IJBAI
International Journal of Bioinformatics and Artificial Intelligence
ISSN: XXXX-XXXX
Vol.1 No.1January-June 2026,
Pages 01-37

Author: [1] **Hridoy Ranjan Bairagya,**

Affiliation:
[1] *Maulana Abul Kalam Azad University of Technology, West Bengal, India*

Corresponding author:
hbairagya@gmail.com.
Doi: 10.32332/ijbai3i2.9950

Dates:
Received 15 December, 2025
Revised 20 December, 2025
Accepted 29 March, 2025
Published 15 Jun, 2026

Read Online:

Scan this QR code with your mobile device or smart phone to read online

Conformational Transition of the Mpro Enzyme: A Computational Approaches for Predicting Alternative Binding Sites of an Anti-COVID Molecule

Abstract : The Mpro enzyme has gained popularity as a target in the life cycle of COVID-19 for the discovery of anti-COVID molecules. The current hypothesis focuses on seventy-two crystal structures of Mpro, both in Conformation A (ConA) and Conformation AB (ConAB) forms. However, no studies have yet evaluated the following aspects: (i) a comparative analysis of ligand-bound crystal structures versus their ligand-free counterparts in both ConA and ConAB forms; and (ii) the identification of alternative binding sites for anti-COVID molecules within the crystal structure of Mpro. The native state is more dynamic than the ligand-bound form, stabilizing after binding to the corresponding ligand. Moreover, Asn142 and Leu141 in ConAB-native, Asn277 and Thr304 in ConAB ligand-bound, Phe305 and Asn72 in ConA native and Ser46 and Glu47 in ConA ligand-bound structures are primarily flexible. During the transition from native to ligand-bound form, Val73, Lys100, Ser123, Cys128, Lys137, Cys156 and Phe294 residues have changed their structural position from buried to exposed in the ConA and Leu50, Arg60, Asn214, Ala285 and Phe305 in the ConAB, respectively. Similarly, Phe305 changed its structural conformation from exposure to burial in the ConA. Investigation on the multi-conformation analysis of 72 crystal structures highlighted that the residues His41, Glu189, Asn142 and Arg188 (apart from His163, Glu166 and Gln189) might act as a catalytic partner along with the Cys-His catalytic dyad and they may be considered alternative binding site of the anti-COVID molecule of the Mpro enzyme.

Key Words: *SARS-CoV-2; Mpro crystal structures; Multi-conformations analysis; Alternative binding sites of Mpro; Anti-COVID molecules*

IJBAI
International Journal of Bioinformatics and Artificial Intelligence
ISSN: XXXX-XXXX
Vol.1 No.1January-June 2026,
Pages 01-37

Introduction

A novel coronavirus was discovered at the end of 2019 when few patients were found with pneumonia in the airway of epithelial cells and the cause was unknown [1]. Covid-19 is linked to the newly found pathogen known as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The SARS-CoV-2 is the third notable coronavirus of the modern era. To reduce the globalization of this certain disease, quarantine and isolation were put in place in many countries around the globe. Experts have encouraged the production of efficient antiviral medication to stop this disease [1,2]. The corona-viridae family, corona-virinae subfamily and nidovirales order are from where CoV got started. The corona-virinae has been classified into four genera that are α-, β-, γ- and δ-CoV. These four genera contain viruses with major veterinary and medicinal importance. Global, social, political and economic unrest are profoundly affected by Covid-19 disease. In that present situation, the application of repurposed medications and care are necessary options for patient therapy [3]. Coronaviruses mainly responsible for mild to moderate upper-respiratory tract infection in humans, in which the common cold is consistent. Various variations of the SARS-CoV-2 HCoV-229E (Alphacoronavirus genera), HCoV-NL63 (Human novel coronavirus), HCoV-HKU1 (Betacoronavirus hongkonense), HCoV-OC43 (Betacoronavirus species), severe acute respiratory syndrome and MERS (Betacoronavirus genera) have been identified after December 2019, when initial Covid-19 infection began [4]. There are different characteristics of this virus, so the delta form was categorized as Variant of Concern (VOC) in 2021. The VOC is the major type of variant that causes higher transmissibility and leads to decreased treatment efficiency and other important aspects, as per the Centers for Disease control and prevention. The mutated SARS-CoV-2 variant that is also known as the omicron variant (B.1.1.529), is the most severe and the World Health Organization (WHO) has now classified it as a VOC in 2021 [5]. The omicron form is growing more quickly and spread to among all the nations. The omicron variant has been transmitted majorly in different places till 2021 [6].

1.1. Biochemical mechanism of Mpro enzyme:

SARS-CoV-2 genome is approximately 30,000 nucleotides and is notable that the SARS-CoV 2 genome encodes four structural and two overlapping poly-proteins, providing an important function in host invasion and maintenance of the life-cycle of the Covid-virus [7]. Two self encoded cysteine proteases, papain-like protease (nsp3) and the main protease (nsp5), have the capability to break down these polyproteins into Non-Structural Proteins (NSPS). The Main Protease (Mpro) is also referred to as 3C-like proteinase. The Mpro processing essential stages are just before viral assembly and maturation. In substrates, the amino acids are numbered mainly as P4-P3-P2-P1 and P1'-P2'-P3' particularly from the N-terminus to the C terminus of the main protease became a conserved protein in all coronaviruses [8]. The site of cleavage is located between P1 and P1' positions and residue of glutamine is needed at the P1 position [9]. In a similar study, SARS-CoV-2 Mpro exhibited a profound interest in catalyzing the process of breakdown of the polypeptide Ac-Val-Lys-Leu-Gln-ACC.

In the mechanism pathway at starting point, a single proton is transported from Cys145 to His41, which is a catalytic dyad and at the very same time, the sulphur atom of Cys145 shows engagement in a nucleophilic attack on the carbonyl carbon of the peptide bond to

IJBAI
International Journal of Bioinformatics and Artificial Intelligence
ISSN: XXXX-XXXX
Vol.1 No.1January-June 2026,
Pages 01-37

obtain an intermediate that is Thio-hemi-ketal [10-13]. The acyl-enzyme complex intermediate produced by the proton transfer from the residue His41 to the atom (nitrogen) of the substrate affects the breakage of the peptide bonds of the polyprotein. This initiates the release of ACC. The initial ACC released from the active site of the enzyme is produced by this event to yield the first product. By the release of ACC, a proton is at the same time transferred to His41 and

an activated water molecule attacks particularly the carbonyl carbon atom of residue Gln5 in the peptide and then the second kind of product is obtained by this mechanism [14].

1.2. Structural and functional role different domains in Mpro enzyme

The active site of Mpro enzyme is present at a shallow crevice in the inter-domain region between I (res. Id. 10-99) and II (res. Id. 100-182). The catalytic residue Cys145 belongs to the oxyanion loop, which forms an S-shaped loop (Gly138-Gly146) in particularly domain II and the amide nitrogen of Cys145 and Gly143 link together to produce the oxyanion hole, which binds the carbonyl group of the peptide bond (scissile) present in the substrates. A β-strand segment (His163-Pro168), adjacent to the oxyanion loop, connects with other residues that play a major role in binding with the substrate (e.g., His163, Glu166). Later, they are close to a loop segment (Gly183-Ala193) situated at the starting of the domain II and III (res. Id. 198-303) linker that plays a role in contributing to make the border of the location of the enzyme's active site. However, during the activation of the nucleophile, the catalytic dyad that contains Cys145 and His41 are assumed to act as a base.

His41 is a kind of residue that belongs to a small helix that is located within a long and highly helical connection loop present in domain I. The positioning of His41 is governed by a particular network of H-bonds. This network includes an interaction between His41-Asp187, initiated by a conserved water molecule and a salt bridge between the Arg40 guanidinium and the Asp187 carboxylate at the linker of domain II and III of the enzymes. There are numerous inter-domains non polar interactions occur among Pro39, Met49 and Leu27, specifically in domain I, as well as His164 and Met165 in domain II, that help to constitute a hydrophobic pocket at site S2 to bind the peptide substrate, leading to complete the formation of the active site2. Basically, domain III is involved in the dimerization process of the protease. Therefore, the resulting homo dimer is assumed to be the active form of the enzyme. Removal of fragments is a vital role of domain III in the Mpro dimerization has been identified. When the catalytic activity of this truncated form of the enzyme was performed against a 14-mer peptide substrate, it was very weak because not more than 20% of the substrate was fragmented after 20h of the experimental period. However, serial truncation experimental data have given insights that in deletion of residues, particularly Gln299 and Arg298, lowers both dimerization and enzymatic activity mainly (1-2%) indicates that the last C-terminal helix in domain III is required for dimerization as well as enzymatic activity.

The N-finger, composed of residues 1-7 in domain I, is a crucial structural component of main protease needed for the activity of SARS-CoV Mpro. In the interaction with a 12-mer peptide, the complete loss of the N-finger has little effect in the case of dimerization, but it eliminates the enzyme's activity. It was discovered in another study that the first three residues of the N

finger have a minor impact on the activity and stability of the enzyme. It has been

IJBAI
International Journal of Bioinformatics and Artificial Intelligence
ISSN: XXXX-XXXX
Vol.1 No.1January-June 2026,
Pages 01-37

reported that 76% of the catalytic activity is conserved during the deletion of the first three residues of N-finger, but the KD (dissociation constant) value for the dimer dissociation increases from 0.28-3.40M. Furthermore, the serial truncation experimental sights have shown that substantial change in activity and structure of enzyme during deletion of Arg4. Crystal structures have clarified the connections between enzymatic activity and dimerization. It has been seen that the N-finger of one protomer is squeezed between its counterpart's domains I and III. This condition enables the substrate-specificity pocket at the S1 site and the oxyanion loop in the next protomer to be shaped by Ser1 and Arg4 from one protomer. Curiously, the oxyanion loop is partially turned as a 310-helix in some monomers (mutants S139A, G11A and R298A) of crystal structures. This prevents access to the active site, preferably to the S1 subsite of the enzyme. Therefore, for both protein dimerization and proper placement of N-finger residues, open, active site is required [2].

1.3. Mpro: outstanding pharmacological targets for developing drugs and vaccines

The Mpro enzyme has bec3ome a more popular target for the discovery of antiviral medications due to its functional effectively in the lifecycle of virus. Mpro has a highly conserved and lack of human homologs, which made it a target among clinical variants. With the usage and advancement of computational modeling, development and efforts focusing on drugs are becoming more effective. However, it depends on numerous variables, such as the geometry of the ligand and protein, protonation sate, chemical reactions, hydration, quantum effects and some other limitations.

1.4. Current status of computational drug development of Mpro enzyme

Computer-Aided Drug Design (CADD) has showed significant potential in discovering lead compounds from traditional medicinal herbs, therefore bridging the gap between traditional and contemporary medicine. Various compounds found through virtual screening and drug repurposing utilizing computational techniques have been created to inhibit the active site of the Mpro enzyme. Furthermore, integrating Molecular Dynamics (MD) simulations, Artificial Intelligence (AI) and other computational tools improves medication efficacy prediction and emphasizes the significance of in vivo investigations to confirm findings. A computer screen of 227 terpenes compounds indicated that six of them, particularly erylosides B, exhibited a significant inhibitory potential against Mpro, with a binding affinity up to -51.9 kcal/mol compared to -33.6 kcal/mol for Lopinavir based on a 100 ns MD simulation. Remdesivir's binding energy against the SARS-CoV-2 Mpro (PDB Id: 6Y84) is -9.85 kcal/mol [15]. Astilbin and Engeletin have showed binding energy against the SARS-CoV-2 Mpro is -15.1 and -14.5 kcal/mol [16]. The in-silico analysis revealed a molecule, schaftoside from Eleusineindica, which demonstrated stronger binding affinity (-8.7 kcal/mol) than N3(N-[(5- METHYLISOXAZOL-3-YL)CARBONYL]ALANYL-L-VALYL-N~1~-((1R,2Z)-4-

(BENZYLOXY)-4-OXO-1-{[(3R)-2-OXOPYRROLIDIN-3-YL]METHYL}BUT-2-ENYL)-L LEUCINAMIDE); 7.8 kcal/mol) [17].

To determine the therapeutic potential of four different classes of newly synthesized imidazole derivatives against SARS-CoV-2 Mpro, a thorough investigation was carried out.

IJBAI
International Journal of Bioinformatics and Artificial Intelligence
ISSN: XXXX-XXXX
Vol.1 No.1January-June 2026,
Pages 01-37

According to the findings, these compounds had exceptional binding affinities and provide the stability with the target proteins. It is noteworthy that all compounds showed significant binding affinities for Mpro, with imidazolyl–methanone C11 andimidazolyl–methanone C12 showing the higher affinity [18]. The molecular docking results showed that following phytochemicals like koenigicine, mukonicine, o-methylmurrayamine A, koenine and girinimbine had higher binding affinities and lower inhibition constants than the reference inhibitor (3WL; Baicalein). The compound Koenine, o-methylmurrayamine A, mukonicine and girinimbine from Murrayakoenigii, like other bioactive natural compounds, present potential hits that may be further structurally changed and investigated in vitro and in vivo for the identification of new Mpro inhibitors [19]. Another study looked at the protease Mpro, one of the main molecular targets of SARS-CoV-2. The PDB Id. 7TOB crystal structure was a crucial target in their in-silico molecular repositioning experiments with a variety of biological functions. Following the experiment of the in-silico data, it has been concluded that Ensitrelvir (binding energy: S-217622−7.3kcal/mol) with substantial promise as antiviral agents against SARS-CoV-2 (Table 1) [20].

Table 1: *Proposed inhibitors with binding energy (kcal/mol) for the Mpro enzyme predicted by various computational methods*.

Type	Name of the proposed inhibitors	Binding energy (kcal/mol)	Type of computational method	References
I.	Erylosides B, Lopinavir, Remdesivir, Astilbin, Engeletin	-10.00 <	MD simulation	[15,16]
II.	Schaftoside, 2'-O-methylcajanone, Isoschaftoside, Cajaflavonone,Thiophenyl–imidazole C1,Thiophenyl–imidazole C2, Thiophenyl–imidazole C3 Pyridyl–imidazole C5, Imidazolyl–methanones C 11, Imidazolyl–methanones C 12, Imidazolyl–methanones C13	-8.99 to -8.00	Molecular docking	[17,18]
III.	Betacyanin, Isorhamnetin, Betavulgarin, Pinostrobin, Cajanone, N3, Mukonicine, O-Methylmurrayamine A, Koenine, Girinimbine, Ensitrelvir (S-217622), Thiophenyl–imidazole C4, Pyridyl–imidazole C7, Pyridyl–imidazole C8, Pyridyl–imidazole C9, Quinoline–imidazole C14, Quinoline–imidazole C15, Quinoline–imidazole C16, Quinoline–imidazole C17, Quinoline– imidazole C18			
IV.	Koenigicine, Ribalinidine, Cajanol, Betagarin, Longistyline, Indicaxanthin, Betaxanthin, (E)–Beta ocimene, Atazanavir (BMS-232632)			
V.	Spartein, Vulgaxanthin-I, Vulgaxnthin II, Juliflorine			

IJBAI
International Journal of Bioinformatics and Artificial Intelligence
ISSN: XXXX-XXXX
Vol.1 No.1January-June 2026, Pages 01-37

-7.99 to -7.00 Molecular docking [17-20] -6.99 to -6.00 Molecular docking [17-20] -5.99 to -5.00 Molecular docking [17]

VI. Hordenine -4.99 to -4.00 Molecular docking [17]

VII. Tetracosanoicacid, Hydoxygenistein, cis-Methyl 11-eicoscnoate, (Z)-Hetriacontane, Hentriacontanol, Octadecadienoicacid, Dotetracont-15-en-9-ol, Methyl-18-methylnonadecanoate, 2-Lachnophyllum ester, Methyl stearate, Oxalate -3.99 to 3.00 Molecular docking [17]

1.5. Current status of experimental development of inhibitors for Mpro enzyme

The experimental drug development involves the utilization of small molecule to treat a range of infections and illnesses. The anti-Mpro properties of few compounds from High Throughput Screening (HTS) have been experimentally analyzed using a combination of the Fluorescence Polarization (FP) and Fluorescence Resonance Energy Transfer (FRET) assays.

