IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 8 Issue 1, January 2021
ISSN (Online) 2348 – 7968 | Impact Factor (2020) – 6.72
www.ijiset.com

Bioinformatics Analysis of *B. abortus strain 2308* protein and its drug docking.

Rajan Keshri, Student at lovely professional university

Harpreet Kaur working as a Director at wel 'n' wil life science health and research centre

Dhanashri R Patil Bachelor of Pharmacy from Sinhgad College of Pharmacy, Savitribai Phule Pune University, Pune 411041.

Nilesh N. Sonawane Bachelor of Pharmacy from Sinhgad College of Pharmacy, Savitribai Phule Pune University, Pune 411041

Abstract

Brucellosis is among the fast-spreading disease on Earth causing casualties in livestock as well as in humans. It's an alarming situation to kick off the study of Brucellosis causing agent. Brucellosis is an infectious disease caused by Brucella abortus. Here, we have carried out the Insilco analysis of Brucella abortus strain 2308. This strain is mainly responsible for the disease. Here, we have tried to study the B. abortus strain 2308 by the means of bioinformatics methodology. We have run several Insilco tools to predict its structure and function. Moreover, we have carried out the methodology of protein homology modelling on this strain. Furthermore, we have also carried out several protein chain analyses, protein-protein interface, structure alignment, structure prediction and active site prediction.To make our study more productive we have also performed drug docking. These results will add on a little info in the emerging bioinformatics data regarding Brucella abortus.

Keyword: - Brucellosis, Brucella abortus, bioinformatics, protein homology modelling, structure prediction, infectious diseases, drug docking.

INTRODUCTION

Miscarriage is the leading cause of brucellosis in cattle, leading to miscarriage and infertility in adult animals. Brucella S.P.P. Gram-negative, phospholytic intracellular bacteria. The genus Brucella is mostly involved in mammalian pathogens. Brucellosis is a major infectious disease in humans and animals. Brucella species cause zoonotic disease brucellosis. Brussels sprouts can cause miscarriage in cattle, sheep, goats, pigs and dogs. Furthermore, it has been found that there is evidence in humans that brucellosis can cause spontaneous abortion (Khan, M. Y., Mahi et al. 2001). In cattle, miscarriage occurs at 6-8 months of gestation and in cattle the hygroma of the bull or the organs or reproductive tract of the bull is infected. The disease is transmitted from animal to animal in a variety of ways, including exposure to infected embryos (Samartino, L. E., & Enrite et 1993). Human brucellosis is an occupational disease of many veterinarians, butchers and workers. Infection usually occurs in the case of conception or maintenance of the fetal membrane, contact with feces, carcasses and vaginal discharge of infected animals. Bacteria can enter through conjunctivitis and skin contact (Hajibemani, A., and Shekhali Islami et al. 2020). The disease is usually asymptomatic in young animals and non-pregnant women. Even in the absence of abortion, mass excretion of the organism into the placenta, fetal fluid, and vaginal discharge occurs. The mammary gland and associated lymph nodes are also infected, and the organism can excrete in milk

(Priyanka, S.B., and S.K., K. et.al 2019). It infects 500,000 people worldwide each year. After entering humans or animals, the bacteria attack the bloodstream and lymph, where they multiply within phagocytic cells and eventually cause septicemia. Symptoms of sudden fever, miscarriage, asthenia, endocarditis and encephalitis. Brucella does not contain well-known bacterial viral agents such as cytolysin, capsules, exotoxins, secretory proteins, fibria, phase-encoded toxins and viral plasmids. Broccoli infects phagocytic macrophages and non-phagocytic epithelial cells (e.g., hela cells) in vivo and in vitro. The vacuole of brucellosis paras macroscopes depends on their ability to survive and replicate in phagocytic compartments. Specifically, the brucellosis life cycle consists of two stages: (i) chronic infection of phagocytic macrophages leading to brucellosis survival and replication, and (ii) acute infection of non-phagocytic epithelial cells, leading to reproductive tract pathology and leading to miscarriage. The spleen and liver are organs that contain many bacterial cells after the brucellosis invasion. After a large percentage of Brucella cells are killed in Vivo, the remaining Brucella cells survive and remain in Vivo longer (he, y 2012). Intracellular replicas in Brucella phagocytes. Systemic infection occurs after invasion of the lymphatic vessels, colonization of the uterus, male genitals and mammary gland. Miscarriage in the last trimester of pregnancy causes strong irritation in the uterus, which is thought to be due to high concentrations of erythritol and steroid hormones. In 1954, congressional funding was first approved to eradicate the disease from the country for the Cooperative State - Federal Brucellosis Aidification Program. As with other animal disease eradication efforts, the success of the program depends on the support and participation of livestock producers. The basic procedure is to vaccinate the calves, test the cattle and test the domestic buffaloes for infection and slaughter the infected animals. Herd deposits, if funds are available, can be used if herds are severely affected. Identification of animals in the market, monitoring for detection of infected animals, testing of affected herds and vaccination of calves re-established in brucellosis affected areas are important features of the current program. The uniform methods and rules of the program set the minimum standards for states to achieve elimination. Under the Active Surveillance Program (Ragan, VE (2002), states are designated brucellosis-free when their cattle or bison are not found to be infected for 12 consecutive months. Milk from dairy herds is tested two to four times by testing a small sample obtained. Milk and domestic bison herds do not produce milk for sale. Be sure to test for brucellosis. With some exceptions, herd tests must include cattle and buffaloes older than 6 months of age to test for the disease. Sell to Rani. d test. If brucellosis is found in a positive animal surveillance test, the rest of the herd will be tested for livestock and bison (Ragan, VE (2002). Fortunately, there has been a combination of progresses in milk pasteurization and extermination. Cases, some of the 6,400 cases were reported in 1947. Recently, approximately 100 cases of brucellosis in humans have been reported annually. Enters for Disease Control and Prevention (Carbel, MJ (2006). Products of brucellosis-induced ab in general. The fetus, placenta and cervix secrete a large number of organisms. Transmission is usually caused by bursal-induced miscarriage (fetal, placental, uterine discharge) or ingestion of contaminated material by these products. Brucella may be present in cervical discharge starting 2 weeks earlier. The incubation period is very variable, ranging from 2 weeks to 1 year or more. The minimum incubation period from infection to miscarriage is approximately 30 days (YAEGER, M.J. and HOLLER, L.D., 2007). Bowin Bacterial causes of infertility and miscarriage. Current Medicine in Large Veterinary Science (pp. 389–399). WB Saunders .. There is no effective way to identify infected animals by their existence. The most obvious

sign in pregnant animals is miscarriage or the birth of an abandoned calf. Changes in the normal lactation period due to miscarriage and delayed conception reduce milk production. Not all infected cows abort, but those who become pregnant usually get pregnant in the fifth and seventh months. Infected cows usually miscarry once, but one percent of miscarriages occur during additional pregnancies, and later calves may become weak and unhealthy. Infected cow calves may have a latent infection, i.e. infections that are not detected until they become pregnant, aborted or give birth. Although their calves appear to be healthy, infected cows harass and release infected organisms and should be considered a dangerous source of disease. Other signs of brucellosis include significantly reduced fertility with a poor pregnancy rate, resulting in cervical infection, and (sometimes) enlarged, arthritic joints. The incubation period is the period between exposure of the organism to infection and the first appearance of disease symptoms. The incubation period of brucellosis in cattle, bison and other animals is approximately 2 weeks to 1 year and in some cases longer. When miscarriage is the first sign, the minimum incubation period is usually 30 days. Some animals become aborted before developing a positive response to the diagnostic test. Other infected animals can never be aborted. In general, infected animals that do not develop a positive response to the clinical trial within 30 to 60 days of gestation may not develop a positive response for several months to more than a year (Rachel, MP, Wahl, LC and Hill, FI, 2018).

Brucella abortus strain 2308 was actually reported in 1940 as the most severe strain of a cow's fetus recovered from embryo, as demonstrated by B. Chan et al. (2005), the other B. indicator was found in chromosome synthesis with the abortus gene, and there are some differences. In 2009, WGS analysis results from a species known as B. abortion 2308A were traditionally obtained under entry number GCA_000182625.1. No additional information is given about the isolates used in these two WGS projects (Suarez-Esquivel, M., Ruiz-Villalobos et.al. 2016).

The genus Brucella is mostly mammalian pathogens that, due to their low infectious properties, classify aerosol transmission and treatment difficulties as potential bioterrorism agents. Brucellosis is a major infectious disease in humans and animals. Many species of Brucella (B. bortus, B.melitensis, and B. usis) are isolated from a wide variety of animals. All three brucellosis species cause serious human disease, which occurs in the acute stage and causes damage to various organs. Brucellosis is a major problem in the Mediterranean region and parts of Asia, Africa and Latin America. When the infection is localized to the brain or heart, it can lead to malignant meningitis or malignant endocarditis, respectively. Brucella abortus (genus 2308) is very rare in humans, cattle and some other domestic animals. It has 2 chromosomes in its gene. Despite the high protection of the genetic spine shared by the three brucella species, there are several important species-specific differences. A total of 207 pseudogenes were identified in B. bortus (Strain 2308). Sixteen pseudogenes are common in all three species, 9 of which originated from a single event. Although brucellosis contains large flagellar genes, they exhibit species-specific genetic inactivity and as a result they are immobile. Inactive genes consist of single-point mutations or small deletions that are essential for the function or assembly of the flagellum or the major structural proteins of this organ. All three brucellosis species have a complete and functional type IV disruption of the secretory system.

Part of the regulatory phosphorylate system that regulates the growth, division and intracellular survival of the bibortus cell within mammalian host cells. This signaling path

consists of CckA, ChpT, CtrA and CpdR. ChpT effectively and exclusively contains phosphoril groups from the CTP kinase to the receiver domain of both CTRA and CPDR. Does not bind to ATP. Excessive representation of chpT as a result of defect in cell morphology, DNA content and intracellular survival in human bodyguards.

Doxycycline(4,5,5,6,12)-4-(dimethylamino)-3,5,10,12,12-pentahydroxy-6-methyl-1,11-dioxo-1,4,4a,5,5a,6,11,12a-octahydrotetracene- 2-carboxamide) is a drug used for the treatment of infections caused by bacteria and protozoa. Doxycycline is contraindicated during pregnancy, breastfeeding, infant and childhood stages in human's because it crosses into breastmilk and disrupts bone and tooth development in these stages upto the age of eight years. Drug are observing resistance to microbes due to long use. We are docking the doxycycline with to check if they will have better activity and probably show no resistance to microbes (Ikpeazu, O. V., Otuokere, I. E., & Igwe, K. K. 2017).

Aminoglycosides are the best characterized class of antibiotics that bind directly to ribosomal RNA. Aminoglycosides cause decreases in translational accuracy and inhibit translocation of the ribosome. Aminoglycoside antibiotics bind to a conserved sequence of rRNA that is near the site of codon
subunits. Aminoglycoside binding stabilizes the tRNA
by decreasing tRNA dissociation rates, which interferes with proofreading steps that ensure translational fidelity .Besides their medical importance, aminoglycoside antibiotics have provided insights into ribosome function (Yoshizawa, S., Fourmy et.al 1998)

In this study we mainly focused on protein patterns and interactions, also carried out several protein chains analyses, protein-protein interface, structure alignment, structure prediction and active site prediction and a little info in the emerging bioinformatics data regarding Brucella abortus.

METHODOLOGY

1. **Selection of query protein.**

   As we have explained above that we have decided to work on B. abourtus. The query protein is B. abortus strain 2308. This protein is responsible for abortion in Cattel's. For acquiring this protein, we have used Uniprotkb. Uniprotekb is a protein database.

2. **Local alignment search.**

   We are running Blast. Blast is a local alignment search tool. It detects parts that resemblance between biological sequences. This is an online based platform equates protein sequences to sequence databases and calculates the statistical significance of matches.

   For our work we have used Blastp. Blastp is a specific blast type that is performed when are handling a protein sequence and we have to search it against protein database.

3. Secondary Structure prediction.

Secondary structure concerns to the fixture, recurring arrangements in space of neighbouring amino acid residues in a polypeptide chain. It is preserved by hydrogen bonds between amide hydrogens and carbonyl oxygens of the peptide backbone. The major secondary structures are α-helices and β-structures.

