human viral diseases were surveyed across the available literature and the Matrix protein VP40 of the Ebola virus was selected based on a defined set of criteria. These criteria included maximum possible metabolic non-overlap between the host and the pathogen, contemporary demand of disease-specific therapeutics and other factors. A database of AMPs were created from available peptide structures from the Protein Data Bank and a series of blind docking experiments were conducted in the CLUSPRO web-server using these AMPs (ligand) to be docked onto the Ebola Matrix protein (receptor). The top-ranked docked poses returned by CLUSPRO were then accumulated and re-ranked using the structure- based binding energetic ($\Delta G_{binding}$) in the EnCPdock web-server. The lowest-energy poses for each ligand (AMP) were then compared across, to lead to one finally proposed AMP as a plausible binder of the Ebola Matrix protein. We plan to extend our pilot project in the near-future using Molecular Dynamic Simulations of the top-ranked docked pose(s).

Introduction

"Selecting drug targets in contemporary Human pathogenic diseases and screening them across a database of antimicrobial peptides for plausible therapeutic interventions" This is the main topic of this research work. Anti-infective drugs play a critical role in the modern world by dramatically lowering the rates of infectious disease-related death worldwide [1]. Anti-Microbial Peptides (AMPs) are multi-functional effectors of the innate immune system on mucosal surfaces and display antimicrobial activity against a variety of pathogenic viruses, bacteria, and fungi [1,2]. Based on a predetermined set ofcriteria, the Matrix protein VP40 of the Ebola virus was chosen as a crucial therapeutic target among a wide range of human viral infections. These requirements included the least amount of metabolic overlap between the host and the pathogen, the current need for disease-specific treatments, and other elements. via order to dock these AMPs (ligands) onto the Ebola Matrix protein (receptor), several blind docking experiments were carried out via the CLUSPRO web-server using the library of AMPs created from readily available peptide structures from the Protein Data Bank. The EnCPdock web-server's structure-basedbinding energetics (Gbinding) were then used to compile and rerank the top-ranked docked poses that CLUSPRO had returned. The lowest energy poses for each ligand (AMP) were then compared, which led to the suggestion of AMP as a likely Ebola Matrix protein binder.

- ***Introducing Human Pathogenic Disease (Ebola Virus)***

The Ebola virus has been chosen in this instance. Since its outbreak in 1976, the Ebola virus disease (EVD), one of the deadliest epidemic viral illnesses, has posed a challenge to our efforts to comprehend its ecology, epidemiology, path physiology, and pathogenesis, as well as the best course of treatment and prevention[3]. The 2013–2015 West African Ebola outbreaks, by far the worst in the disease's history, offers a fresh chance to comprehend this serious ailment in ways that might eventually result in a sizable drop in death[4].

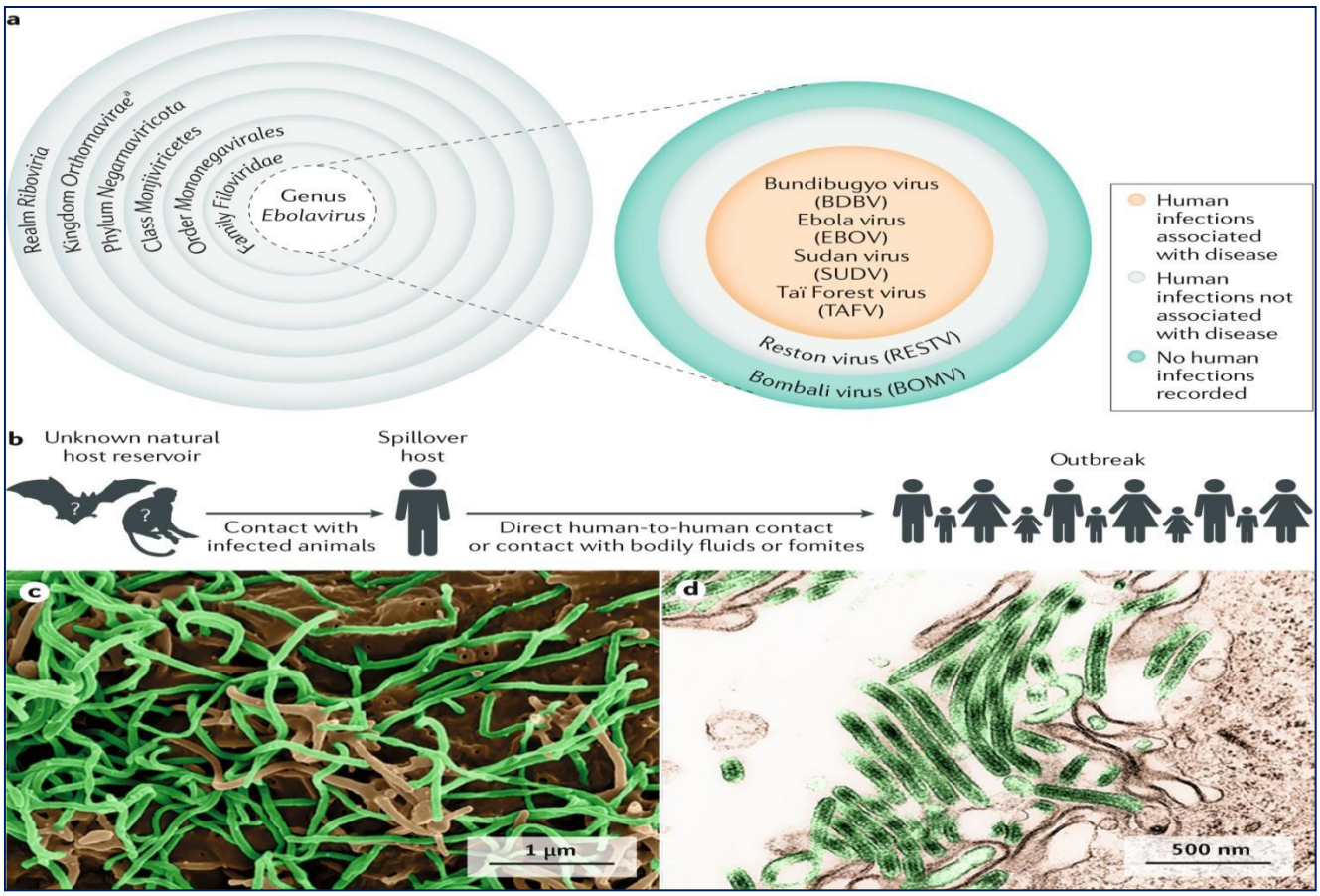


Figure 1: <u>Filovirus taxonomy and Ebola virus transmission</u>

To comprehend this serious ailment in ways that might eventually result in a sizable drop in death. The acute viral disease known as Ebola hemorrhagic fever (EHF) causes fever and a subsequent bleeding disorder that is highly lethal in both humans and nonhuman primates [3,4]. The mortality rate ranges from 50% to 100%. This filovirus is categorized as a biological class 4 pathogen because of its deadly nature [5,6]. Unknown is where the virus naturally lives. Because of this, nothing is known about how the Ebola virus spreads or how it reproduces in its host. The epidemiologic route of transmission is clearly defined, despite the fact that the primary source of the virus is unclear. Several tests have shown to be accurate and helpful for identifying the Ebola virus. EHF does not have an FDA-approved antiviral therapy. From 2 to 21 days pass during incubation. Patients who are able to establish an immune response to the virus will start their recovery and lengthy convalescence in 7 to

10 days. The main form of treatment for infected individuals is supportive management, with special focus on blood pressure control, blood volume maintenance, and theadministration of extra oxygen. Containment of this potentially fatal virus is essential because there is no particular treatment other than supportive management and palliative care. Health care professionals with confirmed illnesses had a greater mortality rate than non-health care employees in practically all EHF outbreaks[6].

- ***Introducing Antimicrobial peptides(AMPs)***

The innate immune system uses antimicrobial peptides (AMPs) to defendthe host against harmful bacterium invasion. Cationic AMPs are currently being studied as possible antibiotic substitutes to help combat the issue of antimicrobial resistance [7]. Despite having a wide range of lengths, amino acid compositions, and secondary structures, all peptides have a unique membrane-bound amphipathic conformation that can be reached. Recent research has shown that they accomplish their antimicrobial activity by interfering with a number of crucial cellular functions. Even more than one mechanism can be used by some peptides [8,9]. Additionally, it is now known that a number of whole proteins or protein fragments have innate antibacterial activity. To enable the rational design of new antimicrobial medicines, a deeper comprehension of the structure-activity connections of AMPs is necessary. With powerful antimicrobial action and distinctive antimicrobial processes, antimicrobial peptides (AMPs) are a good replacement for current antimicrobials and offer an edge over conventional antibiotics in the fight against drug-resistant bacterial illnesses [9,10].

- ***Introducing The Drug Target***

Innovative drug targets are the foundation of novel treatments in regions with significant unmet medical needs. 'Biological' have increased the range of druggable molecules, yet there are still only a finite number of suitable drug targets. The identification and evaluation of a drug target's potential therapeutic benefit is based not only on experimental, mechanistic, and pharmacological studies, but also on theoretical assessments of molecular drug ability, early assessments of potential side effects, and considerations regarding commercialization opportunities [11]. From the viewpoint of a pharmaceutical business, this article explains the essential characteristics of a good pharmacological target. Currently, the main issues with systemic drug administration include even biodistribution of medications throughout the body, a lack of drug specific affinity towards a pathological site, the requirement of ahigh total dose of a drug to achieve a high local concentration, non- specific toxicity, and other adverse side-effects brought on by high drug doses. Many of these issues may be solved by drug targeting, which is the predominance of drug accumulation in the target zone regardless of the technique and route of drug administration. The main methods of drug targeting currently used are direct drug application into the affected area, passive drug targeting (passive drug accumulation due to leaky vasculature or enhanced permeability and retention-EPR-effect), "physical" targeting (based on abnormal pH values and/or temperatures in the pathological zone), and genetic targeting [12]. Targeting employing certain "vector" molecules (ligands with a higher affinity towards the area of interest) and magnetic targeting (or targeting of a drug immobilised on paramagnetic materials under the action of an external

magnetic field). The most opportunities are offered by the final strategy. Targeted drug delivery in vivo has been successfully carried out using pharmaceutical carriers like soluble polymers, microcapsules, microparticles, cells, cell ghosts, liposomes, and micelles. The use of microreservoir-type systems offers clear advantages, Targeted drug delivery in vivo has been successfully carried out using pharmaceutical carriers like soluble polymers, microcapsules, microparticles, cells, cell ghosts, liposomes, and micelles. The use of microreservoir-type systems offers clear advantages, such as high loading capacity, possibility to control size and permeability of drug carrier systems, and use of a relatively small number of vector molecules to deliver substantial amounts of a drug to the target, though direct conjugation of a drug molecule with a targeted moiety is also possible (immunotoxin)[12,13]. We'll think about how the systems and methods indicated above can actually be used to administer therapeutic and diagnostic chemicals [13].

Subsequently checking the efficacy of those peptides' binding to the selected human pathogenic disease's therapeutic target. According to a predetermined set of criteria, the Matrix protein VP40 of the Ebola virus was chosen as a crucial therapeutic target among a wide range of human viral infections. These requirements included the least amount of metabolic overlap between the host and the pathogen, the current need for disease-specific treatments, and other elements.

Methods and Materials

The methods and materials used to arrive at the conclusions of this research work are described below.

I. Select Human Pathogenic disease

A Human pathogenic disease must be chosen before work can begin. At the beginning of the work was to find a pathogenic human disease with some contemporary biomedical relevance. The human pathogenic Ebola virus was utilized in this instance.

II. Selected Ebola virus

I discovered that the Ebola virus is a very pathogenic disease after choosing numerous pieces of information.

The greatest outbreak of Ebola hemorrhagic fever since the first case in 1976 occurred in West Africa in 2014 and is now known as the Ebola virus illness. Already, the number of illnesses and fatalities has surpassed the total number of cases reported in all prior epidemics put together. The WHO report from November 21, 2014 states that the overall number of confirmed or suspected 15,351 suspected cases are associated with 5,459 recorded deaths in the current outbreak. The most impacted nations are Guinea, Liberia, and Sierra Leone are the most affected countries [3,4,14] . This project's goal is to choose therapeutic targets for modern human infections including the Ebola virus and evaluating them for logical therapeutic intervention using a database of antimicrobial peptides.

III. Introducing Ebola Virus

There are six recognized species in the genus Ebola virus, which is a member of the Filoviridae family. The International Committee on Taxonomy of Viruses' (ICTV's) schematic taxonomy categorization o

the Filoviridae is shown in. Yang and colleagues have recently suggested a new genus, Dianlovirus, which includes the virus known as Mnglà virus (MLAV) that is present in Chinese bats. The Zaire Ebola virus species of the Ebola virus genus is what is currently referred to be the Ebola virus. Except for Reston virus (RESTV) and Bombali virus (BOMV), most members of the genus Ebola virus produce severe and frequently deadly hemorrhagic fever in humans and non-human primates (NHPs).As opposed to RESTV, which is known to be harmful for humanized mice but has not yet been linked to human infections, BOMV exclusively infects bats.

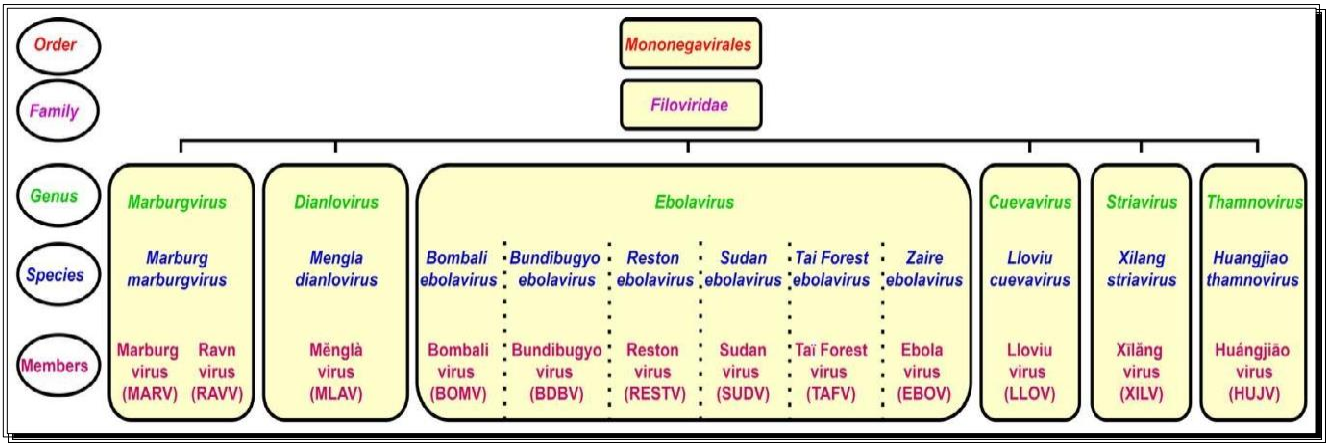


Figure 2: *Taxonomical classification of Ebola viruses.*

IV. Description of Ebola virus's proteins

a) Matrix protein VP40:

The most widely expressed protein, VP40, plays a crucial role in viral assembly and budding [16,22]. Since a mutation in the VP40 aa sequence from amino acids 292-295 changed these functions, it was claimed that these amino acids are crucial for the synthesis of VLPs and for controlling theinhibition of viral transcription. According to a recent study, aa 326 is important for the stability of VP40 because it participates in the SUMO- VP40 interaction [23]. Additionally, VP40 has two late budding domains (L- domains) that interact with host proteins at amino acids 7–10 (PTAP) and 10–13 (PPEY). While PPEY binds to ubiquitin ligase, neural precursor cell-expressed developmentally downregulated 4(Nedd4)[15],

as well as to ITCH E3 ubiquitin ligase, PTAP forms a complex with the tumor susceptibility gene 101 protein (tsg101). While the PPEY- Nedd4 complex covalently ubiquitinates viral matrix proteins, which is necessary for virus budding, the PTAPtsg101 association aids in the recruitment of VP40 into lipid raft domains on the plasma membrane. These results strongly imply that L-domains are limited in viral replication but necessary for budding [16,22,23,24].

