

Molecular Modeling and Virtual Screening of Molecular Inhibitors for Leptospiral Collagenase

Vikram Kumar

Madurai Kamaraj University

Nagesh Srikakulam

Madurai Kamaraj University

Veeranarayanan Surya Aathmanathan

Bharathidasan University

Padikara K Satheeshkumar

Banaras Hindu University

Madanan Gopalakrishnan Madathiparambil

Regional Medical Research Centre Port Blair

Muthukalingan Krishnan

Bharathidasan University

Jebasingh Tennyson (✉ jebasinghs@gmail.com)

Madurai Kamaraj University <https://orcid.org/0000-0001-9703-6090>

Research Article

Keywords: Leptospira, Collagenase, Threading, Docking, Structure-based drug designing, Molecular dynamics, Pharmacophore

Posted Date: March 3rd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-259091/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Collagenase is a virulence factor which facilitates the invasion of pathogenic *Leptospira* into the host. In the present study, the model of Leptospiral collagenase was constructed by employing threading method with the crystal structure of collagenase G. Three ligand binding sites at N-terminus, catalytic site and C-terminus were predicted by Metapocket server. Among sixty seven inhibitors from the ChEBI and Zinc databases, Protohypericin is predicted as the best inhibitor since it binds at the catalytic site of Leptospiral collagenase. Molecular dynamic simulation studies validated the stability of interaction between the active site of Leptospiral collagenase and Protohypericin. The docking and molecular simulation studies corroborated the potential of the ligand to curb leptospiral infection.

1. Introduction

Leptospirosis is a zoonotic disease caused by pathogenic spirochetes *Leptospira*, which has very high impact on human and animal health worldwide [1]. Clinical symptoms range from a self-resolving acute undifferentiated febrile illness to severe health conditions, such as acute renal failure [2], jaundice, haemorrhage [3] and vascular collapse [4]. *Leptospira* gets transmitted directly through the infected tissues and body fluids (blood, milk from the infected mother, sexual contacts), and indirectly through contaminated water [5]. In nature, *Leptospira* colonizes, survives (weeks to months in moist soil and water), forms aggregates and biofilm on the surface of water and soil [6]. When a healthy person comes in contact with *Leptospira* by various sources, the pathogen enters through skin abrasion and mucous membrane of eyes, throat and nose and subsequently infects various organs of the body by travelling through bloodstream [5].

The progression of infection starts with the adhesion of *Leptospira* to the extracellular matrix (ECM) of the host tissue [7] followed by invasion into various organs of the host [8]. The invasion process is known to involve crossing the ECM barrier which is composed of collagen, proteoglycans and glycoproteins. Collagen serves as the major structural component of nearly all mammalian tissues [9,10,11]. Type I, III, and IV collagens are involved in the formation of epithelial and endothelial barriers against pathogen invasion. It has been reported that collagenase of *Clostridium histolyticum* damaged these cells and subsequently invaded into host tissues [12]. Collagenase of *Vibrio parahaemolyticus* [13], *Fusobacterium nucleatum* [14] and *Clostridium perfringens* [15] act as virulence factors for invasiveness and tissue injury. However, the exact mechanism by which *Leptospira* invasion through extracellular barriers takes place is not fully understood.

Results from comparative genome analysis between pathogenic (*L. interrogans*) and non-pathogenic (*L. biflexa*) strains of *Leptospira*, identified the comprehensive array of pathogenic genes associated with pathogen establishment in the host. In the analysis, majority of pathogenic genes are found to be of unknown function. Interestingly, collagenase encoding gene is absent in the genome of non-pathogenic *L. biflexa* [17] but a collagenase-encoding gene (*colA*) is present in the pathogenic strains of *Leptospira* [18]. It has been reported that *colA* is an important virulence factor responsible for the invasion and

transmission of *L. interrogans* strain Lai [18]. Janwitthayanan et al. 2013[19] have shown immunoreactivity of recombinant ColA protein with the *Leptospira* infected patient sera. Wet lab experimental data related to *in vivo* transcriptome obtained during the host adaption within the peritoneal cavities of *Leptospira* infected rat revealed that collagenase (*colA/LIC12760*) expression increased 49 folds (*p* value of 2.63E-51) [20].

In the present study, we have predicted the molecular structure of Leptospiral collagenase using *ab-initio* method. Docking studies have also been performed and identification of Protohypericin as the best inhibitor for Leptospiral collagenase was established. Molecular dynamics results show a strong interaction and better stability of Protohypericin with the active site of the Leptospiral collagenase substantiating the docking results.

2. Methods

2.1 Modeling, refinement and evaluation of predicted model protein

The complete gene sequence of collagenase (Q72NR9) was retrieved from the whole genome annotated sequence of *L. interrogans Icterohaemorrhagiae serovar copenhageni* (Fiocruz L1 – 130) using Uniprot database (<http://www.uniprot.org/>) [21]. Protein family of Leptospiral collagenase was searched against protein family database (pfam) (<http://pfam.sanger.ac.uk/>) [22]. Secondary structure of the Leptospiral collagenase was predicted using PSIPRED online server (<http://bioinf.cs.ucl.ac.uk/psipred/>) [23]. I-Tasser (<http://zhanglab.ccmb.med.umich.edu/I-TASSER/>) [24] was used to predict the 3-D model based on threading. The structures predicted in I-Tasser were ranked based on C-score. Swiss PDB viewer (<http://spdbv.vital-it.ch/>) [25] was used to bring amino acids from disallowed region to allowed region of the Ramachandran plot with loop building method and energy minimization of the predicted model by using GROMOS96 force field in vacuum [26]. Model was further refined by using 3D^{refine} server (<http://sysbio.rnet.missouri.edu/3Drefine/>) [27] which improved the global and local structure quality by optimizing hydrogen bond network and atomic level energy minimization based on composite physics and knowledge-based force field. Quality and stereo-chemical parameter of the model structure were evaluated using PROCHECK [28], ProSA [29] and ERRAT [30]. PyMOL [31] was used as molecular visualization of protein 3-D structures.

2.2 Prediction of ligand binding site using Metapocket

The active site of *Leptospira* collagenase was predicted using Metapocket 2.0 server (<http://projects.biotec.tu-dresden.de/metapocket/>) [32] which combines the predictions of individual algorithms like LIGSITE, PASS, Q-SiteFinder, SURFNET, Fpocket, GHECOM, ConCavity and POCASA. The top three ligand binding cavities were obtained from the predicted sites by each algorithm. The predicted sites were clustered and they were ranked based on z-scores obtained for each predictor. The z-score of predicted binding sites helps to calculate the mass centre for each cluster and the potential binding amino acid residues in the binding pockets.

2.3 Ligand based- virtual screening

The low molecular weight compound, Funalenone (Mw: 288.25 Da), which is a microbial collagenase inhibitor (EC 3.4. 24.3) was retrieved from ChEBI database (<http://www.ebi.ac.uk/chebi/chebiOntology.do?chebield=CHEBI:67189>) [33]. Structural analogue of Funalenone and its identical molecule from ChEBI were searched against Zinc database (<http://zinc.docking.org/>) [34]. In ChEBI database, Pinoquercetin (ChEBI 8224) was found to be a structurally similar compound showing 70% similarity against funalenone. These two molecules were again searched in the ZINC database. 65 (4 molecules for Funalenone and 61 for Pinoquercetin) structurally similar compounds were found (Table S1) with 70% and 90% similarity respectively. Overall 67 (2+4+61) filtered structural analogue were subjected to virtual screening protocol with Autodock vina [35]. Top three ligand binding pockets were used as docking sites to find the best potential inhibitor which binds and inhibits the Leptospiral collagenase activity.

2.4 Docking and scoring

Molecular docking was performed by using the Autodock vina to perform parallel computing to speed up the execution and accuracy [35]. Autodock tool python scripts [36] were used for ligand and receptor preparation. The *mol2* format ligand files were downloaded from ZINC database. The receptor was prepared by adding polar hydrogens, merging non-polar charges and removing water molecules as well as lone pair electrons. During ligand and receptor preparation, Gasteiger partial charges were added. Autodock vina was used for rigid body docking protocol. The grid box was generated for three different ligand binding sites (A, B and C) as predicted by the Metapocket server. The points of grid box at each dimension (x, y and z) for A, B and C sites were kept at 32 X 30 X 36; 40 X 40 X 40 and 42 X 42 X 36 with a spacing 1Å respectively. The grid box center for binding sites A, B and C were position at 255.1, 104.6, 132.9; 265.9, 45.6, 225.6 and 266.1, 29.0, 249.1 respectively. The configuration files for above binding pockets were created for receptors. The maximum number of binding pockets of the ligand was defined as 9. The exhaustiveness was given as 8. The rest of the parameters were kept as default. The docking studies were carried out by parsing the ligand and configuration files to Autodock vina program. The output of the docking was generated in pdbqt format for each ligand in a single file with multiple conformations. It also generates a log file, which consists of minimum binding energy in ascending order and RMSD values for each conformation of ligand molecules. The best ligand molecule was chosen based on minimum binding energy required for binding of ligand to the receptor molecule. PyMOL was used to visualize ligands in the pdb format.