GOR (Garnier-Osguthorpe-Robson) is one of the popular Secondary structure prediction tool utilizing information theory. GOR method is established on probability parameters deduced from empirical studies of known protein tertiary structures solved by X-ray crystallography.

4. Homology-modelling.

A protein is known by its structure. Is its structure that describe its function and its creates various levels of interaction limits. We have tried to create the protein model on the basis of homology modelling. The homologs structure is the relationship between structures or genetic sequence gained from the most recent common ancestor.

Here, homologs modelling is carried out by the application of SWISS-MODEL. SWISS-MODEL is a fully fashioned protein structure homology-modelling server.

5. Running PDBsum

PDBsum database is a pictorial database which focuses on furnishing at-a-glance summary of the subject matter of each 3D structure situated in the Protein Data Bank (PDB).It demonstrates the molecule(s) that constitute the structure and schematic diagrams of their interactions.

a) PROCHECK analyses**.**

The PROCHECK analyses furnish a thought of the stereochemical quality of all protein chains in a given PDB structure. The highlighted parts of the proteins which appear to have strange geometry and furnish an overall assessment of the structure as a whole.

b) PROTEIN CHAIN
c) WIRING DIAGRAM
d) TOPOLOGY DIAGRAM
e) PROMOTIF

6) Alignment of 4qpk with 4qpj

Protein alignment is carried out Because protein structure is more evolutionarily conserved than sequence, structural alignments can be more authentic between sequences that are very longitudinally related and that have diverged very widely that sequence equivalence cannot authentic discover their resemblance.

So, we have aligned out two protein sequence that are 4qpk (from SWISS-model) and the 4qpj (from Blastp) by using Pymole.

7) Finding Active Site.

The **active site** is the area of a functional protein where substrate molecules bind and perform a chemical reaction.

The active site of the modelled protein has been located by using Pymole.

8) Retrieval and preparation of Ligands molecules for docking with proteins:

The structures of drugs (ligands) including doxycycline and gentamicin were obtained from the PubChem database (https://pubchem.ncbi.nlm.nih.gov/) and drawn in ChemDraw Ultra 12.0. The ligand structures were transformed into most stable conformations utilizing the Vega ZZ program using the method of conjugate gradient and SP4 force field. The final ligand preparation was done using AutoDock Tools 1.5.6 and saved in pdbqt format.

9) Retrieval and preparation of Proteins for docking with ligands:

The 3D structure of protein including (PDB ID: 4QPK & 4QPJ) were obtained from protein data bank (http://www.rcsb.org/pdb/home/home.do) and saved in PDB format. The active sites of respective proteins were determined using Discovery Studio Visualizer 4.0 program. The protein was prepared using AutoDock Tools 1.5.6 by removing water molecules, adding polar hydrogen, and Kollman charges. The grid box was placed in the center of protein structure using the auto grid method and the x, y, z coordinates were saved, and the protein structure was exported in pdbqt format.

10) Docking Experiment:

The final docking of the ligands and proteins was performed by utilizing AutoDock Vina 1.0 at an exhaustiveness of 80. The docking results were simulated and visualized in Discovery Studio Visualizer 4.0 and Pymol 1.8.6.0.

Results

1) Selection of query protein.

The query protein amino acid sequence are as follows: -

>sp|Q2YQA5|CHPT_BRUA2 Protein phosphotransferase ChpT OS=Brucella abortus (strain 2308) OX=359391 GN=chpT PE=1 SV=1

MSLPVTLSALDLGALLCSRICHDIISPVGAINNGLELLEEGGADEDAMALIKSSARNA
SARLQFARIAFGAAGSAGVQIDTGDAQNVATEYFRNEKPEFTWEGARVLLPKNKVK
LLLNMLLIGNGAIPRGGSLAVRLEGSDTDPRFVITVKGRMLRVPPKFLELHSGAAPEE
PIDAHSVQPYYTLLLAEEAGMKISIHATAEDIVFSAE

2) Local alignment search.

Table: -1.1

Search Parameters	
Program	Blastp
Word size	6
Expect value	0.05
Hitlist size	100
Gapcosts	11,1
Matrix	BLOSUM62
Filter string	F
Genetic Code	1
Window Size	40
Threshold	21
Composition-based stats	2

Table: -1.2

Database	
Posted date	Oct 8, 2020 3:33 AM
Number of letters	31,533,171
Number of sequences	119,997
Entrez query	None

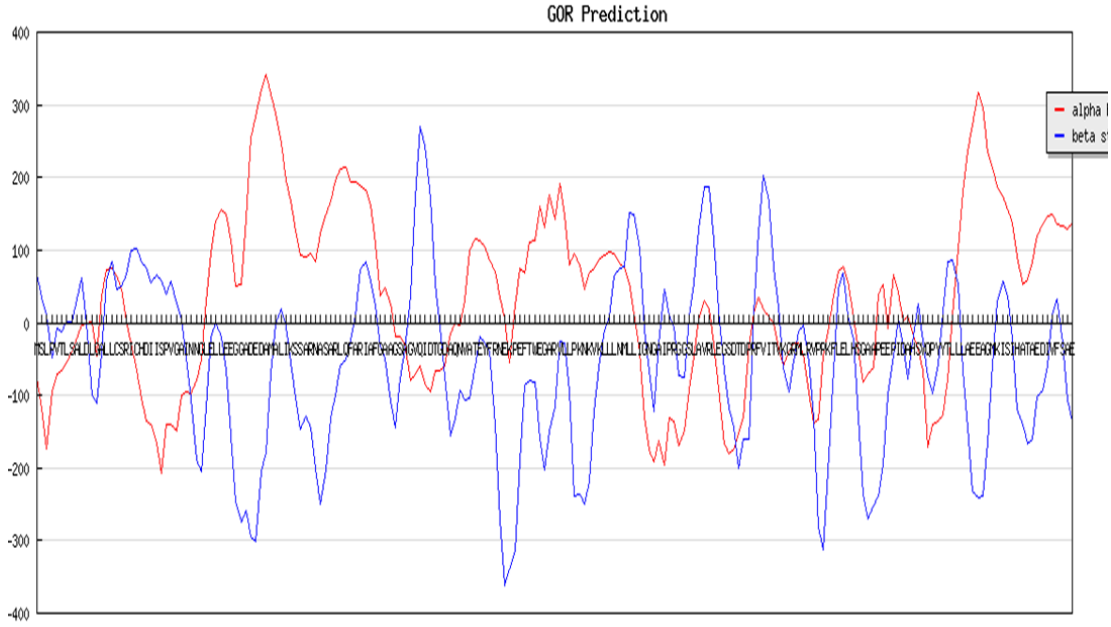
Table: -1.3

Karlin-Altschul statistics		
Lambda	0.317689	0.267
K	0.135153	0.041
H	0.383751	0.14
Alpha	0.7916	1.9
Alpha_v	4.96466	42.6028
Sigma		43.6362

Table: -1.4 Blastp results(against pdb database)

S.no	Select for downloading or viewing reports	Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession
1)	Select seq pdb|4QPJ|A	Chain A, Phosphotransferase [Brucella abortus 2308]	424	424	100%	5e-153	99.52%	4QPJ_A

2)	Select seq pdb|4FMT|A	Chain A, ChpT protein [Caulobacter vibrioides CB15]	99.4	99.4	89%	2e-25	36.27%	4FMT_A
3)	Select seq pdb|4FPP|A	Chain A, Phosphotransferase [Caulobacter vibrioides CB15]	99.4	99.4	89%	4e-25	36.27%	4FPP_A



3) Secondary structure prediction.

Fig: -1.1 alpha helixes and beta sheets arrangements shows by graph.

GOR Algorithm
Protein name:

		Ia	Ib	It	Ic	St
1	M	-78	65*	-15	-10	E
2	S	-118	35	5	50*	C
3	L	-174	10	-227	147*	C
4	P	-96	-49	-45	23*	C
5	V	-71	-8*	-111	-25	E
6	T	-65	-12	-117	-5*	C
7	L	-55	2*	-106	-10	E
8	S	-40	1*	-38	-12	E
9	A	-22	33*	-35	-10	E
10	L	-4	61*	-73	0	E
11	D	-1	-1*	-25	-17	E
12	L	2	-99	41*	-35	T

13	G	-47	-111	77*	-17	T
14	A	37*	-42	-60	-10	H
15	L	72*	57	-83	-75	H
16	L	74	83*	-15	-122	E
17	C	63*	45	62	-137	H
18	S	47	50*	-18	-130	E
19	R	2	67*	7	-105	E
20	I	-27	99*	-10	-67	E
21	C	-65	102*	5	-17	E
22	H	-104	83*	-70	25	E
23	D	-135	74*	-45	30	E
24	I	-140	55*	-55	28	E
25	I	-166	65	-72	66*	C
26	S	-207	58	-269	180*	C
27	P	-141	39	-99	60*	C
28	V	-141	57*	-49	25	E
29	G	-150	25	-22	50*	C
30	A	-100	3	-85	55*	C
31	I	-95	-48	-100	48*	C
32	N	-97	-105	27	43*	C
33	N	-79	-190	127*	45	T
34	G	-50	-206	87*	50	T
35	L	30*	-118	-23	25	H
36	E	95*	-19	-10	-77	H
37	L	140*	-2	-88	-130	H
38	L	155*	-19	-135	-147	H
39	E	150*	-65	-40	-147	H
40	E	114*	-156	90	-125	H
41	G	50	-248	151*	-45	T
42	G	52	-275	46	60*	C
43	A	142*	-258	-78	73	H
44	D	254*	-294	-74	15	H
45	E	284*	-302	-23	-55	H
46	D	322*	-205	-24	-100	H
47	A	341*	-178	-96	-147	H
48	M	315*	-71	-162	-147	H
49	A	287*	0	-173	-145	H
50	L	248*	18	-210	-155	H
51	I	197*	-5	-190	-124	H
52	K	168*	-56	-73	-92	H
53	S	127*	-105	-5	-50	H
54	S	94*	-146	-5	-17	H
55	A	91*	-129	-76	-30	H
56	R	95*	-145	-18	-42	H
57	N	83*	-200	31	-53	H
58	A	122*	-250	59	-47	H

59	S	149*	-203	2	-47	H
60	A	166*	-130	-47	-80	H
61	R	197*	-100	-48	-100	H
62	L	212*	-60	-95	-120	H
63	Q	213*	-50	-43	-122	H
64	F	194*	-24	-87	-130	H
65	A	194*	10	-163	-115	H
66	R	189*	72	-153	-120	H
67	I	182*	85	-194	-122	H
68	A	163*	57	-130	-127	H
69	F	108*	24	-58	-100	H
70	G	38*	-32	-92	-45	H
71	A	49*	-48	-95	0	H
72	A	25*	-108	-12	-10	H
73	G	-18	-144	27*	10	T
74	S	-20	-78	-35	45*	C
75	A	-30	-36	22*	15	T
76	G	-80	37*	5	20	E
77	V	-70	172*	-145	30	E
78	Q	-60	270*	-156	20	E
79	I	-85	243*	-192	28	E
80	D	-95	171*	-140	63	E
81	T	-65	53	-84	75*	C
82	G	-65	-8	4	90*	C
83	D	-60	-72	-17	90*	C
84	A	-15	-155	-5	35*	C
85	Q	-2	-132	53*	10	T
86	N	-4	-93	64*	3	T
87	V	30*	-107	-15	-10	H
88	A	100*	-104	-65	-20	H
89	T	115*	-60	-107	-35	H
90	E	114*	-18	-72	-83	H
91	Y	104*	-29	-76	-100	H
92	F	87*	-39	-60	-87	H
93	R	71*	-139	35	-65	H
94	N	34	-270	152*	-40	T
95	E	5	-363	205*	25	T
96	K	-54	-339	-34	152*	C
97	P	24	-315	83*	18	T
98	E	74*	-190	71	-42	H
99	F	69*	-87	-29	-65	H
100	T	110*	-80	-162	-75	H
101	W	114*	-81	-134	-103	H
102	E	160*	-161	39	-124	H
103	G	131*	-202	68	-122	H
104	A	175*	-150	-70	-91	H

