SRI VENKATESWARA INTERNSHIP PROGRAM FOR RESEARCH IN ACADEMICS (SRI-VIPRA)

2022

Project Report on
Viroinformatics to study emerging viral diseases (SVP-2203)

1961 - 2021
Tirumala Tirupati Devasthanams
Sri Venkateswara College
(University of Delhi)

Sri Venkateswara College
University of Delhi
Dhaula Kaun
New Delhi, Delhi- 110021

SRI VIPRA PROJECT (2022)

Viroinformatics to study emerging viral diseases

Name of mentor:	**Prof. N. Latha**
Department:	**Biochemistry**
Designation:	**Professor**

Students involved in the given project

S.No.	Name	Course	Photograph
1.	Zeaan Mehernosh Pithawala	B.Sc. (Hons.) Biochemistry	
2.	Aditi Singh	B.Sc. (Hons.) Zoology	
3.	Tapasya Khanna	B.Sc. (Hons.) Biochemistry	

4.	Chirag Pareek	B.Sc. (Hons.) Biochemistry	
5.	Vanshika Rana	B.Sc. Life Sciences	
6.	Kareena Gusain	Bsc(Hons) Biochemistry	-

CERTIFICATE

This is to certify that the students from Sri Venkateswara College have participated in the summer project entitled Viroinformatics to study emerging viral diseases. The participants have carried out the research project work under my guidance and supervision from June 2022 to September 2022. The work carried out is original.

Prof. N. Latha
Department of Biochemistry
Sri Venkateswara College
University of Delhi
Benito Juarez Road
Dhaula Kuan Delhi-110016, India
Email: nlatha@svc.ac.in

ACKNOWLEDGEMENT

We are grateful to the Almighty God for keeping us in good health throughout the work. We are also thankful to our family for being the eternal strength and providing unconditional love.

We would like to express sincere gratitude to our teacher and guide Professor Latha Narayanan for her expert guidance with valuable suggestions and ideas during the entire tenure of our work. Her communication and encouragement increased our inquisitiveness in our work by many folds. We are extremely grateful to her for providing us her valuable time and unconditional help. We are obliged to work under her mentorship to achieve this
endeavour.

We would also like to thank our other guide, Ms. Mansi Pandit for helping us throughout this, for answering all our minor queries and reaching out to us whenever required.

We are immensely grateful to all those involved in the project for helping and rectifying all the basic errors as without them, it would not have been possible to develop the project within the prescribed time.

Lastly, we would thank our college for providing us all the resources, knowledge and a lifetime experience for the project.

INDEX

Contents

ABSTRACT

An orthopox virus of the Poxviridae family that is now causing concern on a global scale is the monkeypox virus (MPV). It has two viral clades known to be unique to Central and Western Africa. The natural reservoirs are probably different African rodents and primates. Direct contact with diseased animals is how zoonotic transmission happens. Close contact with infected people (such as by respiratory droplets, skin-on-skin contact, or sexual contact) or fomites can cause human to human transmission. Human MPV illness often starts with a fever prodrome and lymphadenopathy, then progresses to a diffuse maculopapular to vesiculopustular skin or mucosal lesion eruption. The febrile prodrome may be missing and the skin/mucosal lesions may be restricted to the genital and anal areas in the current 2022 epidemic, which is now primarily affecting males who have sex with men (MSM). Hence, we aim to fathom the phylogeny of this virus better and develop therapeutic drugs in-silico that would efficaciosly inhibit the growth or kill the pathogen. To achieve this, we first did an extensive literature review, followed by performing a phylogenetic analysis of the orthopox viruses. For in-silico drug design, we had to first develop a 3D structure of the drug target, P37, followed by its docking with various therapeutic agents.

Keywords: Monkeypox, Phylogenetics, Drug Design

INTRODUCTION

At the conclusion of a ten-year rigorous surveillance and vaccination effort, the 33rd Planet Health Assembly formally declared smallpox to have been eradicated from the world in 1980. Few people under 30 years old are currently sufficiently protected against smallpox because the worldwide immunisation programme was stopped in the late 1970s. It is considered that older people who received vaccinations decades ago have little or no residual immunity. Additionally, those with compromised immune systems, such as those who have the human immunodeficiency virus (HIV), are particularly susceptible to contracting orthopoxvirus infections, including those caused by the vaccinia virus, the vaccine used to prevent smallpox (VAC).

Monkeypox (MPX) cannot be eliminated, unlike smallpox, which had an animal reservoir. There is evidence to support the claim that widespread smallpox vaccination in the 1970s was able to lower the incidence of sickness following known MPX exposure by as much as 90%. The reintroduction of vaccination poses significant challenges due to the high incidence of HIV infection in central Africa today and the associated elevated risk of consequences from VAC injection.

Antiviral medication would undoubtedly be advantageous in the absence of immunisation for the treatment of MPX patients as well as for containing the infection.

A taxonomically diverse range of mammalian species can contract monkeypox, however the virus has only ever been isolated from a wild animal once, a Funisciurus squirrel in the Democratic Republic of the Congo. Human illnesses have been related to animal contact, although it might be challenging to identify the exact animal contact that led to a case in regions where bushmeat from a variety of species is often hunted or prepared, as well as where mouse infestations in homes are common. It is thought that saliva, respiratory excretions, or contact with lesion exudate or crust material are the routes of transmission. Another exposure source may be the shedding of viruses through faeces. The obstacles that surveillance systems face include inadequate infrastructure, a lack of funding, inadequate diagnostics, a lack of specimen collection, and clinical difficulties in identifying monkeypox sickness.

Civil war and population dislocation raise worries about the virus spreading to a region lacking monkeypox or about people moving to more heavily forested areas where they are more likely to come into contact with wildlife and contract a variety of zoonoses. To better comprehend the variety of elements involved in disease transmission and distribution, the recorded increase in the frequency of human disease requires deeper investigation and analysis. Advances in our knowledge of this essential zoonosis will assist better direct prevention tactics and lessen human disease. There are still numerous mysteries concerning human disease, animal reservoirs, and the virus itself.