2.5 Generation of Pharmacophore model

Pharmacophore model was determined by using LigandScout [37] for the best pose of ligand in the predicted active sites of the Leptospiral collagenase. It generates the 3-D model for the given ligands by superimposition and merges the chemical features and functionalities such as H-bond donors, acceptors, ionizable groups, etc. This aligned chemical features in 3-D space shows the essential interactions of small organic ligands with a macromolecular receptor.

2.6 Molecular dynamic simulations

The stability of the docked protein-ligand complex was subsequently studied under optimum thermobaric conditions using molecular dynamic studies. GROMINGEN MACHINE for CHEMICAL SIMULATIONS (GROMACS) molecular dynamics package v 5.1.4 [1] was employed to study the stability of the protein-ligand complex. The gromacs96 43a1 force field [6] and PRODRG online server [38] were used to generate topology and determine the intermolecular interactions for protein and ligands, respectively. The protein-ligand complex was placed in rhombic dodecahedron box. SPC water model was used to fill the protein-ligand complex with water and ions. The steric clashes and bad contacts between protein-ligand complex and the system were reduced by performing energy minimization, which uses a steep descent algorithm approach. Further, the system was equilibrated at 300 K temperature and 1 bar pressure for 100 ns, using two steps NPT and NVT ensemble through leap frog algorithm. After equilibration, the molecular dynamics was performed for 20 ns to analyze the trajectory of protein-ligand complex through the system. The RMSD of respective protein-ligand complexes were analyzed using GROMACS in-built analysis tools.

3. Results

3.1 Molecular Modeling, refinement and evaluation of Leptospiral collagenase model

Collagenase plays a pivotal role in the pathogenesis and it is known to be an important virulence factor responsible for the invasion and transmission [13, 22, 28, 33]. Therefore, it is important to develop drugs or inhibitors against Leptospiral collagenases to inhibit the pathogenesis of *Leptospira*. We decided to search a molecular inhibitor for the collagenase to inhibit the pathogenesis of *Leptospira*. However, no crystal structure could be found for Leptospiral collagenase. When we searched sequence of collagenase from *Leptospira interrogans Icterohaemorrhagiae serovar copenhageni* (Fiocruz L1 – 130) against PDB database using different PAM and BLOSUM matrices, we did not find significant similarity with the crystal structure. The sequence similarity and query coverage identified for Leptospiral collagenase was found to be less than 30%, which makes it not suitable for the homology modeling. Therefore, we computationally modeled the structure of Leptospiral collagenase by employing the threading method. Secondary structure of the model was predicted by using PSIPRED online server [31]. The model of Leptospiral collagenase was found to have 24 α helices, 19 β sheets and 50 turns (Figure 1). I-TASSER server [37] was used for threading the molecular structure of Leptospiral collagenase with the crystal structures of collagenase G (2Y3U and 4ARE) from *Clostridium histolyticum* as templates. Based on C-value generated from I-TASSER server, ranking of models was worked out. Model1 which showed the highest C-values was ranked as first. This model was used for further refinement of structure and evaluation. Model1 was subjected to loop building by using SPDB viewer software [16] to bring the amino acids from disallowed region to allowed region in Ramachandran plot. SPDB viewer was also used for energy minimization of the refined model of Leptospiral collagenase using GROMOS96 force field in vacuum. This structure was further re-refined using 3D^{refine} server [5] (Figure 2). The z-score was found to be -9.5 (Figure 3A) which indicated that the model lying in the range of available structures in the PDB that were resolved by X-ray

crystallography. The residue energy was found to be largely negative (Figure 3B) indicating that the model has significantly less error prone regions. The amino acid distribution of Leptospiral collagenase in the Ramachandran plot was assessed by PROCHECK [24]. This analysis revealed that, phi-psi torsion angle for 81.8% of residues of Leptospiral collagenase are in the most favorable region (A, B and L), 5.5 % in additionally allowed region (a, b, l, p), 2.3 % in the generously allowed region (~a,~b,~l,~p) and 0.5% in the disallowed region of the Ramachandran plot (Figure 3C). Verify 3D [8] score was found to be 0.67 (Figure S1), which was greater than zero indicated that the environment profile of the model falls in the satisfactory level. The overall quality of the model was analyzed based on correctly and incorrectly positioned amino acids distribution by ERRAT2 [8]. The results showed that the value of 74.54 (Figure 3D), which indicates that the quality of model is good. The refined predicted structure of Leptospiral collagenase (model1) was superimposed with crystal structures of colG (2Y3U and 2Y50) crystal structures (Figure 4A, B and C). The RMSD values for 2Y3U, 2Y50 and top ranked model1 were found to be 0.99, 1.06 and 0.63, respectively.

3.2 Prediction of ligand binding sites in Leptospiral collagenase

Leptospira is known to secrete variety of extracellular proteases in order to digest the extracellular matrix of the host cells. Among them, collagenase is an important protease that targets and digests collagen which is the most abundant protein in the extracellular matrix of the host cells. Moreover, it is also found to be conserved among the different strains of *Leptospira*. Therefore, collagenase might act as a best drug target for *Leptospira*. In order to work out suitable inhibitors for the Leptospiral collagenase, we predicted ligand binding sites by using Metapocket 2.0 server [17]. From the predicted binding cavities, top three ligand binding sites were chosen, each one resided at N- terminal, catalytic site and C-terminal which were named as A, B and C, respectively (Figure 5).

3.3 Lead identification

From the available information in literature, it was found that Funalenone (ChEBI 65932) [18] could be a potent type-1 microbial collagenase inhibitor. Therefore, we looked for the structurally similar compound to Funalenone in ChEBI and ZINC databases. In ChEBI database, Pinoquercetin (CHEBI 8224) was found to be a structurally similar compound showing 70% similarity. These two molecules were further searched in the ZINC database to evaluate the similarity indices. 65 (4 molecules for Funalenone and 61 for Pinoquercetin) structurally similar compounds were found (Table S1) with 70% and 90% similarity, respectively. Overall 67 (2+4+61) molecules were taken further for docking studies to find out the potential inhibitor for Leptospiral collagenase.

3.4 Docking of validated lead molecules using Autodock vina

In order to find out the best lead molecules, 67 molecules were employed for docking study. The docking program was set to generate the ten best poses for each molecule at each predicted binding site (A, B and C) for the predicted *Leptospira* collagenase structure. After docking, the best poses with the lowest binding energy was chosen for each small molecule. However, the docking analysis of the selected small

molecules revealed that there are variations in their binding energies. Our results indicate that four selected small molecules (Protohypericin, Hypericin, Protopseudohypericin and Pseudohypericin) efficiently bind to the Leptospiral collagenase (Figure 6 & 7) with minimum negative binding energy (kCal/mol) of -8.4, -8.2, -8.0 and -7.9 at A-site; -9.7, -9.6, -8.7 and -8.4 at B-site; -9.6, -9.2, -9.0 and -8.5 at C-site (Table S1). For Funalenone, the binding energy values are -5.9, -6.7 and -8.0 respectively for A, B and C-sites for best binding poses. The surface electrostatic potential surface of *Leptospira* collagenase as shown in figure 7 revealed the ligand binding cavities of superimposed four ligands at each predicted ligands binding pockets (Figure 7). The selected ligands occupied and were found to bind with the surrounding residues in the space between Cys23, Ser30, Asn32, Leu34, Thr39, Ala42, Gln43, Gln46, Gln47, and Glu62 at A-site; Asn481, Gly483, Arg497, Ser502, Ile503, Glu508, Leu509, His 512, Glu544, Ser563, Leu564, Glu567, and Tyr592 at B-site; Thr405, Tyr407, Asp410, Ala696, Phe697, Gly698, Lys838, Leu839, Gly841, Glu842, Leu843, and Leu846 at C-site (Figure 8). Interestingly, one of the predicted binding sites (B-site) was found to be the active site of the Leptospiral collagenase with HEXXH and EXXXE domains of GluZincin superfamily [39]. The docked molecule at the B-site has shown interaction with catalytic residues of Glu544 and Glu548 which are surrounded by His and Glu at 512th and 513th position [40] (Figure 8).

All four selected molecules were subjected to further analysis on the basis of binding energy and drug-like properties (Table S1). The selected molecules were found to be cyclic and possessed significant properties (xlogP, Apolar desolvation (kcal/mol), Polar desolvation (kcal/mol), H-bond donors, H-bond acceptors, net charge, tPSA (Å²), molecular weight (g/mol), and rotatable bonds) at pH 7. After comparative binding energy analysis at the three predicted binding sites, protohypericin was predicted to be the best inhibitor (-9.7 kCal/mol) among the 67 molecules that were screened. All those four lead molecules followed some of the Lipinski's rule of five (Lipinski, 2004) at satisfactory level. Therefore, the rule of five can be bypassed by delivering these molecules through non-oral routes (Dermal, intravenous or pulmonary) [27]. It has been reported that, pulmonary permeability is less sensitive to the polar hydrogen bonding because generally pulmonary drugs have higher polar surface area [41] Therefore, it would be possible to increase the permeability nature of the selected inhibitors to act as drug by modifying their functional groups.