105	R	143*	-116	-91	-132	H
106	V	191*	-26	-107	-150	H
107	L	150*	-27	-133	-85	H
108	L	80*	-92	-256	45	H
109	P	95*	-238	35	-29	H
110	K	79	-236	101*	-25	T
111	N	46	-250	70*	5	T
112	K	68*	-219	17	-5	H
113	V	75*	-122	-42	-45	H
114	K	88*	-51	-33	-72	H
115	L	93*	-9	-74	-113	H
116	L	97*	7	-80	-118	H
117	L	95*	63	-60	-132	H
118	N	81*	74	-61	-141	H
119	M	75	77*	-103	-130	E
120	L	55	151*	-126	-107	E
121	L	13	150*	-142	-99	E
122	I	-40	103*	-85	-59	E
123	G	-132	-11	-10	8*	C
124	N	-177	-66	54	85*	C
125	G	-192	-122	53	105*	C
126	A	-163	-27	-70	130*	C
127	I	-196	47	-359	175*	C
128	P	-130	11	-133	38*	C
129	R	-136	-6	39*	20	T
130	G	-170	-73	132*	55	T
131	G	-148	-76	68	100*	C
132	S	-92	8	-23	75*	C
133	L	-45	55*	-107	-25	E
134	A	6	128*	-83	-97	E
135	V	31	187*	-170	-140	E
136	R	20	188*	-118	-115	E
137	L	-30	114*	-68	-80	E
138	E	-90	17	125*	-42	T
139	G	-167	-65	185*	35	T
140	S	-180	-117	102	130*	C
141	D	-173	-144	75	130*	C
142	T	-154	-201	38	155*	C
143	D	-131	-160	-154	202*	C
144	P	-31	-161	-13	40*	C
145	R	12	14*	-13	-35	E
146	F	35	124*	-95	-80	E
147	V	18	203*	-223	-112	E
148	I	8	167*	-240	-92	E
149	T	0	75*	-172	-92	E
150	V	-35	22*	-120	-95	E

151	K	-59	-62	22*	-58	T
152	G	-39	-95	-20*	-25	T
153	R	-29*	-53	-85	-32	H
154	M	-29	-13*	-102	-63	E
155	L	-44	-4*	-90	-67	E
156	R	-97	-52	0*	-22	T
157	V	-138	-130	-209	115*	C
158	P	-133	-281	-127	150*	C
159	P	-45	-313	116*	50	T
160	K	-12	-206	104*	-8	T
161	F	38*	-64	-45	-65	H
162	L	70*	46	-153	-87	H
163	E	77*	68	-32	-122	H
164	L	58*	11	-32	-122	H
165	H	8*	-23	-40	-110	H
166	S	-40	-130	95*	-42	T
167	G	-81	-237	175*	18	T
168	A	-70	-270	93	110*	C
169	A	-61	-251	-169	182*	C
170	P	40	-239	40	58*	C
171	E	53	-195	84*	45	T
172	E	-7	-100	-132	117*	C
173	P	65*	-46	-67	-25	H
174	I	41*	5	-169	-37	H
175	D	4*	-32	-82	-52	H
176	A	8*	-77	-40	-50	H
177	H	-15	-25	35*	-7	T
178	S	-37	25	100*	26	T
179	V	-67	-25	60*	38	T
180	Q	-172	-75	-139	172*	C
181	P	-141	-98	30	75*	C
182	Y	-136	-57	16	70*	C
183	Y	-127	14	-40	40*	C
184	T	-77	85*	-50	5	E
185	L	10	87*	-60	-55	E
186	L	103*	53	-63	-92	H
187	L	181*	-65	-112	-104	H
188	A	237*	-145	-107	-122	H
189	E	272*	-232	-64	-150	H
190	E	318*	-240	-15	-187	H
191	A	296*	-238	27	-180	H
192	G	235*	-165	-35	-175	H
193	M	214*	-57	-93	-110	H
194	K	188*	27	-150	-133	H
195	I	174*	58	-190	-117	H
196	S	155*	37	-144	-91	H

197 I 138* -22 -170 -99 H
198 H 90* -120 -139 -79 H
199 A 53* -141 -78 -58 H
200 T 60* -167 -61 -12 H
201 A 85* -160 -55 -17 H
202 E 120* -102 -49 -45 H
203 D 135* -94 -124 -80 H
204 I 147* -60 -238 -112 H
205 V 148* 12 -210 -147 H
206 F 136* 33 -128 -115 H
207 S 133* -28 -102 -80 H
208 A 129* -107 -35 -70 H
209 E 138* -135 25 -77 H

4) Homology-modelling.

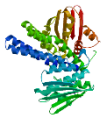
5 model and 48 templates were found

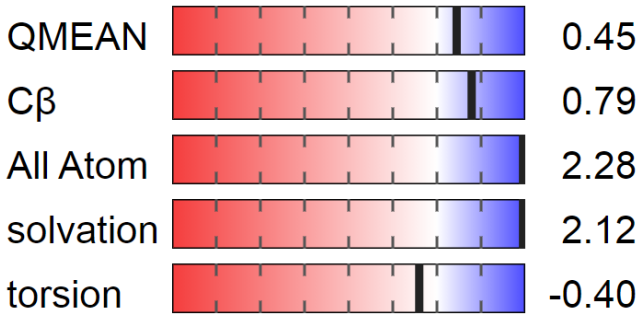
The selected model results are listed below: -

Models

The following models were built (see Materials and Methods "Model Building"):

Table: -2.3

Model #01	File	Built with	Oligo-State	Ligands	GMQE	QMEAN
	PDB	ProMod3 3.1.1	homo-dimer (matching prediction)	None	0.94	0.45



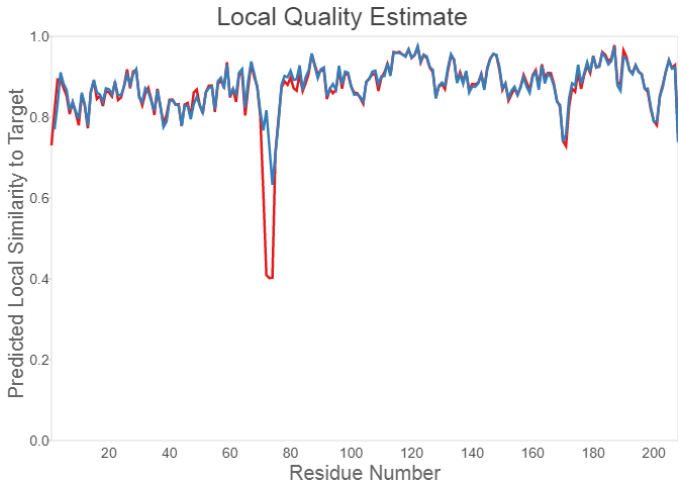


Fig: -1.3

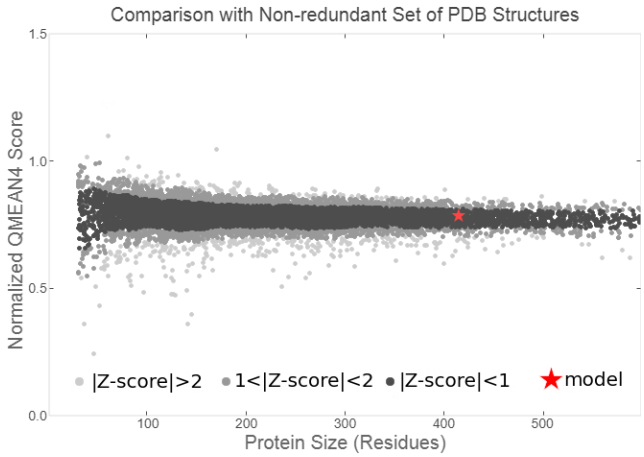


Fig: -1.4

Table: -1.5

Template	Seq Identity	Oligo-state	QSQE	Found by	Method	Resolution	Seq Similarity	Range	Coverage	Description
4qpk.1.A	99.52	homo-dimer	0.86	BLAST	X-ray	1.66Å	0.60	2 - 209	1.00	PHOSPHOTRANSFERASE

Table: -2.5; Excluded ligands

Ligand Name.Number	Reason for Exclusion	Description
GOL.1	Not biologically relevant.	GLYCEROL
GOL.2	Not biologically relevant.	GLYCEROL
GOL.4	Not biologically relevant.	GLYCEROL

Ligand Name.Number	Reason for Exclusion	Description
NA.3	Not biologically relevant.	SODIUM ION
PO4.5	Not biologically relevant.	PHOSPHATE ION

Target
MSLPVTLSALDLGALLCSRICHDIISPVGAINNGLELLEEGGADEDAMALIKSSARNA
SARLQFARIAFGAAGSAGVQID
4qpk.1.A
MSLPVTLSALDLGALLCSRICHDIISPIGAINNGLELLEEGGADEDAMALIKSSARNAS
ARLQFARIAFGAAGSAGVQID

Target
TGDAQNVATEYFRNEKPEFTWEGARVLLPKNKVKLLLNMLLIGNGAIPRGGSLAVR
LEGSDTDPRFVITVKGRMLRVPPK
4qpk.1.A
TGDAQNVATEYFRNEKPEFTWEGARVLLPKNKVKLLLNMLLIGNGAIPRGGSLAVR
LEGSDTDPRFVITVKGRMLRVPPK

Target FLELHSGAAPEEPIDAHSVQPYYTLLLAEEAGMKISIHATAEDIVFSAE
4qpk.1.A FLELHSGAAPEEPIDAHSVQPYYTLLLAEEAGMKISIHATAEDIVFSAE

Target
MSLPVTLSALDLGALLCSRICHDIISPVGAINNGLELLEEGGADEDAMALIKSSARNA
SARLQFARIAFGAAGSAGVQID
4qpk.1.B
MSLPVTLSALDLGALLCSRICHDIISPIGAINNGLELLEEGGADEDAMALIKSSARNAS
ARLQFARIAFGAAGSAGVQID

Target
TGDAQNVATEYFRNEKPEFTWEGARVLLPKNKVKLLLNMLLIGNGAIPRGGSLAVR
LEGSDTDPRFVITVKGRMLRVPPK
4qpk.1.B
TGDAQNVATEYFRNEKPEFTWEGARVLLPKNKVKLLLNMLLIGNGAIPRGGSLAVR
LEGSDTDPRFVITVKGRMLRVPPK

Target FLELHSGAAPEEPIDAHSVQPYYTLLLAEEAGMKISIHATAEDIVFSAE
4qpk.1.B FLELHSGAAPEEPIDAHSVQPYYTLLLAEEAGMKISIHATAEDIVFSAE

5) Running PDBsum

The colouring/shading on the plot constitutes the different regions (Morris et al. (1992)). The darkest areas (here shown in red) correspond to the "core" regions representing the most favourable combinations of phi-psi values.
Ramachandran plot regions;
The different regions on the Ramachandran plot are as described in Morris et al. (1992).

The regions are labelled as follows:

A - Core alpha	L - Core left-handed alpha
a - Allowed alpha	l - Allowed left-handed alpha
~a - Generous alpha	~l - Generous left-handed alpha
B - Core beta	p - Allowed epsilon
b - Allowed beta	~p - Generous epsilon
~b - Generous beta	

The different area was brought from the detected phi-psi distribution for 121,870 residues from 463 known X-ray protein structures. The two most preferred area are the "core" and "allowed" regions which correspond to 10° x 10° pixels having more than 100 and 8 residues in them, respectively. The "generous" regions were defined by Morris et al. (1992) by extending out by 20° (two pixels) all round the "allowed" regions. In fact, the authors found very few residues in these "generous" regions, so they can probably be treated much like the "disallowed" region and any residues in them investigated more closely.