1. Literature Review

Introduction

The Monkeypox belongs to the Orthopoxvirus genus in the Poxviridae family.
The linear genome's central section contains the genes necessary for viral replication, while the terminal regions contain the genes that code for host response modifiers and other host-interacting proteins. Both in terms of gene content and sequencing, the centre part of the genome is relatively conserved, whereas the terminal sections are more varied. Monkeypox virus, together with other poxviruses, is considered one of the largest and most complex viruses
They are brick-like shaped particles with a size ranging from 220 nm to 450 nm in length and 140 nm to 260 nm in width.
The orthopox virion consists of four major elements-

1. Core
2. Lateral bodies
3. Outer membrane
4. Outer lipoprotein envelope.

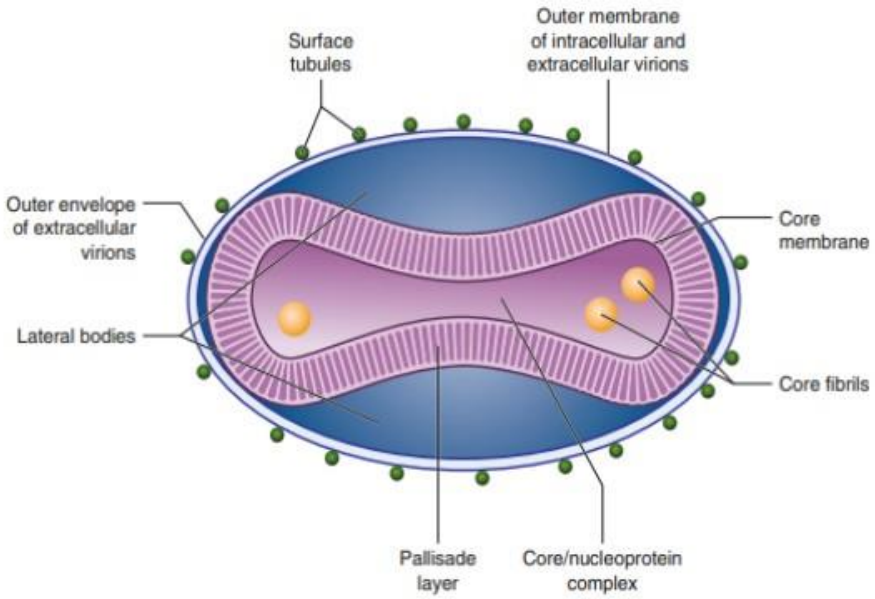


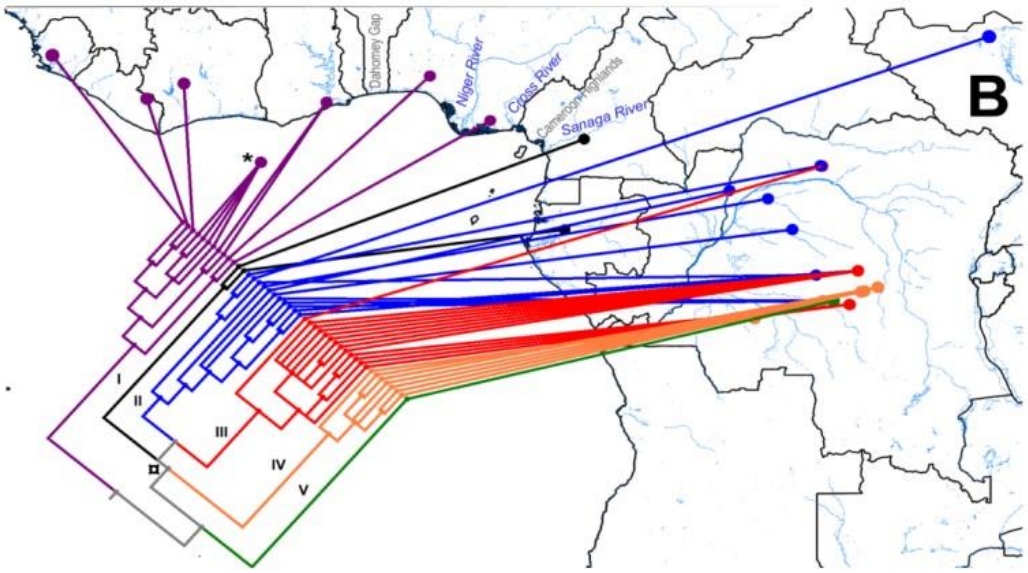
Figure- Poxvirus Particle

Phylogeographic analysis of the previous Human Monkeypox outbreak

Phylogeographic Investigation of previous Human Monkeypox outbreak shows that there are two distinct genetic clades of monkeypox, the Central African (or Congo Basin) clade and the West African clade. Most cases were infected with the Central

African clade, which was found in the CAR, DRC, and South Sudan. The outbreak in the US (2003) and the outbreak in Nigeria (which started in 2017) cover the largest part of the West African clade cases.Central African clade was also found in Sierra Leone and the travel-related cases in Israel and Singapore. Preliminary sequencing data of two UK cases were also determined to be consistent with the West African clade. Between Niger River and Cross River is the eastern most isolate of West African clade while the western most isolate of the Central African clade is south of the Sanaga River. Cross and Sanaga rivers are said to be the biogeographic barriers for mammal species. Rivers and other geographical barriers leads to higher genetic divergence because of less mobility of animals across these barriers. Similar to speciation event.

The current human monkeypox virus belongs to the Western Africa



Phylogeographic Investigation of previous Human monkeypox outbreak

2. Phylogenetic analysis

Introduction

Phylogenetics is the study of evolutionary relationship between different organisms. To understand and predict pandemics, we need to reconstruct the evolutionary history of the organisms that cause them using large amounts of genetic data and complex computational approaches. Its use can be extended for drug repurposing and design (if there is a high similarity between two organisms).