3.5 Pharmacophore Modeling

It is pertinent to mention that Leptospiral collagenase does not possess any crystal structure. Therefore, we used ligand-based Pharmacophore modeling using LigandScout [42]. Pharmacophore (active ligand) was generated based on alignment algorithm by superimposition of top four lead molecules binding at B site. Superimposition of ligands based on conformation of H-bond donors, H-bond acceptors, aromatic, hydrophobic and ionizable groups in 3-D space are shown in figure 9.

3.6 Molecular dynamics of collagenase with Protohypericin, Hypericin, Protopseudohypericin and Pseudohypericin

GROMACS molecular dynamics package was used to validate the stability and interaction of the collagenase with protohypericin, hypericin, protopseudohypericin and pseudohypericin. The lower RMSD values indicate the stability of ligands with collagenase. The RMSD of collagenase – hypericin complex started to fluctuate at 0.24 nm and completed at 1 nm for the span of 20 ns with average of 1 nm RMSD throughout the simulation duration (Figure 10a). The RMSD of collagenase- protohypericin complex started to fluctuate at 0.24 nm and completed at 1.4 nm for the span of 20 ns with average of 1.25 nm throughout the simulation period (Figure 10b). The RMSD of collagenase- protopseudohypericin complex started to fluctuate at 0.24 nm and completed at 1.8 nm for the span of 20 ns with average of 1.5 nm throughout the simulation period (Figure 10c). The RMSD of collagenase-pseudohypericin complex started to fluctuate at 0.24 nm and completed at 1.3 nm for the span of 20 ns with average of 1 nm throughout the simulation period (Figure 10d). The RMSD values of protein – ligand complexes revealed stronger binding of protohypericin, hypericin, protopseudohypericin and pseudohypericin with collagenase. These results are also inconcordance with the docking results of collagenase with protohypericin, hypericin, protopseudohypericin and pseudohypericin ligands.

4. Discussion

Although histotoxicity of bacterial pathogens is mostly dependent on the production of specific toxins, extracellular proteolytic enzymes are thought to play key roles during the initial phase of host colonization; they are known to assist the spread of the pathogen in the host tissues. Among the plethora of proteolytic enzymes, microbial collagenases are strongly linked to bacterial pathogenesis. One third of human skin is composed of type IV collagen. Human blood vessels has extracellular basal lamina, which is mainly composed of extracellular matrix proteins, laminin, fibronectin, heparan sulphate and type IV collagen. Proteolysis results in the dissolution of ECM which results in leakage of blood vessels leading to haemorrhage. Therefore, inhibiting the interaction of pathogen with the host system is an ideal way of inhibiting the pathogenesis of *Leptospira*.

Collagenase is known to play a pivotal role in the pathogenesis [13, 14, 21, 28, 33]. Therefore, it is essential to develop drugs or inhibitors against Leptospiral collagenases so as to inhibit the pathogenesis of *Leptospira*. In the present study, we decided to search a molecular inhibitor for the collagenase in order to inhibit the pathogenesis of *Leptospira*. But, as of now, no crystal structure of *Leptospira* collagenase is available. Therefore, we predicted its molecular structure by using threading methods with crystal structure of collagenase G (2Y3U and 4ARE) of *Clostridium histolyticum* using I-Tasser server. SPDB viewer was used for loop building and model was further refined by 3D^{refine} server. After modeling, the superimposition of Leptospiral collagenase with crystal structure of collagenase G (2Y3U and 4ARE) of *Clostridium histolyticum* was done. The RMSD values for 2Y3U, 2Y50 and initial model (model1) were found to be 0.99, 1.06 and 0.63, respectively. Further, we predicted the ligand binding sites of the Leptospiral collagenase using Metapocket server and subsequently, we screened a molecular inhibitor for collagenase at the three top predicted binding sites. Accordingly, 67 lead molecules were identified against Leptospiral collagenase from Zinc and ChEBI databases (Table S1). These lead molecules were

docked at three predicted binding sites with ten different poses individually. Binding site B (Figure 7) is found to be the best binding site which is present at the active site of the Leptospiral collagenase. Protohypericin, Hypericin, Protopseudohypericin and Pseudohypericin inhibitors were found to have lowest negative binding energy which binds at this site in chronological order. Based on alignment algorithms, these inhibitors were superimposed and pharmacophore model was generated (Figure 9). Molecular dynamics was performed for protohypericin, hypericin, protopseudohypericin and pseudohypericin with Leptospiral collagenase. The RMSD of protein-ligand complexes were found to be relatively low for all the protein-ligand complexes which further validates our docking results. The lower RMSD indirectly proves the stable interaction of ligands with the active site of the collagenase. Thus molecular dynamic simulation shows a stronger binding of protohypericin, hypericin, protopseudohypericin and pseudohypericin with Leptospiral collagenase, which might help in drug discovery for leptospiral infections.

In future, the four inhibitors that have been characterized could be refined by adding additional interactive groups in order to increase the accessibility and binding efficiency at the catalytic site of the Leptospiral collagenase. Based on the pharmacophore model of Leptospiral collagenase, the inhibitors could also be chemically synthesized. The knowledge of the docking studies can be applied in wet lab studies in order to find out the best inhibitor molecule towards developing a drug.

5. Conclusion

In the present study, the model of *Leptospira* collagenase was built by threading method with the crystal structure of collagenase G from *Clostridium histolyticum* and it was used for *in-silico* drug screening at predicted ligand binding sites. Protohypericin has shown as the best inhibitor molecule due to its least minimum binding energy (-9.7 kCal/mol) value and its interaction with the catalytic residues of the collagenase. The average RMSD value of protohypericin and Leptospiral collagenase was 1.25 nm which proves that protohypericin has better binding with the catalytic site of Leptospiral collagenase. Hence, this molecule could be synthesized as a drug to effectively control and manage pathogenesis of *Leptospira*.

Declarations

Funding

This work was supported by Science and Engineering Research Board (SERB), Department of Science and Technology (DST), Govt. of India under Project No. SR-SO-HS-0078-2012 to TJ and SR/SO/HS/0108/2012 to MGM.

Conflict of interest

The authors declare that they have no conflict of Interests.

Availability of data and material

Not Applicable

Code availability

Not Applicable

Author's contribution

V.K and T.J conceived and designed the research problem. V.K, N.S and V.S.A designed and performed experiments. V.K, N.S, V.S.A and T.J. analyzed the experiment results. T.J, P.K.S, M.G.D and M.K wrote the manuscript. All authors have reviewed and approved the manuscript as submitted.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors

Consent to participate

Not Applicable

Consent for publication

Not Applicable

Acknowledgement

We also thankful to are thankful to Lady Tata memorial Trust, Bharathidasan University and DST-PURSE for providing HPC computational facility and Regional Medical Research Centre (ICMR) in Port Blair, Andaman & Nicobar Islands are gratefully acknowledged.

References

1. Abraham MJ, Murtola T, Schulz R, et al (2015) Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* 1–2:19–25. doi: 10.1016/j.softx.2015.06.001
2. Awad MM, Ellemor DM, Bryant AE, et al (2000) Construction and virulence testing of a collagenase mutant of *Clostridium perfringens*. *Microb Pathog* 28:107–117. doi: 10.1006/mpat.1999.0328
3. Barocchi MA, Ko AI, Galvão Reis M, et al (2002) Rapid translocation of polarized MDCK cell monolayers by *Leptospira interrogans*, an invasive but nonintracellular pathogen. *Infect Immun* 70:6926–6932. doi: 10.1128/IAI.70.12.6926-6932.2002