The Selected protein for procheck analyse is 4QPK

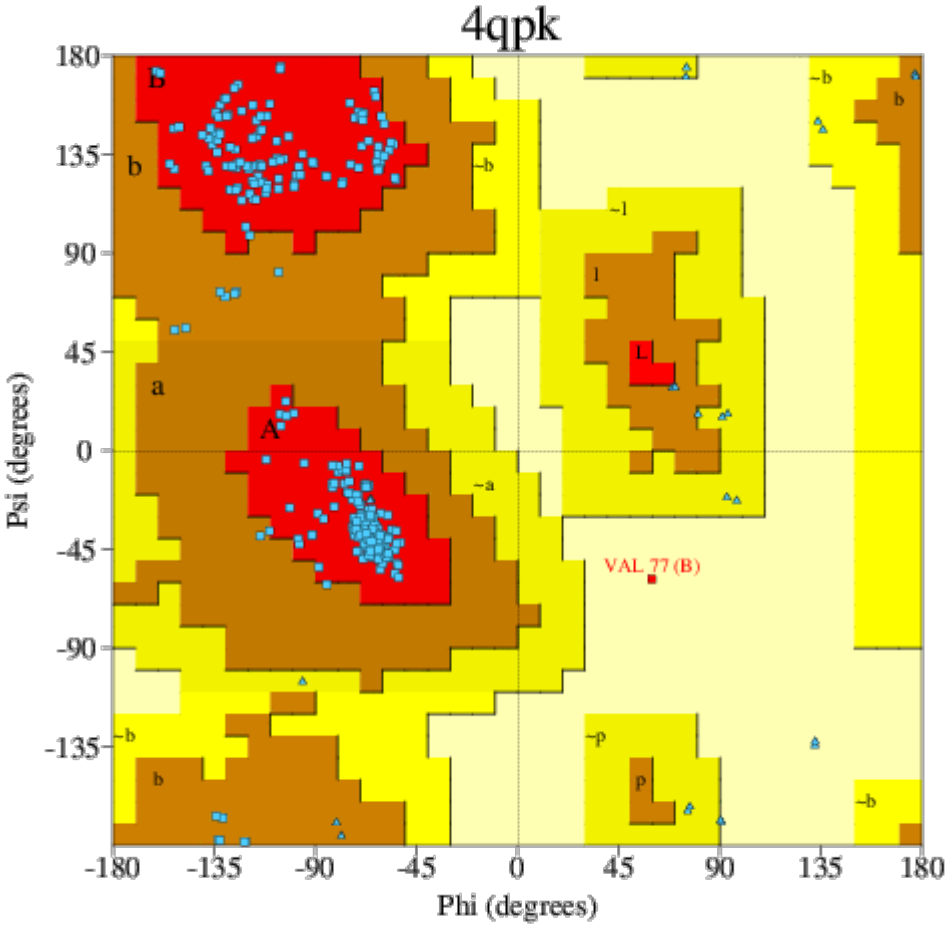


Fig: -1.6

Ramachandra plot analysis

		No. of residues	%-tage
Most favoured regions	[A,B,L]	332	94.9%
Additional allowed regions	[a,b,l,p]	17	4.9%
Generously allowed regions	[~a,~b,~l,~p]	0	0.0%
Disallowed regions	[XX]	1	0.3%*
Non-glycine and non-proline residues		350	100.0%
End-residues (excl. Gly and Pro)		5	
Glycine residues		35	
Proline residues		22	
Total number of residues		412	

Based on an analysis of **118** structures of resolution of at least **2.0** Angstroms and *R*-factor no greater than **20.0** a good quality model would be expected to have over **90%** in the most favoured regions [A,B,L].

2. G-Factors

Parameter	Score	Average Score
Dihedral angles:-		
Phi-psi distribution	0.24	
Chi1-chi2 distribution	0.25	
Chi1 only	0.11	
Chi3 & chi4	0.57	
Omega	-0.20	
		0.12
Main-chain covalent forces:-		
Main-chain bond lengths	0.64	
Main-chain bond angles	0.50	
		0.56
OVERALL AVERAGE		0.30

G-factors provide a measure of how **unusual**, or out-of-the-ordinary, a property is.
Values below -0.5* - unusual
Values below **-1.0**** - highly unusual

Ligands

1) **GOL 301(A)**

Ligand GOL - Glycerol

[Glycerin; propane-1,2,3-Triol]

Formula: C3H83O3

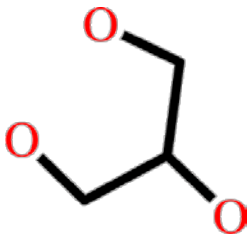


Fig: -1.7 ligand structure

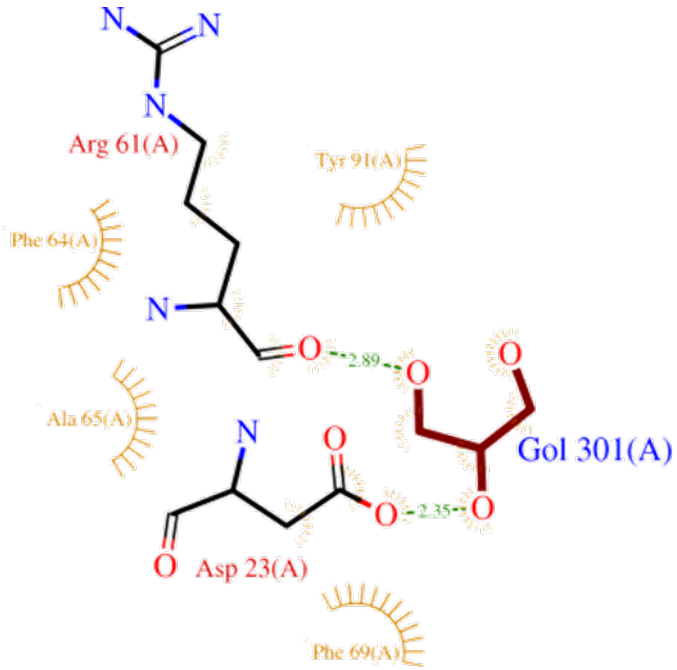


Fig: -1.8; GOL 301(A)

(also representing equivalent ligand GOL 302(B))

2) **GOL 302(A)**

Ligand GOL - Glycerol

[Glycerin; propane-1,2,3-Triol]

Formula: C3H83O3

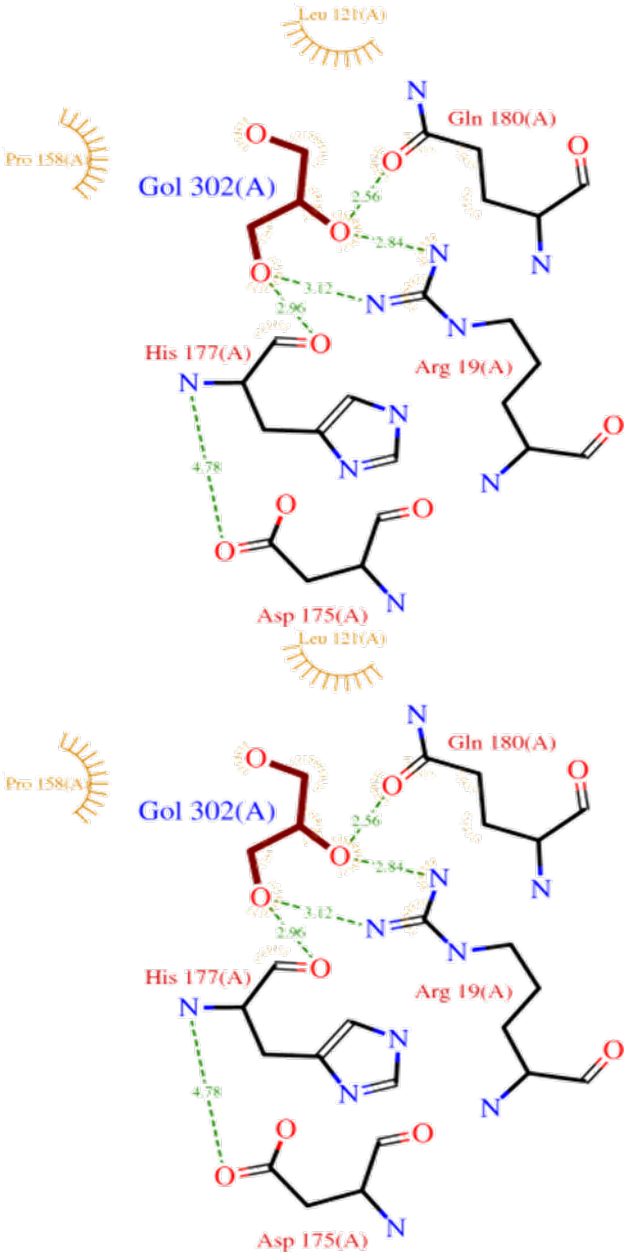


Fig: -1.9; GOL 302(A)

3) **GOL 302(B)**
4) **PO4 303(A)**

Ligand PO4 - Phosphate ion

Formula: O4P3

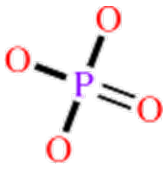


Fig: -2.0

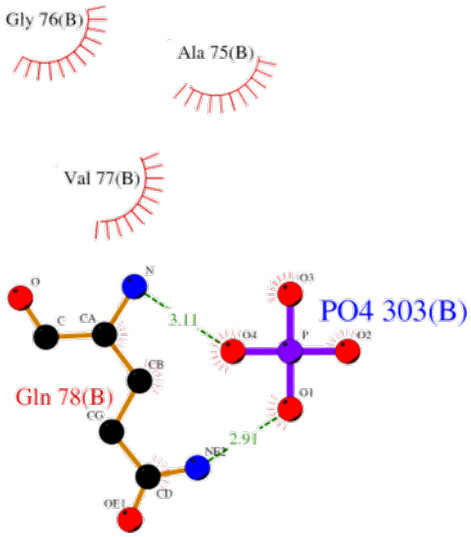
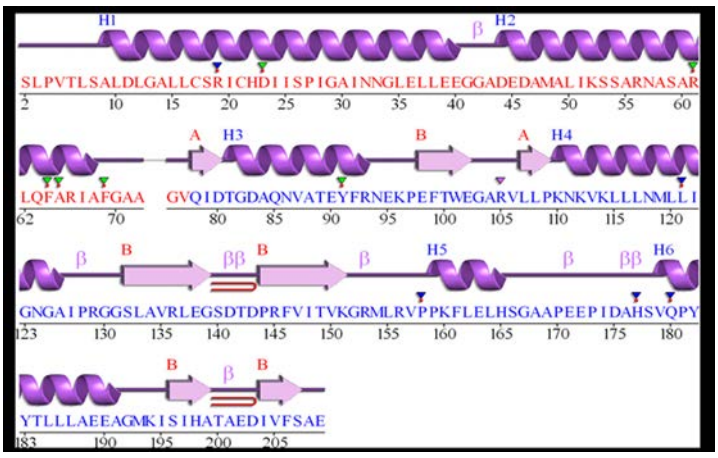


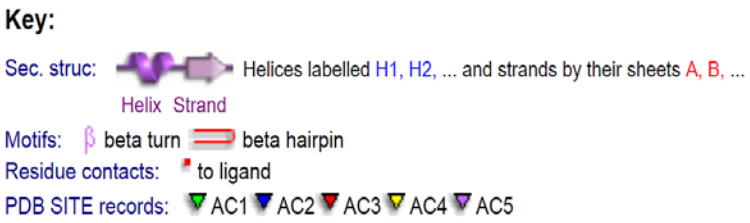
Fig: -2. 1; PO4 303(B);

Too small to validate

Protein Chain

The 'wiring diagram' shows the protein's secondary structure elements (α-helices and β-sheets) together with various structural motifs such as β- and γ-turns, and β-hairpins. The yellow linking bars labelled 1 and 2 represent disulphide bonds. The single-letter amino acid codes showing the protein's sequence are coloured red or blue depending on whether they belong to CATH structural domain 1 or 2, respectively. Catalytic residues are indicated by a box surrounding the amino acid code. Red dots above the single-letter codes signify residues that interact with any bound ligand(s) while coloured lines underneath represent residues belonging to a PROSITE pattern, the redder the colour the more highly conserved the residue in the pattern





Topology diagra diagram exemplifies how -by-side, to figure the domain's central β-sheet. The diagram also demonstrates the relative locations of the α-helices, here presented by the red cylinders. The small arrows indicate the directionality of the protein chain, from the N- to the C-terminus. The numbers within the secondary structural elements correspond to the residue numbering given in the PDB file.