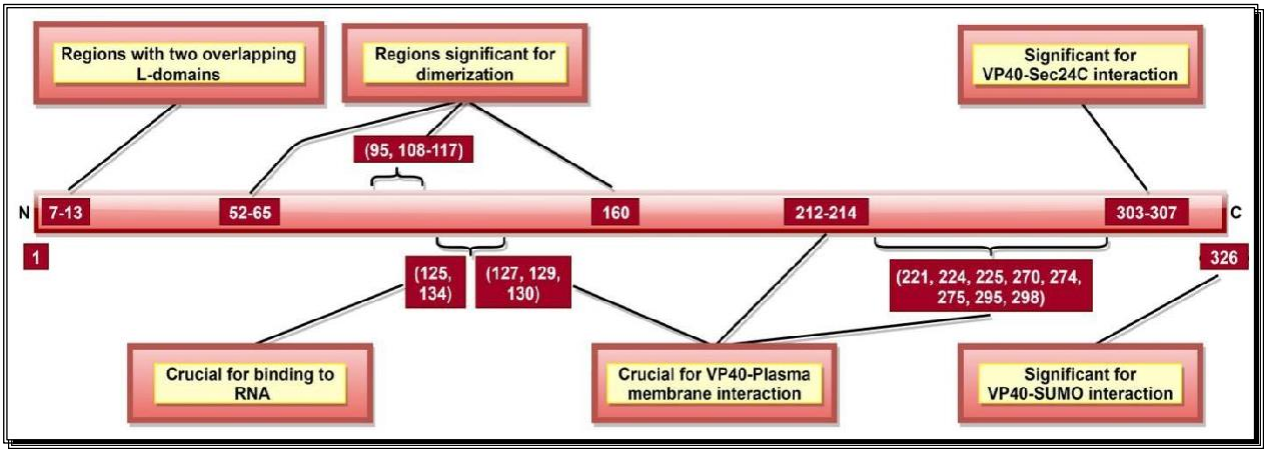


Figure 3: *A schematic representation of EBOV viral protein 40 (VP40).*

b) VP35:

Tetrameric VP35 is similar to other NNS RNA viruses in terms of functionality[16,27]. With its NTPase and helices activities, VP35 is essential for viral transcription and replication. This suggests that it may influence transcription through NTP hydrolysis or NTP-dependent unwinding of RNA helices, respectively. As it binds the monomeric state of NP to stop premature and non-specific encapsidation of viral RNA, VP35 also aids in genome packaging and nucleocapsid assembly [28].
Host immune response evasion, in which host anti-viral defence is hindered in various ways, depends on VP35. In both dsRNA-binding-dependent and dsRNA-binding-independent ways, it can inhibit the host interferon (IFN) response. In this regard, it has been demonstrated that the VP35 CTD region, in particular aa 221-340, serves as an RNA-binding domain (RBD) or an IFN inhibitory domain (IID)[27,28].

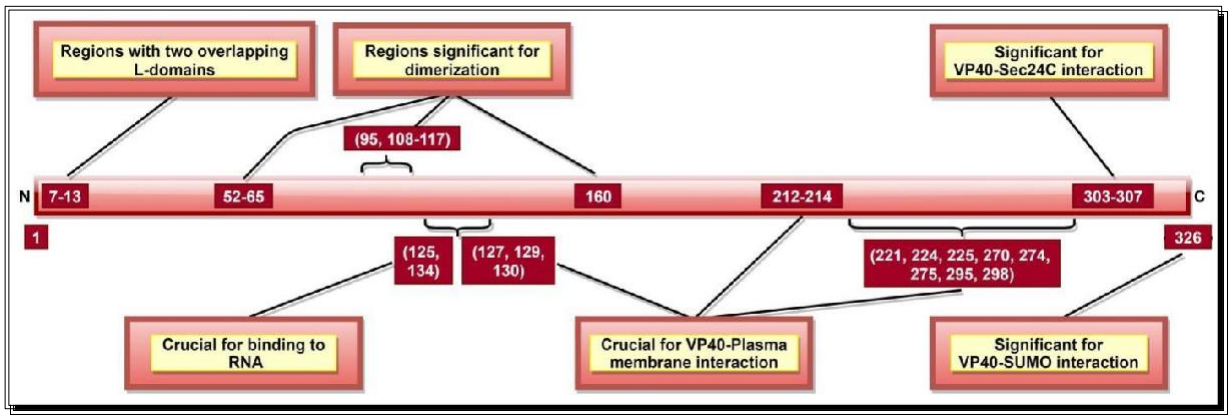


Figure 4: *A schematic representation of EBOV viral protein 35 (VP35)*

c) L protein:

The largest, most complicated, and most functional EBOV protein, L [20], contains 2212 amino acids and is a component of the RNP complex. There are five domains in L protein, namely, (a) RNA-dependent RNA polymerase (RdRp) domain with transcription/replication and polyadenylation activity, (b) capping domain with polyribonucleotidyl transferase (PRNTase) activity, (c) connector domain (CD) with an organisational role, (d) a methyltransferase domain with MTase activity,and e) a small C-terminal domain. In addition, a homo-oligomerization domain is present in residues 1-450 of EBOV L. L-VP35 interaction occurs in a non-competitive way and is not dependent on L homo- oligomerization [16,21] . L protein aa 1-380 are implicated. Re-localization of L into viral inclusion bodies is made possible by L-VP35 binding. The RESTV L protein was demonstrated to interact with VP30 in addition to VP35 [16,20,21].

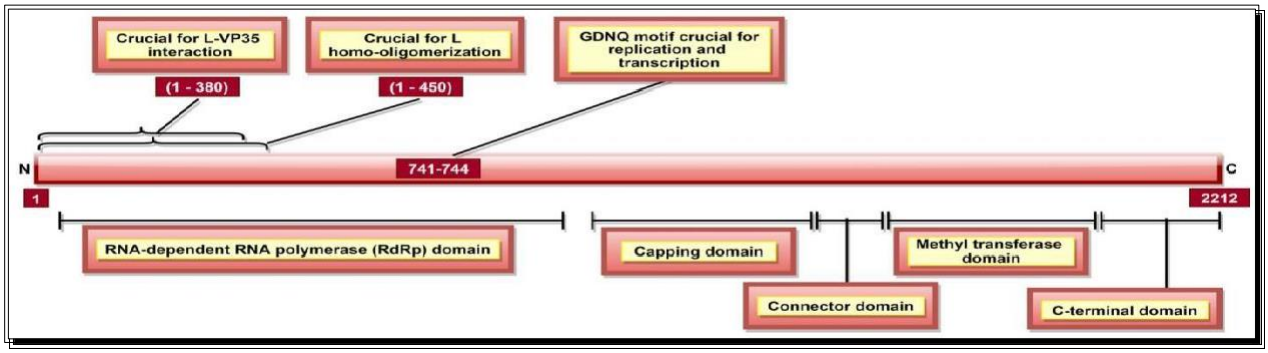


Figure 5 : *A schematic representation of EBOV L protein.*

d) Nucleoprotein (NP):

The synthesis of NC and RNP is aided by the multifunctionality of the EBOV NP protein.The importance of aa 1-600 for NC formation and viral replication has been established[16,19]. Additionally, RNA encapsidation/ssRNA binding and NP oligomerization involve amino acids 1-450. NP-ssRNA interaction, which is necessary for NC formation, is made easier by NP oligomerization. The importance of aa 111 in NP oligomerization, viral transcription, and replication was recently discovered. Another study emphasised the importance of information about inclusion bodies and the development of infectious virus-like particles (VLPs) in the NP C-terminal domain (CTD) aa 641-739. It's interesting to note that lab results show that only point mutations in NP and L are necessary for viral adaptation to various species[25,26].

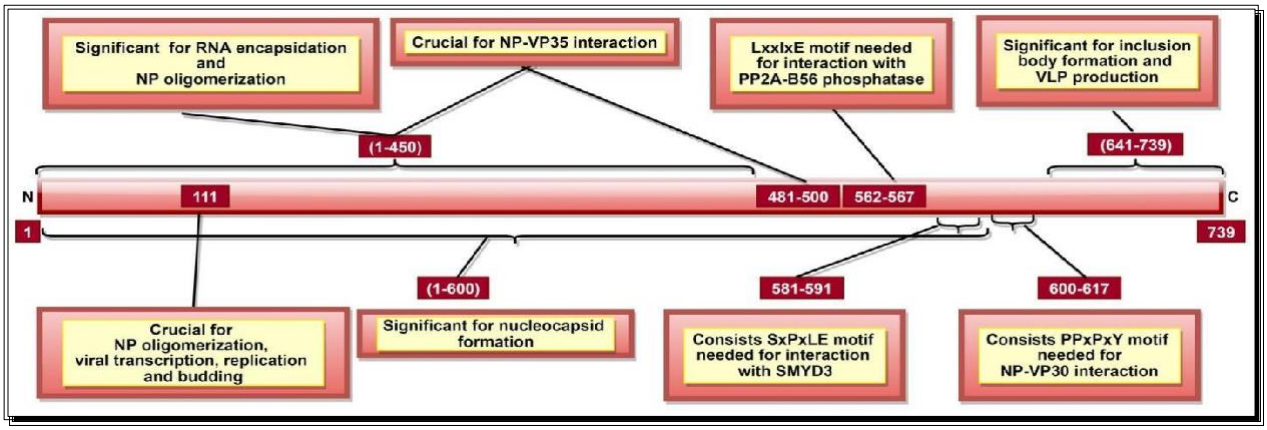


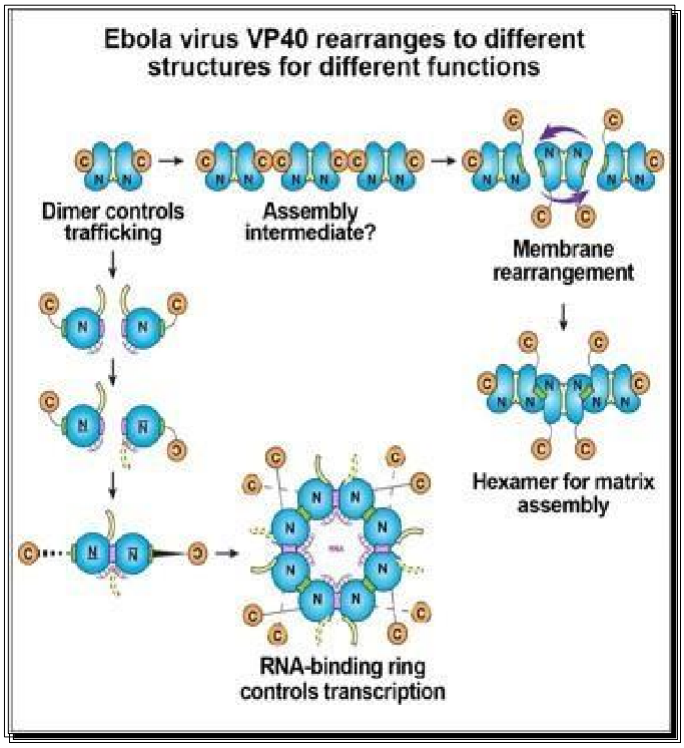
Figure 6: *A schematic representation of EBOV nucleoprotein (NP)*

e) Soluble Secreted Glycoprotein (sGP):

The full-length transmembrane spike glycoprotein (GP), which shares 295 N-terminal amino acids with sGP, is the main GP gene product[17]. sGP has been found to include six N-glycosylation sites (aa 40, 204, 228, 238, and 268) and one C-mannosylation site (aa 288). Although a structural role was considered for it in a study where sGP was substituted for GP1, creating a functional sGP-GP2 protein, it is a non-

structural protein (NSP). Bradley et al. speculated that sGP may have a role in the spread of viruses by showing that it inhibits the production of pro-inflammatory cytokines by uninfected macrophages and hinders the chemotaxis of activated macrophages. Additionally, sGP has the ability to reduce inflammation and reestablish the endothelial cell barrier function that GP had disrupted [16]. It was also demonstrated that sGP serves as a pawn for anti-GP antibodies, aiding in host immune evasion. A recent study proposed serum sGP detection as a biomarker for the diagnosis of Ebola virus disease (EVD), as significant amounts of this protein are present in blood during the early stages of the illness[16,17,18].

- *The proteins found in the Ebola virus include VP40, nucleoprotein, VP35, glycoprotein, VP24, VP30 and others. The protein matrix protein VP40 was used in this study.*



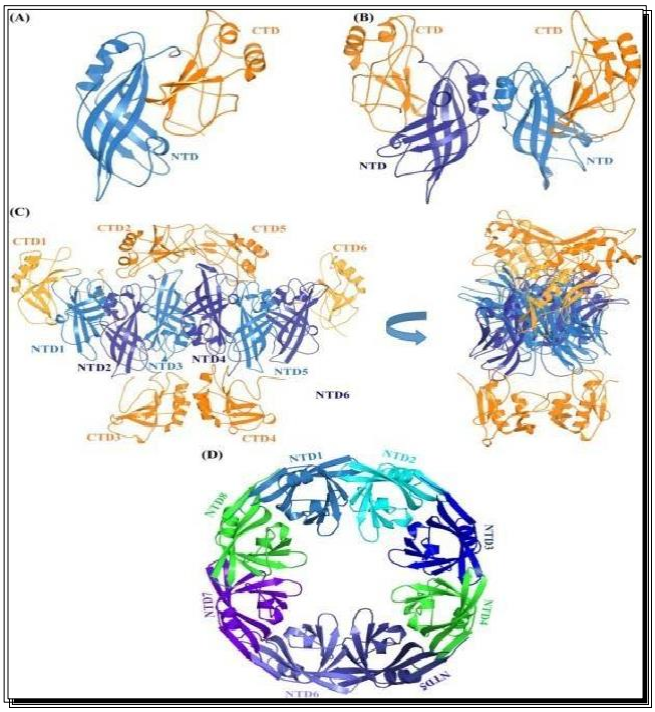


Figure 7: *Different structures for different Functions Of Ebola virus VP40*

- *The matrix protein VP40's experimental structure.7jzj is the PDB ID for this experimental structure.*

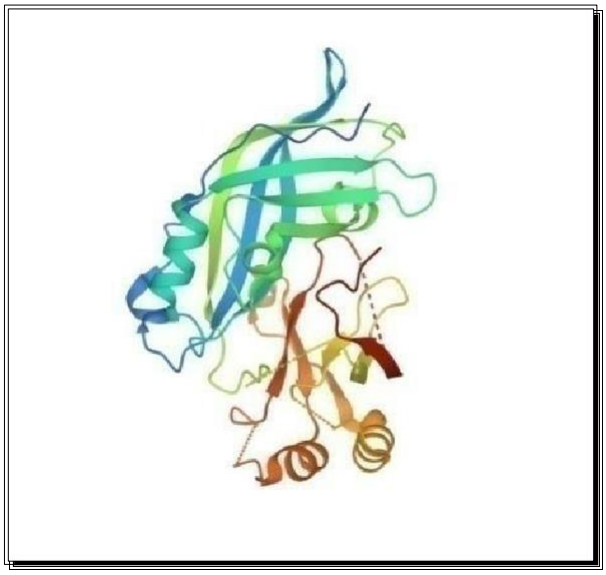


Figure 8: *A chain of 7JZJ PDB ID*

- ***Using Servers:*** *Below is a discussion of all the servers that assist in obtaining the results listed below.*

a. ***CLUSPRO:*** *Link - https://cluspro.org/help.php ,*
Version – ClusPro 2.0,
Purpose – blind docking(Protein protein Docking),
Mode - GPU mode [30]