Softwares that are used for carrying out Phylogenetic analysis

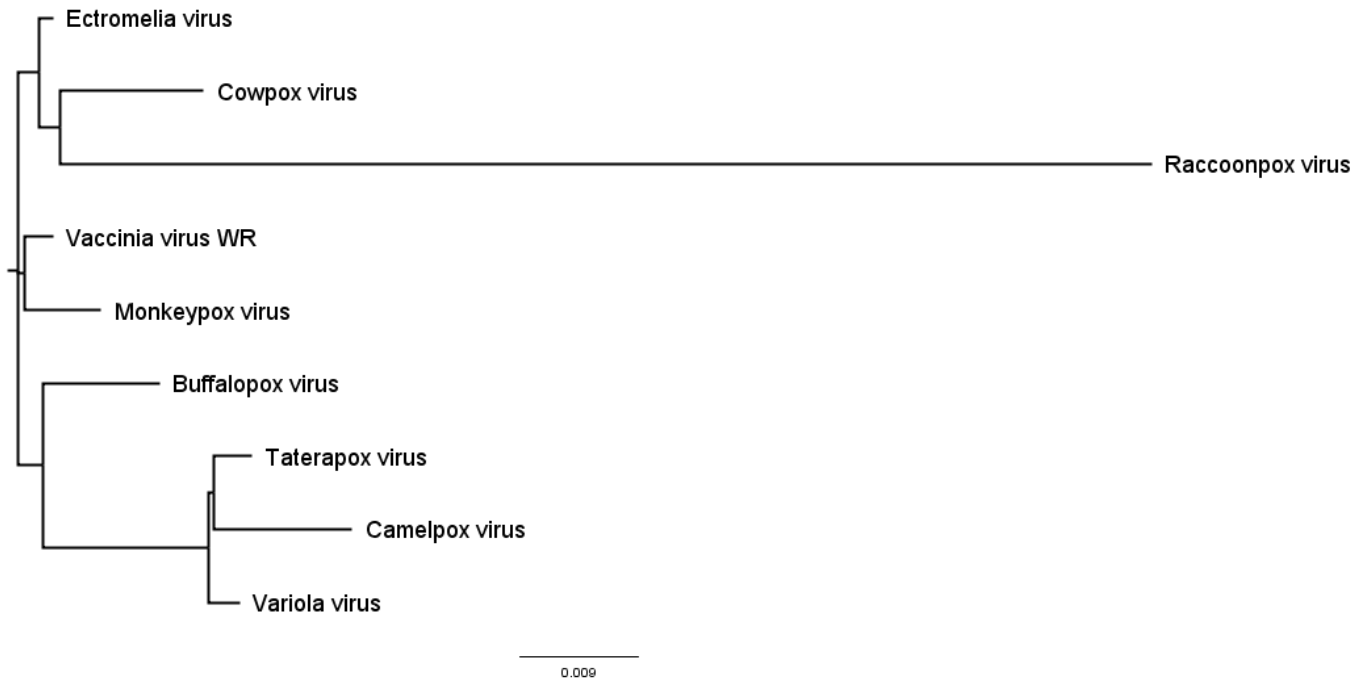
Name	Purpose
BLAST	Sequence Alignment
CLUSTALW	Sequence Alignment
MAFFT	Sequence Alignment
IQTREE	Tree Generation
EvolView	Visualization Software
SeaView	Visualization Software
MEGA	Sequence Alignment
CACTUS	Sequence Alignment
iTOL	Visualization Software
FigTree	Visualization Software

BEAST	Sequence Alignment

Results of Phylogenetic analysis

The amino acid sequences of P37 Protein in different Orthopox viruses (Ectromelia virus, Vaccinia virus, Monkeypox virus, Cowpox virus, Raccoonpox virus, Buffalopox virus, Taterapox virus, Camelpox virus, Variola virus) were downloaded and were used as an input for MAFFT and CLUSTALW for carrying out Multiple Sequence Alignment (MSA).
MAFFT uses the approach of identifying the homologous regions by fast Fourier transform and has a simplified scoring system that substantially alleviates CPU time and improves the accuracy. CLUSTALW takes an approach of progressive alignment wherein it aligns most similar sequences first. Both of these softwares are highly accurate and have been widely used to carry out MSA.

After MSA was done, the aligned sequences were downloaded and their evolutionary tree was generated with the help of FigTree software. Both, MAFFT and CLUSTALW gave the same results.



Phylogenetic Tree of P37 protein present in various Orthopox viruses

From the phylogenetic tree, we can infer that P37 protein of Monkeypox virus is closely related to Vaccinia virus. This indicates that the mechanism of P37's action

would also be similar to the Monkeypox virus, thereby opening doors of novel therapeutic drug candidates that can be repurposed to cure monkeypox virus infections.

3. Current Therapeutics of Monkeypox virus

Introduction

There are no treatments specifically for monkeypox virus infections. However, monkeypox and smallpox viruses are genetically similar, which means that antiviral drugs and vaccines developed to protect against smallpox may be used to prevent and treat monkeypox virus infections.

Vaccines

Two vaccines licensed by the U.S. Food and Drug Administration (FDA) are available for preventing monkeypox infection – JYNNEOS (also known as Imvamune or Imvanex) and ACAM2000.

JYNNEOS
It contains a live virus that does not replicate efficiently in human cells. It is administered as two subcutaneous injections four weeks apart. The immune response takes 2 weeks after the second dose for maximal development. It is licensed for use in the prevention of smallpox or monkeypox in people ages over 18.

ACAM2000
It is a live Vaccinia virus vaccine that is replication competent. It is administered as one percutaneous dose via multiple puncture techniques with a bifurcated needle. The immune system response takes 4 weeks for maximal development. Following the inoculation, a lesion will develop at the site of vaccination. The lesion may take up to 6 weeks or more to heal.

Antiviral Drugs

Tecovirimat

Tecovirimat is an orally available small molecule with activity against orthopoxviruses. Tecovirimat is an inhibitor of viral p37 and blocks the ability of virus particles to be released from infected cells.

Tecovirimat is an Orthopoxvirus VP37 Envelope Wrapping Protein Inhibitor. The mechanism of action of tecovirimat is TPOXX prevents virus spread by inhibiting the function of the major envelope protein (F13L), thereby preventing the virus from leaving an infected cell.

Although TPOXX has shown efficacy against multiple OPXVs in various animal model systems, it has also been noted that a single amino acid alteration to the orthopoxviral F13L protein allows for resistance to TPOXX treatment

Aurintricarboxylic acid (ATA)

ATA blocks the phosphorylation of extracellular signal-regulated kinase 1-2, which is essential for OPV replication. It also inhibits the phosphatase activity of viral enzymes required for viral transcription. Nonetheless, there is only in-vitro data available.

Cidofovir (CDV)

It is an ANP analog of deoxycytidine and is highly efficacious against all Orthopox Viruses. It has been used topically to treat human molluscum infections and intravenously in humans and several animal models to treat other orthopoxvirus infections. Acyclic Nucleoside phosphonates (ANPs) are converted to 5′-nucleotides by various kinases. Generally, the nucleoside analogs require triphosphorylation before they can function in an antiviral capacity. Although, intravenous administration requires attentive clinical care and may lead to Nephrotoxicity and incorporation into host DNA.