Yan et al. suggested dieckolmay act as a new competitive inhibitor against SARS-CoV-2 with an IC50 value of 4.5 μM by integrating fluorescence polarization technology with a biotin– avidin system in a unique screening strategy [21]. Similarly, evaluation based on Fluorescence Resonance Energy Transfer (FRET) revealed that aloes in inhibited SARS-CoV-2 [22]. Another example of a repurposed medication that is recognized as a strong inhibitor of other coronaviruses was the 5-hydroxytryptamine receptor antagonist cinanserin, which was also identified in vitro screening investigation as a Mpro binder [23]. Few research groups suggest that the N3analogue 1was discovered to be a nanomolar inhibitor of SARS-CoV-2 Mpro (IC50 value 0.15 μM) capable of inhibiting replication in VeroE6 cells (EC50=2.88 μM) by in vitro screening of effective SARS-CoV Mpro inhibitors [24]. Another study (VeroE6 cell assay method) found that GRL-0820 (initially developed by National Center for Global Health & Medicine) had a Mpro IC50 of 0.073 μM and an acceptable anti-SARS-CoV-2 activity (EC50 = 15 μM, VeroE6) among a screen of structural analogues [25]. Other experimental analysis (in vitro method) revealed CCF981(2-(benzotriazol-1-yl)-N-[(3-chlorophenyl) methyl]-N-[4-(1H imidazol-5-yl) phenyl] acetamide/CCF0058981) emerged as the lead drug with a significant antiviral action (EC50 = 0.56 μM; VeroE6) and a good Mpro inhibition capacity (IC50 = 0.068 μM) [26]. In the field of finding of Mpro inhibitors, the crystal structure (PDB Id. 7VH8, 7VLQ) emphasizes thioimidate production or covalent connection. The anti-inflammatory drug ebselen, which has been in the inhibitor race since early 2020, is the subject of an unusual instance of covalent inhibition of SARS-CoV-2 Mpro.

IJBAI
International Journal of Bioinformatics and Artificial Intelligence
ISSN: XXXX-XXXX
Vol.1 No.1January-June 2026,
Pages 01-37

Ebselen's antiviral activity and significant inhibition were demonstrated by enzymatic tests (IC50 = 0.67 µM, EC50 = 4.67 µM, VeroE6) [27]. The inhibitor MUT056399 (4-(4-ethyl-5-fluoro 2-hydroxyphenoxy)-3-fluorobenzamid) also exhibited considerable inhibition of SARS-CoV 2 viral activity (EC50 = 38.2 µM, CC50> 100 µM) (Table 2) [28]. The chemical compound 2 Dihydroxyphenyl (3,4-)Trihydroxy-8-(p-tolylselanyl) 3,5,7-in a vero cell model experiment, 4H-chromen-4-one demonstrated a safe and efficient inhibitory action (IC50 = 8 µM) against the replication of SARS-CoV-2 and it was a potent (IC50 = 11 µM) and reversible (Ki = 3.8 µM) Mproinhibitor [29-33]. Zhonget al. found that oridonin effectively inhibited SARS-CoV-2 activity and bound to it via covalent bonding, inhibiting SARS-CoV-2 in Vero E6 cells with an IC50 of 4.95 µM. This suggests that oridonin prevented SARS-CoV-2 replication by inhibiting 3CLpro [34]. Myricetin has also been demonstrated to diminish lung inflammation in a mouse model of bleomycin-induced pulmonary inflammation and inhibit viral infection in vitro (EC50 = 8 µM (VeroE6); 0.9 µM), suggesting that it may be helpful in treating COVID-19 symptoms [35]. Thorough in vitro research found that α-chloroacetamide 6 is a strong inhibitor of SARS-CoV-2 Mpro (IC50 = 0.4 µM), suggesting that structural design should focus on more specialized peptide-mimetics [36].

Table 2: *Details on the crystallographic structure of the conformation A (ConA) Mpro enzyme and its native form conformation.*

Resolu valuef Crystall Mean
No. of
S R Chain s /ographic
parameters
ldb Id reque nce t method Isotro r
. n ee/ R t R athe w's S eflect K
Iion Sefinemenolv ent iatio
Ph ions
o. Y2 ue work co effic ient (Å²) water)
(Å) nt ent len pic (pr
16 valgt h co B factorotein/

A/30
E1.750.22/0 .17
6REFMAC 7 59.94 3.07 55.46 7.56 27171
27D
A/30 N
AV1.770.25/0 .21 R1.450.20/0
REFMAC/P
A/30 6 A38.03 1.98 24.68
37DJ HENIX 8.96 25365
6PHENIX6.536.32 1.93 21 5.29 85995
.17
47N
XH2.10.27/0 .22 57V7

IJBAI
International Journal of Bioinformatics and Artificial Intelligence
ISSN: XXXX-XXXX
Vol.1 No.1January-June 2026, Pages 01-37

M2.08$^{0.30/0}$ .21

6^{7VJ}

Y1.9$^{0.28/0}$ .23

7^{7VJ}

Z1.9$^{0.25/0}$ .22

8^{7V}

K11.93$^{0.27/0}$ .22

9^{7V}

K22$^{0.28/0}$ .23

A/30

6PHENIX$^{7.5}$37.87 1.98 46.55 39.47 28805

A/30

6PHENIX 6 37.43 1.97 21.8 16.52 15800

A/30

6PHENIX 6 38.01 1.98 38.62 39.45 21030

A/30

6PHENIX 6 41.73 2.11 44.21 43.04 22355

A/30

6PHENIX 6 38.01 1.98 45.65 21.13 17608

A/30

6PHENIX 6 40.66 2.07 48.85 59.18 18845

1

K82.4$^{0.33/0}$

7V

0

.27

1

7W

O32.01$^{0.24/0}$ 1

.20

1

7W

OF1.72$^{0.20/0}$ 2

.20

A/30

6PHENIX 6 41.93 2.12 31.23 35.15 11028 A/30

6PHENIX^{N}A44.3 2.21 28.99 18.02 18872

A/30

6PHENIX 6 54.51 2.7 30.95 9.22 37363

Numerous comparative studies were conducted periodically to identify an inhibitor of Mpro. The in vitro assay revealed that apixaban (IC50=0.01 μM) was 21 times more active than GC 376 (dipeptidyl bisulfite adduct salt) and superior to rivaroxaban [37]. Chiou et al. found 20 drugs that inhibit SARS-CoV-2 Mproin silico and in vitro. Ethacrynic acid was the most effective inhibitor, with an IC50 value of 1.11 μM. Naproxen's anti-inflammatory and immunosuppressive properties (IC50 = 3.45 μM) may be beneficial in COVID-19 therapy [38]. In one investigation, Mpro inhibitors were identified using plaque reduction analysis. The substance S-217622 (ensitrelvir) showed a high antiviral action in the plaque reduction experiment (EC50 = 0.37 μM) and a very strong binding to Mpro (IC50 = 0.013 μM) [39]. In studies for cytopathic effect inhibition (EC50 = 0.497 μM) and plaque reduction (EC50 = 0.558

IJBAI
International Journal of Bioinformatics and Artificial Intelligence
ISSN: XXXX-XXXX
Vol.1 No.1January-June 2026,
Pages 01-37

μM), CCF0058981 (2-(benzotriazol-1-yl)-N-[(3-chlorophenyl)methyl]-N-[4-(1H-imidazol-5- yl)phenyl]acetamide), a new drug derived from ML300 (second probe of molecular libraries probe production centers network) (SARS-CoV2 inhibitor), demonstrated enhanced antiSARS-CoV-2 properties and nanomolar activity against SARS-CoV-2 3CLpro [40]. Additionally, researchers discovered that compounds 15b (4,4-Difluorocyclohexyl) (phenyl)methyl ((S)-4-methyl-1-oxo-1-(((S)-1-oxo-3-((S)-2- oxopyrrolidin-3-yl)propan-2-yl)amino)pentan-2-yl)carbamate) and 15C (Sodium (2S)-2-((2S)-2-((((4,4-

difluorocyclohexyl)(phenyl)methoxy)carbonyl)amino)-4-methylpentanamido)-1-hydroxy-3- ((S)-2-oxopyrrolidin-3-yl)propane-1-sulfonate) were the most effective against SARS-CoV-2 3CLpro, with IC50 values of 0.13 and 0.17 μM, respectively, after employing the fluorine-walk technique to investigate the binding mechanisms of the F-substituted phenyl ring (Table 3) [41].

Table 3: *Proposed inhibitors for the Mpro enzyme predicted by various experimental methods.*

Type of experiment	Name of proposed inhibitors	IC_{50} (μM)	References
(I) Fluorescence Polarization technology (FP)	Dieckol (FP with a biotin–avidin system)	4.5	[21-22]
	Aloesin (FRET)	NA	
(II) VeroE6 cell assay	Cinanserin	125	[23-34]
	N3 analogue 1	0.15	
	GRL-0820 (Indole-chloropyridinyl-ester derivative)	chlorophenyl) methyl]-N-[4-(1H-imidazol 5-yl) phenyl] acetamide)	
	CCF981 (2-(benzotriazol-1-yl)-N-[(3-	0.073	0.068
	Ebselen	0.67	
	MUT0563994 (4-ethyl-5-fluoro-2-hydroxyphenoxy)-3-fluorobenzamid	trihydroxy-8-tolylselanyl)-4H-chromen-4-one (p-	
	2-(3,4-Dihydroxyphenyl)3,5,7-	NA	11
	Oridonin	4.95	
(III) In vitro	Myricetin	NA	[35-38]
	α-chloroacetamide	6	0.4
	Apixaban	0.01	
	Ethacrynic acid	1.11	
	Naproxen	3.45	
(IV) Plaque reduction assay	S-217622 (Ensitrelvir)	0.013	[39-40]
	(benzotriazol-1-yl)-N-[(3- phenyl] acetamide)		
CCF0058981 (2-	chlorophenyl) methyl]-N-[4-(1H-imidazol 5-yl)	0.068	

IJBAI
International Journal of Bioinformatics and Artificial Intelligence
ISSN: XXXX-XXXX
Vol.1 No.1January-June 2026, Pages 01-37

(V) Fluorine-walk	15b (4,4-Difluorocyclohexyl) (phenyl)methyl ((S)-4-methyl-1-oxo-1-(((S)-1-oxo-3-((S)-2- oxopyrrolidin-3-yl) propan 2-yl) amino) pentan-2-yl) carbamate	(S)-2-((S)-2-((S)-2-((4,4-difluorocyclohexyl) (phenyl)methoxy) carbonyl) amino)-4-methylpentanamido)-1-hydroxy-3-((S)-2-oxopyrrolidin-3-yl) propane-1-sulfonate
	15c Sodium (2S)-2-((2S)-2-((((4,4-difluorocyclohexyl) (phenyl)methoxy)	0.13 0.17 [41]

1.6. Role of water molecules in Mpro enzyme

The SARS-CoV-2 Main Protease (Mpro) is crucial for viral replication and requires structural stability and interactions from twenty-four conserved water molecules in its active sites and two corresponding domains. These water molecules, occupying domains I (DI), II (DII) and the I-II interface and stabilize structure. The W1 in DI domain is dynamic and exposed near catalytic His41, supports transient interactions but lacks long-term stability. W2, associated with W1, also shows dynamic behavior. The W3, stable and buried, forms a H-bond network that supports enzyme integrity. The drug ability of W3 is indicated by its displacement by ligand ZINC2870664. W4 interacts with Glu55, suggesting a context-dependent structural role. W5, with low stability and minimal hydrogen bonds, is suitable for water displacement in water-based drug design. W6, bonded to Asn95 and Lys97, enhances structural stability. W7 forms salt bridges at the DI-DII interface, demonstrating its importance through high protein-water interaction energy. W8 stabilizes the structure through interactions with Arg40 and Met82. W9 is located near Gln19 and Thr21, but its functional role is to be investigated. W10 stabilizes the DI-DII interface via hydrogen bonds, while W11 and W12 maintain local structures through interactions with acidic residues. W13, with high entropy, is suitable for water displacement, influencing catalytic His41. W14 and W15 exhibit instability and minimal binding. W16 and W18 show limited structural roles, while W17 is not ideal for drug design due to stability. Unstable W19 and W24 have potential in water-mediated drug design. W20 and W21, stable at the interface, are key structural components. W22 and W23 demonstrate stability, with W23 showing strong binding to residues.

1.7. Crystallographic studies of Mpro enzyme

As per availability in the protein databank of Mpro crystal structures, it has been categorized into four distinct subcategories to enhance detail understanding:

• The ConA (conformation A) of Mpro in its native state, without any ligand (Table 2). • The ConA of Mpro in its ligand-bound form (Table 4).

• The ConAB(conformation AB) of Mpro in its native state, also without ligand (Table 5). • The ConABof Mpro when it is bound to a ligand (Table 6).

IJBAI
International Journal of Bioinformatics and Artificial Intelligence
ISSN: XXXX-XXXX
Vol.1 No.1January-June 2026,
Pages 01-37

The Conformation A (ConA) represents the dimeric conformation of Mpro, incorporating only molecule A in the crystal structures. In contrast, Conformation AB (ConAB) specifically refers to the dimeric form that includes both molecules A and B. The present analysis explores the crystal structures of the Mpro enzyme across two conformations, various space group specifically focusing on those with a resolution cutoff of less than 2.00 Å. The findings include twelve native Conformation A (ConA) of Mpro, thirty conformation Astructures of Mpro in complex with various ligands, nine Conformation AB (ConAB) of Mpro and twenty-one conformation AB of Mpro with diverse ligand complexes. The ligands (O6K, GQU, TRS, EOF, H60, H6F, H6R, MYC, J7O, 2XI, 35J, 5P9, XTP, V46, W48, I54, Q56, Y48, N63, S4L, K36, U88, XNJ, G65, 5IL, 4WI, ODN, 3XI, G7L, QNC and 9FF) were utilized to crystallize the ConA form of the Mpro enzyme, while ligands (7XB, 7YY, 80X, G7O, BOV, 4WI, UQW, URK, UQZ, USZ, BTB, HER, J7R, 2RI, O6K, H63, H6L) were employed to create a complex for the ConAB form of the Mpro enzyme.