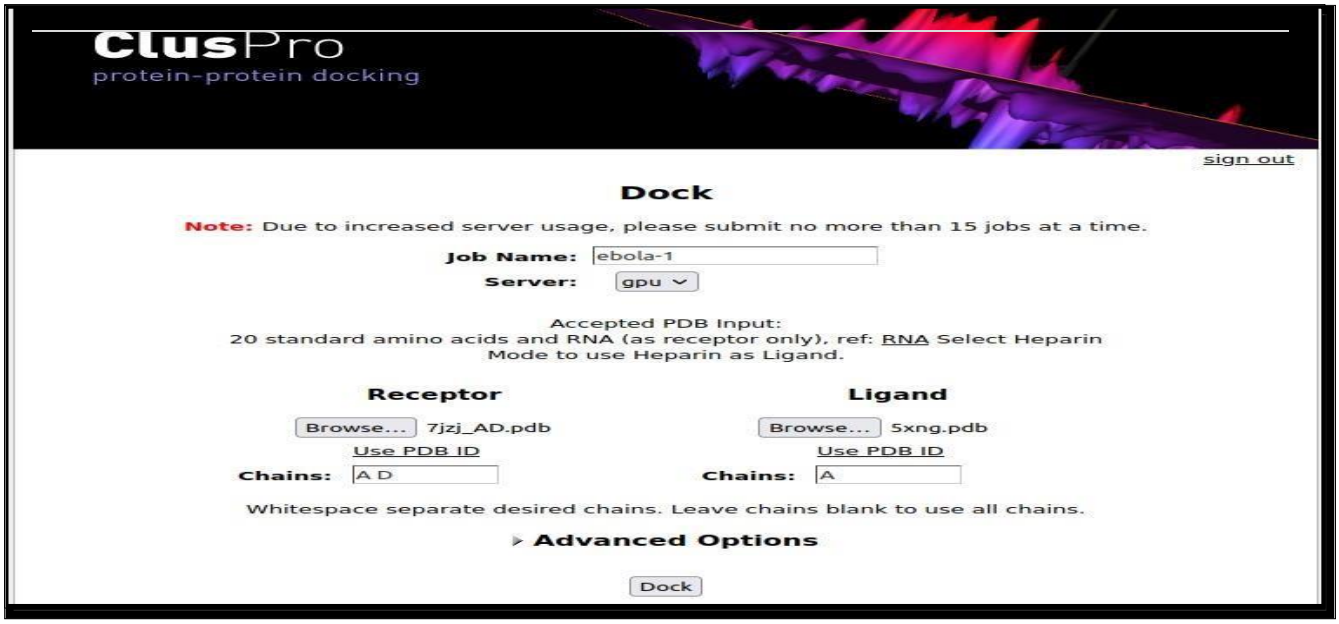


Figure 9: *ClusPro Docking Server*

b. ***EnCPdock:*** *Link -* https://scinetmol.in/EnCPdock/ [33]

Purpose – Re-ranking of the docked poses by structure based binding energetics

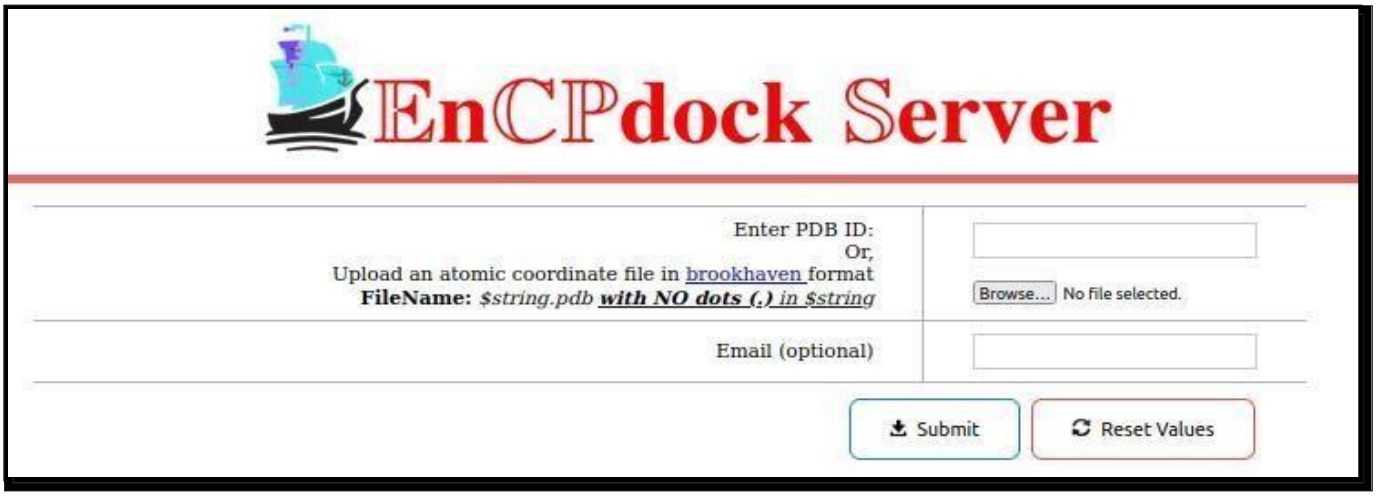


Figure 10: *EnCPdock Server*

c. ***PyMol:*** *Link –* https://pymol.org/2/ [32]

Version – PyMOL 2.5.5

Purpose - Structural analyses, superposition, creating figures and visual survey of the structures.

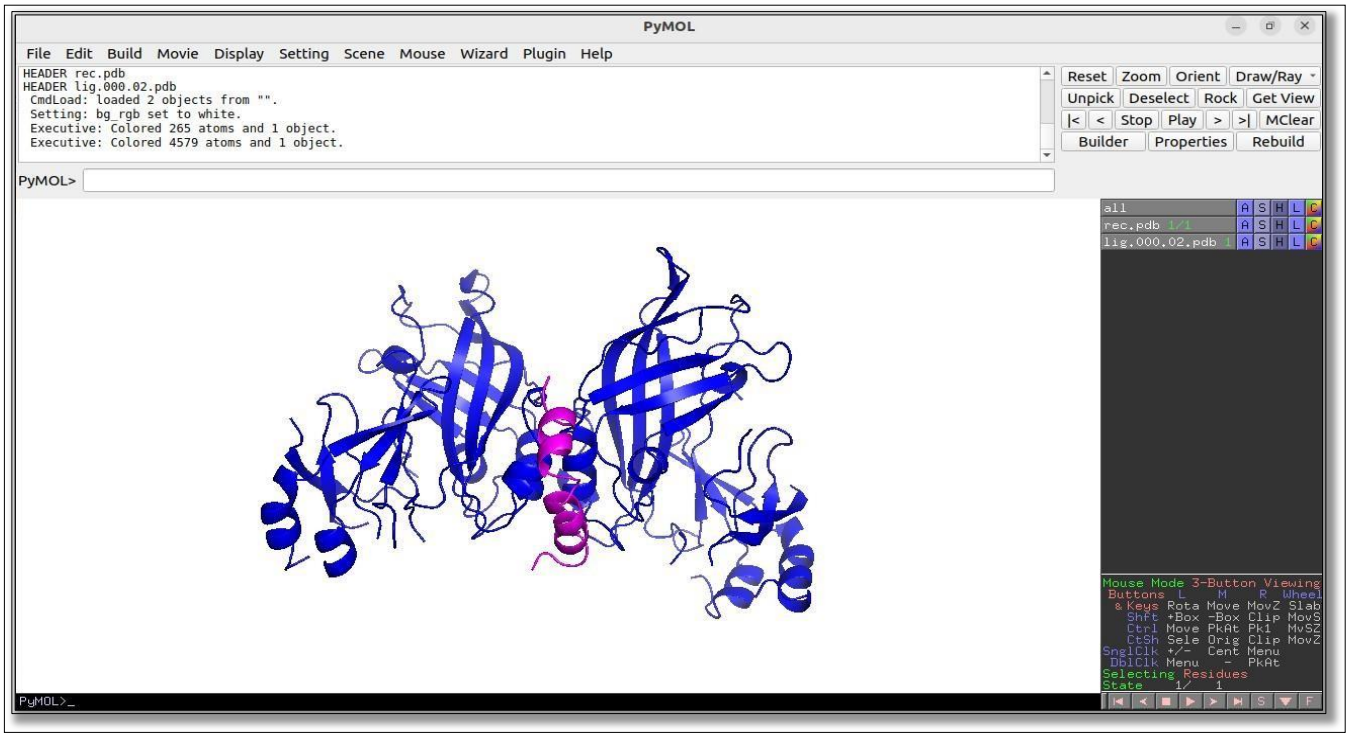


Figure 11: ***PyMOL***

Results and Discussion

One of the most effective therapeutic groups has been antibiotics, which have made many of the biggest strides in contemporary medicine possible. However, with recent reports of bacterial strains resistant to all recognized drugs, antibiotic-resistant bacteria are now recognized as serious dangers to public health. Antimicrobial peptides (AMPs) are naturally occurring compounds that have the potential to become the foundation of a new class of anti-infective that target these challenging microorganisms. With an emphasis on the clinical significance of priming the antibiotic pipeline and the potential role AMPs can play in the fight against drug-resistant bacteria, the special functions and characteristics of AMPs are reviewed. One of the most significant medical advancements of the twentieth century, antibiotics are becoming less and less effective as a wide variety of bacteria develop resistance to both common and last resort commercial medications. A newer class of compounds has been postulated as a potential source of novel anti-infective, although the conventional focus of antibiotic discovery programmers in large pharmaceutical corporations (and elsewhere) has been centered on small molecule medicines. Citation. A complex class of chemicals known as antimicrobial peptides (AMPs), also known as host defense peptides, serve as the first line of defense against microbial threats.

- *The Matrix protein VP40 of the Ebola virus was chosen as a critical therapeutic target in this time-constrained master's thesis after being reviewed in the available literature along with a number of other human viral infections. These requirements included the least amount of metabolic overlap between the host and the pathogen, the current need for disease-specific treatments, and other elements.*

- *The Protein Data Bank was used to construct a database of AMPs, and utilising this database, a series of blind docking experiments*

were carried out in the CLUSPRO web server employing the AMPs (ligands) to be docked onto the Ebola Matrix protein (receptor).

- ***Table 1: <u>Table of the Ebola virus's target proteins</u>***

Organism (Viral Pathogen)	*Protein Details*	*Experimental Structure (Yes/No)*	*Homologue in Human host (Yes/No)*	*Sequence Similarity (%)*
Ebola virus	• *Matrix protein VP40* • *Polymerase cofactor VP35* • *RNA-directed RNA polymerase L* • *Membrane-associated protein VP24* • *Nucleoprotein* • *polymerase complex protein* • *RNA-dependent RNA polymerase* • *sGP*	*Yes (**7JZJ**)*	*Little chance of human homology*	

- ***Table 2: <u>Table of Anti-microbial peptide filtered</u>***

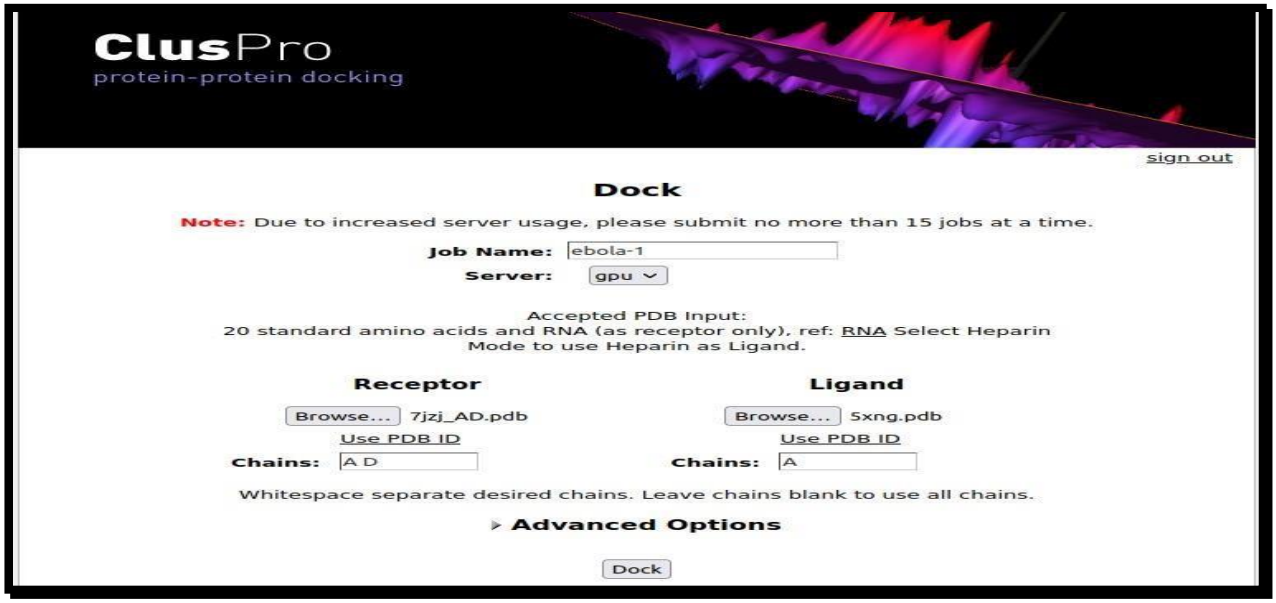
PDB ID	Chain ID	Organism	Generic Name	Experimental Method	Resolution (Å)	R-Factor (Free)	Length (aa)
6S6M	A	Human	LL37(17-29)	X-ray	1.35	0.268	13
6S6N	A	Gorilla	LL-37(17-29)	X-ray	1.10	0.179	13
7NPQ	A	Human	LL37(17-29)	X-ray	1.50	0.192	13
2K6O	A	Human	LL37	NMR	N/A	N/A	37
2NA3	A	Human	LL37	NMR	N/A	N/A	12
5XNG	A	Human	Microgel MAA60	NMR	N/A	N/A	17
5XRX	A	Human	Microgel MAA60	NMR	N/A	N/A	17
2N8D	A	synthetic construct	Lavracin	NMR	N/A	N/A	21
2AMN	A	NA	Fowlicidin-1	NMR	N/A	N/A	26
7M79	A	Schistosoma mansoni	SCHISTOCIN-3.1(PRESENCE OF DPC-D38 MICELLES)	NMR	N/A	N/A	14
7M73	A	Schistosoma mansoni	Schistocin-2	NMR	N/A	N/A	18
2G9P	A	NA	latarcin 2a from spider venom	NMR	N/A	N/A	26
7M77	A	Schistosoma bovis	Schistocin-3	NMR	N/A	N/A	14
7M67	A	Schistosoma mansoni	Schistocin-1	NMR	N/A	N/A	21
2L1Q	A	Human	Liver	NMR	N/A	N/A	40
2RLH	A	NA	RP-1 bound to DPC micelles	NMR	N/A	N/A	18
2RLG	A	NA	RP-1 bound to SDS micelles	NMR	N/A	N/A	18
2MHW	A	Bombina maxima	maximin-4 in SDS micelles	NMR	N/A	N/A	27

I. Docking analysis

- *The Ebola virus target protein 7jzj pdb id has been docked with antimicrobial peptides.*

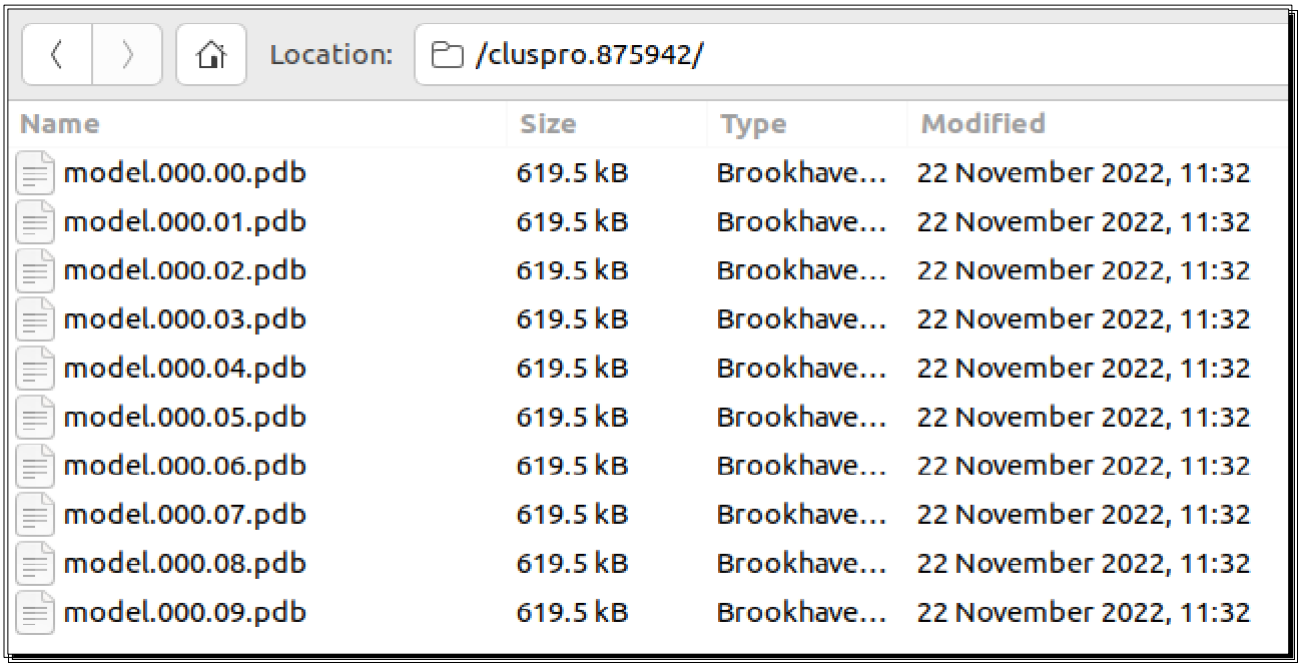
7JZJ_A_chain	Docked	2AMN_A_chain(AMP)
7JZJ_A_chain	Docked	2K6O_A_chain (AMP)
7JZJ_A_chain	Docked	2N8D_A_chain (AMP)
7JZJ_A_chain	Docked	2G9P_A_chain (AMP)
7JZJ_A_chain	Docked	2MHW_A_chain (AMP)
7JZJ_A_chain	Docked	2NA3_A_chain (AMP)
7JZJ_A_chain	Docked	2RLG_A_chain (AMP)
7JZJ_A_chain	Docked	2RLH_A_chain (AMP)
7JZJ_A_chain	Docked	5XNG_A_chain (AMP)
7JZJ_A_chain	Docked	6S6M_A_chain (AMP)

- ***Using server: Cluspro Docking server***



- ***Cluspro docking results:***

Ten results are returned for each docking after docking the Cluspro docking server.