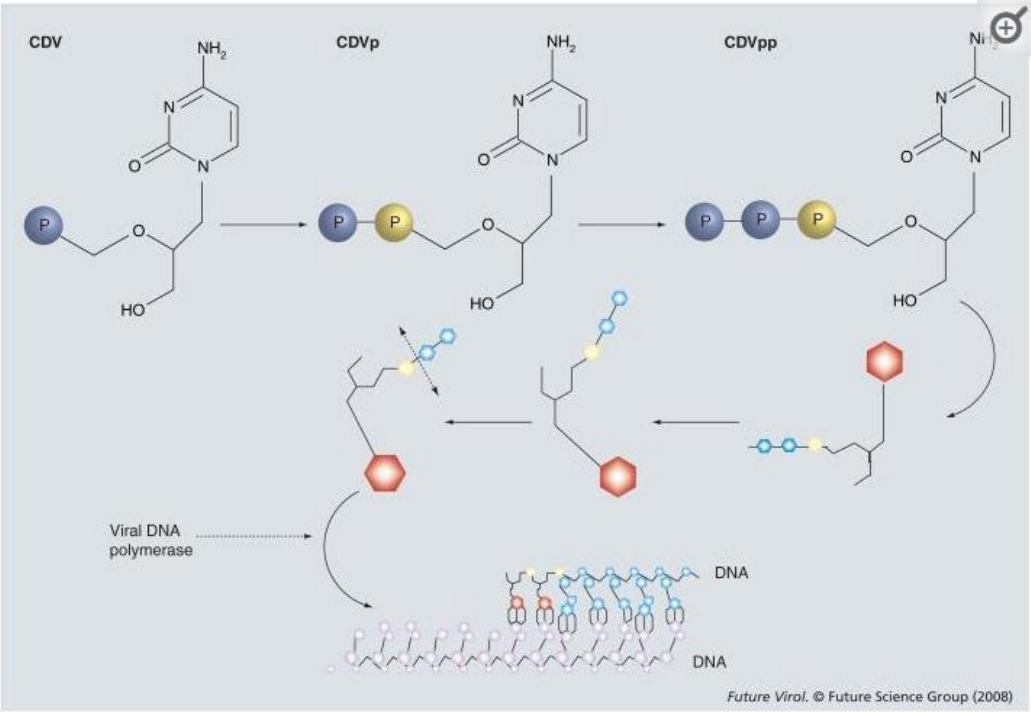
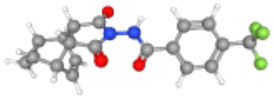
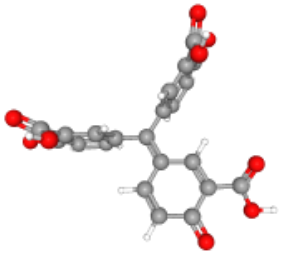
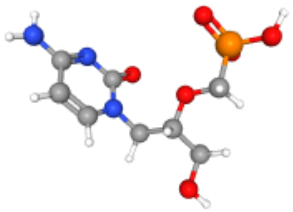


Figure- Mode of action of Cidofovir

Drug	Structure
Tecovirimat	
Aurintricarboxylic acid (ATA)	

Cidofovir (CDV)	

4. Drug Target

P37 Protein

P37 interacts with the Rab9 GTPase and TIP47, which are components of late endosome-derived transport vesicles. leads to the formation of the virus-specific wrapping complex for enveloped virions.

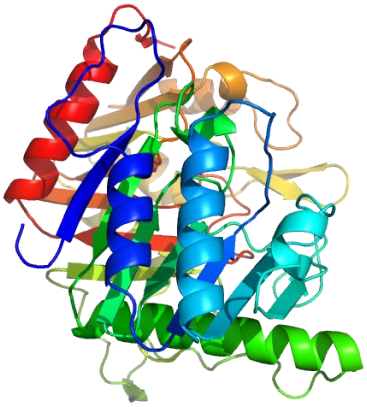
Structure prediction of P37 Protein

The structure of P37 was not available and therefore we used in-silico tools for the prediction of P37 protein's 3 dimensional conformation. First, we retrieved the FASTA sequence of the protein from the NCBI database or uniprot by searching for the appropriate protein with correct accession number. We then proceeded to use the various online softwares available to predict the protein tertiary structure. SWISS-Modeller is one such common online software which we had used to predict the protein structure. It uses the principle of homology modelling where the query sequence is matched with a template sequence already present in the database and a

model is built around that. Another algorithm which we used was AlphaFold, an artificial intelligence system that accurately predicts protein structures by novel neural network architecture taking into account the evolutionary, physical and geometric constraints of protein structures.

Results obtained from SWISS-Modeller

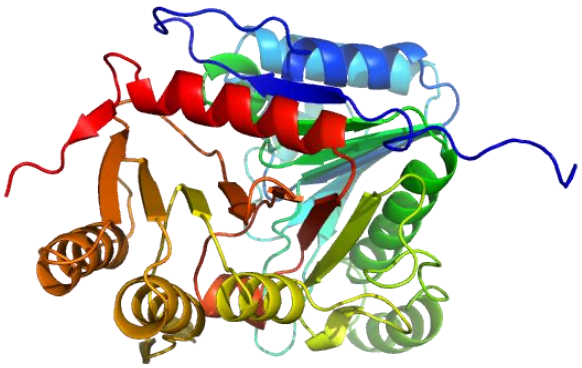
SWISS-Modeller is easy to use software and the process was not that computer intensive. We gave our FASTA Sequence as an input and then searched for a template. The template with the highest global model quality estimator (GMQE) was selected and a model was built.



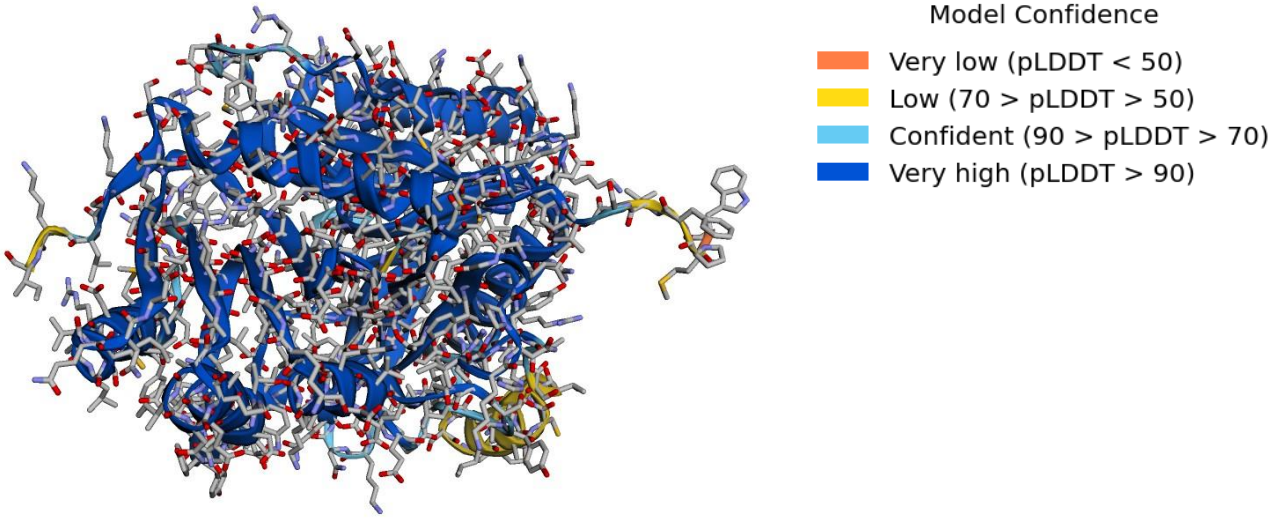
Structure of P37 as predicted by SWISS Modeller