Table 4: *Details on the crystallographic structure of the conformation A (ConA) Mpro enzyme and its ligand-bound conformation.*

Sl. No.	Pdb Id	R-value Resolution (Å)	Chain Refinement	m Crystallographic parameters of protein/water)	No. of Reflections
1	6Y2F	1.95 0.22/0.18	A/306 REFMAC	6 BUSTERNA 35.3 1.9 34.73 8.29 17469	
2	7D3120.21/0.17A/30	9 55.28 2.75 36.38 13.23 26755		3 7DG B 4 7DG F 5 7DG H 6 7DH J 1.68 0.24/0.20 A/306 PHENIX 6 55.04 2.74 38.93 12.38 40758 1.64 0.22/0.20 A/306 PHENIX 6 54.3 2.69 28.72 9.51 42425 1.97 0.24/0.20 A/306 PHENIX 6 45.24 2.25 24.93 13.76 21944 1.96 0.23/0.19 A/306 PHENIX 6 55.38 2.76 35.49 11.94 25421	
7	7DP N	P 2.1 0.22/0.19 A/301 PHENIX		56.83 2.85 47.03 33.28 21684	
		A			

IJBAI
International Journal of Bioinformatics and Artificial Intelligence
ISSN: XXXX-XXXX
Vol.1 No.1January-June 2026, Pages 01-37

8 7EN 9 1.9 0.23/0.20 A/306 PHENIX 5 56.01 2.8 26.04 9.32 29292

9 7FA Y 10 7NT Q C

2.1 0.23/0.19 A/306 BUSTER 1.5 0.21/0.17 A/306 REFMA A

N A N 37.47 1.97 23.69 10.96 42065

38.33 1.99 39.55 12.32 15634

11 7P51 1.47 0.22/0.18 A/306 REFMA A

N 38.74 2.01 23.9 9.34 45078

C

12 7TIA 1.64 0.19/0.17 A/306 PHENIX 5 55.77 2.78 46.69 13.6 44624 13 7TIU 1.65 0.20/0.17 A/306 PHENIX 5 56.71 2.84 47.33 12.07 45012 14 7TIV 2.08 0.22/0.18 A/306 PHENIX 5 56.66 2.84 48.24 21.37 22684

15 7TI W 1.68 0.20/0.17 A/306 PHENIX 5 55.55 2.77 50.19 14.1 41662

16 7TIX 2 0.22/0.17 A/306 PHENIX 5 56.7 2.84 49.07 16.91 25428 17 7TIY 1.79 0.20/0.17 A/306 PHENIX 6 56.84 2.85 39.53 11.28 35542 18 7TIZ 1.55 0.19/0.17 A/306 PHENIX 5 55.64 2.77 41.71 10.26 52603 19 7TJ0 2.17 0.22/0.17 A/306 PHENIX 5 56.48 2.83 51.3 22.18 19875

20 7TO B A 1.85 0.27/0.25 A/306 PHENIX

2.05 0.21/0.17 A/306 REFMA N A 37.37 1.96 30.74 11.93 15614 57.86 2.92 33.55 12.24 32747

N 21 7TU U

C

2 9 A / 3 0 6 P H E N I X 7. 5

2 7 T V S 1. 8 9 0. 2 2 / 0. 1

55.14 2.74 45.41 50.48 28672

23 7TV X 2.09 0.21/0.18 A/306 PHENIX 55.37 2.76 60.39 68.03 20506

5. 5

2 0. 2 5 / 0. 1 9 A /

4 7 V 1 T 2. 5 6

IJBAI
International Journal of Bioinformatics and Artificial Intelligence
ISSN: XXXX-XXXX
Vol.1 No.1January-June 2026, Pages 01-37

3 0 6 P H E N I X N A

38.65 2 58.68 195.42 8656

25 7VH 8 5. 6 1.59 0.20/0.18 A/306 PHENIX 45.23 2.25 28.47 6.23 41814

26 7VIC 2.1 0.24/0.22 A/306 REFMA 56.8 2.85 31.13 22.28 22123

N C A 27 7WO 1

2.15 0.22/0.18 A/306 PHENIX 6 43.53 2.18 29.8 17.66 16392

28 7WY M N A 2.05 0.22/0.19 A/306 PHENIX 38.44 2 53.99 25.35 16257

Int J Bioinfor Intell Comput, Vol 4, Issue 2, August 2025 143

ISSN 2816-8089

2 9 7X6J 1.5 0.20/0.18 A/306 PHENIX N A

55.49 2.76 19.49 10.82 53236

30 7X6 K N A 2.34 0.27/0.23 A/306 PHENIX 45.63 2.26 39.5 NIL 12993

Table 5: *Details on the crystallographic structure of the conformation AB (ConAB) Mpro enzyme and its native conformation.*

Sl. No.	Pdb id	Resolution (å)	R work value / R free value	Chains / sequence length	Refinement method	Solvent content	Crystallographic parameters Matthew's coefficient	Mean isotropic B factor	Ratio of protein/water	No. of reflections
1	7NT T	1.74	0.26/ 0.21	A/B/306	REFMAC	NI	38.6	2	29.06	17.06	52172

IJBAI

International Journal of Bioinformatics and Artificial Intelligence

ISSN: XXXX-XXXX

Vol.1 No.1January-June 2026, Pages 01-37

L				
2 7VJ			2.2 0.28/ A/B/30	PHENIX 8.5 56.98 2.86 56.08 73.68 39287
W .22	0	6		
3 7VJ			2.2 0.29/ A/B/30	PHENIX 8.5 57.21 2.87 62.04 47.86 39396
X .26	0	6		
4 7VK			2.1 0.24/ A/B/30	PHENIX 8.5 57.21 2.87 58.4 44.63 45326
0 .24	0	6		
5 7VK			2.1 0.28/ A/B/30	PHENIX 8.5 57.31 2.88 53.55 40.72 42380
3 .24	0	6		
6 7VK			2.1 0.27/ A/B/30	PHENIX 8.5 56.86 2.85 64.73 659.71 44742
4 .23	0	6		
7 7VK			2.17 0.27/ A/B/30	PHENIX 8.5 57.17 2.87 57.49 58.76 41092
5 .24	0	6		
8 7VK			2.25 0.27 A/B/30	PHENIX 8.5 57.06 2.86 60.36 45.07 36768
6 0.23	/	6		
9 7VK			2.4 0.24/ A/B/30	PHENIX 8.5 57.06 2.86 77.57 76.1 30465
7 .21	0	6		

Table 6: *Details on the crystallographic structure of the conformation AB (ConAB) Mpro enzyme and ligand–bound conformation.*

Sl.n .	Pdb od	Resolu tio i (å)	R v alue free/r value work	/ s equen ce	Chains m l nt ength	Refine eethod h Solve nt m	Crystal Fonte ct	lographic parameters of athew 's n	No. \oefficient cean	pic factor M isotro	b R

IJBAI
International Journal of Bioinformatics and Artificial Intelligence
ISSN: XXXX-XXXX
Vol.1 No.1January-June 2026, Pages 01-37

atio	(protein/water)	reflectio ns						
	2.2 0.25/0. 19	REFMA C	40.89 14.92 37448					
1 6Y2 G	A/B/3 06	8.5 53.92 2.67						
2 7DG G	1.9 0.21/0. 18	A/B/3 06	6.5 53.4 2.64 39.17 10.78 56431					
3 7DG I	1.9 0.21/0. 19	PHENIX 6 46.75 2.31						
4 7DK 1	A/B/3 06	26.85 13.35 37990 PHENIX 6 61						
2 0.22/0. 18	A/B/3 06	3.15 26.66 7.28 67801 BUSTER						
5 7DP U	1.75 0.20/0. 17	A/B/3 06	PHENIX NI L	38.13 1.99 23.83 11.35 52603				
6 7EN 8	1.83 0.27/0. 23	A/B/3 06	PHENIX 4.6 46.61 2.3 15.61 8.67 55655					
7 7FA Z	PHENIX NI L	8 7NT 1	REFMA C					
2.1 0.24/0. 19	50.53 2.49 40.63 12.05 39929	2.85 0.28/0. 20	6.5 54.29 2.69 56.79 87.13 17333					
A/B/3 06		A/B/3 06						
9 7NT 2	2.15 0.25/0. 20	A/B/3 06	6.5 53.71 2.66					
10 7NT 3	2.33 0.27/0. 21	REFMA C	40 26.15 39220 6.5 53.63					
	A/B/3 06	REFMA C	2.65 49.32 37.23 31556					
11 7N W2	2.1 0.22/0. 19	A/B/3 06	BUSTER 6.5 52.25 2.58 36.83 13.01 39471					
12 7TL L	A/B/3 06	39.54 2.03	18 7VU 6	NI				
13 7U2 8	A/B/3 06	27.23 14.21 53153 39.48	2 0.26/0. 21	45.76 2.27 17.8 12.88 52687 L				
14 7U2 9	A/B/3 06	2.03 23.93 14.86 46261	1.8 0.28/0. 22	19 7VV T				
15 7VL P	A/B/3 05	40.49 2.07 30.31 23.58	A/B/3 11	2.51 0.25/0. 23				
16 7VL Q	A/B/3 00	28383 45.85 2.27 23.3	A/B/3 08	A/B/3 06				
1.63 0.25/0. 21	BUSTER NI L	18.62 94010 54.56 2.71	REFMA C	PHENIX NI L				
1.68 0.25/0. 21	BUSTER NI L	26.58 22.27 51406	REFMA C	47.35 2.34 50.7				
2.09 0.27/0. 20	BUSTER NI L		NI	77.75 22010				
1.5 0.22/0. 20	PHENIX NI L	17 7VT H	37.79 1.98					
1.94 0.23/0. 20	PHENIX NI L		18.55 20.34 34366 L					
20 7WY P	2.3 0.27/0. 22	A/B/3 06	PHENIX 7.5 53.01 2.62 55.3 57.09 32303					
21 7XA R	A/B/3 06	NI						
1.6 0.21/0. 17	REFMA C	45.68 2.26 22.4 10.01 76114 L						

Moreover, no hypothesis studies have been reported to evaluate the following investigations: • The sub-category of available Mpro crystal structures (ligand free or ligand

IJBAI
International Journal of Bioinformatics and Artificial Intelligence
ISSN: XXXX-XXXX
Vol.1 No.1January-June 2026,
Pages 01-37

bound and conformation A or AB form) with resolutions below 2.00Å in the protein databank.
• A comparative analysis of ligand-bound crystal structures against its ligand-free forms in both Conformation A (ConA) and Conformation AB (ConAB).

• The identification of alternative binding sites for the anti-COVID molecule in crystal structure of Mpro.

Therefore, we have conducted a multi-conformations analysis on seventy-two crystal structures of the Mpro enzyme to explore the details of the structural transition from native Mpro to its ligand-bound form, which are essential insights for structure-based drug design. To the best of our knowledge, this is the first computational study that highlights the identification of flexible residues of Mpro crystal structures and trace the conformational changed (structural position from buried to exposed or vise-versa) of some residues of the enzyme in ConA and ConAB state. Multiple analyses of crystal structures of the native and ligand-bound form of Mpro protein provide suggestions to identify the specific important residues in Mpro enzyme that can be targeted as an alternative binding site of anti-COVID molecule. The results from this computational study could be of interest to the experimental community and provide new insights for experimental validation.

2. Methodology

2.1. Structure collection

The quality of a protein structure is essential for interpreting the flexibility and accessibility of its residues. This quality is often assessed using the structural parameters such as resolution and R-values, which include R-factor and R-free. Generally, higher resolution and lower R values are associated with more accurate and reliable interpretations of residue-level

flexibility and accessibility [42]. Well-defined electron density allows for precise positioning of atoms, including side chains and water molecules, which in turn leads to high-confidence B-factors. When the R-free value is close to the R-factor, it indicates that the model is reliable and not overly interpreted. The resolution, R-factor, R-free and residue-level fit statistics are crucial for a robust interpretation of residue-level flexibility and accessibility [43]. The seventy-two crystal structures of the Mpro enzyme were obtained from the RCSB database and categorized into four types:

• The native Conformation A (ConA) of Mpro.

• The ligand-bound Conformation A (ConA) of Mpro.

• The native Conformation AB (ConAB) from of Mpro (without ligand). • The ligand bound form with conformation AB (ConAB) of Mpro.

The findings encompass twelve native ConA structures of Mpro (PDB Ids: 6Y2E, 7DAV, 7DJR, 7NXH, 7V7M, 7VJY, 7VJZ, 7VK1, 7VK2, 7VK8, 7WO3 and 7WOF) and thirty ConA structures of Mpro complexes with various ligands (PDB Ids: 6Y2F, 7TIU, 7VH8, 7D31, 7TIV, 7VIC, 7DGB, 7TIW, 7WO1, 7DGF, 7TIX, 7WYM, 7DGH, 7TIY, 7X6J, 7DHJ, 7TIZ, 7X6K, 7DPP, 7TJ0,

IJBAI
International Journal of Bioinformatics and Artificial Intelligence
ISSN: XXXX-XXXX
Vol.1 No.1January-June 2026,
Pages 01-37

7EN9, 7TOB, 7FAY, 7TUU, 7NTQ, 7TVS, 7P51, 7TVX, 7TIA and 7V1T). Additionally, there are nine native ConAB structures of Mpro (PDB Ids: 7NTT, 7VJW, 7VJX, 7VK0, 7VK3, 7VK4, 7VK5, 7VK6, 7VK7) and twenty-one ConAB structures with various ligand complexes (PDB Ids: 6Y2G, 7U28, 7DGG, 7U29, 7DGI, 7VLP, 7DK1, 7VLQ, 7DPU, 7VTH, 7EN8, 7VU6, 7FAZ, 7VVT, 7NT1, 7WYP, 7NT2, 7XAR, 7NT3, 7NW2 and 7TLL). The ligands used for crystallizing the ConA form of the Mpro enzyme were identified by the following PDB Ids: O6K, GQU, TRS, EOF, H60, H6F, H6R, MYC, J7O, 2XI, 35J, 5P9, XTP, V46, W48, I54, Q56, Y48, N63, S4L, K36, U88, XNJ, G65, 5IL, 4WI, ODN, 3XI, G7L, QNC and 9FF. For creating a complex for the ConAB form of the Mpro enzyme, the following ligands were utilized: PDB Ids 7XB, 7YY, 80X, G7O, BOV, 4WI, UQW, URK, UQZ, USZ, BTB, HER, J7R, 2RI, O6K, H63 and H6L. Among these, the crystal structures 7DJR for the native ConAMpro, 7P51 for the ligand-bound ConAB form, 7NTT for the native ConABMpro and 7VLP for the ligand-bound ConABMpro hold higher resolutions and serve as templates or references for the current hypothesis.