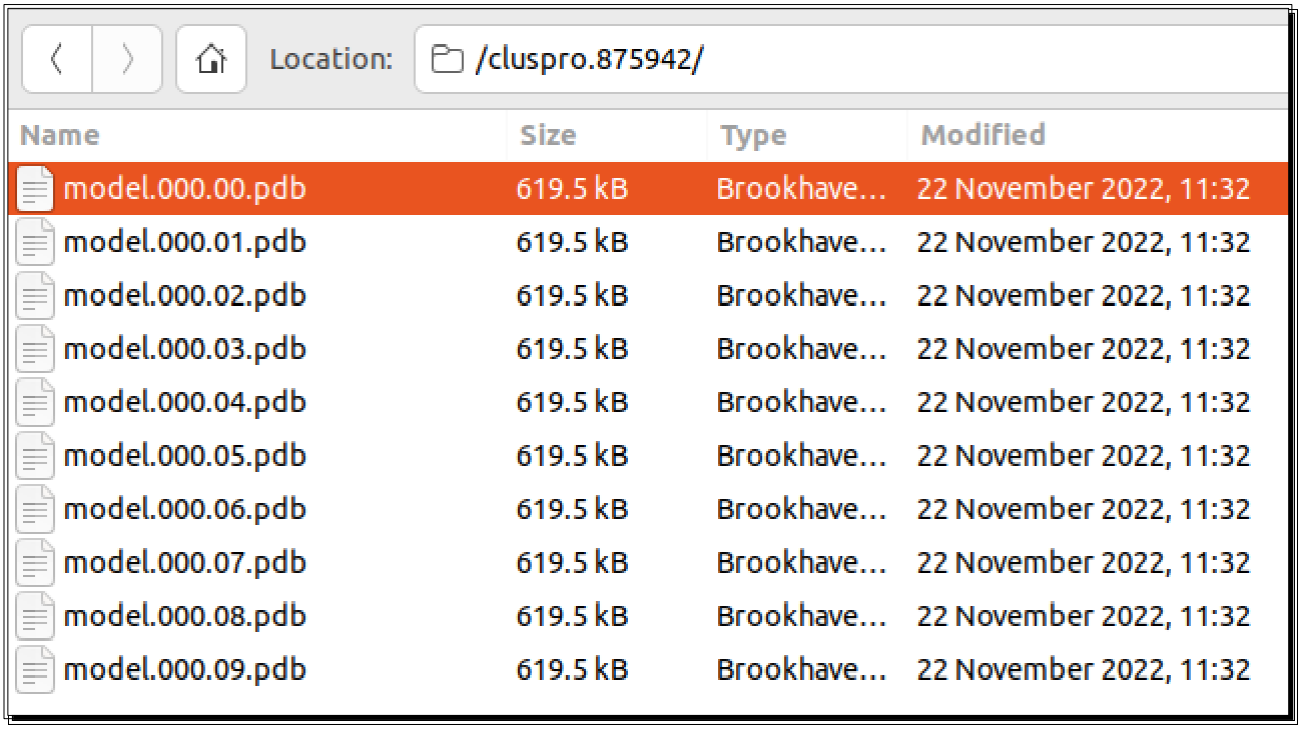


Figure 12: *ClusPro docking model*

- *A cross-platform molecular graphic tool called PyMol was used to display each docking model.*

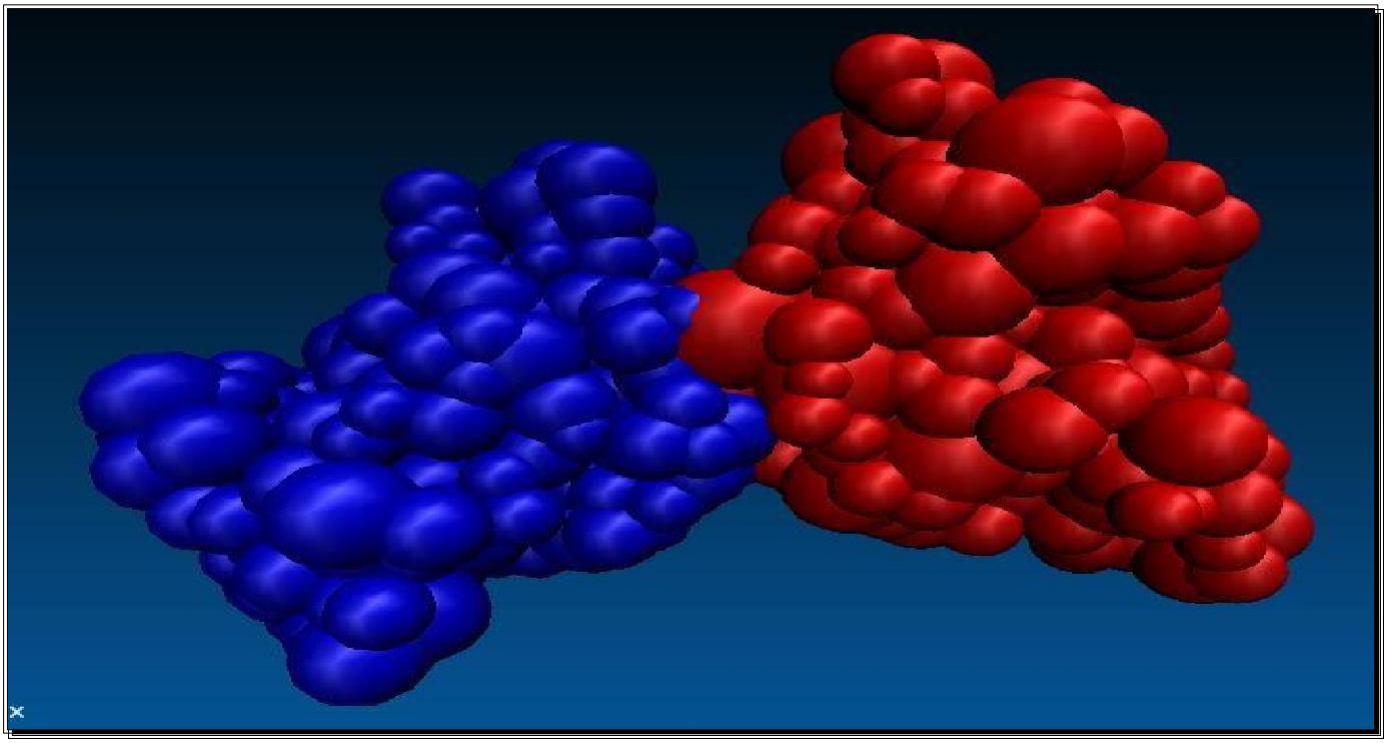


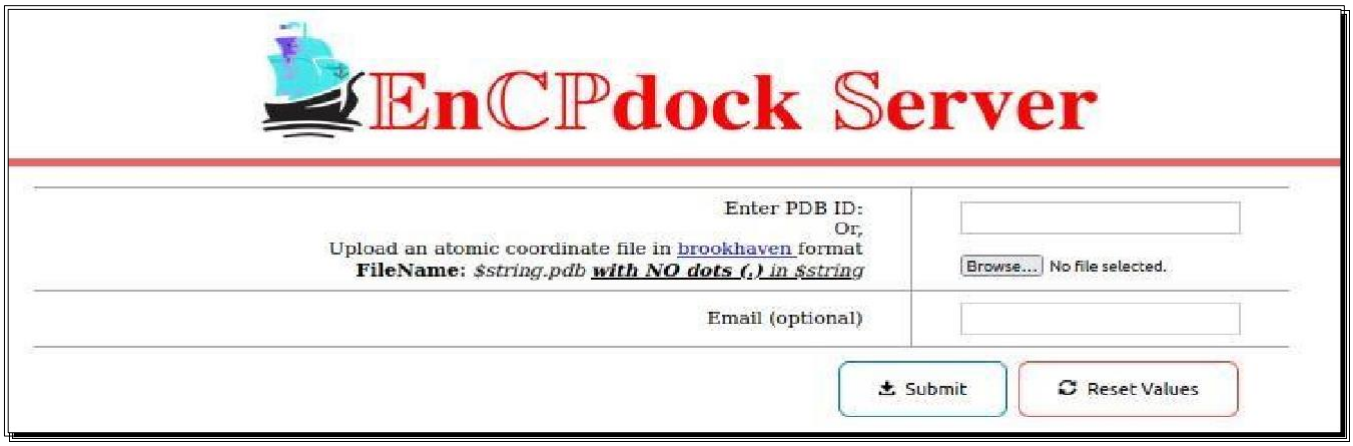
Figure 13: *Displayed Docking model by PyMOL*

II. Post Docking Analysis :

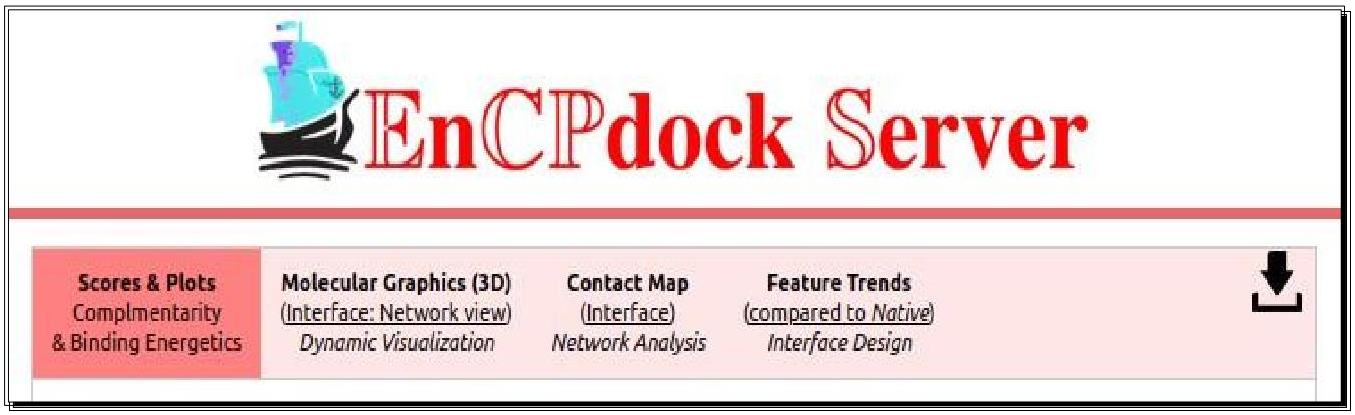
- *Models produced by docking ten antimicrobial peptides (as ligands) with the target protein of the Ebola virus (7JZJ PDB ID as receptor).Using an online web server, some contingency analysis is performed on each of these models.*

- ***Using Server : EnCPdock Server***

  The direct conjoint comparative assessments of complementarity and binding energetics in proteins are performed by the EnCPdock service. The Complementarity Plot (CP), one of its key components, is a recognised validation technique for protein structures that may be used to analyse both globular proteins (folding) and protein-protein complexes (binding). Complementarities can be calculated for both the overall interfacial surfaces of the molecular partners and for docked/bound protein-protein complexes in a residue-by-residue fashion. A variation of CP called CPdock computes the overall shape and electrostatic complementarities (Sc, EC) at the protein-protein interface and displays them as an ordered pair (Sc, EC) in a two-dimensional graphic. Indirectly determining their affinity and stability, CPdock is constructed with knowledge-based probabilistic quality assessments of the bound (or, feasibility of the docked) PPI complex.In EnCPdock, binding free energies (Gbinding) are predicted via supervised learning employing an ideal combination of these complementarity-based and other carefully selected high-level structural descriptors. Along with the estimated energetics for protein-protein interaction, the server enables one to compute, correlate, and analyse complementarity and other high-level structural descriptors.



- *Utilising the EnCPdock server, I discovered how reliable the bindings are. The models' estimates were then verified using a few EnCPdock server tools.*



III. Visualized by EnCPdock :

1. *Picking Up top ranked docked models and computing the following parameters in this peptide protein complexes*
    - *Shape complementarities*
    - *Electrostatic complementarities*
    - *Buried and assessable surface area*
    - *Network parameters(link density, average contact intensity, degree distribution profile)Pertaining to the interfacial contact networks*
2. *Relative an absolute size of ligand and reseptor*
3. *Binding free energy(Δg/binding and K/d)*

4. *The most stable peptide protein complex will be taken toa short molecular dynamics simulation for analysis of dynamic stability.*

- *A table of every model according to deltaG binding (Total) is created after verification.*

- ***Table 3.1: <u>Table of Reranked by EnCPdock</u>***

Descriptors / Energetics	***7JZJ_A + 2AMN_A***				***7JZJ_A + 2K6O_A***				***7JZJ_A + 2N8D_A***				***7JZJ_A + 2G9P_A***				***7JZJ_A + 2MHW_A***			
	Rank-1	***Rank-2***	***Rank-3***	***Rank-4***	***Rank-1***	***Rank-2***	***Rank-3***	***Rank-4***	***Rank-1***	***Rank-2***	***Rank-3***	***Rank-4***	***Rank-1***	***Rank-2***	***Rank-3***	***Rank-4***	***Rank-1***	***Rank-2***	***Rank-3***	***Rank-4***
<PDB ID>_<Chain-ID>	***model 1***	***model7***	***model2***	***model5***	***model8***	***model0***	***model1***	***model6***	***model1***	***model9***	***model0***	***model 7***	***model 2***	***model 5***	***model 6***	***model 4***	***model 7***	***model 2***	***model 9***	***model0***
N_{intres}	39	29	36	35	43	34	30	34	27	25	26	21	33	29	29	31	26	21	22	17
Sc	0.664	0.726	0.616	0.706	0.674	0.734	0.730	0.668	0.727	0.700	0.749	0.600	0.714	0.655	0.677	0.609	0.616	0.704	0.673	0.735
EC	0.037	0.029	-0.013	-0.321	0.040	-0.123	0.138	0.240	0.102	0.365	0.046	0.107	0.216	0.040	0.224	0.228	-0.015	0.182	0.069	-0.080
nBSA	0.101	0.074	0.091	0.082	0.126	0.086	0.084	0.095	0.077	0.080	0.078	0.065	0.079	0.086	0.075	0.075	0.068	0.062	0.078	0.056
$nBSA_p$	0.093	0.068	0.109	0.093	0.124	0.050	0.043	0.068	0.083	0.075	0.077	0.080	0.077	0.091	0.074	0.074	0.049	0.052	0.095	0.040
$nBSA_{np}$	0.105	0.077	0.083	0.076	0.126	0.104	0.105	0.109	0.074	0.082	0.078	0.058	0.080	0.084	0.075	0.075	0.078	0.067	0.070	0.064
fracI	0.129	0.098	0.117	0.113	0.150	0.103	0.109	0.135	0.093	0.099	0.097	0.076	0.106	0.113	0.096	0.092	0.082	0.084	0.098	0.075
Ld	0.087	0.113	0.110	0.115	0.084	0.094	0.109	0.104	0.130	0.167	0.158	0.145	0.135	0.108	0.131	0.129	0.103	0.157	0.137	0.208
ACI	3.188	2.609	2.818	2.758	4.324	2.556	2.833	3.207	4.476	2.840	3.538	5.500	2.324	2.273	3.731	3.074	1.882	2.647	3.188	2.267
$Slope_{dd}$	-1.878	-1.701	-1.797	-1.549	-1.977	-2.012	-1.973	-1.522	-1.911	-1.412	-0.904	-1.869	-1.610	-2.118	-1.417	-1.453	-2.161	-1.324	-2.159	-1.341
$Yinter_{dd}$	-0.422	-0.386	-0.437	-0.491	-0.379	-0.324	-0.335	-0.443	-0.392	-0.556	-0.646	-0.348	-0.531	-0.298	-0.494	-0.485	-0.214	-0.436	-0.273	-0.493
CCp_{dd}	-0.802	-0.975	-0.924	-0.979	-0.858	-0.946	-0.951	-0.975	-0.990	-0.939	-0.914	-0.998	-0.959	-0.973	-0.913	-0.890	-0.999	-0.817	-0.880	-0.969
logN	2.717	2.717	2.717	2.717	2.726	2.726	2.726	2.726	2.712	2.712	2.712	2.712	2.717	2.717	2.717	2.717	2.718	2.718	2.718	2.718
log_{asp}	1.280	1.280	1.280	1.280	1.126	1.126	1.126	1.126	1.394	1.394	1.394	1.394	1.280	1.280	1.280	1.280	1.263	1.263	1.263	1.263
$\Delta G_{EnCPdock\_per\text{-}res}$ ***(kcal. mol^{-1})***	-0.155	-0.204	-0.15	-0.121	-0.171	-0.172	-0.179	-0.141	-0.229	-0.215	-0.185	-0.214	-0.25	-0.212	-0.17	-0.146	-0.187	-0.198	-0.181	-0.205
$\Delta G_{enCPdock\_total}$*** ***(kcal. mol^{-1})***	**-6.025**	**-5.909**	**-5.413**	**-4.248**	**-7.374**	**-5.846**	**-5.37**	**-4.785**	**-6.196**	**-5.367**	**-4.812**	**-4.484**	**-8.25**	**-6.138**	**-4.927**	**-4.535**	**-4.852**	**-4.148**	**-3.972**	**-3.492**