4. Bharti AR, Nally JE, Ricaldi JN, et al (2003) Leptospirosis: A zoonotic disease of global importance. *Lancet Infect. Dis.* 3:757–771
5. Bhattacharya D, Cheng J (2013) 3Drefine: Consistent protein structure refinement by optimizing hydrogen bonding network and atomic-level energy minimization. *Proteins Struct Funct Bioinforma* 81:119–131. doi: 10.1002/prot.24167
6. Bonvin AMJJ, Mark AE, Van Gunsteren WF (2000) GROMOS96 benchmarks for molecular simulation. *Comput Phys Commun* 128:550–557. doi: 10.1016/S0010-4655(99)00540-8
7. Caimano MJ, Sivasankaran SK, Allard A, et al (2014) A Model System for Studying the Transcriptomic and Physiological Changes Associated with Mammalian Host-Adaptation by *Leptospira interrogans* Serovar Copenhageni. *PLoS Pathog* 10:. doi: 10.1371/journal.ppat.1004004
8. Colovos C, Yeates TO (1993) Verification of protein structures: Patterns of nonbonded atomic interactions. *Protein Sci* 2:1511–1519. doi: 10.1002/pro.5560020916
9. Daher E, Zanetta DMT, Cavalcante MB, Abdulkader RCRM (1999) Risk factors for death and changing patterns in leptospirosis acute renal failure. *Am J Trop Med Hyg* 61:630–634. doi: 10.4269/ajtmh.1999.61.630
10. De Paepe A, Malfait F (2004) Bleeding and bruising in patients with Ehlers-Danlos syndrome and other collagen vascular disorders. *Br. J. Haematol.* 127:491–500
11. Degtyarenko K, De matos P, Ennis M, et al (2008) ChEBI: A database and ontology for chemical entities of biological interest. *Nucleic Acids Res* 36:. doi: 10.1093/nar/gkm791
12. DeLano WL (2002) Pymol: An open-source molecular graphics tool. *CCP4 Newsl Protein Crystallogr*
13. Diener B, Carrick L, Berk RS (1973) In vivo studies with collagenase from *Pseudomonas aeruginosa*. *Infect Immun* 7:212–217. doi: 10.1128/iai.7.2.212-217.1973
14. Esther Schönauer, Andreas M Kany, Jörg Haupenthal, Kristina Hüsecken, Isabel J Hoppe, Katrin Voos, Samir Yahiaoui, Brigitta Elsässer, Christian Ducho, Hans Brandstetter, Rolf W Hartmann (2017) Discovery of a Potent Inhibitor Class with High Selectivity toward Clostridial Collagenases. *J Am Chem Soc* 139:12696-12703. doi: 10.1021/jacs.7b06935.
15. Gouveia EL, Metcalfe J, De Carvalho ALF, et al (2008) Leptospirosis-associated severe pulmonary hemorrhagic syndrome, Salvador, Brazil. *Emerg Infect Dis* 14:505–508. doi: 10.3201/eid1403.071064
16. Guex N, Peitsch MC, Schwede T (2009) Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophoresis* 30:. doi: 10.1002/elps.200900140
17. Huang B (2009) Metapocket: A meta approach to improve protein ligand binding site prediction. *Omi A J Integr Biol.* doi: 10.1089/omi.2009.0045
18. Inokoshi J, Shiomi K, Masuma R, et al (2010) ChemInform Abstract: Funalenone, a Novel Collagenase Inhibitor Produced by *Aspergillus Niger*. *ChemInform.* doi: 10.1002/chin.200020197

19. Irwin JJ, Shoichet BK (2005) ZINC - A free database of commercially available compounds for virtual screening. *J Chem Inf Model*. doi: 10.1021/ci049714+
20. Janwitthayanan W, Keelawat S, Payungporn S, et al (2013) In vivo gene expression and immunoreactivity of *Leptospira* collagenase. *Microbiol Res*. doi: 10.1016/j.micres.2012.12.005
21. Jelena Konstantinović, Samir Yahiaoui¹, Alaa Alhayek, Jörg Haupenthal, Esther Schönauer, Anastasia Andreas, Andreas M Kany, Rolf Müller, Jesko Koehnke, Fabian K Berger, Markus Bischoff, Rolf W Hartmann, Hans Brandstetter, Anna K H Hirsch (2020) *N*-Aryl-3-mercaptosuccinimides as Antivirulence Agents Targeting *Pseudomonas aeruginosa* Elastase and *Clostridium* Collagenases. *J Med Chem* 63:8359-8368. doi: 10.1021/acs.jmedchem.0c00584.
22. Kassegne K, Hu W, Ojcius DM, et al (2014) Identification of collagenase as a critical virulence factor for invasiveness and transmission of pathogenic leptospira species. *J Infect Dis* 209:1105–1115. doi: 10.1093/infdis/jit659
23. Kühn K (1995) Basement membrane (type IV) collagen. *Matrix Biol*. 14:439–445
24. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr* 26:283–291. doi: 10.1107/s0021889892009944
25. Laskowski RA, MacArthur MW, Thornton JM (2012) PROCHECK: validation of protein-structure coordinates
26. Levett PN (2001) Leptospirosis. *Clin Microbiol Rev* 14:296–326
27. Lipinski CA (2004) Lead- and drug-like compounds: The rule-of-five revolution. *Drug Discov. Today Technol*. 1:337–341
28. Lyerly DM, Kreger AS (1983) Importance of serratia protease in the pathogenesis of experimental *Serratia marcescens* pneumonia. *Infect Immun* 40:113–119. doi: 10.1128/iai.40.1.113-119.1983
29. Magrane M, Consortium UP (2011) UniProt Knowledgebase: A hub of integrated protein data. *Database*. doi: 10.1093/database/bar009
30. Matsushita O, Koide T, Kobayashi R, et al (2001) Substrate Recognition by the Collagen-binding Domain of *Clostridium histolyticum* Class I Collagenase. *J Biol Chem* 276:8761–8770. doi: 10.1074/jbc.M003450200
31. McGuffin LJ, Bryson K, Jones DT (2000) The PSIPRED protein structure prediction server. *Bioinformatics*. doi: 10.1093/bioinformatics/16.4.404
32. Miyoshi SI, Nitanda Y, Fujii K, et al (2008) Differential gene expression and extracellular secretion of the collagenolytic enzymes by the pathogen *Vibrio parahaemolyticus*. *FEMS Microbiol Lett* 283:176–181. doi: 10.1111/j.1574-6968.2008.01159.x
33. Moncrief JS, Obiso R, Barroso LA, et al (1995) The enterotoxin of *Bacteroides fragilis* is a metalloprotease. *Infect Immun* 63:175–181. doi: 10.1128/iai.63.1.175-181.1995
34. Nascimento ALTO, Ko AI, Martins EAL, et al (2004) Comparative Genomics of Two *Leptospira interrogans* Serovars Reveals Novel Insights into Physiology and Pathogenesis. *J Bacteriol*

- 186:2164–2172. doi: 10.1128/JB.186.7.2164-2172.2004
35. Trott O, Olson AJ (2010) Software news and update AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem*. doi: 10.1002/jcc.21334
 36. Steffen C, Thomas K, Huniar U, et al (2010) Software news and updates TmoleX-a graphical user interface for TURBOMOLE. *J Comput Chem* 31:2967–2970. doi: 10.1002/jcc.21576
 37. Yang J, Zhang Y (2015) I-TASSER server: New development for protein structure and function predictions. *Nucleic Acids Res* 43:W174–W181. doi: 10.1093/nar/gkv342
 38. Schüttelkopf AW, Van Aalten DMF (2004) PRODRG: A tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr Sect D Biol Crystallogr*. doi: 10.1107/S0907444904011679
 39. Barret AJ, Rawlings ND (1995) Families and clans of serine peptidases. *Arch Biochem Biophys* 318:247-50. doi: 10.1006/abbi.1995.1227.
 40. Ulrich E, Esther S, Dorota N, Hans B (2011) Structure of collagenase G reveals a chew-and-digest mechanism of bacterial collagenolysis *Nat Struct Mol Biol* 18:1109-1114. doi: 10.1038/nsmb.2127.
 41. Tronde A, Nordén B, Marchner H, et al (2003) Pulmonary absorption rate and bioavailability of drugs in vivo in rats: Structure-absorption relationships and physicochemical profiling of inhaled drugs. *J Pharm Sci*. doi: 10.1002/jps.10386
 42. Wolber G, Langer T (2005) LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. *J Chem Inf Model* 45:160–169. doi: 10.1021/ci049885e