2.2. Analysis of x-ray structures of Mpro

The seventy-two crystal structures of Mpro were analyzed in details. The best five crystal structures (with high resolutions) of Mpro enzyme from each category was considered as Conformation A (ConA) native (PDB Id: 7DJR, 7WOF, 6Y2E, 7DAV, 7VJY); Conformation AB (ConAB) native (PDB Id: 7NTT, 7VK0, 7VK3, 7VK4, 7VK5); ConA ligand-bound (PDB Id: 7P51, 7NTQ, 7X6J, 7TIZ, 7VH8); ConAB ligand-bound (PDB Id: 7VLP, 7XAR, 7TLL, 7U28, 7DPU). These crystal structures were considered for measuring RMSF (Root Mean Square

Fluctuation) value of each residue of Mpro. The experimental B-factors of each backbone "Cα" atoms of crystal structure of Mpro enzyme was converted to RMSF (root mean square fluctuation) value by using Debye-Waller formula: (RMSF=3B/8π2)½ using in-house TCL script. In addition, the Accessible Surface Area (ASA) was determined by get area program where the side-chain Accessibility Surface Area (ASA) of reference residues of four Mpro structures 7DJR for the native ConA Mpro, 7P51 for the ligand-bound ConAB form, 7NTT for the native ConAB Mpro and 7VLP for the ligand-bound ConAB Mpro hold higher were used, moreover the H-bonding interactions between ligands with Mpro protein was measured by Swiss-PDB Viewer program. The thirty ligand-bound ConA and the twenty-one ConAB ligand-bound forms of Mpro were utilized for the calculations. In the GATEAREA program, residues of the protein are classified as exposed or buried based on their Relative Solvent Accessibility (RSA) [44].

(SASA) of each residue to its "random coil" value. Residues are considered solvent-exposed if the RSA ratio exceeds 50% and they are classified as buried if the ratio is less than 20% [45]. The side-chain accessible surface area of reference residues from four Mpro structures was analyzed: 7DJR for the native ConA Mpro, 7P51 for the ligand-bound ConAB form, 7NTT for the native ConAB Mpro and 7VLP for the ligand-bound ConAB Mpro.

3. Results

3.1. Analysis of crystallographic data of Mpro enzyme

The seventy-two X-ray structures of the Mpro enzyme were obtained in various conformations either Conformation A (ConA) with native or ligand-bound states and

IJBAI
International Journal of Bioinformatics and Artificial Intelligence
ISSN: XXXX-XXXX
Vol.1 No.1January-June 2026,
Pages 01-37

Conformation AB (ConAB) with native or ligand-bound states from the protein data bank. Several research groups have independently solved these structures using PHENIX, REFMAC and BUSTER refinement programs at different (4.6–8.5) pH. The Crystallographic Structural Parameters (CSP) such as the Matthews's coefficient, solvent content and calculated mean B factors of proteins, are maximized in 6Y2E of the ConA native, 7TVX in the ConA ligand bound, 7VK4 in the ConAB native and 7NT1 in the ConAB ligand-bound conformation. This suggests that the atoms of these protein structures exhibit "wiggles" due to thermal motion, indicating greater flexibility in that region of the molecule. In contrast, the Matthews's coefficient, solvent content and calculated mean B-factors of proteins are minimum in 7DJR for the ConA native, 7X6J for the ConA ligand-bound, 7NTT for the ConAB native and 7EN8 for the ConAB ligand-bound form. This suggests that the atoms of these protein structures exhibit stability due to low thermal motion, indicating less flexibility in that region of the molecule. Moreover, the observed ratio of the number of protein atoms concerning water molecules (nprot/nhoh) in Mpro-native and ligand-bound form are different ranges that suggest 7VK2 of ConA native, 7V1T in ConA ligand, 7VK7 in ConAB native and 7NT1 of the ConAB ligand-bound ensemble are random compared to remaining structures. Most Conformation AB (ConAB) native-Mpro crystal structures belong to space group P 1 21 1, whereas ligand-bound structures are from P 21 21 21 or P 1 21 1. The native structure has been observed to have over 11,000 reflections in ConA and 30,000 in ConAB, while the ligand bound structure has been found to have over 8,000 in ConA and 17,000 in ConAB. The ligand ID N63 (PDB Id: 7TIZ) and ligand O6K (PDB ID: 6Y2F) form strong hydrogen bonds with their respective crystal structures of the Mpro enzyme in the ligand-bound ConA (Table 7. Similarly, the ligand H6L (PDB ID: 7DGI) and BOV (PDB Id: 7XAR) also interact with the Mpro protein in ConAB (Table 8).

Table 8: *Information regarding the ligands present in the X-ray structure of the Mpro enzyme in ConA.*

Sl no.	Ligand id. (pdb id.)	Chemical name / formula	Smiles	Pub chem Id.	No. of H bonds with protein
1	O6K (6Y2F)	~{tert}-butyl ~{N}-[1-[(2~{S})-3-cyclopropyl-1-oxidanylidene-1-[[(2~{S},3~{R})-3-oxidanylidene-1-[(3~{S})-2-oxidanylidenepyrrolidin-3-yl]-4-[(phenylmethyl)amino]butan-2-yl]amino]propan-2-yl]-2-oxidanylidenepyridin-3-yl]carbamate	/C31 H41 N5 O7	CC(C)(C)OC(=O)NC1=CC=CN(C1=O)[C@@H](CC2CC2)C(=O)N[C@@H](C[C@@H]3CCC[C@@H]3[C@H]2C(=O)N[C@@H](C[C@@H]4CCNC4=O)C=O	14601870	8
2	GQU (7D3I)	(3~{S},3~{a}~{S},6~{a}~{R})-2-[3-[3,5-bis(fluoranyl)phenyl]-3,3~{a},4,5,6,6~{a}-hexahydro-1~{H}-cyclopenta[c]pyrrole-3-carboxamide	/C24 H29 F2 N3 O4	c1c(cc(cc1F)F)CCC(=O)N2C[C@@H]3[C@H]2C(=O)N[C@@H](C[C@@H]4CCNC4=O)C=O	15580440	5
3	TRS (7D3I)	2-AMINO-2-HYDROXYMETHYL PROPANE-1,3-DIOL	/ C4 H12 N O3	C(C(CO)(CO)[NH3+])O	8808875	2

International Journal of Bioinformatics and Artificial Intelligence
ISSN: XXXX-XXXX
Vol.1 No.1January-June 2026, Pages 01-37

	4	EOF (7DGB)	[(3~{S})-2-oxidanylidenepyrrolidin-3-yl]propan-2-yl]-2-[[(~{E})-3-phenylprop-2-enoyl]amino]pentanamide / C22 H29 N3 O4	@@H](C(=O)N[C@@H](C[C@@H]1CCNC1=O)C=O)NC(=O)/C=C/c2ccccc2	8		
	5	H60 (7DGF)	(2~{S})-~{N}-[(2~{S})-1-oxidanylidene-3-[(3~{S})-2-oxidanylidenepiperidin-3-yl]propan-2-yl]-2-[[(~{E})-3-phenylprop-2-enoyl]amino]hexanamide / C23 H31 N3 O4	2-oxidanylidenepyrrolidin-3-yl]propan-2-yl]-2-[[(~{E})-3-phenylprop-2-enoyl]amino]pent-4-ynamide / C21 H23 N3 O4	C) C)NC(=O)c2ccc3ccccc3c2		
	6	H6F (7DGH)	~{N}-[(2~{S})-3-methyl-1-[[(2~{S})-4-methyl-1-oxidanylidene-1-[[(2~{S})-1-oxidanylidene-3-[(3~{S})-2-oxidanylidenepiperidin-3-yl]propan-2-yl]amino]pentan-2-yl]amino]-1-oxidanylidene-butan-2-yl]naphthalene-2-carboxamide / C30 H40 N4 O5	3,5,7-TRIHYDROXY-2-(3,4,5-TRIHYDROXYPHENYL)-4H-CHROMEN-4-ONE / C15 H10 O8	C#CC[C@@H](C(=O)N[C@@H](C[C@@H]1CCNC1=O)C=O)NC(=O)/C=C/c2ccccc 2		
	7	H6R (7DHJ)		CCCC[C@@H](C(=O)N[C@@H](C[C@@H]1CCCNC1=O)C=O)NC(=O)/C=C/c2ccccc 2	c1c(cc(c(c1O)O)O)C2=C(C(=O)c3c(cc(cc3O2)O)O)O	NIL	8 NIL 8
	8	MYC (7DPP)	(2~{S})-~{N}-[(2~{S})-1-oxidanylidene-3-[(3~{S})-2-oxidanylidenepiperidin-3-yl]propan-2-yl]amino]pentan-2-yl]amino]-1-oxidanylidene-butan-2-yl]naphthalene-2-carboxamide / C30 H40 N4 O5	CC(C)C[C@@H](C(=O)N[C@@H](C[C@@H]1CCCNC1=O)C=O)NC(=O)[C@H](C(		NIL	6 5281672 2
	9	J7O (7EN9)	5-bromanyl-2,7-~{N}-methyl-3-nitro-2-[(4~{R},5~{S})-2-(7-oxidanylisoquinolin-4-yl)carbonyl-4-phenyl-diazaspiro[4.4]nonan-7-yl]benzamide / C31 H28 Br N5 O5	CNC(=O)c1cc(cc(c1N2CC[C@]3(C2)CN(C[C@@H]3c4cccc4)C(=O)c5cncc6c5ccc(c6)O)[N+](=O)[O-])Br		NIL	4
	10	2XI (7FAY)	(2~{R})-~{N}-[(1~{R})-2-(~{tert}-butylamino)-2-oxidanylidene-1-pyridin-3-yl-ethyl]-~{N}-(4-~{tert}-butylphenyl)-2-oxidanyl-propanamide / C24 H33 N3 O3	C[C@H](C(=O)N(c1ccc(cc1)C(C)(C)C)[C@H](c2cccnc2)C (=O)NC(C)(C)C)O			1560227 33 5
	11	35J	(7NTQ) N-(pyridin-3-ylmethyl)thioformamide / C7 H8 N2 S	c1cc(cnc1)CNC=S			5827964 1 8
	12	5P9(7P51)	N-indene-1-carboxamide / C15 H11 Cl N2 O2				
	13	XTP (7TIA)	(5-chloropyridin-2-yl)-3-oxo-2,3-dihydro-1H-	benzyl [(2S)-3-cyclopropyl-1-({(2S)-1-hydroxy-3-[(3S)-2-oxopyrrolidin-3-			

International Journal of Bioinformatics and Artificial Intelligence
ISSN: XXXX-XXXX
Vol.1 No.1January-June 2026, Pages 01-37

	yl]propan-2-yl}amino)-1-oxopropan-2-yl]carbamate / C21 H29 N3 O5	H](CC2=O)C(=O)Nc3ccc(cn3)Cl	c1ccc2c(c1)[C@@H](CC2CC2)C(=O)N[C@@H](C[C@@H]3CCNC3=O)CO		5189419 3 1560227 97 2 8
	14 V46 (7TIU) (1S,2S)-2-[(N-{[(3-chlorophenyl)methoxy]carbonyl}-L	3-[(3S)-2-oxopyrrolidin-3-yl]propane-1-sulfonic acid / C21 H30 Cl N3 O8 S	leucyl)amino]-1-hydroxy-	CC(C)C[C@@H](C(=O)N[C@@H](C[C@@H]1CCNC1=O)[C@H](O)S(=O)(=O)O)NC(=O)OCc2cccc(c2)Cl	NIL 7
	15 W48 (7TIV) (1S,2S)-2-[(N-{[(3-chlorophenyl)methoxy]carbonyl}-3-cyclohexyl-L-alanyl)amino]-1-hydroxy	yl]-2-[(N-{[(2,4,5-trifluorophenyl)methoxy]carbonyl}-L-leucyl)amino]propane-1-sulfonic acid / C21 H28 F3 N3 O8 S		CC(C)C[C@@H](C(=O)N[C@@H](C[C@@H]1CCNC1=O)[C@H](O)S(=O)(=O)O)NC(=O)OCc2ccc3ccccc3c2 CC(C)C[C@@H](C(=O)N[C@@H](C[C@@H]1CCNC1=O)[C@H](O)S(=O)(=O)O)NC(=O)OCc2cc(c(cc2F)F)F	
	3-[(3R)-2-oxo-2,3-dihydro-1H-pyrrol-3-yl]propane-1-sulfonic acid / C24 H32 Cl N3 O8 S	19 N63 (7TIZ) (1S,2S)-1-hydroxy-3-[(3S)-2- oxopyrrolidin-3-yl]-2-{[N-({[3-(trifluoromethyl)phenyl]methoxy}carbo		CC(C)C[C@@H](C(=O)N[C@@H](C[C@@H]1CCNC1=O)[C@H](O)S(=O)(=O)O)NC(=O)OCc2cccc(c2)C(F)(F)F	
	16 I54 (7TIW) (1S,2S)-2-[(N-{[(2-chlorophenyl)methoxy]carbonyl}-L	nyl)-L-leucyl]amino}propane-1-sulfonic acid / C22 H30 F3 N3 O8 S		c1ccc(cc1)COC(=O)N[C@@H](CC2CCCCC2)C(=O)N[C@@H](C[C@@H]3CCNC3=O)[C@H](O)S(=O)(=O)O	NIL 8
	leucyl)amino]-1-hydroxy-3-[(3S)-2-oxopyrrolidin-3-yl]propane-1-sulfonic acid / C21 H30 Cl N3 O8 S	20 S4L (7TJ0) (1S,2S)-2-({N-[(benzyloxy)carbonyl]-3- cyclohexyl-L-alanyl}amino)-1-hydroxy	3-[(3S)-2-oxopyrrolidin-3-yl]propane-1-	sulfonic acid / C24 H35 N3 O8 S	NIL 6
	17 Q56 (7TIX) N~2~-{[(naphthalen-2-yl)methoxy]carbonyl}-N-{(2S)-1-oxo-3-[(3S)-2-oxopyrrolidin-3-yl]propan-2-yl}-L-leucinamide /C25 H33 N3 O8 S	c1cc(cc(c1)Cl)COC(=O)N[C@@H](CC2CCCCC2)C(=O)N[C@@H](C[C@@H]3C=CNC3=O)[C@H](O)S(=O)(=O)O			NIL 6 NIL 6
	18 Y48 (7TIY) (1S,2S)-1-hydroxy-3-[(3S)-2- oxopyrrolidin-3-	CC(C)C[C@@H](C(=O)N[C@@H](C[C@@H]1CCNC1=O)[C@H](O)S(=O)(=O)O)NC(=O)OCc2ccccc2Cl			NIL 10 NIL 8
	21 K36 (7TOB)	({N-	(1S,2S)-2-	[(benzyloxy)carbonyl]-L	hydroxy-3-[(3S)-2- leucyl}amino)-1-oxopyrrolidin-3- yl]propane-1-sulfonic acid / C21 H31 N3 O8

International Journal of Bioinformatics and Artificial Intelligence
ISSN: XXXX-XXXX
Vol.1 No.1January-June 2026, Pages 01-37