- ***Table 3.2: <u>Table of Reranked by EnCPdock</u>***

Descriptors / Energetics	***7JZJ_A + 2NA3_A***				***7JZJ_A + 2RLG_A***				***7JZJ_A + 2RLH_A***				***7JZJ_A + 5XNG_A***				***7JZJ_A + 6S6M_A***			
	Rank-1	***Rank-2***	***Rank-3***	***Rank-4***	***Rank-1***	***Rank-2***	***Rank-3***	***Rank-4***	***Rank-1***	***Rank-2***	***Rank-3***	***Rank-4***	***Rank-1***	***Rank-2***	***Rank-3***	***Rank-4***	***Rank-1***	***Rank-2***	***Rank-3***	***Rank-4***
<PDB ID>_<Chain-ID>	***model 3***	***model9***	***model 5***	***model 0***	***model 8***	***model 0***	***model 6***	***model 5***	***model 6***	***model 7***	***model 2***	***model 8***	***model 8***	***model 5***	***model 7***	***model 0***	***model4***	***model7***	***model8***	***model2***
N_{intres}	33	27	23	18	29	19	23	17	24	26	23	19	30	20	24	28	20	24	27	16
Sc	0.736	0.745	0.691	0.652	0.641	0.755	0.655	0.685	0.756	0.668	0.652	0.616	0.757	0.717	0.759	0.670	0.777	0.738	0.682	0.802
EC	0.131	-0.003	-0.208	0.093	0.367	0.277	0.238	0.410	0.299	0.187	0.272	0.378	-0.016	0.122	0.171	0.007	0.223	0.179	0.293	0.146
Nbsa	0.096	0.065	0.069	0.059	0.073	0.049	0.056	0.048	0.061	0.066	0.063	0.061	0.071	0.068	0.061	0.068	0.051	0.057	0.067	0.041
$nBSA_{p}$	0.127	0.068	0.080	0.048	0.067	0.045	0.044	0.046	0.055	0.078	0.064	0.051	0.074	0.076	0.057	0.061	0.042	0.071	0.074	0.031
$nBSA_{np}$	0.080	0.064	0.063	0.065	0.076	0.050	0.061	0.048	0.064	0.061	0.063	0.065	0.070	0.065	0.062	0.071	0.056	0.050	0.063	0.046
fracI	0.108	0.085	0.081	0.069	0.086	0.057	0.064	0.051	0.082	0.092	0.066	0.074	0.090	0.078	0.080	0.098	0.057	0.073	0.081	0.049
Ld	0.123	0.158	0.143	0.150	0.162	0.189	0.138	0.171	0.161	0.163	0.121	0.193	0.130	0.162	0.200	0.139	0.198	0.156	0.159	0.283
ACI	3.839	3.333	2.278	5.083	3.031	4.941	2.500	2.417	3.478	3.240	3.125	2.529	3.536	5.188	3.185	2.731	2.105	3.000	3.148	2.176
$Slope_{dd}$	-1.149	-1.254	-1.907	-2.266	-1.216	-1.428	-1.806	-0.515	-1.361	-1.569	-1.678	-1.524	-1.286	-1.939	-0.991	-1.269	- 1.480	- 1.566	- 1.516	- 1.229
$Yinter_{dd}$	-0.558	-0.559	-0.342	-0.220	-0.650	-0.488	-0.369	-0.380	-0.545	-0.552	-0.342	-0.479	-0.539	-0.364	-0.714	-0.533	- 0.540	- 0.494	- 0.515	- 0.629
CCp_{dd}	-0.964	-0.828	-0.982	-0.981	-0.837	-0.904	-0.975	-1.000	-0.834	-0.953	-0.931	-0.881	-0.878	-0.920	-0.931	-0.914	- 0.928	- 0.972	- 0.941	- 0.909
logN	2.705	2.705	2.705	2.705	2.710	2.710	2.710	2.710	2.710	2.710	2.710	2.710	2.709	2.709	2.709	2.709	2.706	2.706	2.706	2.706
log_{asp}	1.615	1.615	1.615	1.615	1.439	1.439	1.439	1.439	1.439	1.439	1.439	1.439	1.464	1.464	1.464	1.464	1.581	1.581	1.581	1.581
$\Delta G_{EnCPdock\_per\text{-}res}$ ***(kcal.mol^{-1})***	-0.233	-0.162	-0.175	-0.197	-0.184	-0.279	-0.198	-0.241	-0.213	-0.164	-0.182	-0.187	-0.178	-0.254	-0.204	- 0.15	- 0.292	- 0.218	- 0.192	-0.264
$\Delta G_{enCPdock\_total}$ *** ***(kcal.mol^{-1})***	***-7.689***	***-4.383***	***-4.027***	***-3.542***	***-5.325***	***-5.308***	***-4.555***	***-4.095***	***-5.115***	***-4.274***	***-4.195***	***-3.554***	***-5.349***	***-5.078***	***-4.899***	***-4.191***	***- 5.841***	***- 5.228***	***- 5.182***	***-4.225***

- ***Abbreviation:***

  The models are organised according to the short form in the table above; the short form's definition is given below.

Files (extensions)	Abbreviations	Descriptions
*.ScEC	Sc	Shape complementarity
	EC	Electrostatic complementarity
*.asaAngsq	sum(ΔASA)	Net change in solvent Accessible Surface Area (upon complexation)
	$ASA_{complex}$	$ASA_{receptor} + ASA_{ligand}$
	nBSA	normalized Buried Surface Area
	nBSAp	normalized Buried Surface Area (polar)
	nBSAnp	normalized Buried Surface Area (non-polar)
	fracI	Fraction of atoms buried upon association
*.Nintres	N_{intres}	number of Interfacial residues
*.netlen	Ld	Link density of the interfacial contact network
	ACI	Average Contact Intensity of the interfacial contact network
	$slope_{dd}$	Slope of degree-distribution profile of the interfacial contact network (plotted in a log-log scale)
	$Yinter_{dd}$	Y-intercept of the same degree-distribution profile
	CCp_{dd} (r)	Goodness of fit (linear) of the same degree-distribution profile
	N_{rec}	Number of residues pertaining to the receptor chain
	N_{lig}	Number of residues pertaining to the ligand chain
	N_{tot}	Total number of residues pertaining to the protein-protein complex
	logN	$\log_{10}(N_{rec}+N_{lig})$
	log_{asp}	$\log_{10}(N_{rec}/N_{lig})$
*.delGp	$\Delta G_{per\text{-}res}$ (kcal.mol^{-1})	per residue free energy of interaction (binding)
*.delGp	ΔG_{tot} (kcal.mol^{-1})	Total free energy of interaction (binding)
*.Pr_fmax	Features	Input feature vectors (**Structural Descriptors**)
*.Pr_fmax	Scores	Scores for each feature obtained by the input PPI complex
*.Pr_fmax	Pr_{fmax}	Relative Probabilities of each feature-score with respect to the event of the highest observed frequency for that feature

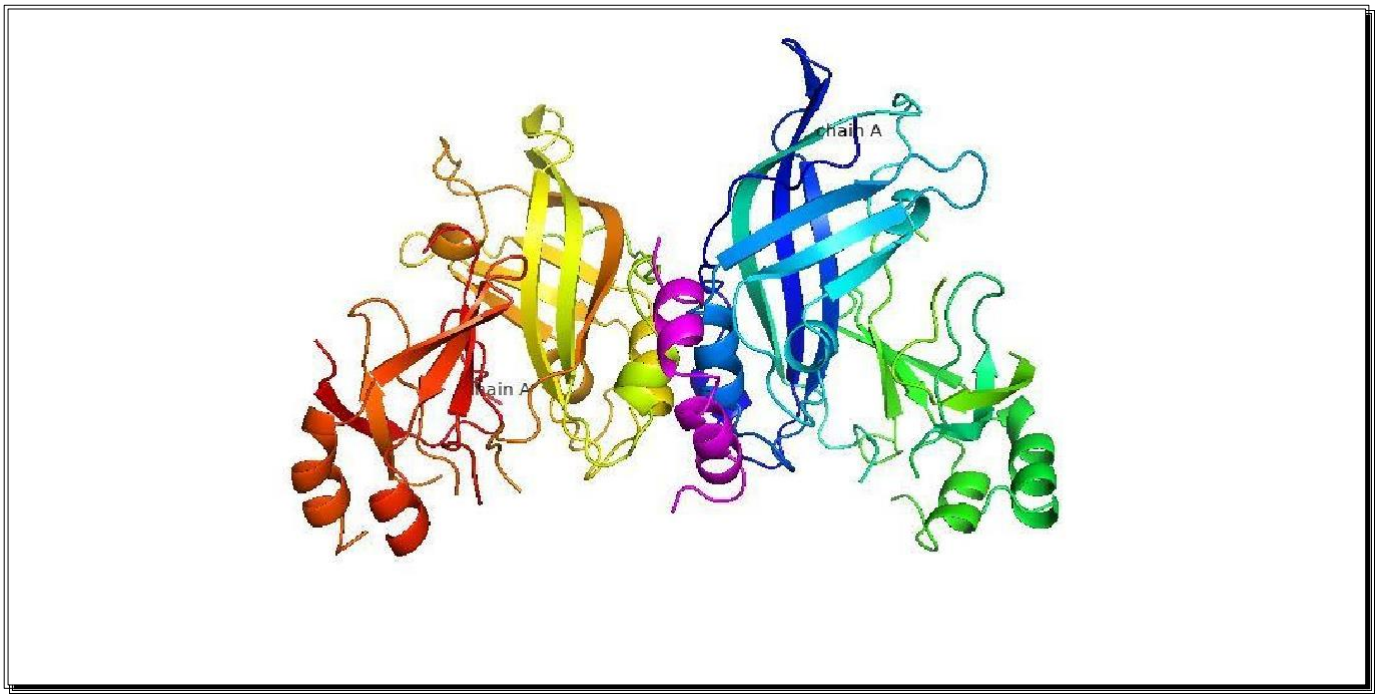
- *Using the total maximum negative free energy of interaction (binding), the top-ranked models were determined.*

- ***Table 4: <u>Table of the top ranked model</u>***

Descriptors / Energetics	**7JZJ_A + 2G9P_A**	**7JZJ_A + 2NA3_A**	**7JZJ_A + 2K6O_A**	**7JZJ_A + 2N8D_A**	**7JZJ_A + 2G9P_A**
	Top_Rank-1	**Top_Rank-2**	**Top_Rank-3**	**Top_Rank-4**	**Top_Rank-5**
<PDB ID>_<Chain-ID>	**Model-2**	**model-3**	**Model-8**	**Model-1**	**Model-5**
N_{intres}	33	33	43	27	29
Sc	0.714	0.736	0.674	0.727	0.655
EC	0.216	0.131	0.040	0.102	0.040
nBSA	0.079	0.096	0.126	0.077	0.086
$nBSA_p$	0.077	0.127	0.124	0.083	0.091
$nBSA_{np}$	0.080	0.080	0.126	0.074	0.084
fracI	0.106	0.108	0.150	0.093	0.113
Ld	0.135	0.123	0.084	0.130	0.108
ACI	2.324	3.839	4.324	4.476	2.273
$Slope_{dd}$	- 1.610	- 1.149	-1.977	-1.911	-2.118
$Yinter_{dd}$	- 0.531	- 0.558	-0.379	- 0.392	- 0.298
CCp_{dd}	- 0.959	- 0.964	-0.858	- 0.990	- 0.973
logN	2.717	2.705	2.726	2.712	2.717
log_{asp}	1.280	1.615	1.126	1.394	1.280
Residence in CP_{dock}	Probable	Probable	Less Probable	Probable	Less Probable
$\Delta G_{EnCPdock\_per-res}$ (kcal.mol^{-1})	- 0.25	- 0.233	-0.171	- 0.229	- 0.212
$\Delta G_{enCPdock\_total}$ (kcal.mol^{-1}) ***	**- 8.25**	**- 7.689**	**-7.374**	**- 6.196**	**- 6.138**

- *Model2 (7JZJ_A + 2G9P_A) is the highest ranked model in the table above.*

- *The highest rated models those with the most negative free energy obtained after ranking with free energy,Below is a description of such structures in clear detail.*
    a) ***Rank-1(Model2):*** *Total maximum negative free energy is(ΔG binding total) = -8.25Kcal.mol^{-1}*



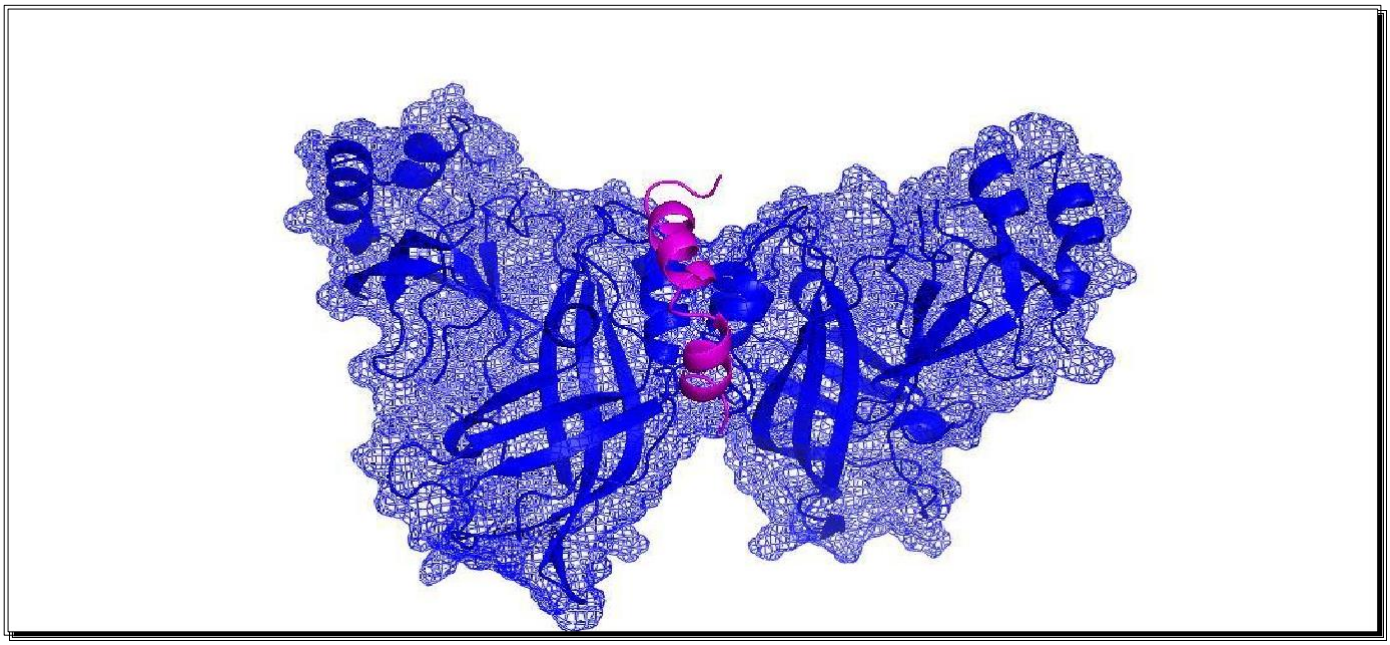


Figure 14: *PyMOL makes it evident that the receptor 7JZJ_A and ligand 2G9P_A are docked.*

b) ***Rank-2(Model3):*** *Total maximum negetive free energy is(ΔG binding total) = -7.689Kcal.mol^{-1}*