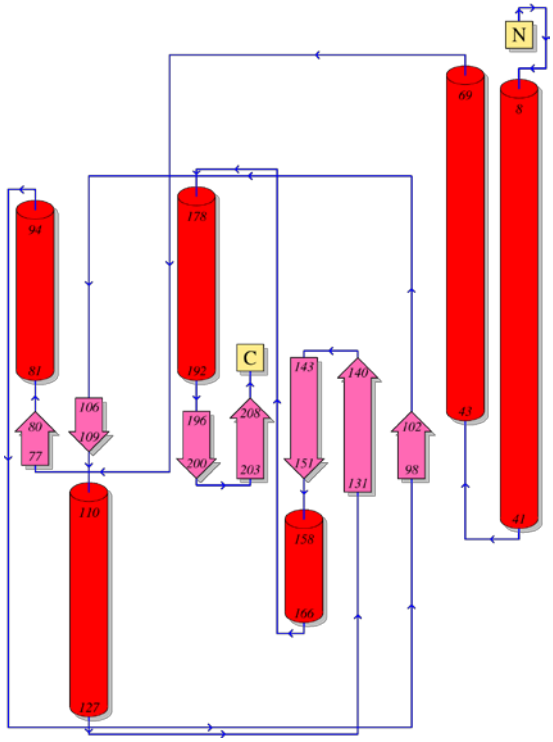


Fig: -2.3; Topology diagram

Promotif

Secondary structure summary

Table: -1.5

Strand	Alpha helix	3-10 helix	Other	Total residues
35 (17.1%)	107 (52.2%)	0 (0.0%)	63 (30.7%)	205

Table: -1.6; 2 beta sheets

Sheet	No. strands	Type	Barrel	Topology
A	2	Antiparall	N	1

				el				
B		5		Mixed		N		1X 1 2X -1

Table: -1.7;2 beta hairpins

Strand 1						**Strand 2**						**Hairpin**
Start		**End**		**Length**		**Start**		**End**		**Length**		**class**
Ser132		Gly139		8		Pro144		Lys151		8		2:4
Ser196		Ala199		4		Ile204		Ser207		4		2:4

Table: -1.8;7 strands

Start		**End**		**Sheet**		**No. resid**
Gln78		Asp80		A		3
Glu98		Glu102		B		5
Leu107		Pro109		A		3
Ser132		Gly139		B		8
Pro144		Lys151		B		8
Ser196		Ala199		B		4
Ile204		Ser207		B		4

Table: -1.9; 6 helices

Start		**End**		**Type**		**No. resid**
Ala9		Glu40		H		32
Asp44		Ala68		H		25
Thr81		Arg93		H		13
Lys110		Ala126		H		17

Pro159		His165		H		7
Val179		Ala191		H		13

Table: -2.0; <u>17 helix-helix interactions</u>

			Helix			Interaction			No. interacting residues		
Helices			**types**			**type**			**Helix 1**		**Helix 2**
A1	A2		H	H		C	N		12		12
A1	A6		H	H		I	I		4		5
A1	B1		H	H		n	n		11		11
A1	B2		H	H		I	I		10		11
A1	B4		H	H		I	I		2		2
A1	B6		H	H		I	I		1		4
A2	A3		H	H		I	I		6		7
A2	A4		H	H		C	N		3		4
A2	A6		H	H		C	C		1		1
A2	B1		H	H		I	I		8		8
A2	B2		H	H		I	I		2		2
A3	A4		H	H		N	I		6		5
A4	A5		H	H		C	N		1		1

A 4	A 6		H	H		I	I		8		8
A 4	B 1		H	H		I	I		3		3
A 5	A 6		H	H		N	N		4		4
A 6	B 1		H	H		I	I		4		1

Table: -2.1; <u>9 beta turns</u>

Turn			
Turn	**Sequence**	**type**	**H-bond**
Gly41-Asp44	GGAD	IV	
Ile127-Gly130	IPRG	I	Yes
Ser140-Asp143	SDTD	I	
Asp141-Pro144	DTDP	IV	
Gly152-Leu155	GRML	VIII	
Pro170-Pro173	PEEP	VIII	
Asp175-Ser178	DAHS	IV	
Ala176-Val179	AHSV	I	
Thr200-Asp203	TAED	I	

Protein-protein interface:

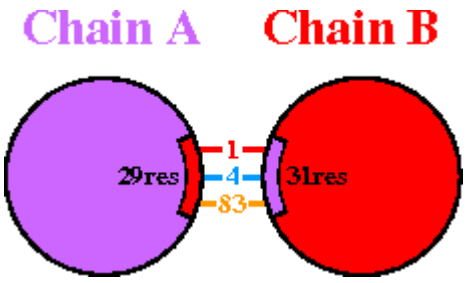


Fig: -2.4

Key: Salt bridges; Disulphide bonds; Hydrogen bonds; Non-bonded contacts

A schematic diagram showing the numbers of interactions across one of the interfaces, namely the A–B interface, and the numbers of residues involved.

Schematic diagram of interactions between protein chains. Interacting chains are joined by coloured lines, each representing a different type of interaction, as per the key above. The area of each circle is proportional to the surface area of the corresponding protein chain. The extent of the interface region on each chain is represented by the black wedge whose size signifies the interface surface area. Statistics for this interface are given below.

Table: -2.2; Interface statistics

Chain	No. of interface residues	Interface area (Å^2)	No. of salt bridges	No. of disulphide bonds	No. of hydrogen bonds	No. of non-bonded contacts
A	29	1519	1		4	83
B	31	1476	1	-	4	83

Fig: -2.5; Residue interactions across interface; Coloured by residue type

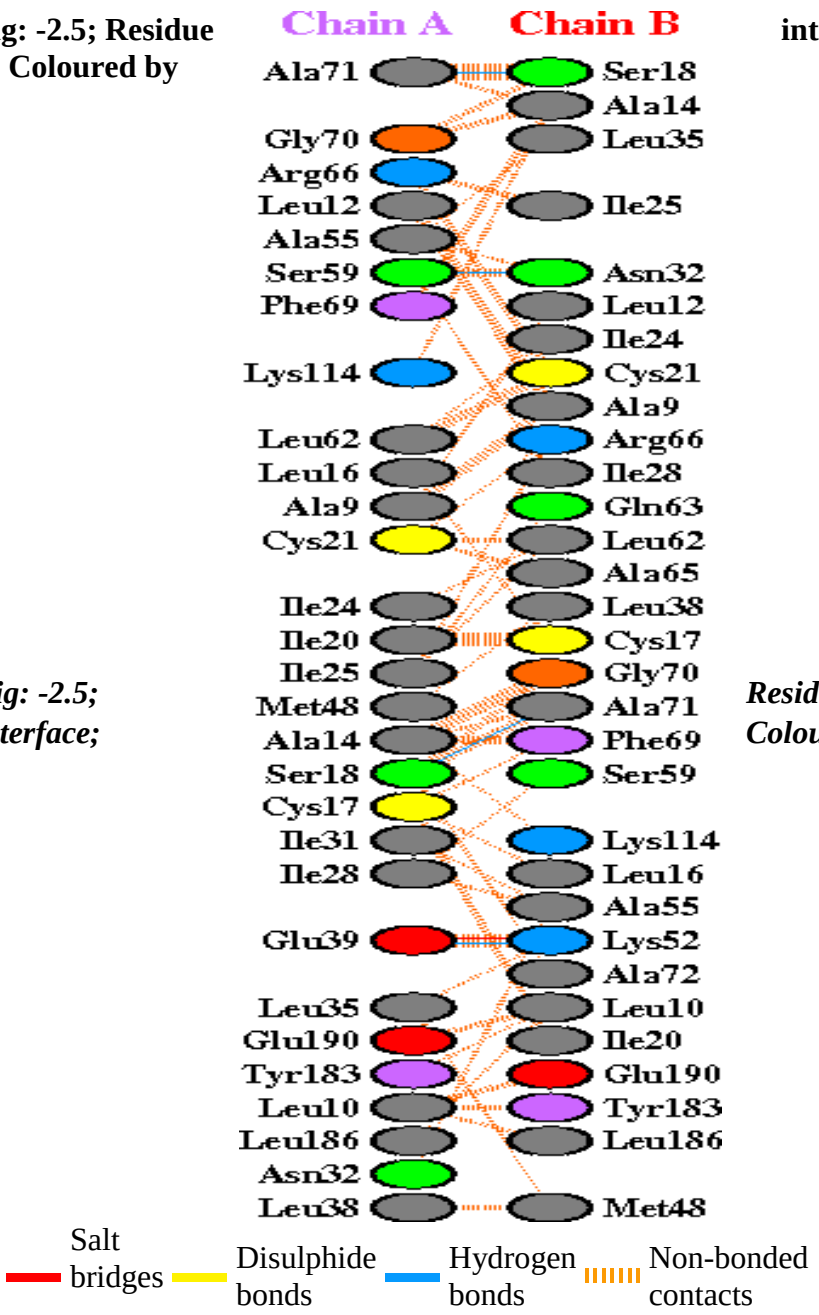


Fig: -2.5; Residue interactions a across interface; Coloured by residue type

Detail of the individual residue–residue interactions across this interface. The colour of the interactions is as above.

The number of H-bond lines between any two residues indicates the number of potential hydrogen bonds between them. For non-bonded contacts, which can be plentiful, the width of the striped line is proportional to the number of atomic contacts.

Residue colours:
Positive (H,K,R); negative (D,E);
S,T,N,Q = neutral;

A,V,L,I,M = aliphatic;
F,Y,W = aromatic;
P,G = Pro&Gly;
C = cysteine.

6) Alignment of 4qpk with 4qpj

Match: read scoring matrix.

Match: assigning 658 x 415 pairwise scores.

MatchAlign: aligning residues (658 vs 415)...

MatchAlign: score 2071.500

ExecutiveAlign: 415 atoms aligned.

ExecutiveRMS: 12 atoms rejected during cycle 1 (RMS=1.40).

ExecutiveRMS: 18 atoms rejected during cycle 2 (RMS=0.85).

ExecutiveRMS: 10 atoms rejected during cycle 3 (RMS=0.73).

ExecutiveRMS: 5 atoms rejected during cycle 4 (RMS=0.68).

ExecutiveRMS: 6 atoms rejected during cycle 5 (RMS=0.67).

Executive: RMS = 0.651 (364 to 364 atoms)