Results obtained from AlphaFold

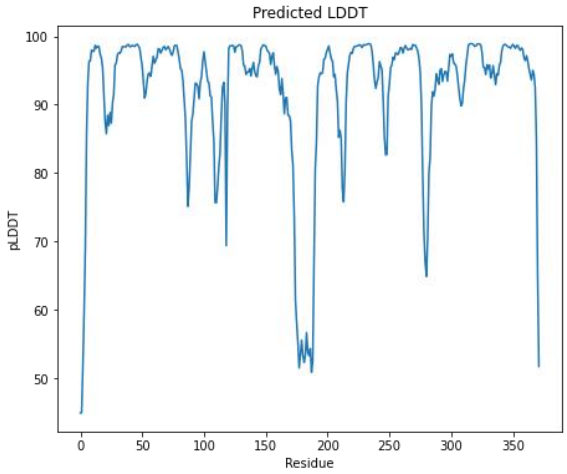
AlphaFold's prediction was computationally intensive and time consuming. We gave our FASTA Sequence as an input and then the program constructed the structure of protein.

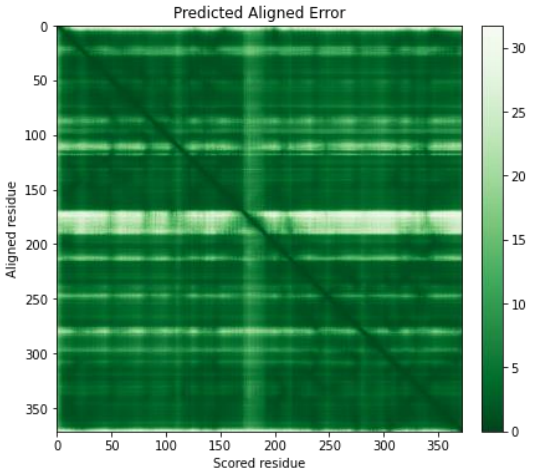


Structure of P37 as predicted by AlphaFold



P37 structure predicted by AlphaFold



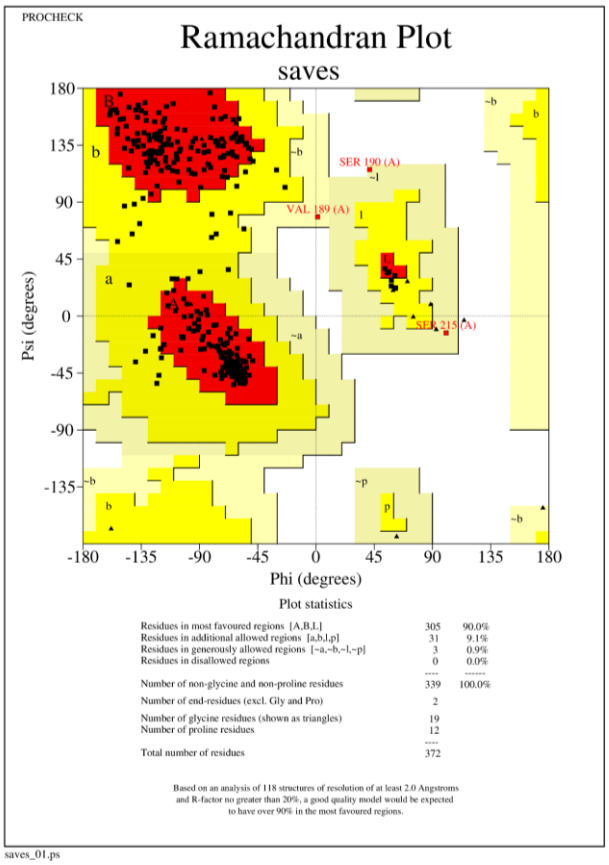


Validation of the P37's predicted structure

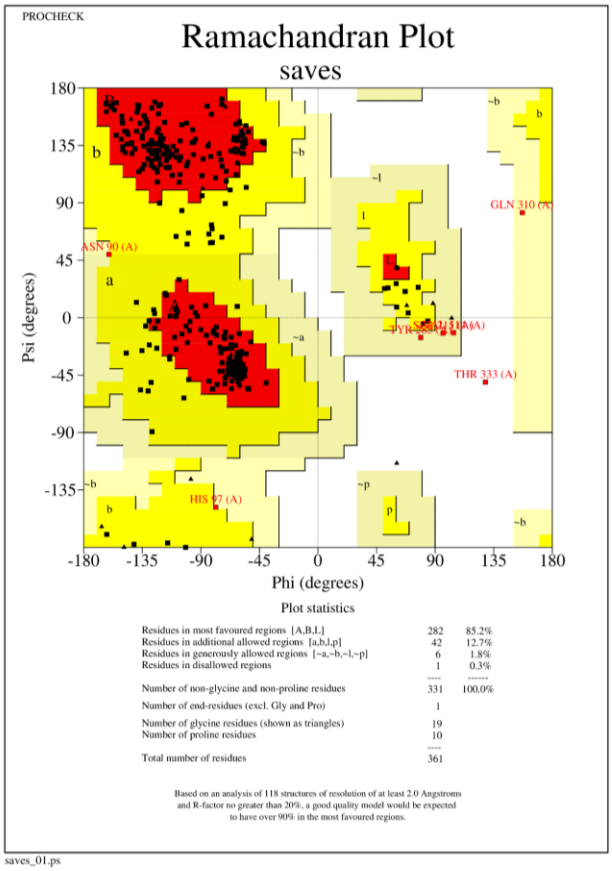
The models which were built was further validated using online validation software SAVES UCLA where we used the PROCHECK program. We used PROCHECK to get the Ramachandran map. For protein validation, the bond length, angles planarity and chirality are checked to see if it is in the allowed regions. All amino acids have an allowed and disallowed region in a plot for phi and psi angles called the Ramachandran plot. The software checks if the amino acids present in the model lies in the allowed and disallowed regions and generates a Ramachandran plot. It also gives information about the H-bond quality and about the resolution of the model that is obtained. More is the resolution, the better is the accuracy of the model in question. Thus, we compared the structures predicted by SWISS-Modeller and came to the conclusion that the structure predicted by AlphaFold was relatively more accurate than that of SWISS-Modeller.