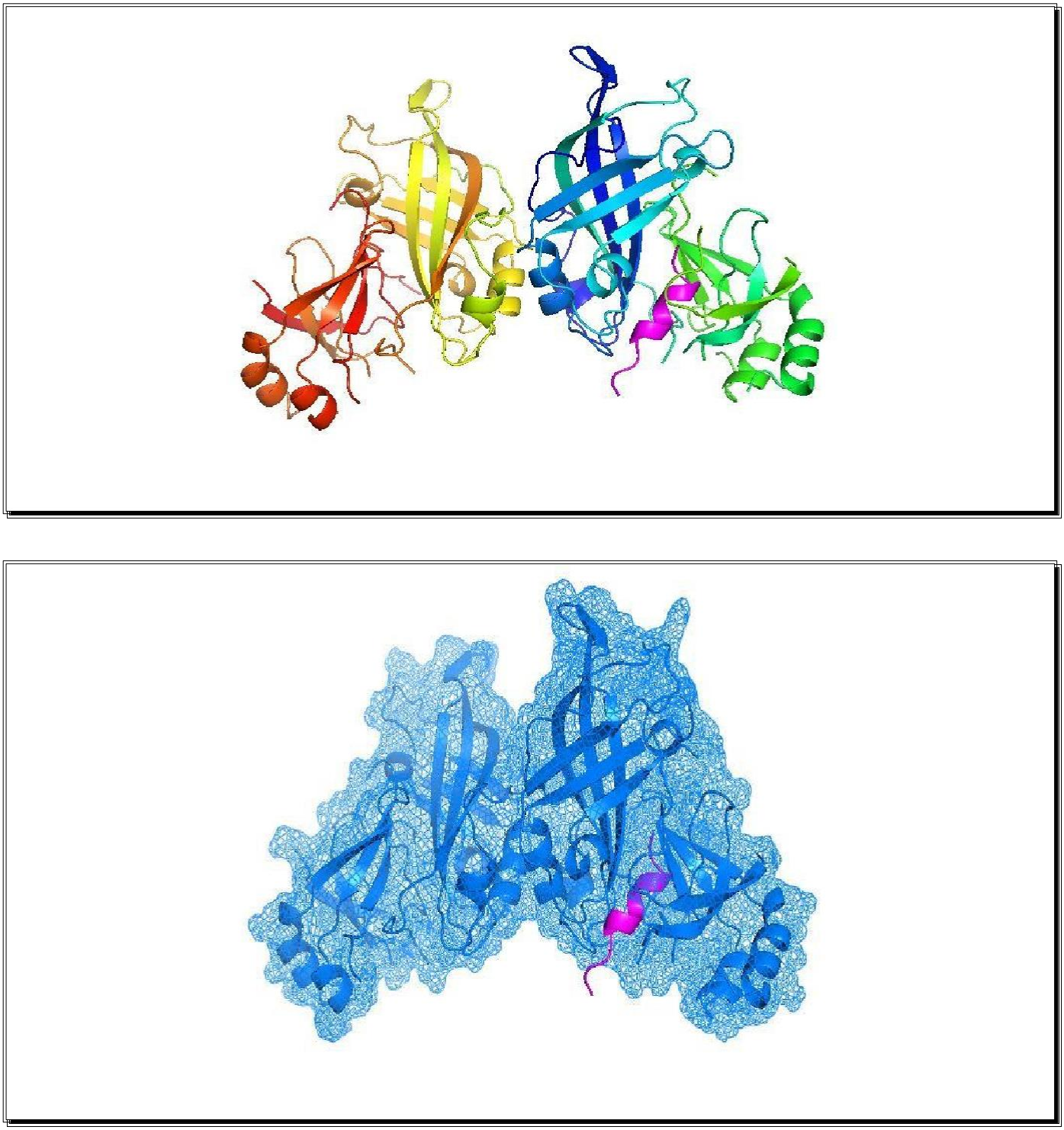


Figure 15: *PyMOL makes it evident that the receptor 7JZJ_A and ligand 2NA3_A are docked.*

c) ***Rank-3(Model8):*** *Total maximum negetive free energy is(ΔG binding total) = -7.374Kcal.mol^{-1}*

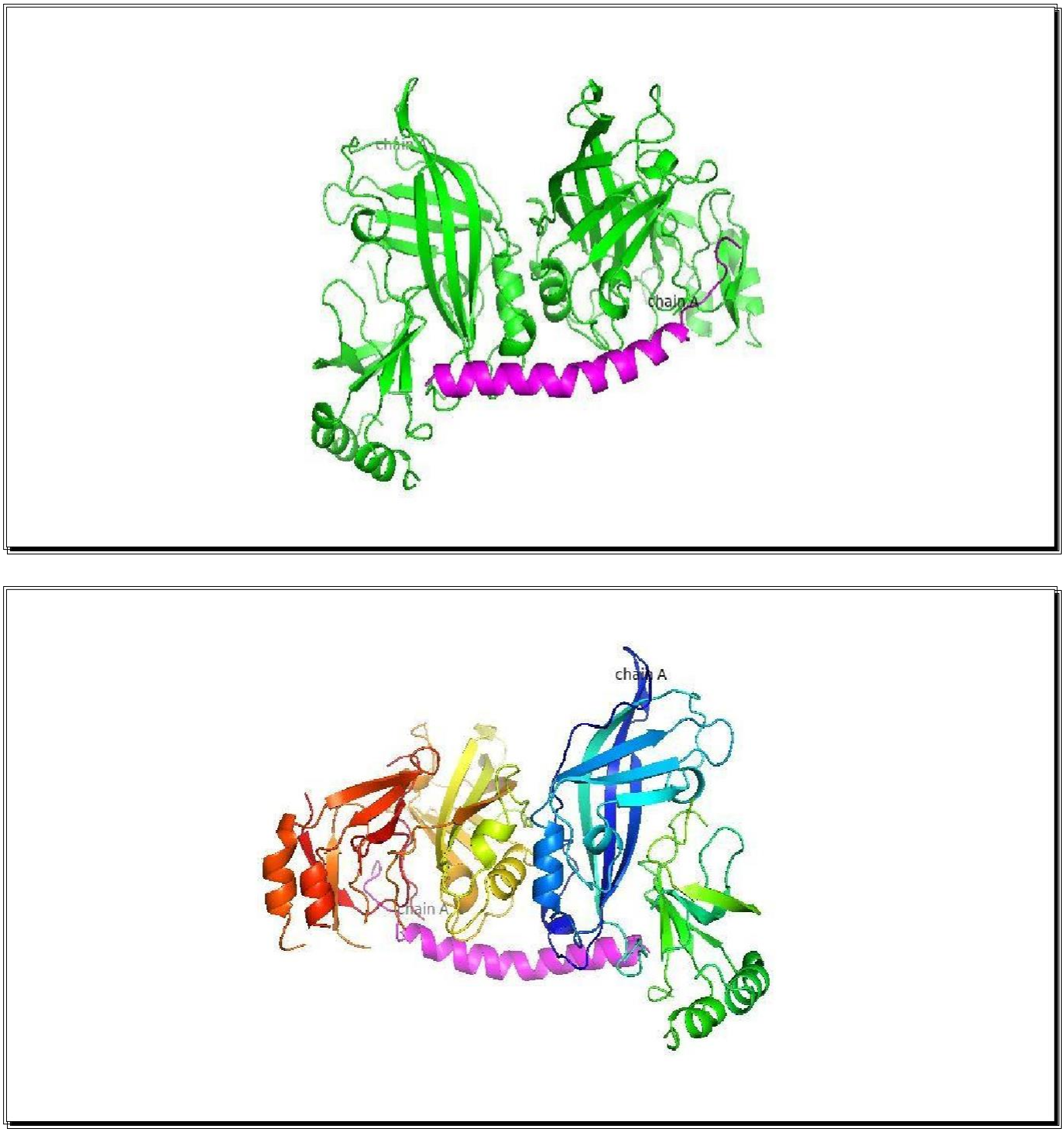


Figure 16: *PyMOL makes it evident that the receptor 7JZJ_A and ligand 2K6O_A are docked.*

d) ***Rank-4(Model1)*** ***:*** *Total maximum negetive free energy is(ΔGbinding total) = -6.196Kcal.mol^{-1}*

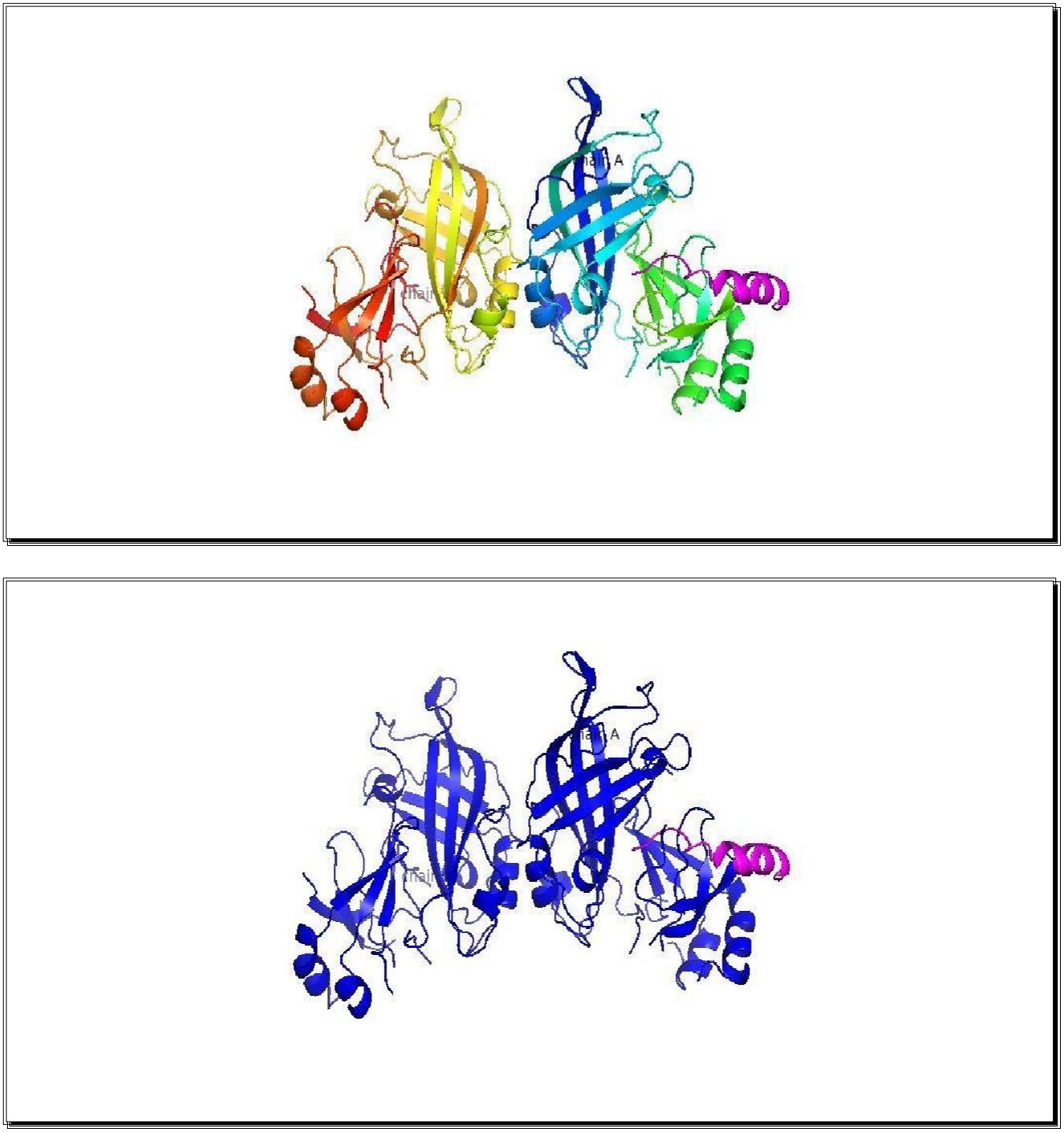


Figure 17: *PyMOL makes it evident that the receptor 7JZJ_A and ligand 2N8D_A are docked.*

e) ***Rank-5(Model5):*** *Total maximum negetive free energy is(ΔG binding total) = -6.138Kcal.mol^{-1}*

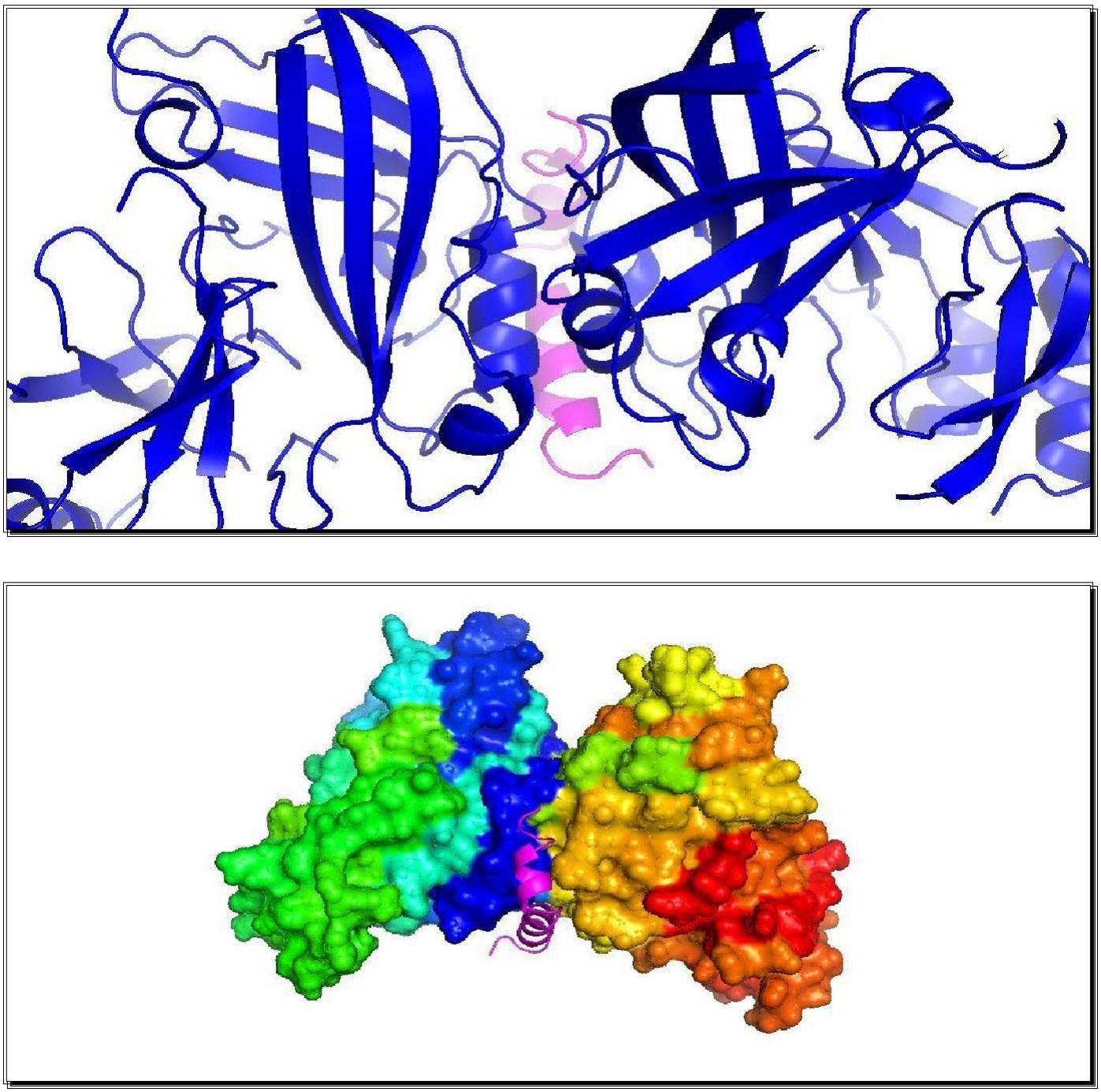


Figure 18: *PyMOL makes it evident that the receptor 7JZJ_A and ligand 2G9P_A are docked.*