Figures

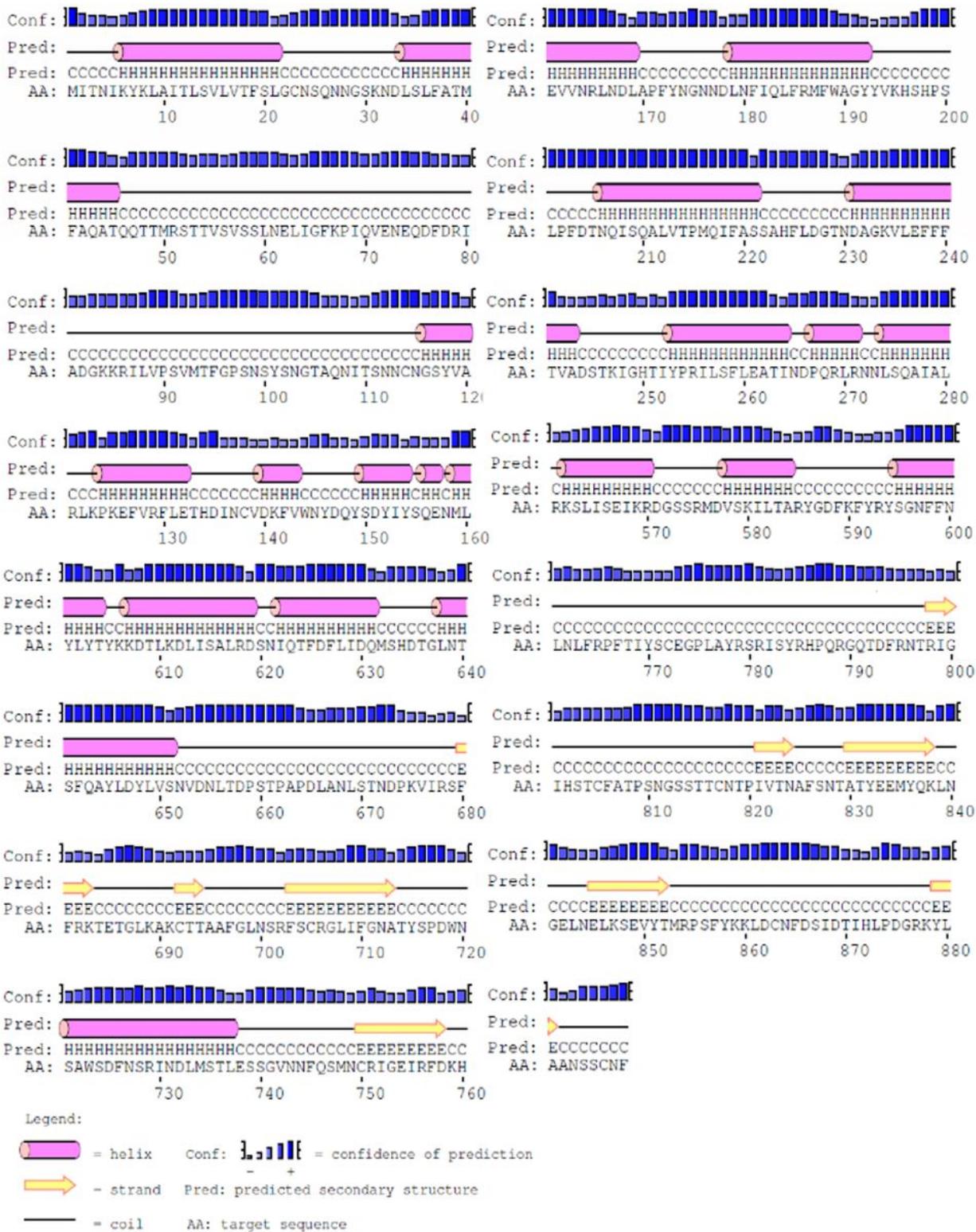


Figure 1

Secondary structure of Leptospiral collagenase: Secondary structure of Leptospiral collagenase showing helix (pink), strand (yellow) and coil regions (black) at the corresponding amino acid positions.

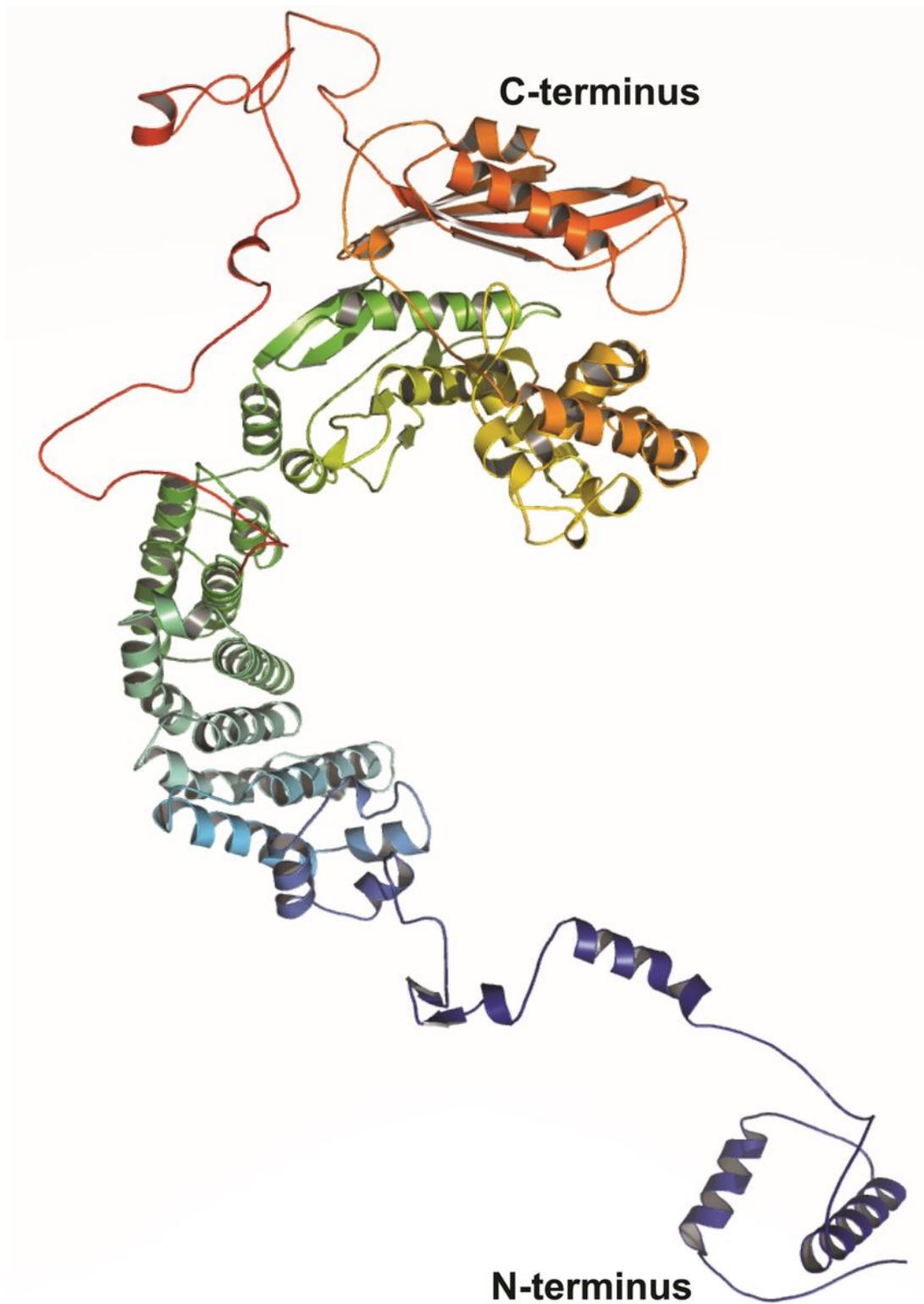


Figure 2

Structural model of Leptospiral collagenase: Ribbon model of Leptospiral collagenase of *Leptospira interrogans* Icterohaemorrhagiae serovar copenhageni (Fiocruz L1 – 130).

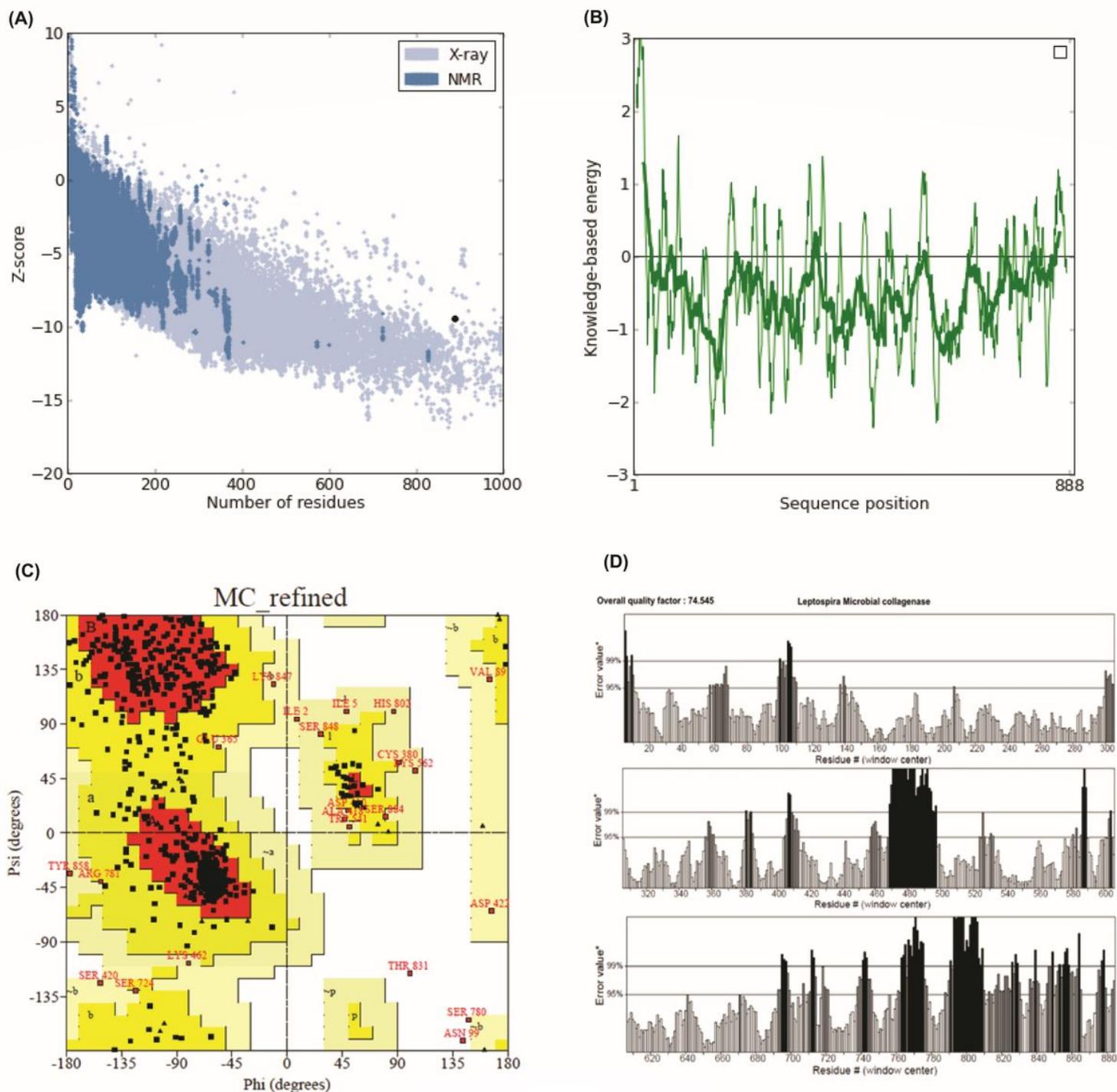


Figure 3

Model quality evaluation: (A) Z- score of Leptospiral collagenase showing Z-score and amino acids residues on X and Y axis respectively. The black dot represents position of the model lying in the range of available structures which are solved by X-ray crystallography in the PDB. (B) Energy plot of Leptospiral collagenase. X axis indicates knowledge based on energy and Y axis indicates position of amino acids in a sequence. The residue energy is found to be largely negative, which shows that the model has less error during the modeling.