Ramachandran Plot for the structure *Ramachandran Plot for the Structure*

predicted by AlphaFold



predicted by SWISS-Modeller



5. Antivirals for treating Monkeypox virus infections

Introduction

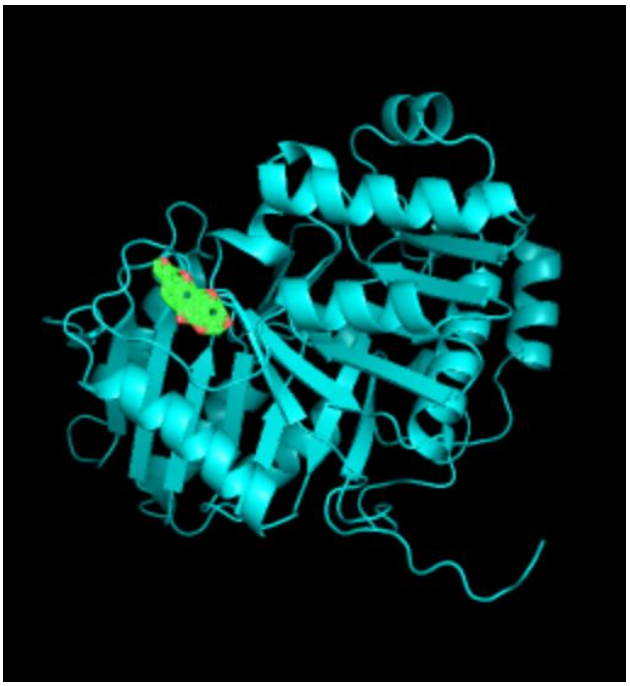
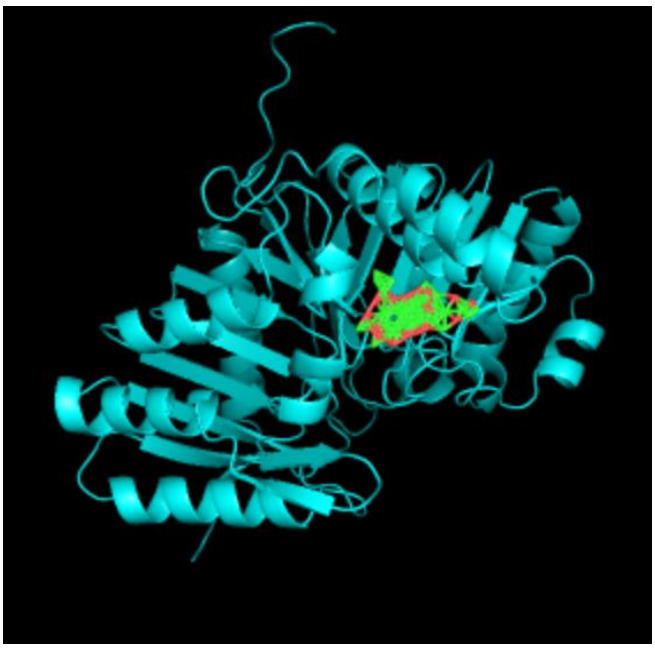
Treatment with antiviral medications for those who have already contracted a virulent orthopoxvirus would offer immediate benefit, in contrast to vaccinations, for which the protective impact is delayed. The effectiveness of antiviral medications in controlling an epidemic has not been examined since they were not accessible during the smallpox eradication campaign.

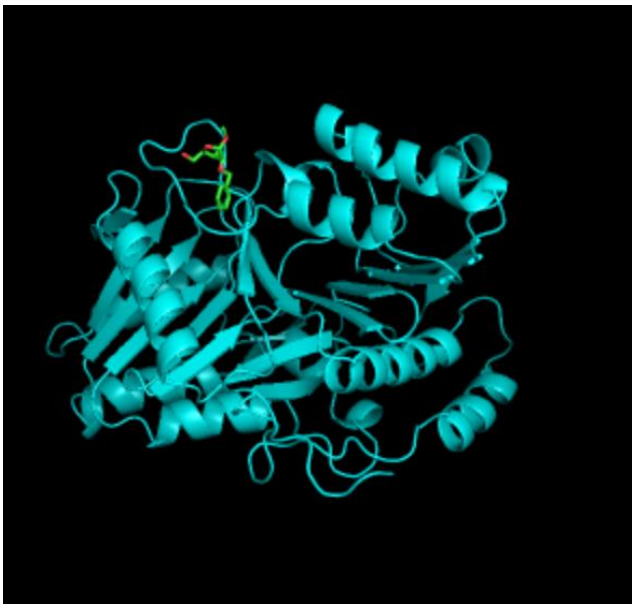
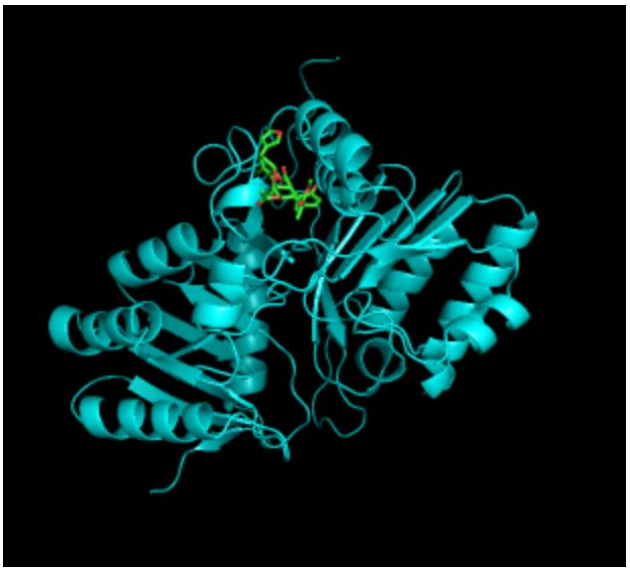
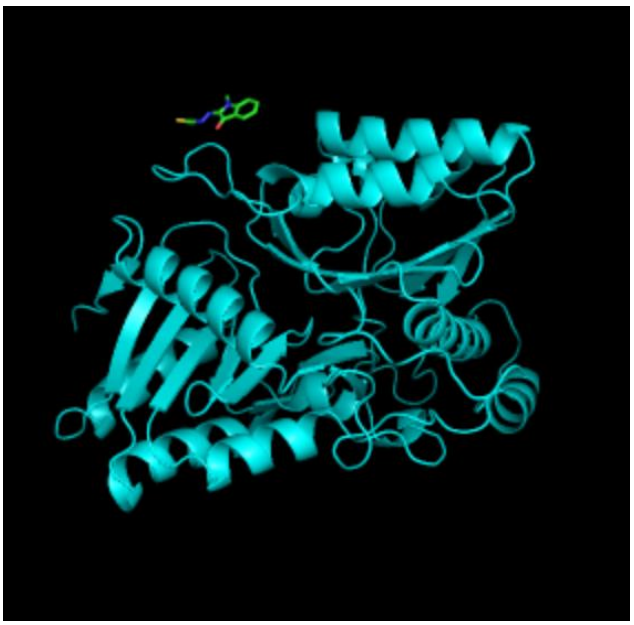
When the assault agent is discovered, early, widespread administration of stored medications might offer life-saving treatment and should be included in the overall medical response. Treatment with an antiviral medicine may attenuate illness, reduce infectivity, and lower mortality depending on the recipient's stage of infection. Antiviral drugs, as opposed to immunizations, may be administered to everyone who was exposed, including those who had underlying immunodeficiency problems.