- *Model2 is the most advanced. In model 2, docking of antimicrobial peptide 2G9P pdb id with Ebola virus target protein 7JZJ pdb id.*
- *Top model analyzed by EnCPdock.All of EnCPdock's tools analysis techniques have been applied to the top model. Below are images of those objects.*

a) Scores & Plot (Complementarity & Binding Energetic)

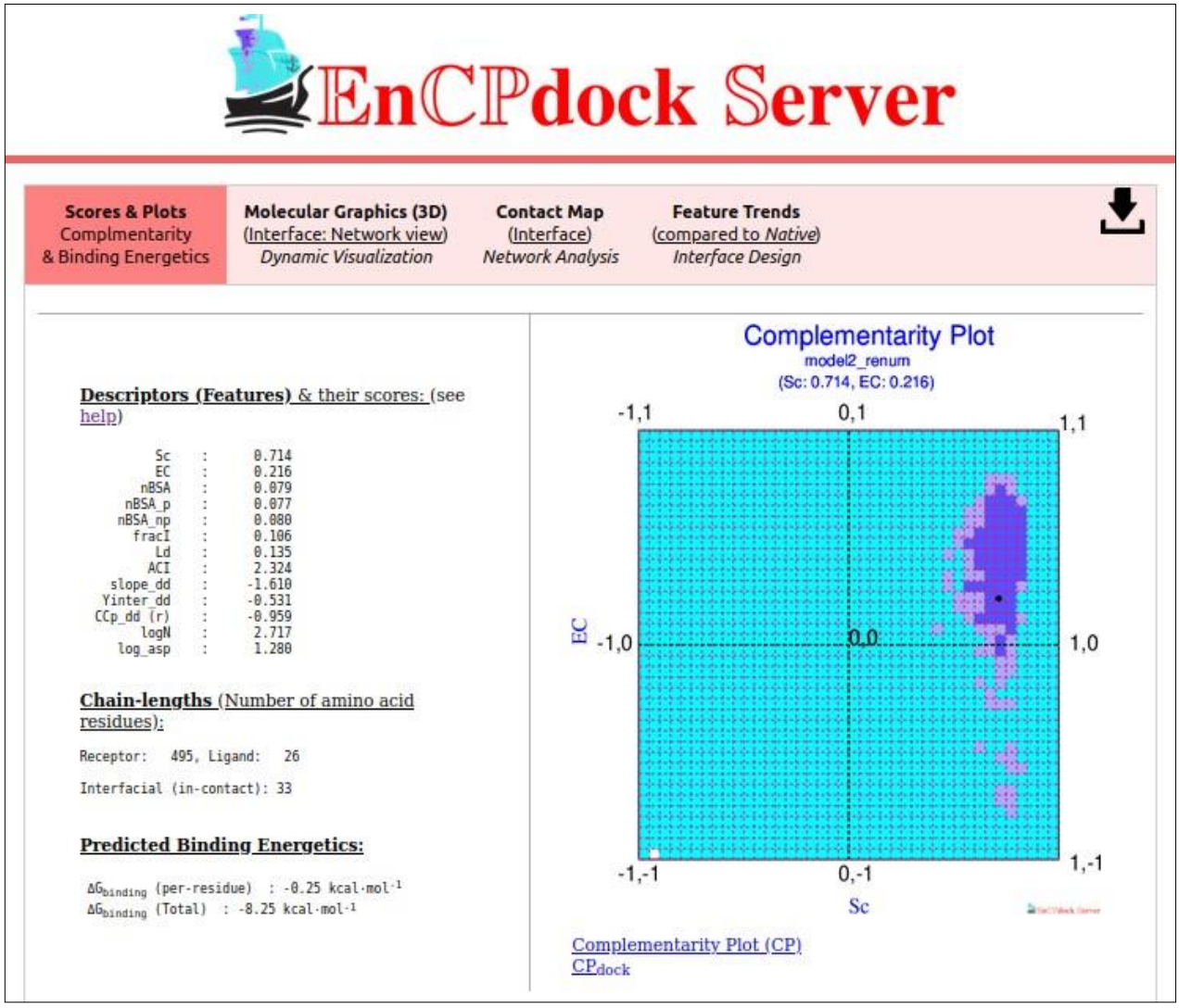


Figure 19: *Scores & Plot of Top ranked model (model2) by EnCPdock*

b) Complementarity Plot

A Complementarity Plot alternative, known as CPdock, is also available in contrast to the residue-wise plots. It allows for the plotting of individual Sc, EC values for the protein-protein interface and the evaluation of the quality of the complex atomic structure (either experimentally solved or computationally built) therein. Peter Colman and colleagues first suggested the concept of "interacting protein-protein surfaces" in the 1990s. Sc, EC are shape and electrostatic complementarities determined for such surfaces. CPdock was originally developed as a scoring function to serve as an initial filter in protein-protein docking.

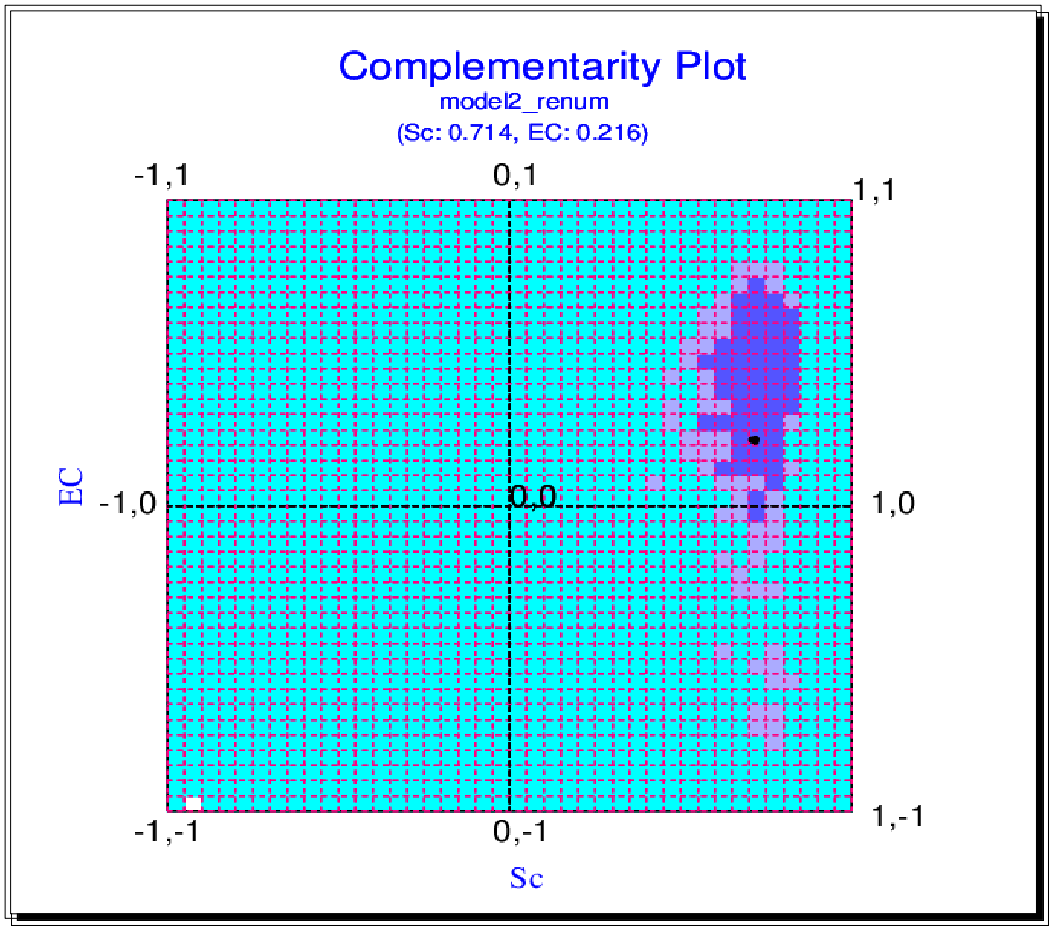
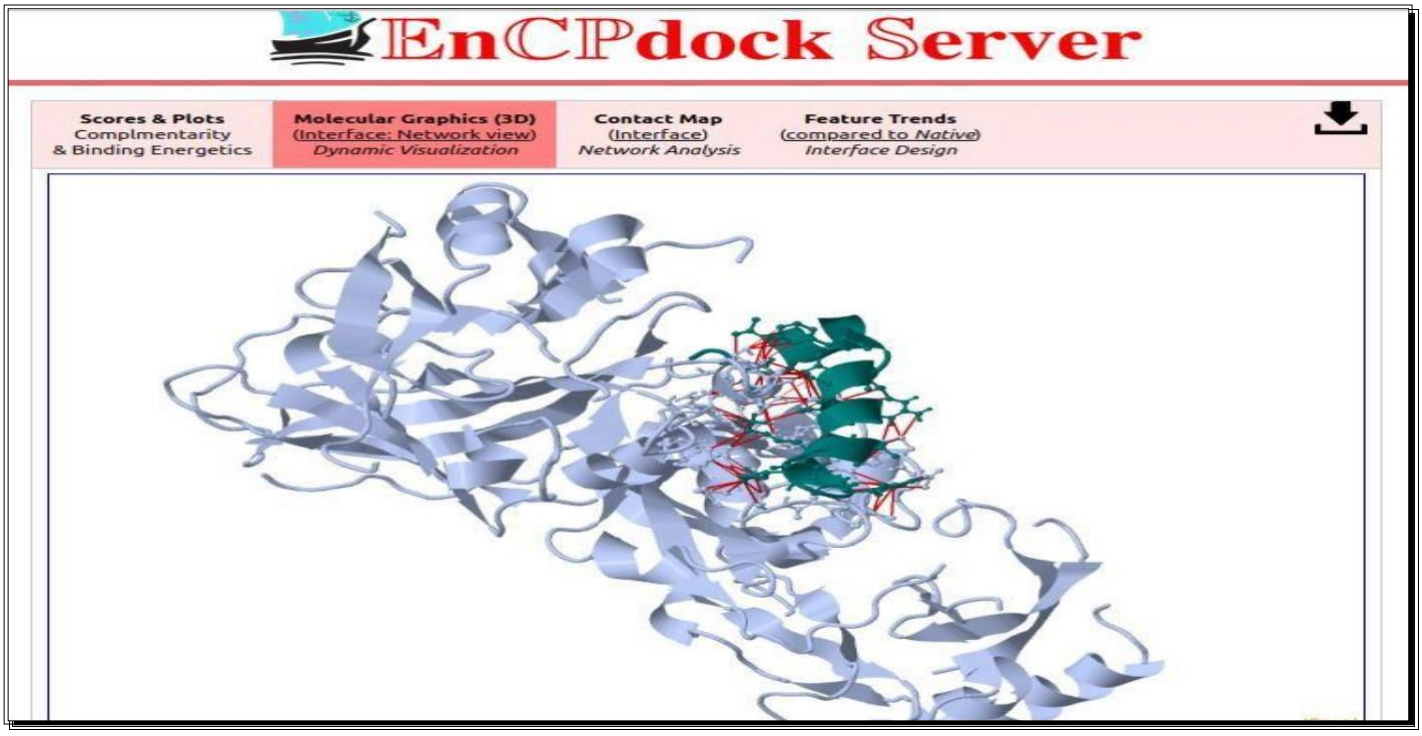


Figure 20: *Complementarity Plot of Top ranked model (model2) by EnCPdock.*

c) Molecular Graphics(3D)Dynamic Visualization



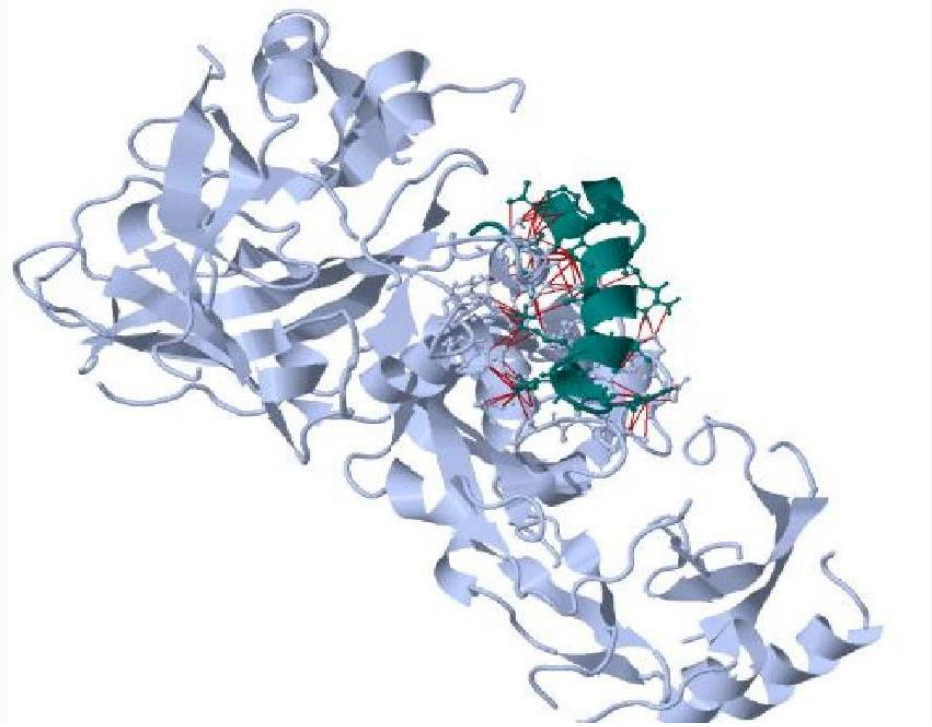
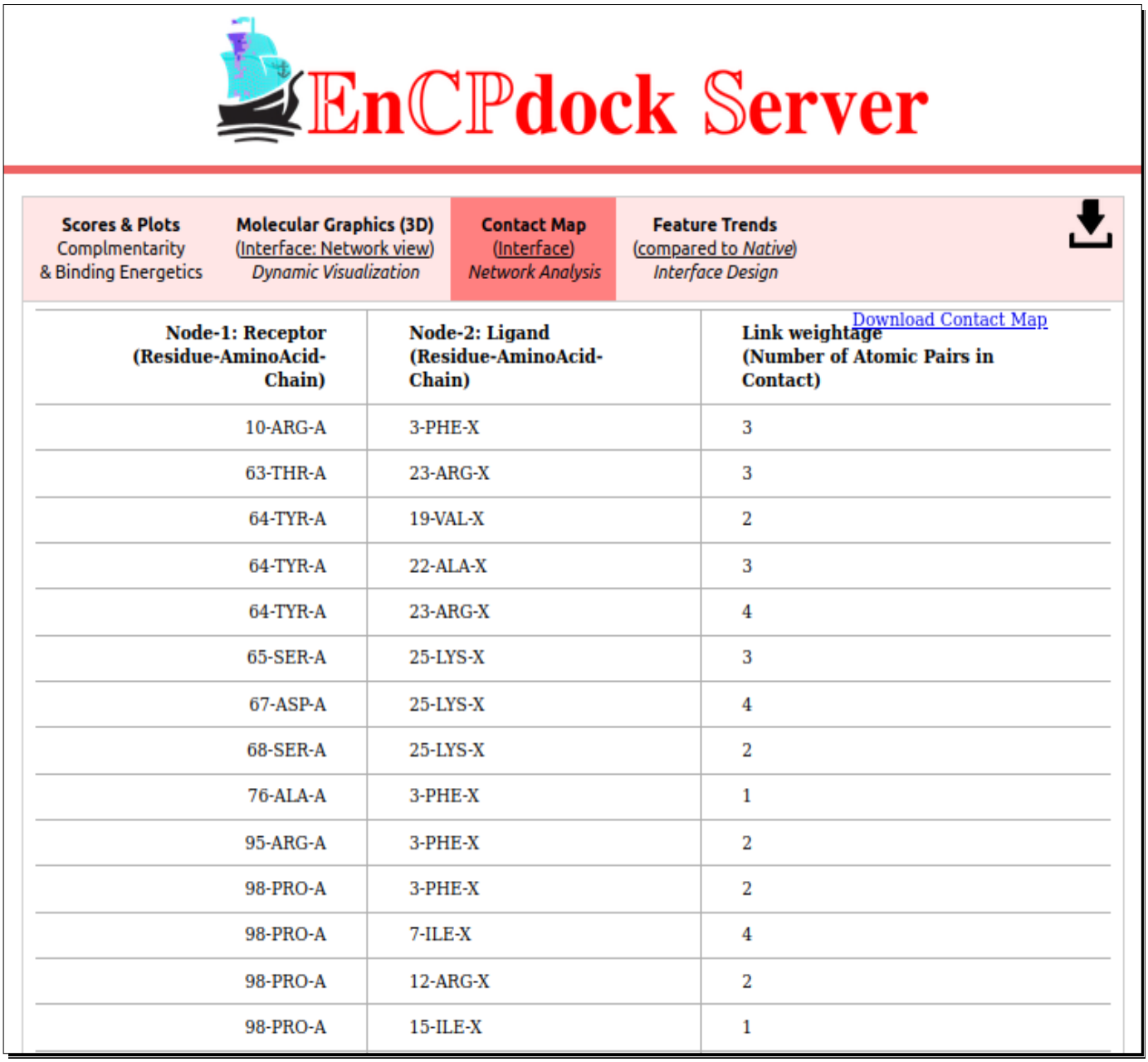