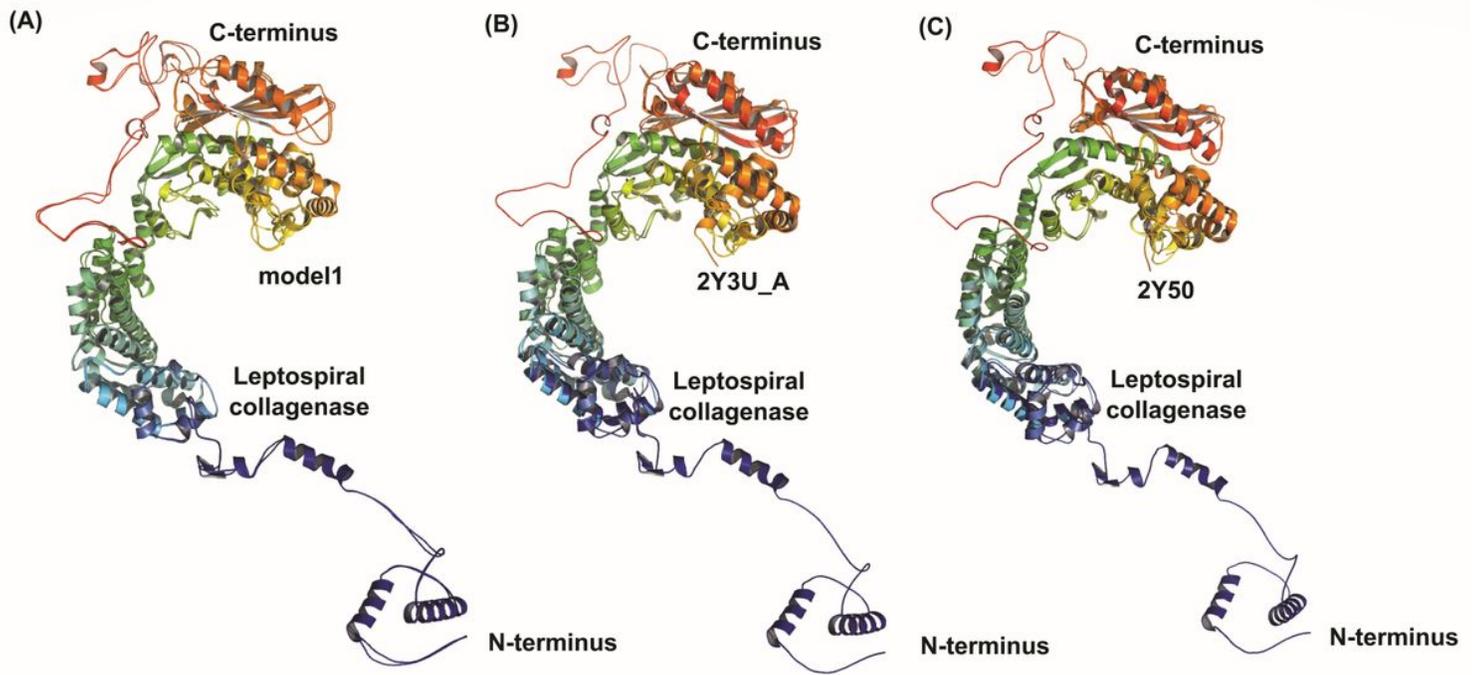


Figure 4

Superimposition of Leptospiral collagenase structures: (A) Superimposition (RMSD value of 0.63) of finally refined Leptospiral collagenase to the model1 template of Leptospiral collagenase obtained from I-tasser server after threading. (B) Superimposition (RMSD value of 0.99) of Leptospiral collagenase to the A chain of collagenase G (PDB ID - 2Y3U) of *Clostridium histolyticum*. (C) Superimposition (RMSD value of 1.06) of Leptospiral collagenase to the crystal structure *Clostridium histolyticum* Collagenase (PDB ID - 2Y50).

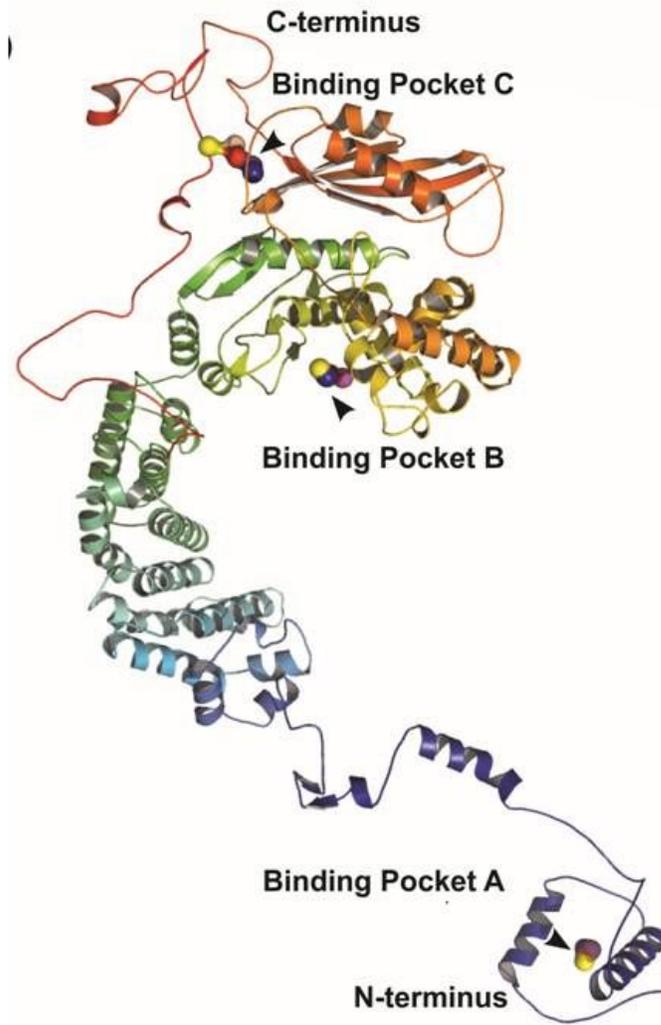


Figure 5

Domains of *Leptospira* collagenase: (B) *Leptospira* collagenase showing ligand binding sites marked with arrow at N-terminus (A site), active site (B site) and C-terminus (C site).