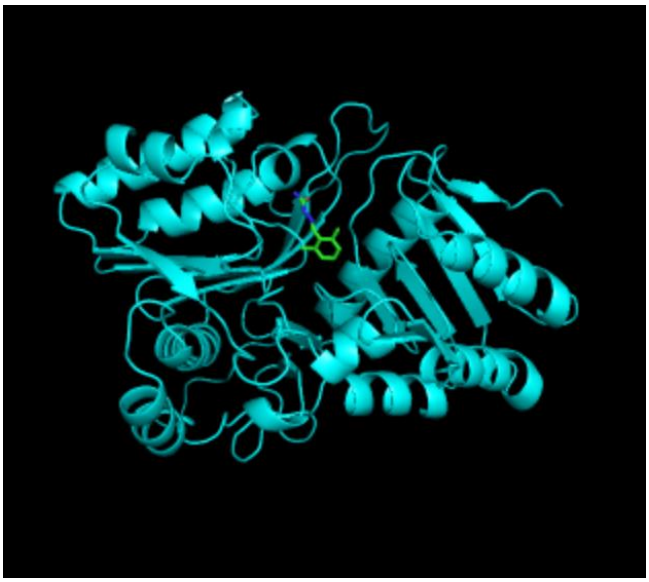
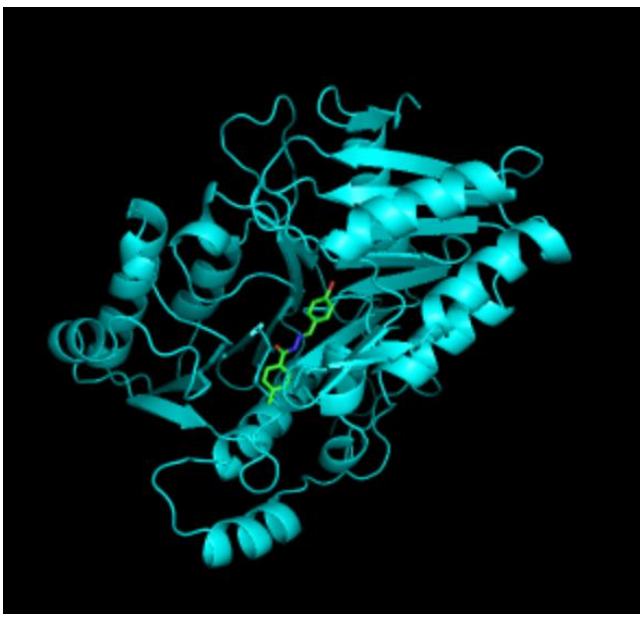
Smallpox was successfully eradicated, and there is limited need for medications that particularly target MPX due to the virus's rarity, remote location, and few reported cases. Antiviral screening tests conducted by the pharmaceutical industry as part of their drug discovery activities typically do not even contain VAC. Recent occurrences have pushed the issue of bioterrorism to the fore of people's minds everywhere. The development, licencing, and stockpiling of adequate supplies of antiviral medications capable of combating aggressive orthopoxviruses has now taken precedence. Live VAC used in smallpox vaccinations can result in disseminated and progressive vaccinia infections, among other significant and perhaps fatal side effects. Another potential area of focus for antiviral therapy is problems from vaccinations. Currently, administering vaccinia immune globulin is the sole method of treatment for these infections (VIG). It has never been proven that VIG treatment is effective for these illnesses. Recent approval of cidofovir for the treatment of adverse responses to the smallpox vaccine under an experimental new drug programme. Because of biosafety concerns, there hasn't been any published research comparing the relative sensitivity of different orthopoxviruses to antiviral medications in clinical usage.

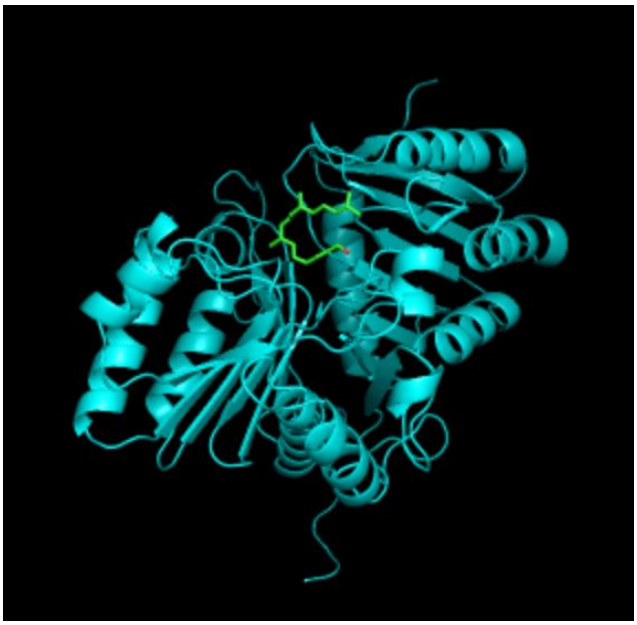
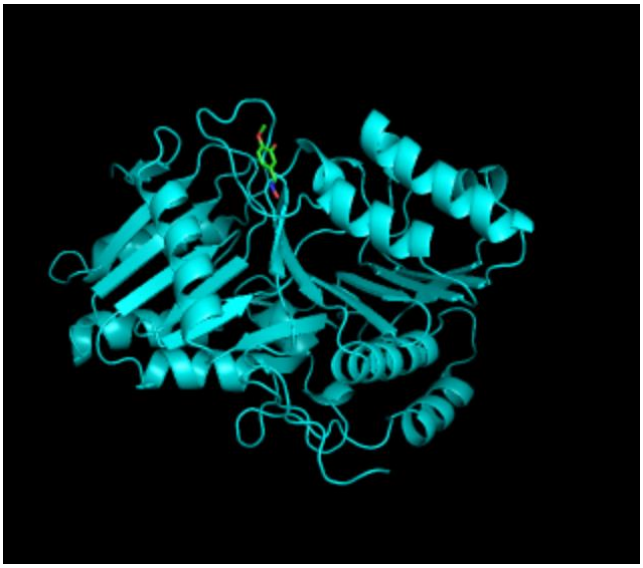
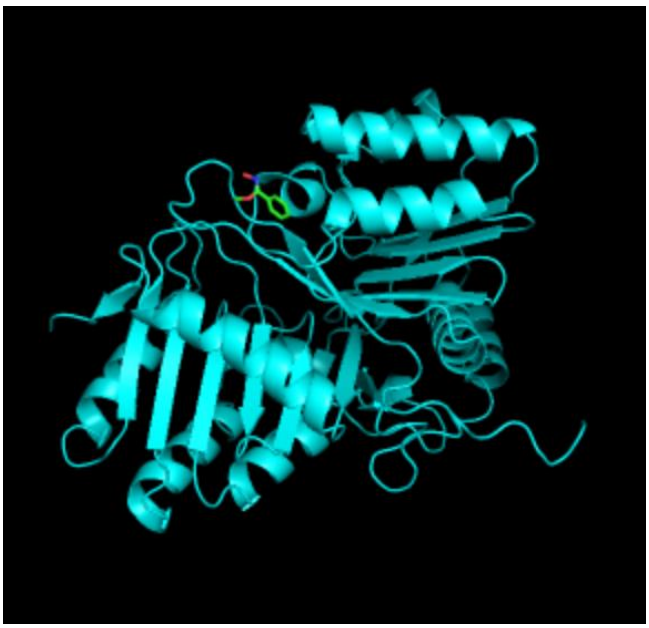
Results of Docking