Figure 21: *Top ranked model's (Model2) dynamic visualization by EnCPdock*

d) Contact map (Network Analysis)



EnCPdock Server

Scores & Plots
Complmentarity
& Binding Energetics

Molecular Graphics (3D)
(Interface: Network view)
Dynamic Visualization

Contact Map
(Interface)
Network Analysis

Feature Trends
(compared to *Native*)
Interface Design

Download Contact Map

Node-1: Receptor (Residue-AminoAcid-Chain)	Node-2: Ligand (Residue-AminoAcid-Chain)	Link weightage (Number of Atomic Pairs in Contact)
10-ARG-A	3-PHE-X	3
63-THR-A	23-ARG-X	3
64-TYR-A	19-VAL-X	2
64-TYR-A	22-ALA-X	3
64-TYR-A	23-ARG-X	4
65-SER-A	25-LYS-X	3
67-ASP-A	25-LYS-X	4
68-SER-A	25-LYS-X	2
76-ALA-A	3-PHE-X	1
95-ARG-A	3-PHE-X	2
98-PRO-A	3-PHE-X	2
98-PRO-A	7-ILE-X	4
98-PRO-A	12-ARG-X	2
98-PRO-A	15-ILE-X	1

Figure 22: *Contact map of Top ranked model (model2) by EnCPdock*

e) Feature trends (Interface Design)

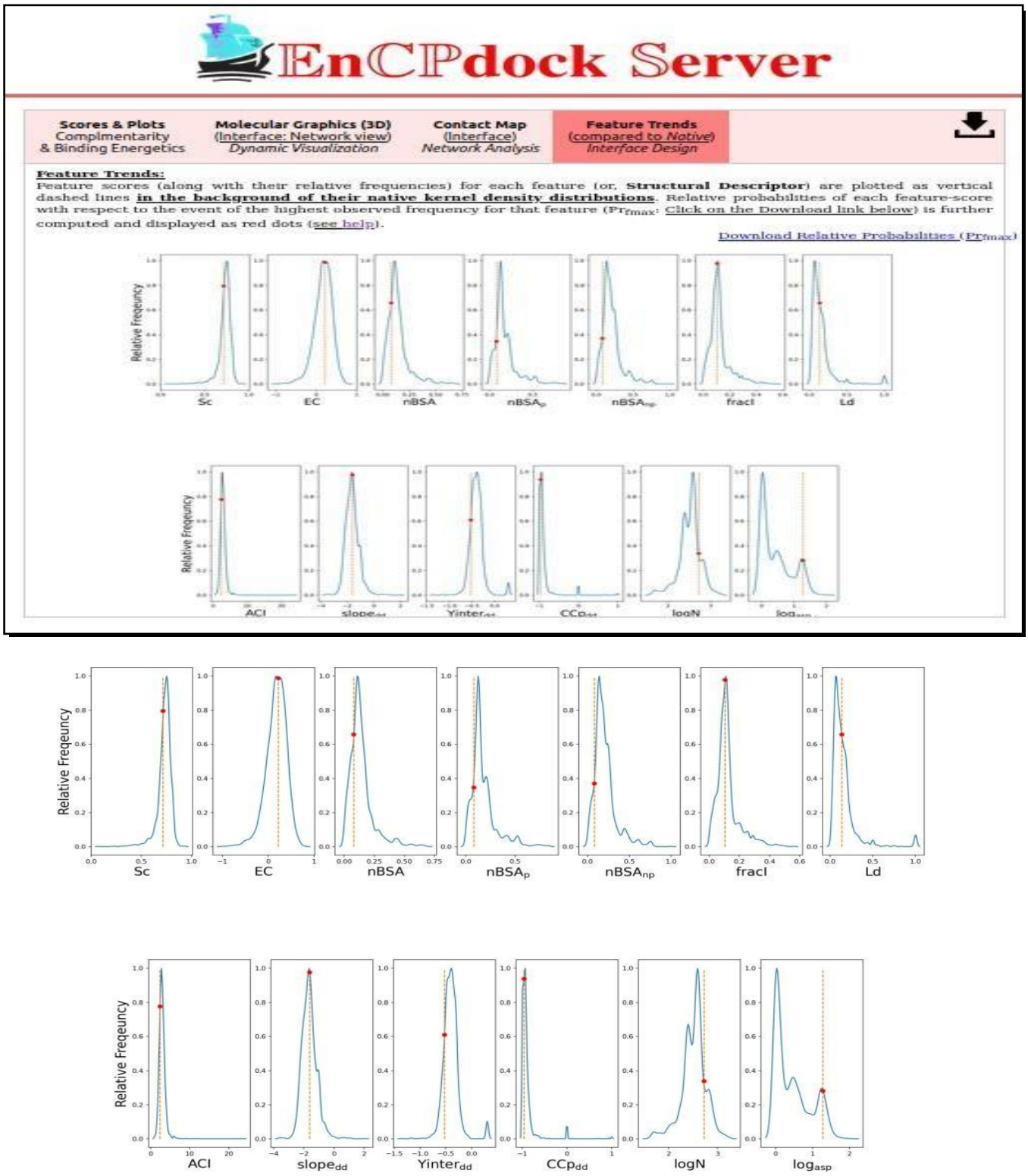


Figure 23: *Features trends of top ranked model (model2) by EnCPdock.*

- *The top-ranked docked poses returned by CLUSPRO were then accumulated and re-ranked using the structure-based binding energetics ($\Delta G_{binding}$) in the EnCPdock web-server. The lowest-energy poses for each ligand (AMP) were then compared across, to lead to one finally proposed AMP as a plausible binder of the Ebola Matrix protein. Then we will do a small molecular dynamics simulation of the peptide protein complex that gave the best results (model 2). Simulations of molecular dynamics will reveal ifit is becoming stablely bonded or separating over time. We plan to extend our pilot project in the near-future using Molecular Dynamic Simulations of the top-ranked docked pose(s).*

Conclusion

In the current era, anti-infective drugs play a critical role in significantly lowering the rates of infectious disease-related mortality worldwide. Anti-Microbial Peptides (AMPs) are versatile innate immune system effectors on mucosal surfaces that have antimicrobial efficacy against a range of pathogenic viruses, bacteria, and fungi. Based on a predetermined set of criteria, the Matrix protein VP40 of the Ebola virus was chosen as a crucial therapeutic target among a wide range of human viral infections. These criteria included the least amount of metabolic similarity between the host and the pathogen, the current need for disease-specific treatments, and other elements. via order to dock these AMPs (ligands) onto the Ebola Matrix protein (receptor), several blind docking experiments were carried out via the CLUSPRO web-server using the library of AMPs created from readily available peptide structures from the Protein Data Bank. The lowest-energy poses for each ligand (AMP) were then compared across, to lead to one finally proposed AMP as a plausible binder of the Ebola Matrix protein. In the near future, we intend to expand our pilot research using Molecular Dynamic Simulations of the top-ranked docked pose(s). Simulations of molecular dynamics will reveal ifit is becoming stablely bonded or separating over time. Future implementation of this plan is being worked on.

References

1. *The next generation of antimicrobial peptides (AMPs) as molecular therapeutic tools for the treatment of diseases with social and economic impacts NB da Cunha, NB Cobacho, JFC Viana, LA Lima… - Drug discovery 2017 – Elsevier*
2. *Antimicrobial peptides: Possible anti-infective agents Author links open overlay panelJayaram Lakshmaiah Narayana a b, Jyh-Yih Chen*
3. *The pathogenesis of Ebola virus disease, Laura Baseler, Daniel S. Chertow, Karl M. Johnson, Heinz Feldmann, and David M. Morens*
4. *The ecology of Ebola virus, Author links open overlay panelAllison Groseth, Heinz Feldmann, James E. Strong*
5. *Understanding Ebola Virus Transmission by Seth Judson, Joseph Prescott and Vincent Munster.*
6. *Ebola Viral Disease Outbreak — West Africa, 2014 Meredith G. Dixon, MD and Ilana J. Schafer, DVM*
7. *Antimicrobial Peptides for Therapeutic Applications: A Review by Min-Duk Seo , Hyung-Sik Won , Ji-Hun Kim ,Tsogbadrakh Mishig-Ochir and Bong-Jin Lee*
8. Hancock, R.E.; Sahl, H.G. Antimicrobial and host-defense Peptides as new anti-infective therapeutic strategies. *Nat. Biotechnol.* 2006, *24*, 1551–1557. [Google Scholar] [CrossRef]
9. Nguyen, L.T.; Haney, E.F.; Vogel, H.J. The expanding scope of antimicrobial peptide structures and their modes of action.

TrendsBiotechnol. 2011, *29*,464–472.[GoogleScholar] [CrossRef]

10. Fjell, C.D.; Hiss, J.A.; Hancock, R.E.; Schneider, G. Designing antimicrobial peptides: Form follows function. *Nat. Rev. Drug Discov.* 2012, *11*, 37–51. [Google Scholar]

11. Drug targeting , Author links open overlay panelVladimir P Torchilin

12. Computational approaches in target identification and drug discovery *Author links open overlay panel Theodora Katsila , Georgios A. Spyroulias , George P. Patrinos , Minos-Timotheos Matsoukas*

13. Drug–Target Kinetics in Drug Discovery , Peter J. Tonge

14. The 2014 Ebola virus disease outbreak in West Africa, Derek Gatherer

15. Structural dissection of Ebola virus and its assembly determinants using cryo-electron tomography, *Tanmay A. M. Bharat, Takeshi Noda, James D. Riches, and John A. G. Briggs.*

16. Structural and Functional Aspects of Ebola Virus Proteins, by Sahil Jain , *Ekaterina Martynova, Albert Rizvanov, Svetlana Khaiboullina and Manoj Baranwal*

17. Mehedi, M.; Falzarano, D.; Seebach, J.; Hu, X.; Carpenter, M.S.; Schnittler, H.J.; Feldmann, H. A new Ebola virus nonstructural glycoprotein expressed through RNA editing. *J. Virol.* 2011, *85*, 5406–5414. [Google Scholar] [CrossRef] [PubMed][Green Version]

18. Ikegami, T.; Calaor, A.B.; Miranda, M.E.; Niikura, M.; Saijo, M.; Kurane, I.; Yoshikawa, Y.; Morikawa, S. Genome structure of Ebola virus subtype Reston: Differences among Ebola subtypes. Brief report. *Arch. Virol.* 2001, *146*, 2021–2027. [Google Scholar] [CrossRef] [PubMed]

19. Dong, S.; Yang, P.; Li, G.; Liu, B.; Wang, W.; Liu, X.; Xia, B.; Yang, C.; Lou, Z.; Guo, Y.; et al. Insight into the Ebola virus nucleocapsid assembly mechanism: Crystal structure of Ebola virus nucleoprotein core domain at 1.8 A resolution. *Protein Cell* 2015, *6*, 351–362. [Google Scholar] [CrossRef][Green Version]

20. Ayub, G.; Waheed, Y. Sequence analysis of the L protein of the Ebola 2014 outbreak: Insight into conserved regions and mutations. *Mol. Med. Rep.* 2016, *13*, 4821–4826. [Google Scholar] [CrossRef] [PubMed][Green Version]

21. Feldmann, H.; Klenk, H.D.; Sanchez, A. Molecular biology and evolution of filoviruses. *Unconv. Agents Unclassif. Viruses* 1993, *7*, 81–100. [Google Scholar] [CrossRef]

22. Baz-Martínez, M.; El Motiam, A.; Ruibal, P.; Condezo, G.N.; de la Cruz-Herrera, C.F.; Lang, V.; Collado, M.; San Martín, C.; Rodríguez, M.S.; Muñoz-Fontela, C.; et al. Regulation of Ebola virus VP40 matrix protein by SUMO. *Sci. Rep.* 2016, *6*, 37258. [Google Scholar] [CrossRef] [PubMed][Green Version]

23. Licata, J.M.; Simpson-Holley, M.; Wright, N.T.; Han, Z.; Paragas, J.; Harty, R.N. Overlapping motifs (PTAP and PPEY) within the Ebola virus VP40 protein function independently as late budding domains: Involvement of host proteins TSG101 and VPS-4. *J. Virol.* 2003, *77*, 1812–1819. [Google Scholar] [CrossRef] [PubMed][Green Version]

24. Harty, R.N.; Brown, M.E.; Wang, G.; Huibregtse, J.; Hayes, F.P. A PPxY motif within the VP40 protein of Ebola virus interacts physically and functionally with a ubiquitin ligase: Implications for filovirus budding. *Proc. Natl. Acad. Sci. USA* 2000, *97*, 13871–13876. [Google Scholar] [CrossRef][Green Version]

25. Chen, J.; He, Z.; Yuan, Y.; Huang, F.; Luo, B.; Zhang, J.; Pan, T.; Zhang, H.; Zhang, J. Host factor SMYD3 is recruited by Ebola virus nucleoprotein to facilitate viral mRNA transcription. *Emerg. Microbes Infect.* 2019, *8*, 1347–1360. [Google Scholar] [CrossRef] [PubMed][Green Version]

26. García-Dorival, I.; Wu, W.; Armstrong, S.D.; Barr, J.N.; Carroll, M.W.; Hewson, R.; Hiscox, J.A. Elucidation of the Cellular Interactome of Ebola Virus Nucleoprotein and Identification of Therapeutic Targets. *J. Proteome Res.* 2016, *15*, 4290–4303. [Google Scholar] [CrossRef] [PubMed]

27. Schümann, M.; Gantke, T.; Mühlberger, E. Ebola virus VP35 antagonizes PKR activity through its C-terminal interferon inhibitory domain. *J. Virol.* **2009**, *83*, 8993–8997. [**Google Scholar**] [**CrossRef**]

28. Leung, D.W.; Ginder, N.D.; Fulton, D.B.; Nix, J.; Basler, C.F.; Honzatko, R.B.; Amarasinghe, G.K. Structure of the Ebola VP35 interferon inhibitory domain. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 411–416. [**Google Scholar**] [**CrossRef**] [**PubMed**][**Green Version**]

29. *https://www.ncbi.nlm.nih.gov/*

30. *https://cluspro.org/help.php*

31. *https://www.rcsb.org/*

32. *https://pymol.org/2/*

33. *EnCPdock* *Reference:* *https://www.biorxiv.org/content/10.1101/2023.02.26.530084v1?rss=1*