S		CC(C)C[C@@H](C(=O)N[C@@H](C[C@@H]1CCNC1=O)[C@@H](O)S(=O)(=O)O)NC(=O)OCc2ccccc2	1373496 27 6			
22	U88 (7TUU)	5-nitro-1,3-thiazole / C3 H2 N2 O2 S	c1c(scn1)N(=O)=O	541661	NIL	
23	XNJ (7TVS)	N-(4-methyl-3-{[4-(pyridin-3-yl)-1,3-thiazol-2-yl]amino}phenyl)-4-[(piperazin-1-yl)methyl]benzamide / C27 H28 N6 O S	Cc1ccc(cc1Nc2nc(cs2)c3cccnc3)NC(=O)c4ccc(cc4)CN5CCNCC5	1303243 11 2		
24	G65 (7TVX)	Masitinib / C28 H30 N6 O S	Cc1ccc(cc1Nc2nc(cs2)c3cccnc3)NC(=O)c4ccc(cc4)CN5CCN(CC5)C	1007464 02		
25	5IL (7V1T)	5,8-bis(oxidanylidene)-7-[(2-piperazin-1- ylphenyl)amino]naphthalene-1-sulfonamide / C20 H20 N4 O4 S	c1ccc(c(c1)NC2=CC(=O)c3cccc(c3C2=O)S(=O)(=O)N)N4CCNCC4		NIL	3
26	4WI (7VH8)	(1R,2S,5S)-N-{(1E,2S)-1-imino-3-[(3S)-2-oxopyrrolidin-3-yl]propan-2-yl}-6,6-dimethyl-3-[3-methyl-N-(trifluoroacetyl)-L-valyl]-3-azabicyclo[3.1.0]hexane-2-carboxamide / C23 H34 F3 N5 O4	[H]/N=C/[C@H](C[C@@H]1CCNC1=O)NC(=O)[C@@H](NC(=O)[C@@H]2[C@@H]3[C@@H](C3(C)C)CN2C(=O)[C@H](C(C)(C)C)NC(=O)C(F)(F)F		NIL	8
27	ODN (P7VIC)	(1beta,6beta,7beta,8alpha,9beta,10alpha,13alpha,14R,16beta)-1,6,7,14-tetrahydroxy-7,20-epoxykauran-15-one / C20 H30 O6	C[C@@H]1[[C@@H]4C(CC[C@@H]2CC[C@H]3[C@@]45CO[C@@]([C@]3([C@@H]2O)C1=O)([C@H](O)C=O)NC(=O)NC(=O)[C@H](C(C) C)NC(=O)C2CCCCC2	1373498 31 1		
28	3XI (7WO1)	N-[(2S)-3-methyl-1-[[(2S)-4-methyl-1-oxidanylidene-1-[[(2S)-1-oxidanylidene-3-[(3S)-2-oxidanylidenepiperidin-3-yl]propan-2-yl]amino]pentan-2-yl]amino]-1-oxidanylidene-butan-2-yl]cyclohexanecarboxamide / C26 H44 N4 O5	CC(C)C[C@@H](C(=O)N[C@@H](C[C@@H]1CCCNC1=O)[C@H](O)C=O)NC(=O)[C@H](C(C)C)NC(=O)C2CCCCC2		NIL	8
29	G7L (7WYM)	N-methyl-N-[[4-(trifluoromethyl)-1,3-thiazol-2-yl]methyl]prop-2-enamide/ C9 H9 F3 N2 O S	CN(Cc1nc(cs1)C(F)(F)F)C(=O)C=C		NIL	1
30	QNC (7X6J)	quinoline-2-carboxylic acid / C10 H7 N O2	c1ccc2c(c1)ccc(n2)C(=O)O	7124	NIL	
31	9FF (7X6K)	1H-indole-2-carbaldehyde /C9 H7 N O	c1ccc2c(c1)cc([nH]2)C=O		NIL	1

Table 9: *Information regarding the ligands present in the X-ray structure of the Mpro enzyme in ConAB.*

Sl. No.	Ligand id (pdb id.)	Chemical name /formula	Smiles	PubCH EM ID		

IJBAI
International Journal of Bioinformatics and Artificial Intelligence
ISSN: XXXX-XXXX
Vol.1 No.1January-June 2026, Pages 01-37

No. of H bonds withprotein

1 O6K (6Y2G)	~{tert}-butyl ~{N}-[1-[(2~{S})-3-cyclopropyl-1-oxidanylidene-1-[[(2~{S},3~{R})-3-oxidanyl-4-oxidanylidene-1-[(3~{S})-2-oxidanylidenepyrrolidin-3-yl]-4-[(phenylmethyl)amino] butan-2-yl] amino] propan-2-yl]-2-oxidanylidene pyridin-3-yl] carbamate / C31 H41 N5 O7	CC(C)(C)OC(O)[C@H](C(=O)NCc4ccc =O)NC1=CC=CN(cc4)O C1=O)[C@@H](CC2CC2)C(=O) N[C@@H](C[C@@H]3C CNC3= CCCC[C@@H](C(= 3-O)N[C@@H](C[C@@H]1CCNC1=O)C=O)N C(=O)/C=C/c2ccccc2	1460187 08 11
2 H63 (7DGG)	~{N}-[(2~{S})-3-amino]-1-[(2~{S})-1-oxidanylidene-3-[(3~{S})-2-oxidanylidenepiperidin-3-yl] propan 2-yl] amino] pentan-2-yl]-2-[[(~{E})-3-phenylprop-2-enoyl] amino] hexanamide / C22 H29 N3 O4	NIL 8				
3 H6L (7DGI)	~{N}-[(2~{S})-3-amino]-1-methyl-1-[[(2~{S})-4-methyl-1-oxidanylidene-1-[[(2~{S})-1-nitro oxidanylidene-butan-2-yl]-4-hydroxy-ethyl)-amino]-2-hydroxymethyl-propane-1,3-diol / C8 H19 N O5	NIL 9				
4 BTB (7DK1)	2-[bis-(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol	C(CO)N(CCO)C(CO)(CO)CO	81462 4			
5 HER (7DPU)	7-methoxy-3,5-bis(oxidanyl)-2-[3,4,5-tris(oxidanyl)phenyl] chromen-4-one / C16 H12 O8	COc1cc(c2c(c1 C(=O)OC(=C(C2=O)O) c3cc(c(c(c3)O)O)O)O	4425963 6 5			
6 J7R (7EN8)	~{N}-[(1~{S},2~{R})-2-[[4-bromanyl-2-(methylcarbamoyl)-6-nitro uinoline -4-carboxamide / C24 H24 Br N5 O4	CCC[C@H]2NC(=O)c3cncc4c3 cccc4)N(=O)=O)Br NIL 3				
7 2RI (7FAZ)	(2~{R})-~{N}-dibenzofuran-3-yl-~{N}-[(1~{R})-2-[[(1~{S})-1-(4-fluorophenyl) ethyl] amino]-2-oxidanylidene-1-pyridin-3-yl-ethyl]-2-oxidanyl propanamide / C30 H26 F N3 O4	CNC(=O)c1cc(cc(c1N[C@@H]2 C[C@@H](c1ccc(cc1) F)NC(=O)[C@@H](c2cccnc2)N(c3ccc4c5cc cc c5oc4c3)C(=O)[C@@H](C)O	1560227 32 5			

IJBAI
International Journal of Bioinformatics and Artificial Intelligence
ISSN: XXXX-XXXX
Vol.1 No.1January-June 2026, Pages 01-37

		[(2R)-1-[2-(1H-indol-3-yl) ethylamino]-1-oxidanylidene-butan-2-yl] prop-2- enoate / C17 H20 N2 O3	oxidanylidene-ethyl]-~{N}-propyl enamide / C22 H24 N2 O5	COc1cccc(c1O)C=C	1560227 83
8	UQW (7NT1)		~{N}-(4-~{tert}-butylphenyl)-~{N}-[(1~{R})-2-[2-(3-fluorophenyl) ethylamino]-2-oxidanylidene-1-pyridin-3-yl-ethyl] propanamide / C28 H32 F N3 O2	CCCN([C@@H](c1cccc(c1)O)C(=O)NCc2ccc3c(c2)OCO3)C(=O)C =C	1559235 25
9	URK (7NT2)	[(1S)-2-[(2,3-dimethoxyphenyl) methylamino]-1-(4-nitrophenyl)-2-oxidanylidene-ethyl] prop-2-enoate / C20 H20 N2 O7		CCC(=O)N(c1ccc(cc1)C(C)(C)C)[C@H](c2cccnc2)C(=O)NCCc3ccc c(c3)F	1560227 87 1 1
10	UQZ (7NT3)		~{N}-[(1~{S})-2-(1,3-benzodioxol-5-ylmethylamino)-1-(3-hydroxyphenyl)-2-oxidanylidene-ethyl] propanamide / C28 H34 F3 N5 O4	CC[C@H](C(=O)NCCc1c[nH]c2c1cccc2)OC(=O)C=C	1560227 82 3 2
11	USZ (7NW2)				
12	4WI (7TLL)	(1R,2S,5S)-N-{(1E,2S)-1-imino-3-[(3S)-2-oxopyrrolidin-3-yl] propan-2-yl}-6,6-dimethyl-3-[3-methyl-N-(trifluoroacetyl)-L-valyl]-3-azabicyclo [3.1.0] hexane-2-carboxamide / C23 H34 F3 N5 O4		[H]/N=C/[C@H](C[C@@H]1CCNC1=O)NC(=O)[C@@H]2[C@@H]3[C@@H](C3(C)C)CN2C(=O)[C@H](C(C)(C)C)NC(=O)C(F)(F) F	NIL 7
13	4WI (7U28)	(1R,2S,5S)-N-{(1E,2S)-1-imino-3-[(3S)-2-oxopyrrolidin-3-yl]propan-2-yl}-6,6-dimethyl-3-[3-methyl-N-(trifluoroacetyl)-L-valyl]-3-azabicyclo [3.1.0] hexane-2-carboxamide / C23 H34 F3 N5 O4		[H]/N=C/[C@H](C[C@@H]1CCNC1=O)NC(=O)[C@@H]2[C@@H]3[C@@H](C3(C)C)CN2C(=O)[C@H](C(C)(C)C)NC(=O)C(F)(F) F	NIL 8 NIL 7
14	4WI (7U29)	(1R,2S,5S)-N-{(1E,2S)-1-imino-3-[(3S)-2-oxopyrrolidin-3-yl] propan-2-yl}-6,6-dimethyl-3-[3-methyl-N-(trifluoroacetyl)-L-valyl]-3-azabicyclo [3.1.0] hexane-2-carboxamide / C23 H34 F3 N5 O4		[H]/N=C/[C@H](C[C@@H]1CCNC1=O)NC(=O)[C@@H]2[C@@H]3[C@@H](C3(C)C)CN2C(=O)[C@H](C(C)(C)C)NC(=O)C(F)(F) F	
15	4WI (7VLP)	(1R,2S,5S)-N-{(1E,2S)-1-imino-3-[(3S)-2-oxopyrrolidin-3-yl] propan-2-yl}-6,6-dimethyl-3-[3-methyl-N-(trifluoroacetyl)-L-valyl]-3-azabicyclo[3.1.0] hexane-2-carboxamide / C23H34 F3 N5 O4		[H]/N=C/[C@H](C[C@@H]1CCNC1=O)NC(=O)[C@@H]2[C@@H]3[C@@H](C3(C)C)CN2C(=O)[C@H](C(C)(C)C)NC(=O)C(F)(F) F	
16	4WI (7VLQ)	(1R,2S,5S)-N-{(1E,2S)-1-imino-3-[(3S)-2-oxopyrrolidin-3-yl] propan-2-yl}-6,6-dimethyl-3-[3-methyl-N-(trifluoroacetyl)-L-valyl]-3-azabicyclo[3.1.0] hexane-2-carboxamide / C23H34 F3 N5 O4 (1R,2S,5S)-N-)[		[H]/N=C/[C@H](C[C@@H]1CCNC1=O)NC(=O)[C@@H]2[C@@H]3[C@@H](C3(C)C)CN2C(=O)[	

IJBAI
International Journal of Bioinformatics and Artificial Intelligence
ISSN: XXXX-XXXX
Vol.1 No.1January-June 2026, Pages 01-37

17 7XB (7VTH) dimethyl-3-[3-methyl-N-(trifluoroacetyl)-L-valyl]-3-azabicyclo [3.1.0] hexane-2-carboxamide / C23 H34 F3 N5 O4	2-[4-[[4-[bis(fluoranyl)methoxy]-2-methyl-phenyl] amino]-2,6-bis(oxidanylidene)-3-[[3,4,5-tris(fluoranyl)phenyl] methyl]-1,3,5- triazin-1-yl]-N-methyl-ethanamide / C21 H18 F5 N5 O4	CC(C)C[C@@H](C(=O)NN(C[C@@H]1CCNC1=O)C(=O)CF)NC(=O)c2cc3c([nH]2)cccc3F	NIL	NIL 9
18 7YY (7VU6) 6-[(6-chloranyl-1,3,5-triazine-2,4-dione / C22 H17 Cl	2-methyl-indazol-5-yl) methyl-1,2,4-triazol-3-yl)methyl]-1-[[2,4,5-tris(fluoranyl)phenyl] methyl]-	Cc1cc(ccc1NC2=NC(=O)N(C(=O)Nc2cccc(c2)Cl	NIL	NIL 5
19 80X (7VVT)	amino]-3-[(1-	Cc1cc2ccc(c(cc2n1)Cl)NC3=NC(=O)N(C(=O)N3Cc4cc(c(cc4F)F)F)	NIL	NIL 5
20 G7O (7WYP)		Cc5ncn(n5)C C=CC(=O)N(Cc1nc2ccccc2s1)C3 CC3	NIL	NIL 5
21 BOV (7XAR) N-(3-chlorophenyl)-2-[(2R)-1-ethanoyl 3-oxidanylidene-piperazin-2-yl]ethanamide / C14 H16 Cl N3 O3 N-(1,3-benzothiazol-2-ylmethyl)-N cyclopropyl-prop-2-enamide / C14 H14 N2 O S 4-fluoranyl-~{N}-[(2~{S})-1-[2-(2-fluoranylethanoyl)-2-[[(3~{S})-2- oxidanylidenepyrrolidin-3-yl]methyl]hydrazinyl]-4-methyl-1- oxidanylidene-pentan-2-yl]-1~{H}- indole-2-carboxamide / C22 H27 F2 N5 O4		CC(=O)N1CCNC(=O)[C@H]1C		

3.2. Analysis of atomic flexibility of crystal structures of Mpro enzyme

The comparative analysis of Root-Mean-Square Fluctuation (RMSF) for Cα atoms in different conformations of the Mpro enzyme's X-ray structure was analyzed (Figure 1). In the native Conformation AB (ConAB), Ser45-Glu46, Asn72-Val73, Gly138-Ser144 and Tyr154-Asp155 exhibit greater flexibility, with their RMSF values exceeding 1.5 Å. In contrast, the ligand

bound form of the ConAB is observed to be stable, with RMSF values below 1 Å. For the native ConA form of Mpro, only one residue (Gln306) is highly flexible, having an RMSF greater than 1.5 Å, compared to its ligand-bound form. These results reveal that the ligand-bound forms, both in the ConA and ConAB are more stable than their native states. Additionally, the native ConAB is predominantly dynamic, as several regions show increased flexibility compared to the remaining conformations. The Root Mean Square Fluctuations (RMSF) derived from crystallographic B-factors illustrate the initial flexibility of the Mpro protein's structure.