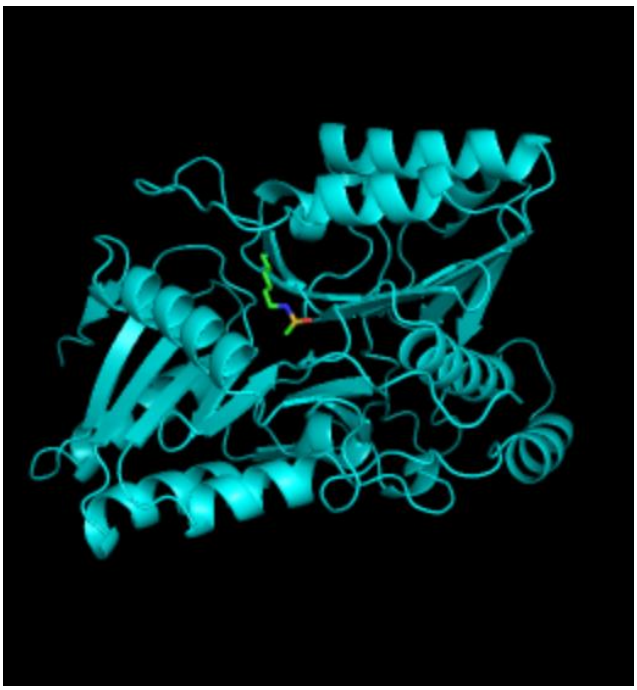
Ligand	Global Energy (KJ)	Image

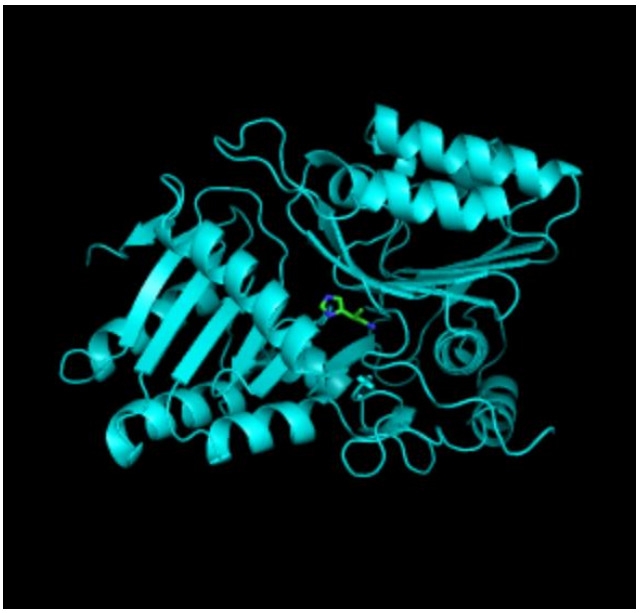
Nimbidine	-63.9	
Azadirachtin	-54.4	

Methyl-2-O-benzyl-d-arabinofuranoside	-37.11	
nimbin	-35.06	
N-methylisatin 3-thiosemicarbazone	-32.73	

2,6-Dimethylbenzaldehyde carbamoylhydrazone	-32.35	
butan-2-one-4-(3-hydroxy-2-methoxyphenyl)	-31.24	
N'-[(E)-(4-Bromophenyl)methylidene]-4-methylbenzohydrazide	-30.08	

Phytol	-28.5	
benzaldehyde-3-hydroxy-4-methoxy	-26.03	
methyl N-hydroxybenzenecarboximidate	-24.23	

Terpinen-4-ol	-23.9	
n-Hexylmethanesulfonamide	-23.32	

Imidazole, 4-(2-amino-1-methylethyl)	-22.08	

Ligand	Global Energy (KJ)	Image
Spongothymidine (ara-T)	-33.24	

Spongouridine Ara-A (9-β-D-arabinosyladenine 3′-0-Acetylara-A	-33.24	

Vidarabine	-42.79	

Sansalvamide fungus, Fusarium sp. Molluscum contagiosum	-45.48	
Didemnaketals A and B (cyclodidemniserinol trisulfate)	-58.23	

Rifampicin (actinomycete Streptomyces mediterranei)	-77.81	

Streptovaricin B Streptomyces spectabilis	-87.31	

Distamycins-	-50.21	
Amphotericin B	-43.51	

Aphidicolin	**45.09**	

Bleomycin	-40.53	

Camptothecin	-44	

Netropsin Streptomyces netropsi	-41.96	

Sinefungin Streptomyces sp.	-43.7	
Cordycepin	-40.12	

Fucoidan	-26.04	
Acyclovir	-36.75	
AcDa 1	-36.3	

Macrolactin A	-37.88	
Azidothymidine	-38.88	
Avarol	-40.05	

Weinbersterol A	-41.12	
Halovir E	-44.51	
Phomasetin	-46.51	

Halovir A	-49.95	
Stachyflin	-50.39	
Clathsterol	-56.52	

Halovir B	-58.02	
Halovir C	-62.4	
Weinbersterol B	-63.03	