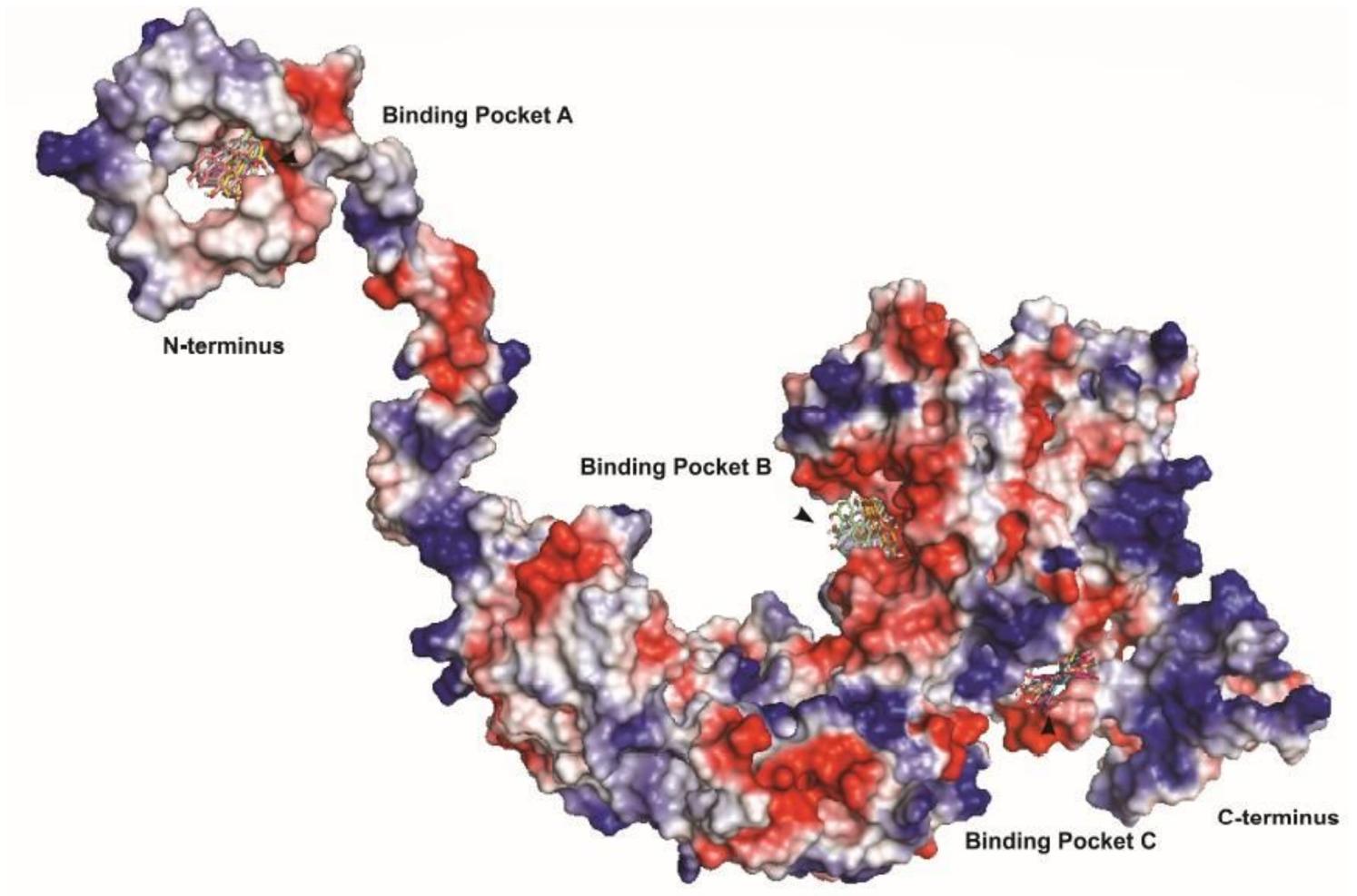


Figure 6

Electrostatic surface potential of Leptospiral collagenase showing ligand binding cavities of four ligands at each predicted ligand binding pockets.

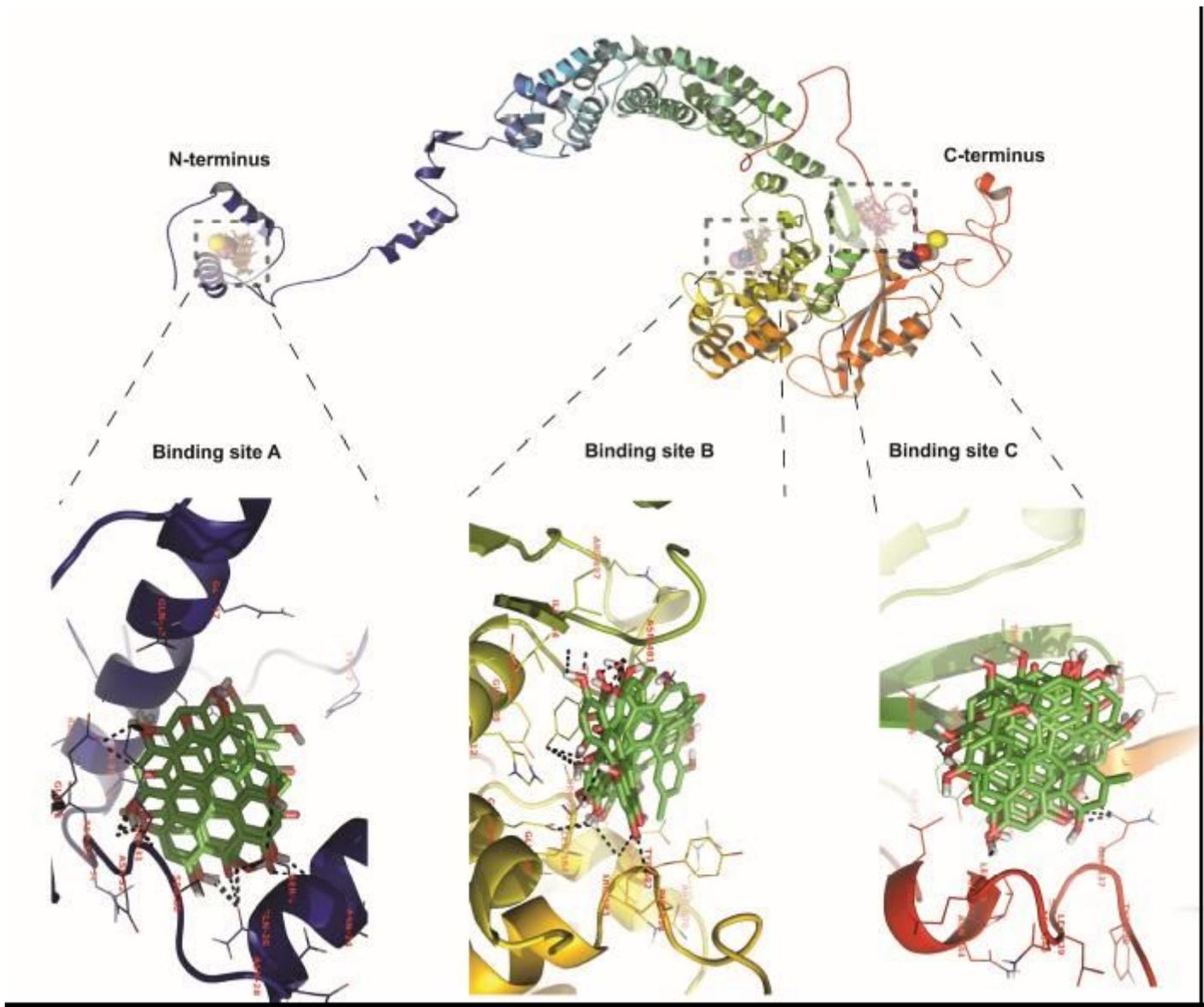


Figure 7

Binding poses of four ligands at three binding sites of Leptospiral collagenase. Model showing binding of superimposed Leptospiral collagenase inhibitors (Protohypericin, Hypericin, Protopseudohypericin and Pseudohypericin) at the binding sites A, B and C.

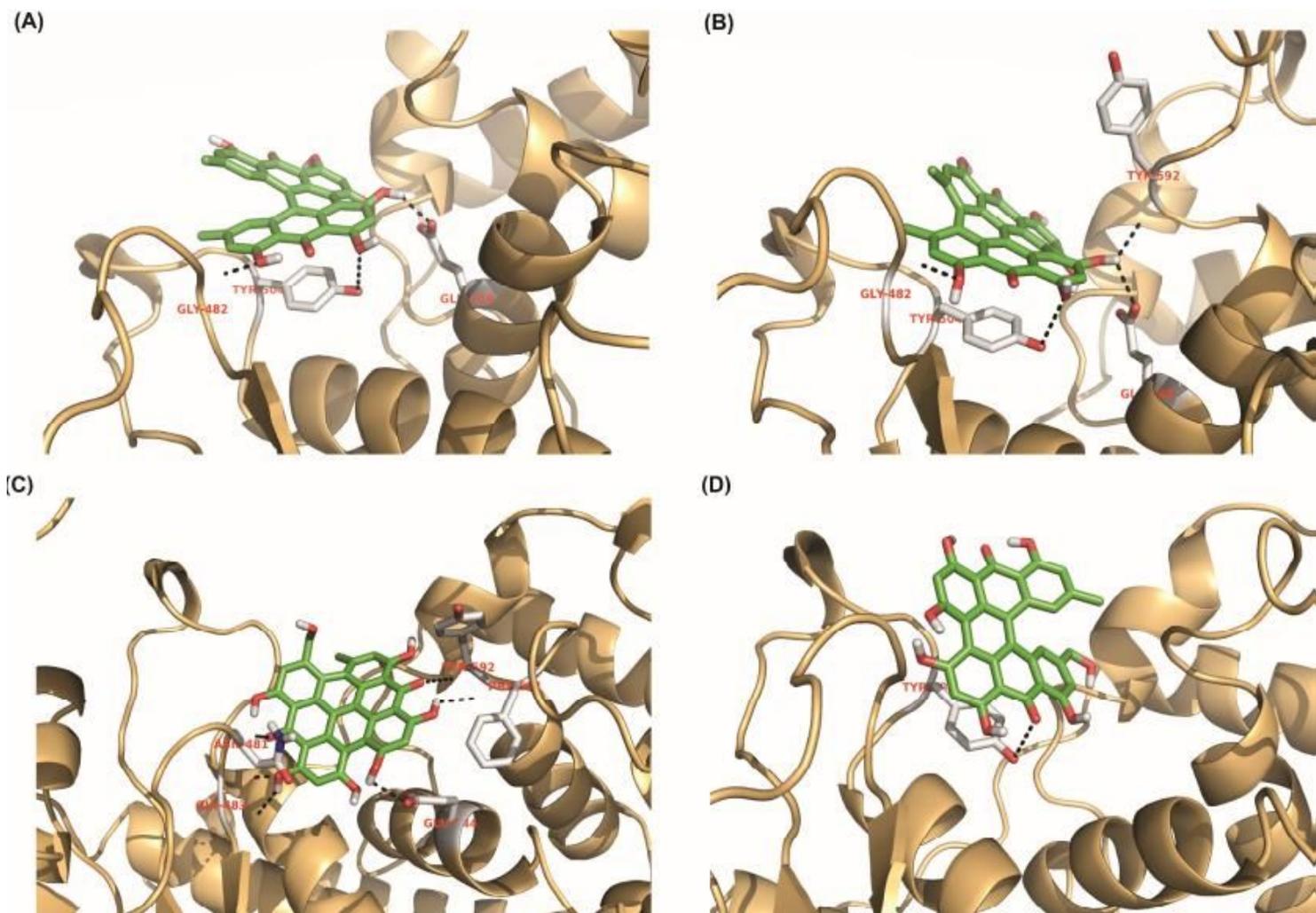


Figure 8

Binding poses of four top ligands at the binding site B of Leptospiral collagenase: Model showing the binding of four inhibitors Protohypericin (A) Hypericin (B) Protopseudohypericin (C) and Pseudohypericin (D) at the binding site B.