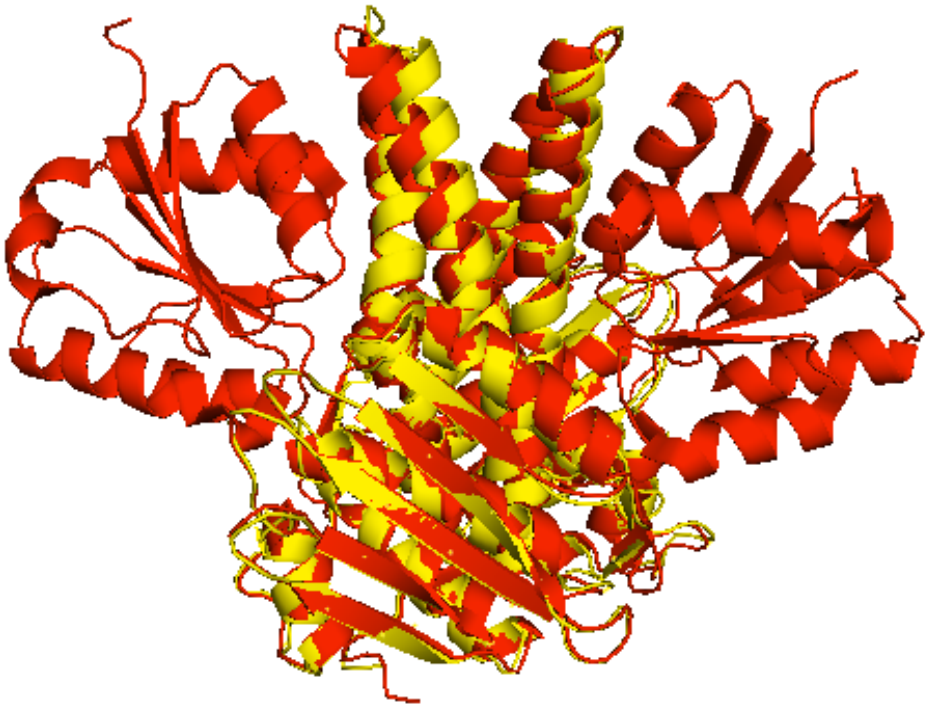


Fig: -2.6. Alignment of 4qpk with 4qpj

7) **Finding Active Site.**

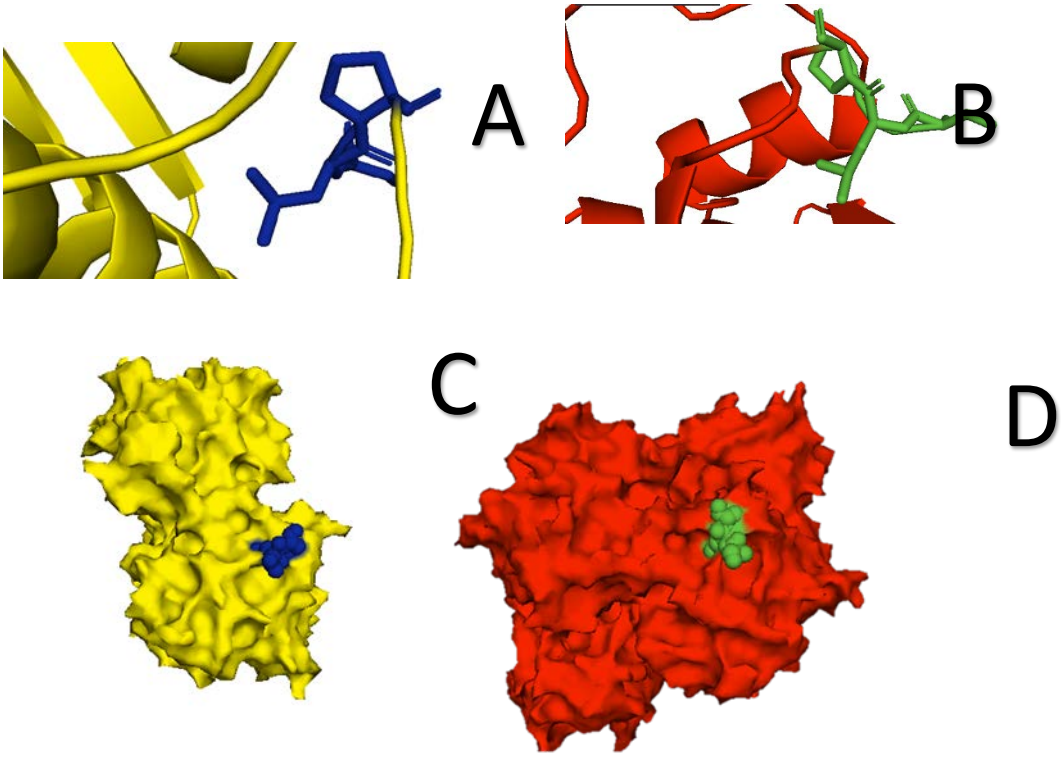


Fig: -2.7; (A) and (C) shows the active site of 4qpk; (B) and (D) shows active site of 4qpj

8) **Results from Docking**

Table 2.3: Docking results of ligands in terms of binding affinity (kcal/mol):

Sr. No	Drugs	Affinity (kcal/mol)	
		4QPJ	4QPK
1.	Doxycycline	-8.3	-8.7
2.	Gentamicin	-6.5	-7.8

1. **Doxycycline with 4QPJ**

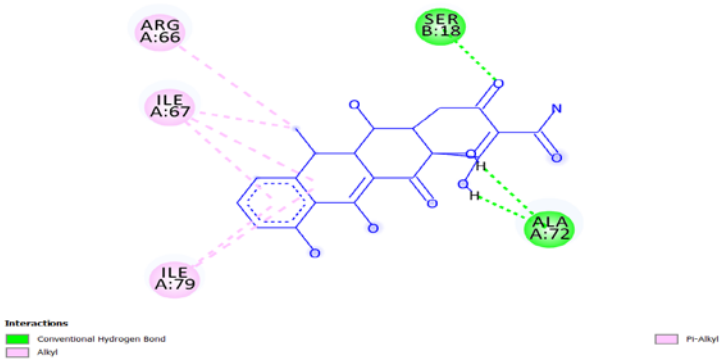


Fig: -2.8; 2D representation of Doxycycline in active site of 4QPJ.

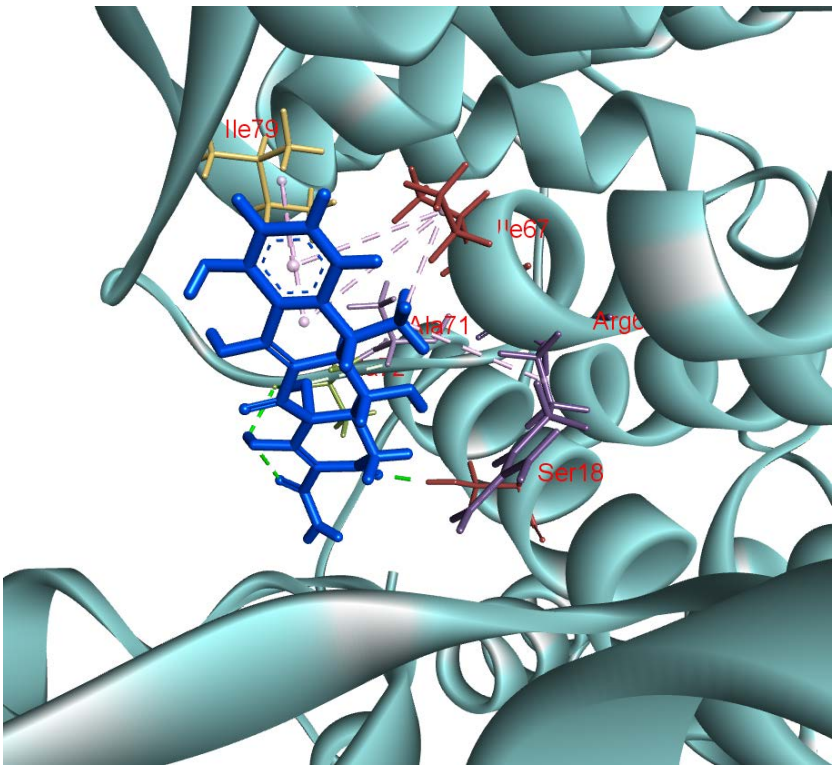


Fig: -2.9; 3D representation of Doxycycline in active site of 4QPJ

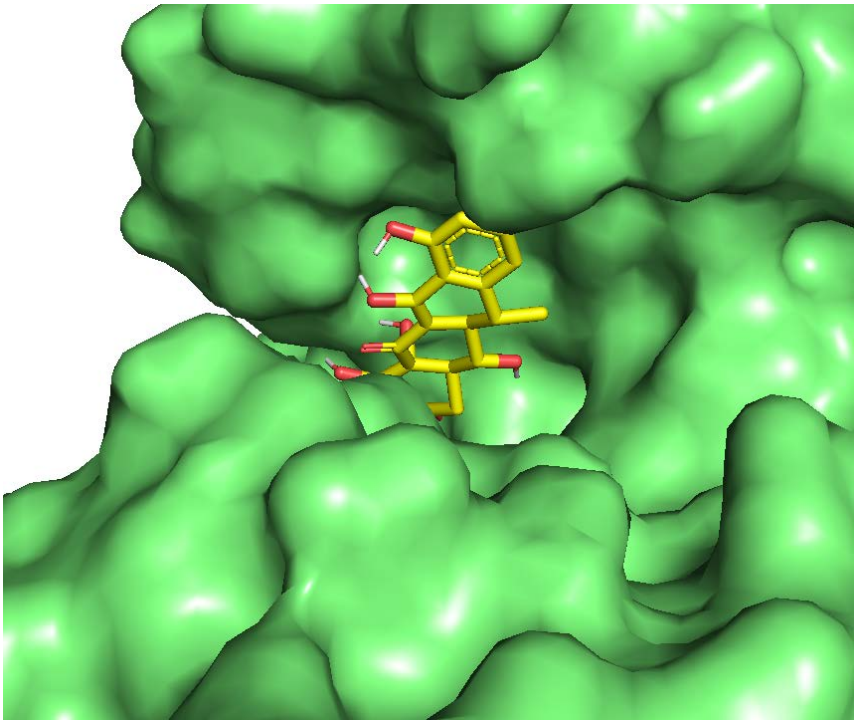


Fig: -3.0; Visualization of docking results showing binding of doxycycline inside the pocket of 4QPJ with ligand as yellow colour sticks