However, they do not reliably reflect actual dynamics of the Mpro protein for several reasons:

- Experimental B-factors incorporate noise, model inaccuracies and effects from crystal packing, rather than solely reflecting atomic movements.
- Atoms may seem artificially rigid or flexible depending on their arrangement within the crystal lattice, which does not represent true dynamics inside the cell.
- B-factors are time-averaged and cannot differentiate between fast and slow dynamics [46- 48].

As a result, B-factors of the residues are subject to artifacts and lack the dynamic behavior. To achieve a more specific fluctuation or flexibility analysis of the Mpro protein, molecular dynamics simulations and NMR-based RMSF investigation are essential, as these methods provide structural insights into the real dynamic motions, correlations and functional transitions of the Mpro protein [48-50].

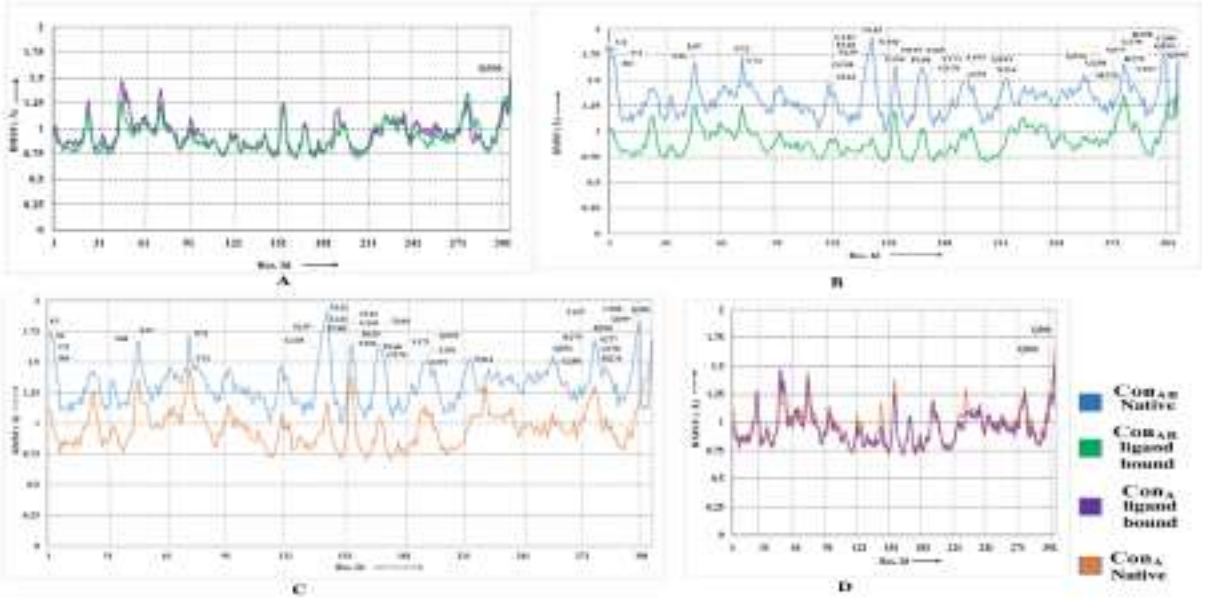

Figure 1: *The graphs compare the root-mean-square fluctuation (RMSF) in Å of Ca atoms across different conformations of the Mpro enzyme's X-ray structure. Panel A shows the superimposition of RMSF (Ca atoms) for the conformation A (ConA) ligand bound state (in violet) and the conformation AB (ConAB) ligand bound state (in green). Panel B compares the ConAB-native state (in blue) with the ConAB-ligand bound state (in green). Panel C illustrates the differences between the ConA-native state (in brown) and the ConAB-native state (in blue). Finally, Panel D displays the comparison between the ConA-native state (in brown) and the ConA-ligand bound state (in violet).*

3.3 Comparative analysis between native ConA and ConAB of Mpro

During the structural transition from the native ConA to the native ConAB state, forty-nine residues undergo conformational changes, shifting from buried to exposed states or vice versa (Figure 2). In panel A, Ala194, Thr198 and Ile213 transition from a buried to a semi-buried state. In panel B, Arg4, Met6, Pro9, Leu50, Arg60, Gly124, Ser139, Pro252, Gln256, Ala260, Ala285 and Gly302 move from an exposed to a buried state, indicating that these twelve residues gain stability in the native ConAB form of Mpro compared to its ConA state. Panel C showsGlu14, Tyr118, Pro122, Val125, Tyr126, Gly138, Gly143, Ser158, Asp216 and Glu290 that residues change from a semi-buried state to a buried state. In panel D, Val73, Lys100, Lys137, Ser254, Arg279 and Phe294 convert from a buried state to an exposed state, suggesting that these residues become less stable in the ConAB native state compared to their ConA form.

Panel E illustrates that residues Ser1, Gly2, Gly11, Thr45, Ser121, Leu141, Tyr154, Asp155, Glu166, Pro168, Ile249, Gly275, Thr280, Leu286, Val297 and Thr304shift from an exposed to a semi-buried state. Additionally, Asp229 and Gly251 in panel F also transition from a semi buried to an exposed state. Overall, the computational results suggest that twelve residues change their conformations from exposed to buried states, while six residues shift from buried to exposed states.

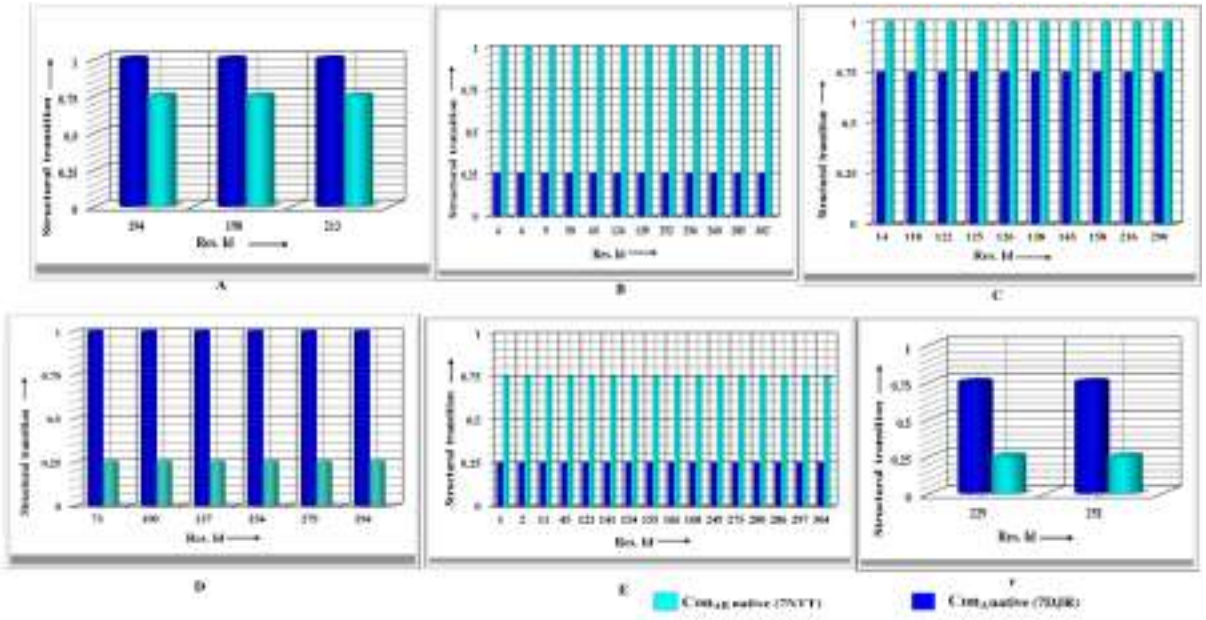

Figure2: *The structural transitions of the residues in the Mpro enzyme are analysed based on the solvent-accessible surface area as it changes in both the native ConA and native ConAB forms. Panel"A" illustrates the structural transition of residues in Mpro from the native ConA to the native ConAB state, indicating a shift from buried to semi-buried. Panel "B" shows the transition from exposed to buried residues. Panel "C" depicts the change from semi-buried to buried. Panel "D" highlights the transition from buried to exposed states. Panel "E" demonstrates the shift from exposed to semi-buried residues, while Panel "F" illustrates the transition from semi-buried to exposed.*

3.4. Comparative analysis between native ConA and its ligand-bound state

During the structural transition from the native ConA to its ligand-bound state, twenty-eight residues undergo conformational changes, shifting from buried, semi-buried or exposed states to either semi-buried or exposed states or vice versa (Figure 3). In panel A, Asn28, Val35, His41, Met49, Ala194, Thr198, Val247 and Ser254 transition from a buried state to a semi

exposed state, indicate a change in their hydration accessibility. In panel B, Gln19, Asp48, Ala70 and Glu288 demonstrate a shift from semi-exposed to buried, suggesting increased a structural stabilization of their environment. Panel C shows Gly124, Asp155, Ile249, Gln256 and Val297 moving from an exposed state to semi-exposed, reflecting a notable change in their positioning. Furthermore, in panel D, Val73, Lys100, Ser123, Cys128, Lys137, Cys156 and Phe294 transition from buried to exposed, suggesting increased surface exposure, while residues Cys145, Gly170 and Val303 shift from semi-exposed to exposed in panel E. Lastly, Phe305 transitions from an exposed state to buried in panel F. The computational findings clearly show that in the structural transition from the native ConA to the ligand-bound state, one residue moves from exposed to buried states, while seven residues transition from a buried

state to being exposed.

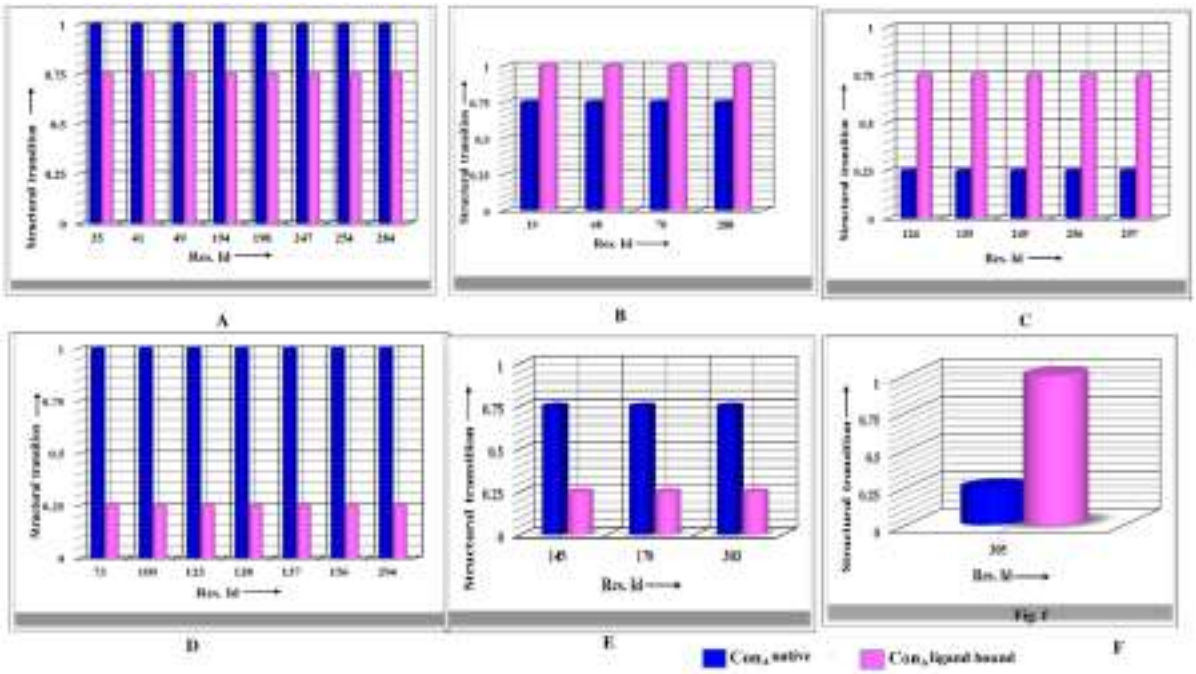

Figure 3: *The structural transitions of the residues in the Mpro enzyme are analysed based on the solvent-accessible surface area as it changes from its native to a ligand-bound state or vice versa, in the ConA form. Panel "A" illustrates the structural transition of residues in Mpro from the native ConA to the ligand bound state, indicating a shift from buried to semi-buried. Panel "B" shows the transition from semi-buried to buried residues. Panel "C" depicts the change from exposed to semi-buried. Panel "D" highlights the transition from buried to exposed states. Panel "E" demonstrates the shift from semi-buried to exposed residues, while Panel "F" illustrates the transition from exposed to buried.*

3.5. Comparative analysis between native ConAB to the ligand-bound

During the structural transition from the native ConAB to its ligand-bound state, a total of fifty-three residues were identified to undergo conformational changes, with shifts occurring between buried, semi-buried and exposed states. This transition is clearly illustrated in Figure 4. In panel A, eighteen residues (Arg4, Val35, Met49, Tyr118, Pro122, Ser123, Gly143, Cys156, Asp216, Asn221, Val233, Val247, Ile259, Met276, Ser284, Leu287, Cys300, Gly302) shift from a buried state to a semi-exposed state. This shift reflects a significant reorganization that may change the stabilization of their structural environment. In contrast panel B, thirteen residues (Gly11, Gln19, Thr21, Thr25, Leu57, Val171, Ala194, Ile213, Trp218, Leu220, Asn231, Ser254, Arg298) exhibit a transition from semi-exposed to buried, indicating a possible retraction into the protein core. Panel C highlights five residues (res. Id. Leu50, Arg60, Asn214, Ala285, Phe305), which move from a buried state to an exposed state, showing a dynamic aspect of the structure. In panel D reveals that nine residues (Leu67, Arg76, Lys100, Gln110, Arg222, Phe223,

Pro241, Gln256, Ser301) transition from exposed to semi-exposed, indicating a change in hydration accessibility. Lastly, in panel E, eight residues (res. Id. Ser1, Gly2, Thr45, Ser121, Asp155, Glu166, Glu270 and Thr280) transition from a semi-buried state to an exposed state, suggesting an increase in surface accessibility. Overall, the computational findings underscore the complexity of the structural transition from the native ConAB to the ligand-bound state, revealing that five residues transition from buried to exposed states.