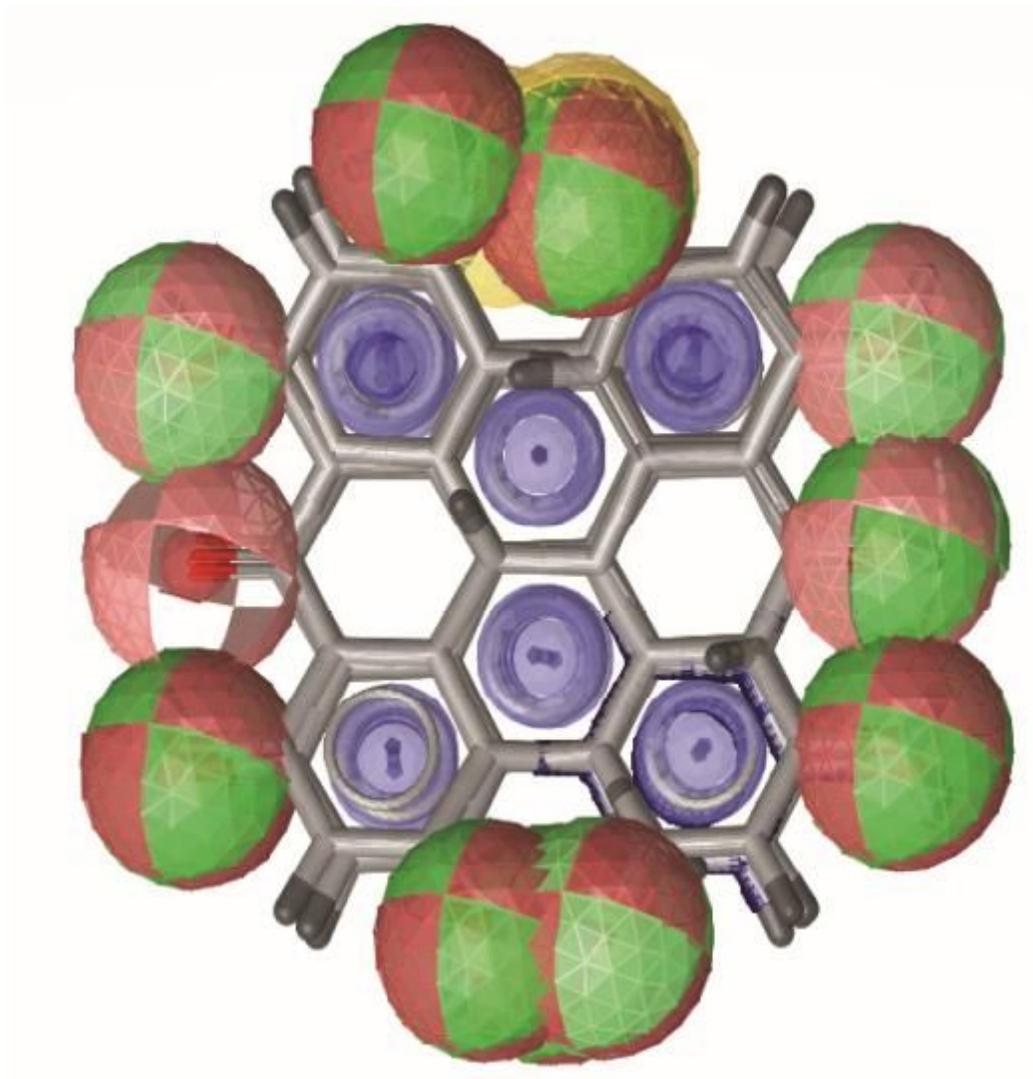


Figure 9

Pictorial representation of pharmacophore model of Leptospiral collagenase: Pharmacophore model showing superimposed conformational structure of Leptospiral collagenase inhibitors. Red, yellow, green vectors indicate electron donor group, hydrophobic features and hydrogen bonding domain features respectively.

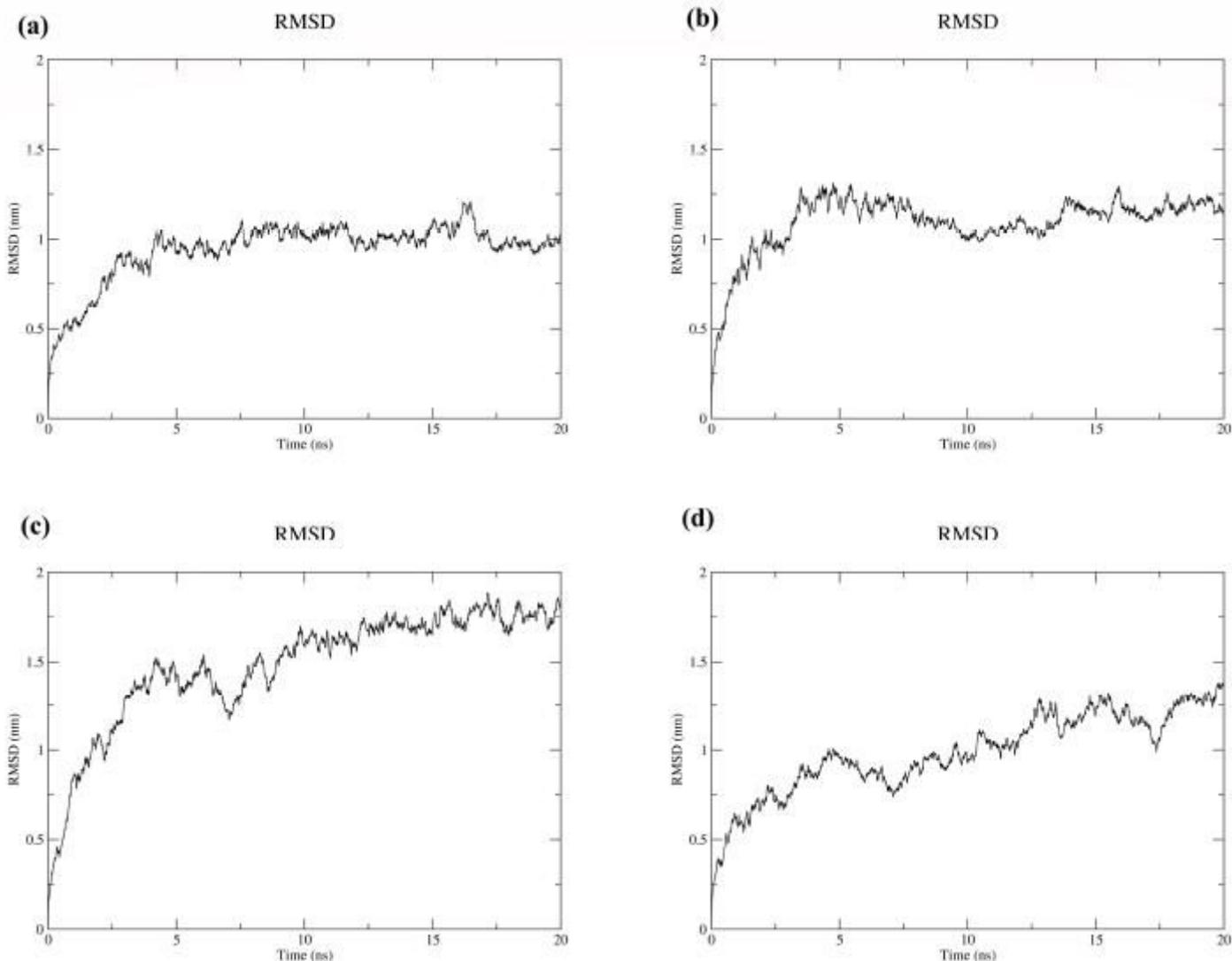


Figure 10

Molecular dynamics simulation study of ligand – protein complex for the time duration of 20ns. (A.) RMSD of docked complex hypericin with Leptospiral collagenase, (B). RMSD of docked complex protohypericin with Leptospiral collagenase, (C). RMSD of docked complex protopsuedohypericin with Leptospiral collagenase and (D). RMSD of docked complex pseudohypericin with Leptospiral collagenase.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [GraphicalAbstract.pdf](#)
- [FigureS1.jpg](#)
- [TableS1.xls](#)