2. Doxycycline with 4QPK

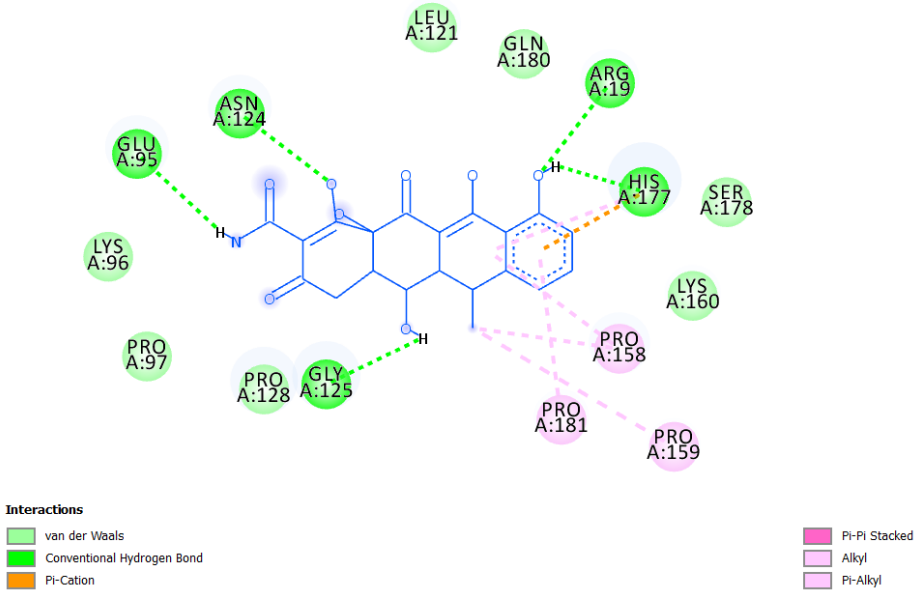


Fig: -3.1; 2D representation of Doxycycline in active site of 4QPK

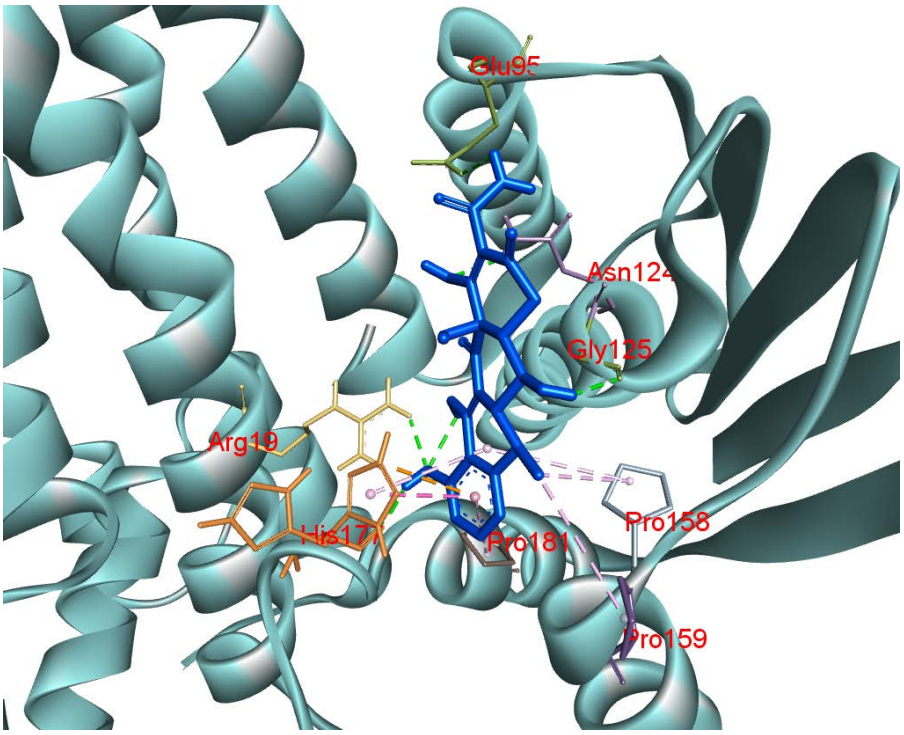


Fig: -3.2; 3D representation of Doxycycline in active site of 4QPK

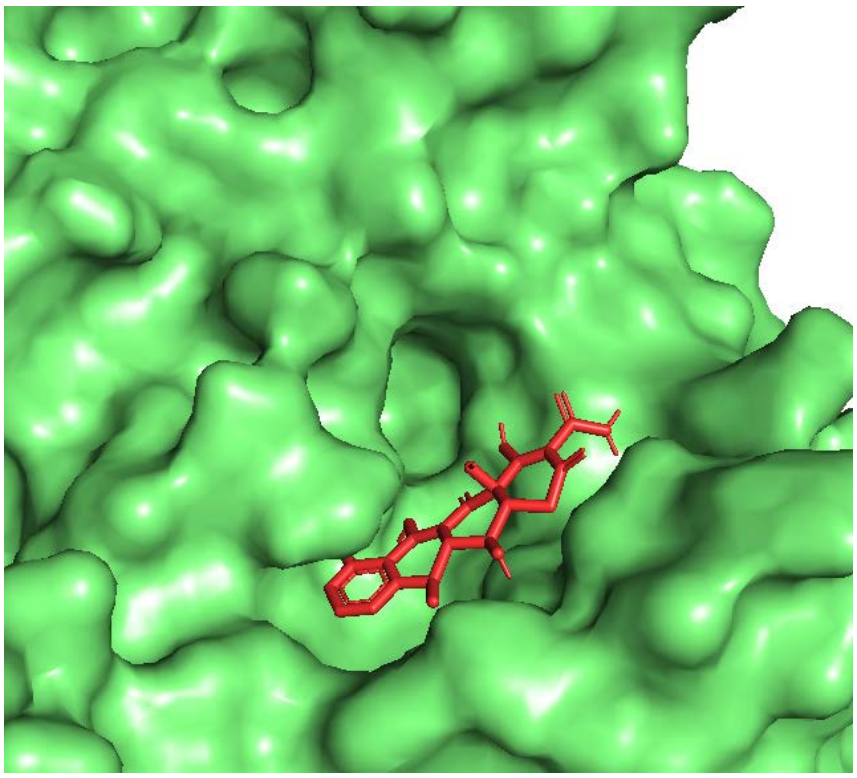


Fig: -3.3; Visualization of docking results showing binding of doxycycline inside the pocket of 4QPK with ligand as red colour sticks

3. Gentamicin with 4QPJ

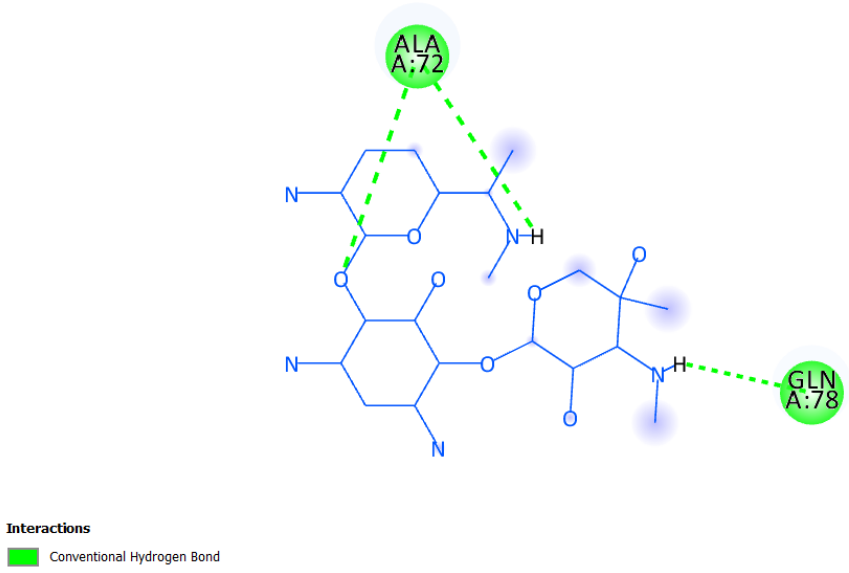


Fig: -3.4; 2D representation of Gentamicin in active site of 4QPJ

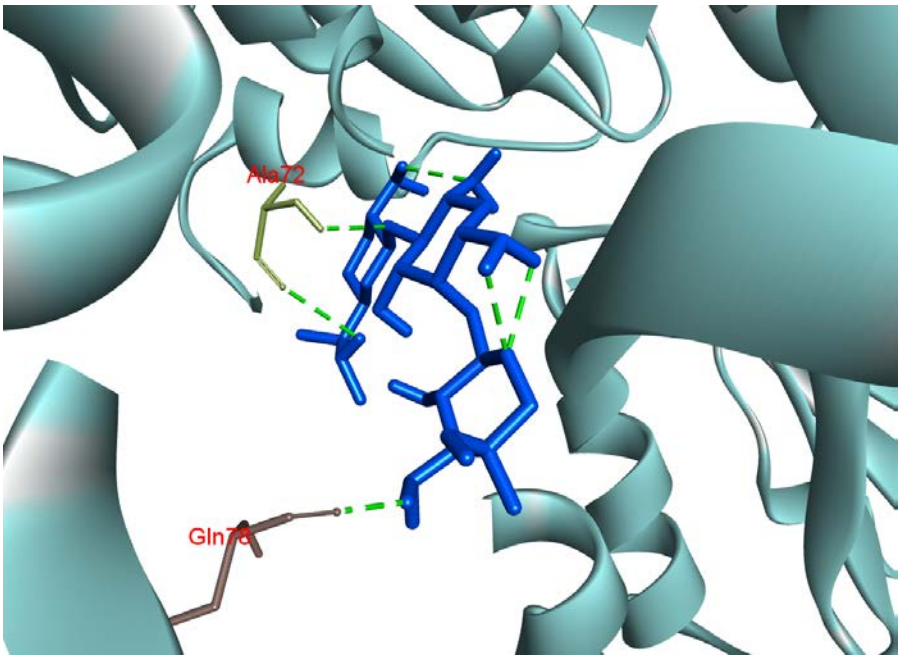


Fig: -3.5; 3D representation of Gentamicin in active site of 4QPJ

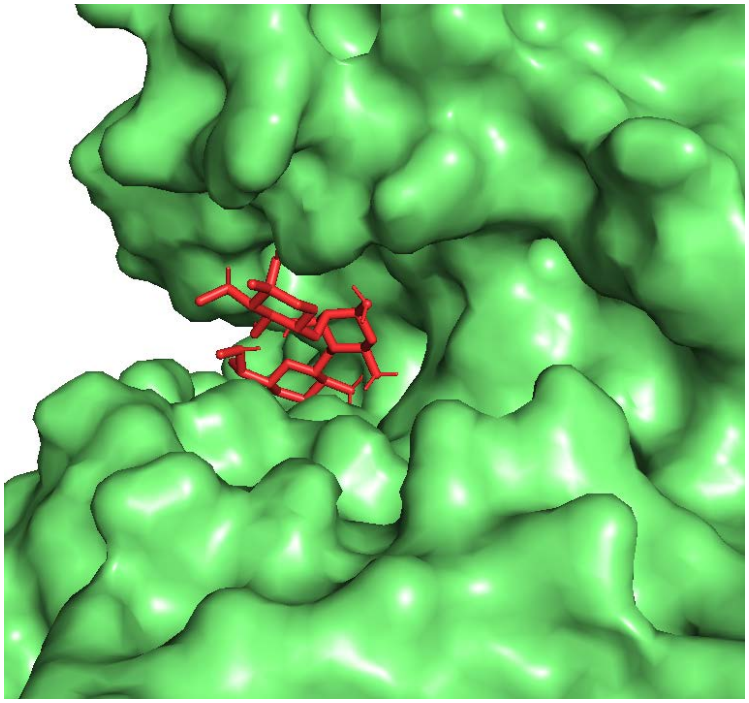


Fig: -3.6; Visualization of docking results showing binding of Gentamicin inside the pocket of 4QPJ with ligand as red colour sticks