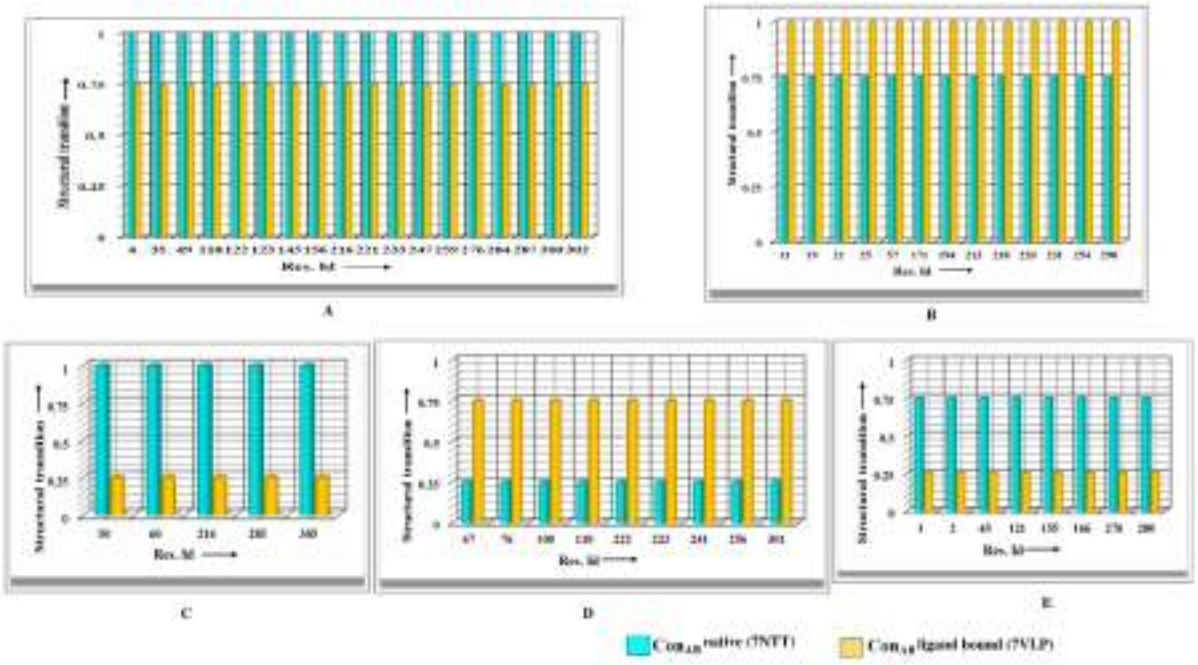

Figure 4: *The structural transitions of the residues in the Mpro enzyme are analysed based on the solvent-accessible surface area as it changes from its native to a ligand-bound state or vice versa, in the ConAB form. Panel"A" illustrates the structural transition of residues in Mpro from the native ConA to the ligand bound state, indicating a shift from buried to semi-buried. Panel "B" shows the transition from semi-buried to buried residues. Panel "C" depicts the change from buried to expose. Panel "D" highlights the transition from exposed to semi-buried states. Panel "E" demonstrates the shift from semi-buried to exposed residues.*

3.6. Comparative analysis between ConA and ConAB ligand-bound state

During the structural transition from the ConA ligand-bound state to the ConAB ligand bound state, a total of fifty-seven residues were identified as undergoing conformational changes. These changes occurred among buried, semi-buried and exposed states. This transition is clearly illustrated in Figure 5. In panel A, two residues (Asn214 and Arg279) shift from a buried state to an exposed state, suggesting an increase in surface accessibility. In panel B, thirteen residues (residue Ids Met6, Pro9, Gly11 and Ser139) transition from exposed to buried, indicating a change in stability. Panel C highlights eighteen residues (residue Ids Glu14, Thr21, Thr25, His41, Leu57,

Gly124, Val125, Tyr126, Gly138, Ser158, Val171, Ala194, Trp218, Leu220, Asn231, Ser254, Glu290 and Arg298) that move from a semi-exposed state to a buried state, indicating a possible retraction into the protein core. In contrast, panel D reveals that ten residues (Asp48, Ala70, Asn221, Val233, Ile259, Met276, Leu287, Glu288, Cys300 and Phe305) transition from a buried state to a semi-exposed state, showcasing a dynamic aspect of the structure. In panel E, nineteen residues (Ser1, Arg4, Leu67, Arg76, Lys100, Gln110, Ser123, Leu141, Tyr154, Gly170, Arg222, Phe223, Pro241, Gly275, Leu286, Ser301, Gly302, Val303 and Thr304) shift from an exposed state to a semi-exposed state. Lastly, in panel F, four residues (Asp155, Ala193, Asp229 and Glu270) transition from a semi-exposed state to an exposed state. Overall, the computational findings underscore the complexity of the structural transition from the ConA ligand-bound state to the ConAB ligand-bound state, revealing that four residues change from exposed to buried states while two residues transition from buried to exposed states.

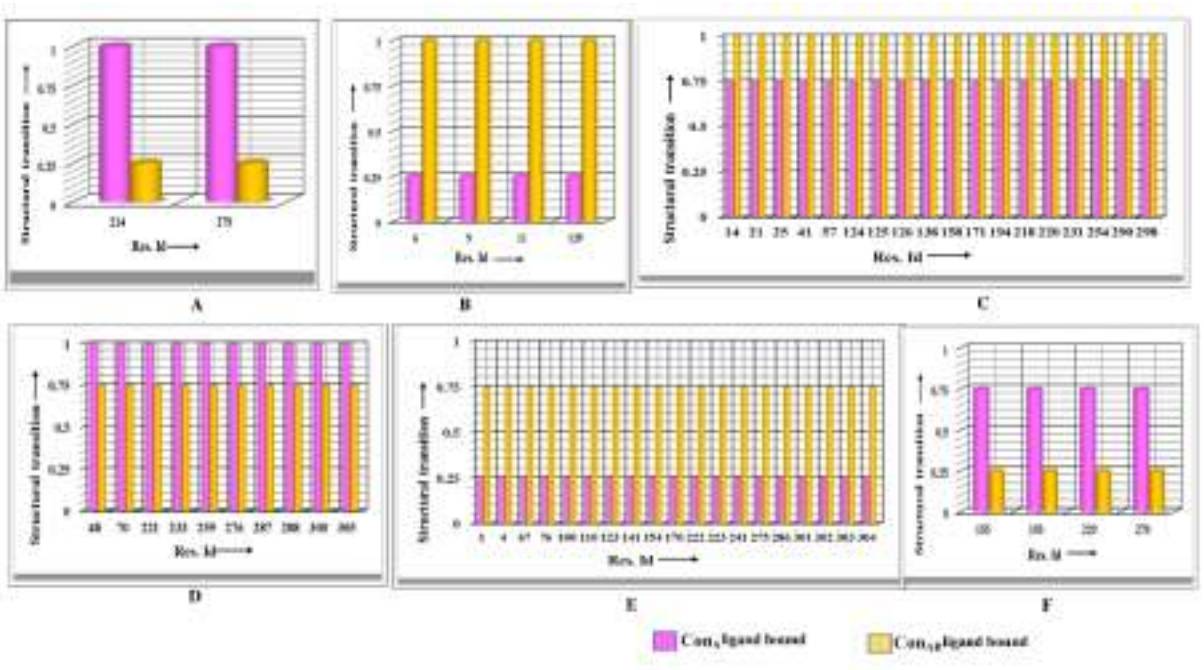

Figure 5: *The structural transitions of the residues in the Mpro enzyme are analyzed based on the solvent-accessible surface area as it changes from its ConA ligand-bound state to ConAB ligand bound state or vice versa. Panel "A" illustrates the structural transition of residues in Mpro from the native ConA to the ligand bound state, indicating a shift from buried to exposed. Panel "B" shows the transition from exposed to buried residues. Panel "C" depicts the change from semi-buried to buried. Panel "D" highlights the transition from buried to semi-buried states. Panel "E" demonstrates the shift from exposed to semi-buried residues, while Panel "F" illustrates the transition from semi-buried to exposed.*

The present study also suggests the potential binding occupancy of ligand-binding residues in both ConA and ConAB forms are very interesting (Figure 6). The catalytic residues His163, Glu166 and Cys145 exhibit a binding propensity of approximately 20% toward various ligands in both the ConA and ConAB forms. In

contrast, His41, Glu189, Asn142 and Arg188 show a lower binding propensity for the ligands. These observations indicate that most ligands in Mpro primarily bind to His163, Glu166 and Cys145, while potential alternative binding sites may include His41, Glu189, Asn142 and Arg188. The alternative binding sites of the Mpro enzyme, which contain the residues His41, Glu189, Asn142 and Arg188, are located away from the canonical active site. The volume of the binding pocket is approximately 770 Å^3, with a depth of 19.00 Å. In terms of functionality, the pocket has the ability to make 23 H bond donors and 57 H-bond acceptors, a hydrophobicity index of 0.42 and a drug score of 8.0. This pocket is particularly important for the binding of inhibitors. Both computational (molecular docking study followed by MD simulation) and experimental (crystallography) studies suggest that novel classes of acridinedione molecules, as well as the compound n- [(1R)-2-(tert-butylamino)-2-oxo-1-(pyridin-3-yl)ethyl]-n-(4-tert butylphenyl) furan-2-

carboxamide (ligand ID: 0EN), bind to the alternative binding site of the Mpro enzyme in the crystal structure noted as PDB Id. 7l0D. The Biochemical and computational experiments suggest and confirm the presence of alternative binding pocket of the Mpro enzyme [51-54].

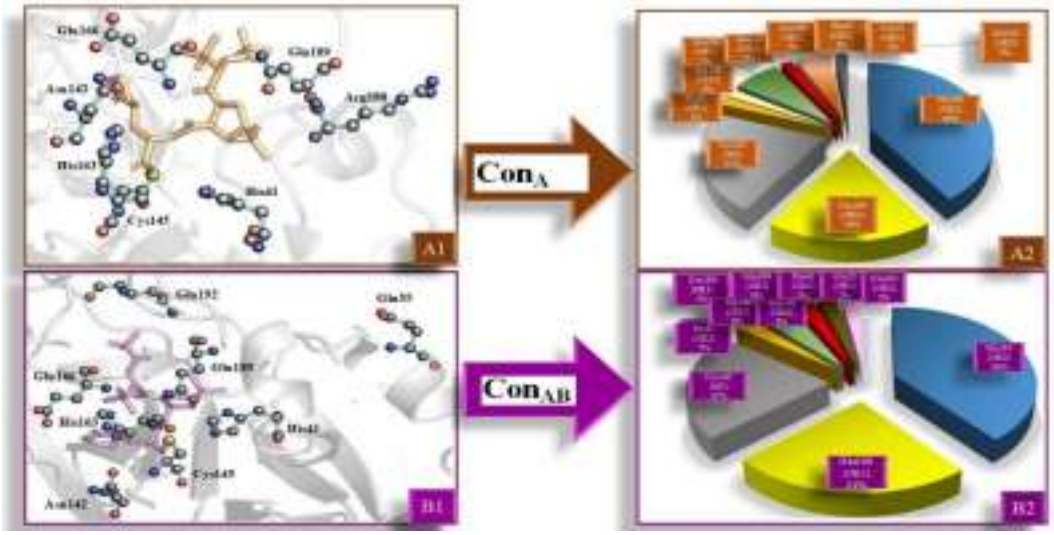

Figure 6: *The binding propensity of ligand-binding residues in the ligand-bound conformation of Mpro, both in its ConA and ConAB forms. In panels A1 and B1, we identify the residues located at the ligand binding sites. Additionally, panels B1 and B2 provide a schematic representation of the average occupancy percentages for these ligand-binding residues, highlighting their significance in the binding process.*

4. Discussion

Conformation A (ConA) refers to the dimeric structure of Mpro, which incorporates one molecule (A) in the crystal structures. In contrast, conformation AB (ConAB) specifically denotes the dimeric form that includes two molecules (A and B) in the crystal structures. When comparing the superimposed complexes of native ConA

and native ConAB, the structural deviation is less than 0.32 Å, while the deviation in their ligand-bound forms is less than 1.00 Å. Therefore, the active site cleft between ConA and ConAB does not show significant structural differences (Figure 7).

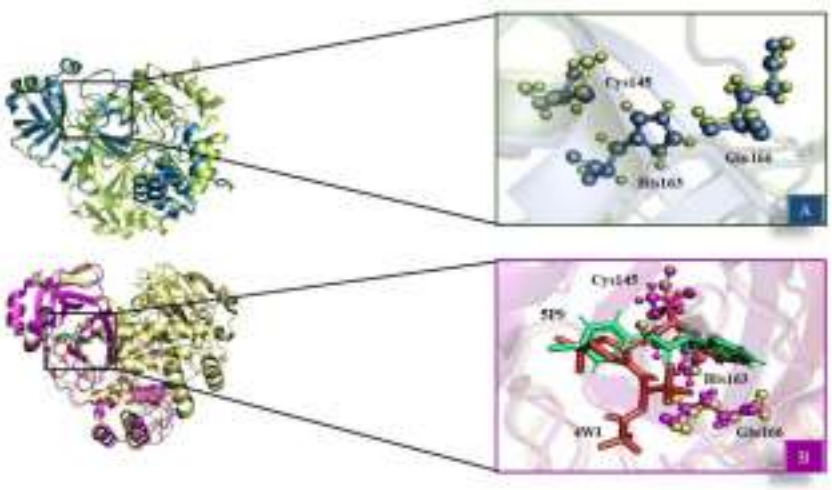

Figure 7: *Panel A illustrates the superimposition of ConA (PDB ID: 7DJR; Blue) with ConAB (PDB ID: 7NTT; Lime) based on their native crystal structures. Panel B depicts the superimposition of ConA (PDB ID: 7P51; Magenta) in a ligand-bound state (ligand ID: 5P9; Green) with ConAB (PDB ID: 7VLP; Pale Yellow) in a ligand-bound state (ligand ID: 4WI; Red). This comparison presents the crystal structures of Mpro, highlighting the active cleft.*

The conformational changes of amino acids in the Mpro enzyme are crucial for determining its stability, folding efficiency and functional specificity. These changes refer to the transition of residues from an exposed state to a buried state during various structural changes. In particular, during the transition from the native ConAof Mpro to its ligand-bound ConAB state, the following residues undergo a change: Arg4, Met6, Pro9, Leu50 and Arg60 from domain D1; Gly124 and Ser139 from domain D2; and Pro252, Gln256, Ala260, Ala285 and Gly302 from domain D3. Additionally, moving from the native ConA to its ligand-bound form also results in residues Val73 from D1, Lys100, Ser123, Cys128, Lys137, Cys156 from D2 and Phe294 from D3 transitioning from exposed to a buried state. In the structural change from native ConAB to its ligand-bound state, residues Leu50 and Arg60 from D1, as well as Asn214, Ala285 and Phe305 from D3, are affected. Furthermore, during the transition from the ligand

bound ConA state to the ligand-bound ConAB state, residues Asn214 and Arg279 of D3 also shift from an exposed to a buried conformation. It is possible that the hydrophilic buried residues contribute to the structural integrity of the Mpro enzyme, leading to infrequent conformational changes as they are tightly packed within the protein core.

Conversely, some residues undergo transitions from a buried to an exposed state, demonstrating increased rotameric flexibility and local conformational changes

driven by solvent accessibility. This dynamic nature enhances their adaptability in binding and function. Notable residues exhibiting this flexibility include Lys100 from D1, Lys137 and Val73 from D2, as well as Ser254, Arg279 and Phe294 from D3, especially in the native state of ConAB compared to their conformations in ConA. Additionally, Phe305 shows a transition from the native ConA state to the ligand-bound ConA state. In the case of conformation A, residues Met6, Pro9, Gly11 and Ser139 from domain D1 transition to the ConAB ligand-bound state, further illustrating the importance of these conformational changes. Possibly, the exposed residues of Domain D3 in conformation A help for dimerization process with another monomer of Mpro.