4. Gentamicin with 4QPK

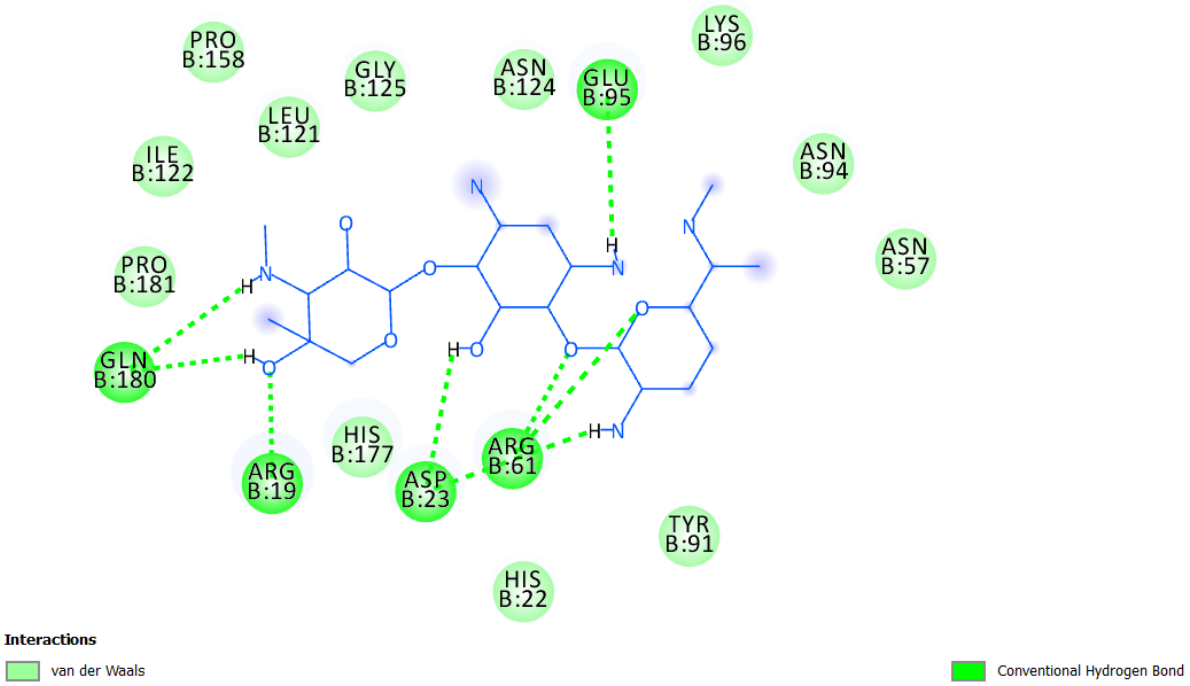


Fig: -3.7; **2D representation of Gentamicin in active site of 4QPK**

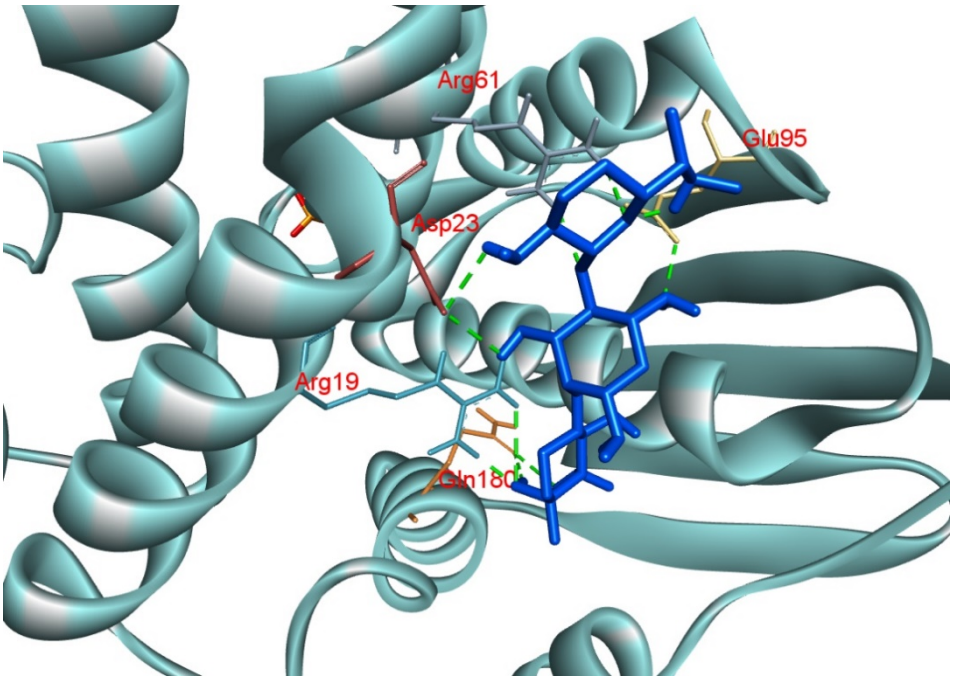


Fig: -3.7; **3D representation of Gentamicin in active site of 4QPK**

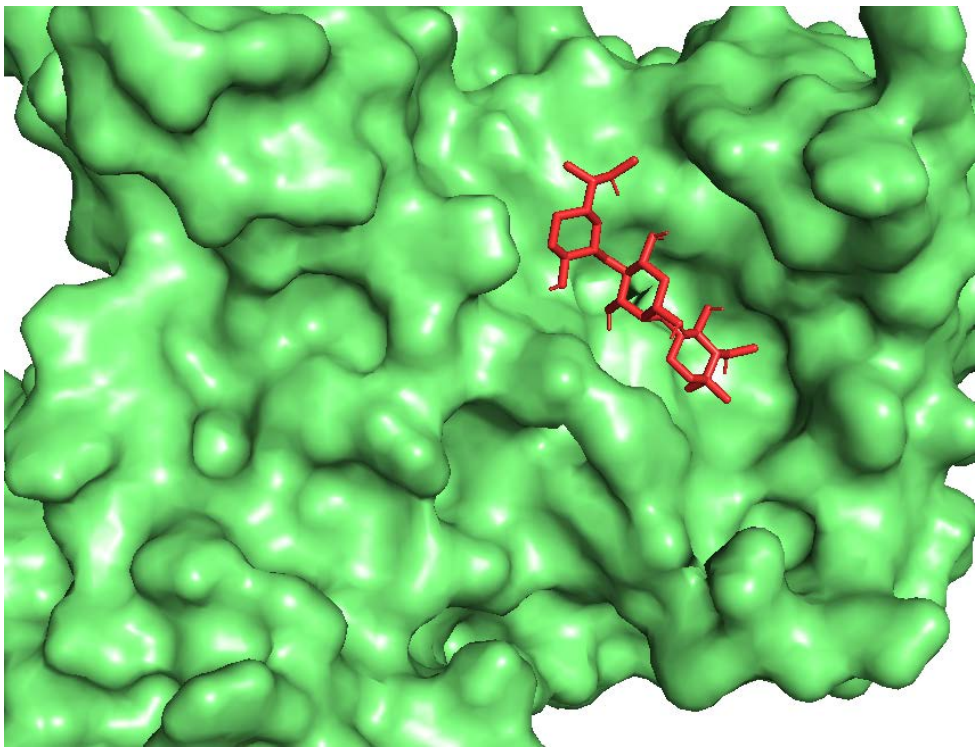


Fig: -3.8; Visualization of docking results showing binding of Gentamicin inside the pocket of 4QPK with ligand as red colour sticks

Conclusion

Insilco analysis of B. abortus strain 2308 protein has been carried out. We have run several tool to understand the viral protein of B.abortus. After running the local alignment search by using Blastp, against pdb database. We have discovered 3 hits. Among those 3 hits 4qpk shows maximum identity of 99.52%, query cover 100% and minimum e value of 5e-153. The same query sequence was put into the secondary structure prediction analysis by using GOR. We have listed above the concentration of alpha helis and beta plated sheets present in the query sequence. Later on a homology modelling of the same query sequence has been done by using Swiss-model where we have discovered 48 templates and 5 model. Among that 5 model the 4qpk shows the maximum possible required identity. Later on we have selected the model from the Swiss- model(4qpk) and analysis it's over the PDBsum protein analysis server. The protein 4qpk have been validated by using Ramachandran plot analysis graph, wiring diagram and topology diagram. Moreover, it has also given use a brief tabulation summary of the protein's secondary structure. Furthermore, we have also discover the protein-protein interface between chain A and Chain B of the modelled protein 4qpk. Later on, we have performed the alignment between 4qpk(modelled protein1) and 4qpj(Blastp seleted hit) by using Pymole. There we have find the RMS value between the two structure is 0.651(364 to 364 atoms). We have also marked the active site of the two structure whose picture has been presented above. Further on we have performed drug docking of the two protein. Here, we have found that the binding energy result obtained from the docking of 4QPJ and 4QPK with ligands Doxycycline and gentamycin was found to be -8.3, -8.7, -6.5, -7.8 kcal/mol respectively.The molecular docking of doxycycline and gentamicin revealed potential interaction with the target protein. Doxycycline has shown a promising binding to

4QPJ with an affinity of -8.3kcal/mol. It is observed to show hydrogen bonding with "chain A" amino acid SER 18 and 2hydrogen bonds with ALA72. It also showing pi-alkyl and alkyl interactions mainly withARG66, ILE67 and ILE79. Doxycycline with an affinity of -8.7, it has also shown 5hydrogen boning interaction with 4QPJ mainly with GLU96, ASN124, GLY125, ARG19,HIS177, ARG19. The van der Waals interaction was encountered with LYS96, PRO97,PRO128. Pi-alkyl interactions are also observed with PRO158, PRO159 and PRO151. Gentamicin is showing hydrogen bonds with ALA72 and GLN78 with an affinity of -6.5 with4QPJ. A prominent binding interaction of gentamicin with 4QPK has been observed with anaffinity of -7.5kcal/mol. Around eight hydrogen bonding interactions have been encounteredwith GLU95, GLN180, ARG19, ASP23 and ARG61. This interaction is also accompanied byvan der Waals interactions with GLY125, LEU121, PRO158, ILE122, PRO181, HIS22 andTYR91

Reference

1. Khan, M. Y., Mah, M. W., & Memish, Z. A. (2001). Brucellosis in pregnant women. Clinical infectious diseases, 32(8), 1172-1177.
2. Samartino, L. E., & Enright, F. M. (1993). Pathogenesis of abortion of bovine brucellosis. Comparative immunology, microbiology and infectious diseases, 16(2), 95-101.
3. Hajibemani, A., & Sheikhalislami, H. Zoonotic pathogens cause of animal abortion and fetal loss.
4. Priyanka, S. B., & SK, K. (2019). Bovine brucellosis, A review on background information and perspective. Journal of Entomology and Zoology Studies, 7(2), 607-613.
5. He, Y. (2012). Analyses of Brucella pathogenesis, host immunity, and vaccine targets using systems biology and bioinformatics. *Frontiers in cellular and infection microbiology*, *2*, 2.
6. Ragan, V. E. (2002). The animal and plant health inspection service (APHIS) brucellosis eradication program in the United States. *Veterinary microbiology*, *90*(1-4), 11-18.
7. Ragan, V. E. (2002). The animal and plant health inspection service (APHIS) brucellosis eradication program in the United States. *Veterinary microbiology*, *90*(1-4), 11-18.
8. Corbel, M. J. (2006). Brucellosis in humans and animals. World Health Organization.
9. YAEGER, M.J. and HOLLER, L.D., 2007. Bacterial causes of bovine infertility and abortion. In Current therapy in large animal theriogenology (pp. 389-399). WB Saunders.
10. Reichel, M.P., Wahl, L.C. and Hill, F.I., 2018. Review of diagnostic procedures and approaches to infectious causes of reproductive failures of cattle in Australia and New Zealand. Frontiers in veterinary science, 5, p.222
11. Suárez-Esquivel, M., Ruiz-Villalobos, N., Castillo-Zeledón, A., Jiménez-Rojas, C., Roop Ii, R.M., Comerci, D.J., Barquero-Calvo, E., Chacón-Díaz, C., Caswell, C.C., Baker, K.S. and Chaves-Olarte, E., 2016. Brucella abortus strain 2308 Wisconsin genome: importance of the definition of reference strains. *Frontiers in microbiology*, *7*, p.1557.
12. Ikpeazu, O. V., Otuokere, I. E., & Igwe, K. K. (2017). In silico structure-activity relationship and virtual screening of monosubstituted doxycycline with Pseudomonas aeruginosa Lipase. J Anal Pharm Res, 5(3), 00139.
13. Yoshizawa, S., Fourmy, D., & Puglisi, J. D. (1998). Structural origins of gentamicin antibiotic action. The EMBO journal, 17(22), 6437-6448.
14. The UniProt Consortium. UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res. 47: D506-515 (2019)

15. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990) "Basic local alignment search tool." J. Mol. Biol. 215:403-410.
16. Gish, W. & States, D.J. (1993) "Identification of protein coding regions by database similarity search." Nature Genet. 3:266-272.
17. Madden, T.L., Tatusov, R.L. & Zhang, J. (1996) "Applications of network BLAST server" Meth. Enzymol. 266:131-141.
18. Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D.J. (1997) "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." Nucleic Acids Res. 25:3389-3402.
19. Zhang Z., Schwartz S., Wagner L., & Miller W. (2000), "A greedy algorithm for aligning DNA sequences" J Comput Biol 2000; 7(1-2):203-14.
20. Zhang, J. & Madden, T.L. (1997) "PowerBLAST: A new network BLAST application for interactive or automated sequence analysis and annotation." Genome Res. 7:649-656.
21. Morgulis A., Coulouris G., Raytselis Y., Madden T.L., Agarwala R., & Schäffer A.A. (2008) "Database indexing for production MegaBLAST searches." Bioinformatics 15:1757-1764.
22. Camacho C., Coulouris G., Avagyan V., Ma N., Papadopoulos J., Bealer K., & Madden T.L. (2008) "BLAST+: architecture and applications." BMC Bioinformatics 10:421.
23. Boratyn GM, Schäffer AA, Agarwala R, Altschul SF, Lipman DJ, & Madden T.L. (2012) "Domain enhanced lookup time accelerated BLAST." Biol Direct. 2012 Apr 17;7:12.
24. Boratyn GM, Thierry-Mieg J, Thierry-Mieg D, Busby B, Madden T.L. (2019) "Magic-BLAST, an accurate RNA-seq aligner for long and short reads." BMC Bioinformatics. 2019 Jul 25;20(1):405.
25. Garnier, J., Osguthorpe, D.J., and Robson, B. (1978) Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. J. Mol. Biol. 120, 97-120.
26. Laskowski, R. A. (2009). PDBsum new things. Nucleic acids research, 37(suppl_1), D355-D359.
27. Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., Lepore, R., Schwede, T. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 46(W1), W296-W303 (2018).
28. Bienert, S., Waterhouse, A., de Beer, T.A.P., Tauriello, G., Studer, G., Bordoli, L., Schwede, T. The SWISS-MODEL Repository - new features and functionality. Nucleic Acids Res. 45, D313-D319 (2017).
29. Guex, N., Peitsch, M.C., Schwede, T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. Electrophoresis 30, S162-S173 (2009).
30. Studer, G., Rempfer, C., Waterhouse, A.M., Gumienny, G., Haas, J., Schwede, T. QMEANDisCo - distance constraints applied on model quality estimation. Bioinformatics 36, 1765-1771 (2020).
31. Bertoni, M., Kiefer, F., Biasini, M., Bordoli, L., Schwede, T. Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. Scientific Reports 7 (2017).
32. Willett, J. W., Herrou, J., Briegel, A., Rotskoff, G., & Crosson, S. (2015). Structural asymmetry in a conserved signaling system that regulates division, replication, and virulence of an intracellular pathogen. Proceedings of the National Academy of Sciences, 112(28), E3709-E3718.
33. Morris AL, MacArthur MW, Hutchinson EG, Thornton JM (1992). Stereochemical quality of protein structure coordinates. Proteins, 12, 345-364.
34. DeLano, W. L. (2002). PyMOL. DeLano Scientific, San Carlos, CA, 700.

35. DeLano, W. L. (2009). The PyMOL Molecular Graphics System; DeLano Scientific: San Carlos, CA, 2002. There is no corresponding record for this reference.
36. DeLano, W. L. (2002). Pymol: An open-source molecular graphics tool.
37. -Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S. and Olson, A. J. (2009) Autodock4 and AutoDockTools4: automated docking with selective receptor flexiblity. J. Computational Chemistry 2009, 16: 2785-91.
38. https://pubchem.ncbi.nlm.nih.gov/