Site-directed mutagenesis and enzymatic activity assays have provided valuable insights into the roles of specific residues in the Mpro protein. Mutation study of His41, Gln189, Asn142 and Arg188 to alanine has been shown to loss of enzymatic activity, highlighting the critical function of these residues in the ligand binding site. Moreover, fluorescence-based live-cell assays utilize engineered reporter systems to generate a detectable signal upon Mpro cleavage. Interestingly, inhibiting Mpro activity effectively restores this signal [54,55].

5. Conclusion

The current hypothesis focuses on analyzing seventy-two crystal structures of Mpro enzyme, whether in conformation AB (ConAB) or conformation A (ConA) form, to better understand the structural changes that occur during the transition from its native to a ligand-bound state. The findings indicate that the ligand-bound conformations, in both forms (ConA and ConAB) of the Mpro enzyme, exhibit stability compared to their native state. Notably, the native ConAB form of Mprodemonstrates significant dynamic behavior, with several regions revealing increased flexibility relative to the other conformations. In the analysis of protein crystal structures, we observe that certain residues demonstrate significant flexibility, which is vital for their function. Possibly, binding of various ligands in the Mpro enzyme may induce conformational changes of the residues from exposed to buried or vice versa. Notably, Asn142 and Leu141 in the ConAB native structure, Asn277 and Thr304 in the ConAB ligand-bound form, Phe305 and Asn72 in the ConA native and Ser46 and Glu47 in the ConA ligand-bound structures all play important roles. The flexibility of the oxyanion loop, particularly due to the

movements of Phe140, Leu141 and Asn142, plays a crucial role in reshaping the S1 pocket of the Mpro enzyme. This pocket transitions from a closed form in its inactive state to an open form when interacting with substrates or inhibitors. These residues contribute to the dynamic nature of the substrate-binding site, which is essential for accommodating substrate and facilitating enzymatic activity. Additionally, these residues are vital for ligand selection, which is critical for the catalytic activity of Mpro and for the effective design of inhibitors.

The transitions from the native ConA to the native ConAB uncover important

insights, as forty-nine residues of Mpro enzyme undergo considerable conformational changes. In the move from the native ConA to the ligand-bound state of Mpro, twenty-eight residues play a significant role in this process. Additionally, fifty-three residues of Mpro enzyme show alterations when transitioning from the native ConAB to the ligand-bound state. Lastly, during the transition from the ConA ligand-bound state to the ConAB ligand-bound state, fifty-seven Mpro residues are involved in changes between buried and exposed states. More specially the hypothesis suggests that during the structural transition from the native ConA to the native ConAB state of the Mpro enzyme, twelve residues change their conformations from exposed to buried states, while six residues shift from buried to exposed states. When transitioning from the native ConA to the ligand-bound state of Mpro, one residue moves from exposed to buried, while seven residues transition from buried to expose. In the transition from the native ConAB to the ligand-bound state, five residues of Mpro transition from buried to exposed states. The conformational transitions observed in the main protease (Mpro) of SARS-CoV-2 upon ligand binding have been analysed through on the basis of structural studies. These transitions involve key and flexible residues of the Mpro protein that influence both catalysis and the drug ability of the new structural pocket in the enzyme. While other viral proteases may exhibit similar structural dynamics, their conformational responses to ligand binding can differ significantly due to variations in protein sequence, active site architecture and regulatory mechanisms. Thus, the conformational transitions described here, from the native to the ligand-bound states, may be considered specifically to SARS-CoV-2 Mpro. Although these conformational transitions may share some similarities within β coronaviruses and restricted to SARS-CoV-2. Finally, during the transition from the ConA ligand-bound state to the ConAB ligand-bound state of Mpro, four residues shift from exposed to buried states, while two residues transition from buried to exposed states. In addition, the present study also suggests that the majority of ligands in Mpro enzyme primarily interact with His163, Glu166 and Cys145. Moreover, potential alternative binding sites may be considered as His41, Glu189, Asn142 and Arg188, which could be significant for the development of anti-COVID molecules targeting the Mpro enzyme.

6. Acknowledgment

HRB and SG would like to express their sincere gratitude to the Department of Bioinformatics at Maulana Abul Kalam Azad University of Technology, West Bengal, for providing the computational facility and other necessary infrastructure.

References

1. Ullrich S, Nitsche C. The SARS-CoV-2 main protease as drug target. Bioorg Med Chem Lett. 2020;30:127377.

2. Suárez D, Díaz N. SARS-CoV-2 main protease: a molecular dynamics study. J Chem Inf Model. 2020;60:5815-31.

3. Sabbah DA, Hajjo R, Bardaweel SK, et al. An updated review on SARS-CoV-2 main proteinase:

protein structure and small-molecule inhibitors. Curr Top Med Chem. 2021;21:442-60.

4. Bzówka M, Mitusińska K, Raczyńska A, et al. Structural and evolutionary analysis indicate that the SARS-CoV-2 Mpro is a challenging target for small-molecule inhibitor design. Int J Mol Sci. 2020;21:3099.

5. Mohapatra RK, Sarangi AK, Kandi V, et al. Omicron (B. 1.1. 529 variant of SARS-CoV-2); an emerging threat: current global scenario. J Med Virol. 2022;94:1780.

6. Araf Y, Akter F, Tang YD, et al. Omicron variant of SARS-CoV-2: genomics, transmissibility and responses to current COVID-19 vaccines. J Med Virol. 2022;94:1825-32.

7. Jin Z, Du X, Xu Y, et al. Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. Nat. 2020;582:289-93.

8. Kim Y, Liu H, Kankanamalage ACG, et al. Reversal of the progression of fatal coronavirus infection in cats by a broad-spectrum coronavirus protease inhibitor. PLoS Pathog. 2016;12:e1005531.

9. Dai W, Zhang B, Jiang XM, et al. Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease. J Sci. 2020;368:1331-5.

10. Bairagya HR, Tasneem A, Sarmadhikari D. Structural and thermodynamic properties of conserved water molecules in Mpro native: a combined approach by md simulation and grid inhomogeneous solvation theory. Proteins: Struct Funct Bioinf. 2024;92:735-49.

11. Bairagya HR, Gupta S, Pal S. Identification of antibiotic drugs against SARS-CoV2 Mpro: a computational approach for drug repurposing. Int J Bioinform Intell Compu. 2024;3:70-90.

12. Reyaz S, Tasneem A, Rai GP, et al. Investigation of structural analogs of hydroxychloroquine for SARS-CoV-2 main protease (Mpro): a computational drug discovery study. J Mol Graph Model. 2021;109:108021.

13. Tasneem A, Rai GP, Reyaz S, et al. The possible molecular mechanism of SARS-CoV-2 main protease: new structural insights from computational methods. Sci Total Environ. 2021;2:108-26.

14. Świderek K, Moliner V. Revealing the molecular mechanisms of proteolysis of SARS-CoV-2 M pro by QM/MM computational methods. Chem Sci. 2020;11:10626-30.

15. Alagawani I, Wang F. In Silico development of SARS-CoV-2 non-covalent Mpro inhibitors: a review. Appl Sci. 2025;15:6544.

16. Ali MA, Sheikh H, Yaseen M, et al. Exploring the therapeutic potential of petiveria alliacea l. phytochemicals: a computational study on inhibiting SARS-CoV-2's main protease. Mol. 2024;29:2524.

17. Alagarsamy V, Shyamsundar P, Narendhar B, et al. Screening of some ayurvedic

phytochemicals to identify potential inhibitors against SARS-CoV-2 Mpro by in silico computational approach. Anti-Infect Agents. 2024;22:20-8.

18. Babalola BA, Adegboyega AE. Computational discovery of novel imidazole derivatives as inhibitors of SARS-CoV-2 main protease: an integrated approach combining molecular dynamics and binding affinity analysis. Covid. 2024;4:672-95.

19. Wadanambi PM, Jayathilaka N, Seneviratne KN. A computational study of carbazole alkaloids from murraya koenigii as potential SARS-CoV-2 main protease inhibitors. Appl Biochem Biotechnol. 2023;195:573-96.

20. Hernández-Serda MA, Vázquez-Valadez VH, Aguirre-Vidal P, et al. In silico identification of potential inhibitors of SARS-CoV-2 main protease (Mpro). Pathogens. 2024;13:887.

21. Yan G, Li D, Lin Y, et al. Development of a simple and miniaturized sandwich-like fluorescence polarization assay for rapid screening of SARS-CoV-2 main protease inhibitors. Cell Biosci. 2021;11:1-4.

22. Hicks EG, Kandel SE, Lampe JN. Identification of aloe-derived natural products as prospective lead scaffolds for SARS-CoV-2 main protease (Mpro) inhibitors. Bioorg Med Chem Lett. 2022;66:128732.

23. Chen L, Gui C, Luo X, et al. Cinanserin is an inhibitor of the 3C-like proteinase of severe acute respiratory syndrome coronavirus and strongly reduces virus replication in vitro. J Virol. 2005;79:7095-103.

24. Iketani S, Forouhar F, Liu H, et al. Lead compounds for the development of SARS-CoV-2 3CL protease inhibitors. Nat Commun. 2021;12:2016.

25. Ghosh AK, Raghavaiah J, Shahabi D, et al. Indole chloropyridinyl ester-derived SARS-CoV-2 3C Lpro inhibitors: enzyme inhibition, antiviral efficacy, structure–activity relationship and X ray structural studies. J Med Chem. 2021;64:14702-14.

26. Han SH, Goins CM, Arya T, et al. Structure-based optimization of ML300-derived, noncovalent inhibitors targeting the severe acute respiratory syndrome coronavirus 3CL protease (SARS CoV-2 3CLpro). J Med Chem. 2021;65:2880-904.

27. Amporndanai K, Meng X, Shang W, et al. Inhibition mechanism of SARS-CoV-2 main protease by ebselen and its derivatives. Nat Commun. 2021;12:3061.

28. Günther S, Reinke PY, Fernández-García Y, et al. X-ray screening identifies active site and allosteric inhibitors of SARS-CoV-2 main protease. Sci. 2021;372:642-6.

29. Chiou WC, Chen JC, Chen YT, et al. The inhibitory effects of PGG and EGCG against the SARS CoV-2 3C-like protease. Biochem Biophys Res Commun. 2022;591:130-6.

30. Du A, Zheng R, Disoma C, et al. Epigallocatechin-3-gallate, an active ingredient of traditional Chinese medicines, inhibits the 3CLpro activity of SARS-CoV-2. Int J Biol Macromol.

2021;176:1-2.

31. Liu SY, Wang W, Ke JP, et al. Discovery of camellia sinensis catechins as SARS-CoV-2 3CL protease inhibitors through molecular docking, intra and extra cellular assays. Phytomed. 2022;96:153853.

32. Park J, Park R, Jang M, et al. Therapeutic potential of EGCG, a green tea polyphenol, for treatment of coronavirus diseases. Life. 20210;11:197.

33. Nguyen TT, Woo HJ, Kang HK, et al. Flavonoid-mediated inhibition of SARS coronavirus 3C like protease expressed in pichia pastoris. Biotechnol Lett. 2012;34:831-8.

34. Hasegawa T, Imamura RM, Suzuki T, et al. Application of acoustic ejection MS system to high throughput screening for SARS-CoV-2 3CL protease inhibitors. Chem Pharm Bull. 2022;70:199- 201.

35. Xiao T, Cui M, Zheng C, et al. Myricetin inhibits SARS-CoV-2 viral replication by targeting Mpro and ameliorates pulmonary inflammation. Front Pharmacol. 2021;12:669642.

36. Stille JK, Tjutrins J, Wang G, et al. Design, synthesis and in vitro evaluation of novel SARS CoV-2 3CLpro covalent inhibitors. Eur J Med Chem. 2022;229:114046.

37. Chaves OA, Sacramento CQ, Fintelman-Rodrigues N, et al. Apixaban, an orally available anticoagulant, inhibits SARS-CoV-2 replication and its major protease in a non-competitive way. J Mol Cell Biol. 2022;14:mjac039.

38. Drayman N, DeMarco JK, Jones KA, et al. Masitinib is a broad coronavirus 3CL inhibitor that blocks replication of SARS-CoV-2. Sci. 2021;373:931-6.

39. Jin Z, Du X, Xu Y, et al. Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. Nat. 2020;582:289-93.

40. Han SH, Goins CM, Arya T, et al. Structure-based optimization of ML300-derived, noncovalent inhibitors targeting the severe acute respiratory syndrome coronavirus 3CL protease (SARS CoV-2 3CLpro). J Med Chem. 2021;65:2880-904.

41. Dampalla CS, Rathnayake AD, Perera KD, et al. Structure-guided design of potent inhibitors of SARS-CoV-2 3CL protease: structural, biochemical and cell-based studies. J Med Chem. 2021;64:17846-65.

42. Berman HM, Westbrook J, Feng Z, et al. The protein data bank. Nucleic Acids Res. 2000;28:235- 42.

43. Laskowski RA, MacArthur MW, Moss DS, et al. PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr. 1993;26:283-91.

44. Levitt M, Greer J. Automatic identification of secondary structure in globular proteins. J Mol Biol. 1977;114:181-239.

45. Rost B, Sander C. Conservation and prediction of solvent accessibility in protein families. Proteins Struct Funct Bioinf. 1994;20:216-26.

46. Carugo O. How large B-factors can be in protein crystal structures. BMC Bioinform. 2018;19:1-9.

47. Ringe D, Petsko GA. The 'glass transition'in protein dynamics: what it is, why it occurs and how to exploit it. Biophys Chem. 2003;105:667-80.

48. Kneller DW, Phillips G, O'Neill HM, et al. Structural plasticity of SARS-CoV-2 3CL Mpro active site cavity revealed by room temperature X-ray crystallography. Nat Commun. 2020;11:3202.

49. Zimmerman MI, Bowman G. SARS-CoV-2 simulations go exascale to capture spike opening and reveal cryptic pockets across the proteome. Biophys J. 2021;120:299.

50. Goyal B, Goyal D. Targeting the dimerization of the main protease of coronaviruses: a potential broad-spectrum therapeutic strategy. ACS Comb Sci. 2020;22:297-305.

51. Kumar V, Dhanjal JK, Kaul SC, et al. Withanone and caffeic acid phenethyl ester are predicted to interact with main protease (Mpro) of SARS-CoV-2 and inhibit its activity. J Biomol Struct Dyn. 2021;39:3842-54.

52. Gahlot D, Mehta D, Sehgal P, et al. Identification of novel alternative binding site for acridinedione derivatives on SARS-CoV-2 main protease: a combined crystallographic and in silico study. J Mol Struct. 2023;1280:135059.

53. Dallakyan S, Olson AJ. Small-molecule library screening by docking with PyRx. Chem Biol: Methods Protoc. 2014;1263:243-50.

54. Hsu MF, Kuo CJ, Chang KT, et al. Mechanism of the maturation process of SARS-CoV 3CL protease. J Biol Chem. 2005;280:31257-66.

55. Froggatt HM, Heaton BE, Heaton NS. Development of a fluorescence-based, high-throughput SARS-CoV-2 3CLpro reporter assay. J Virol. 2020;94:10-128.