Turkish Computational and Theoretical Chemistry

Turkish Comp Theo Chem (TC&TC)

Volume(Issue): 7(1) – Year: 2023 – Pages: 14-36

e-ISSN: 2602-3237

https://doi.org/10.33435/tcandtc.1121985

Received: 26.05.2022 ***Accepted:*** 07.08.2022 *Research Article*

An in silico investigation of allosteric inhibition potential of Dihydroergotamine against Sars-CoV-2 Main Protease (MPro)

M. Murat Yasar[a,f], Ekrem Yasar[b], Nuri Yorulmaz[c], M. Emin Tenekeci[d], Ismail H. Sarpun[e], Erol Eroglu[f,1]

[a]Vocational School of Health Services, Harran University, Sanliurfa, Turkey, 63300.
[b]Medical Faculty, Department of Biophysics, Akdeniz University, Antalya, Turkey, 07058.
[c]Faculty of Arts and Sciences, Department of Physics, Harran University, Sanliurfa, Turkey, 63300.
[d]Faculty of Engineering, Computer Engineering Department, Harran University, Sanliurfa, Turkey, 63300.
[e]Faculty of Sciences, Department of Physics, Akdeniz University, Antalya, Turkey, 07058.
[f]Faculty of Education, Department of Mathematics and Science Education, Akdeniz University, Antalya, Turkey, 07058.

Abstract: Possible allosteric inhibitors of MPro were investigated using in silico methods. To this end, FDA-approved drugs in the DrugBank database were subjected to a virtual screening, and drugs that strongly bind distant from the catalytic site of MPro were identified using molecular docking. Among the identified drugs, Dihydroergotamine (DHE) was chosen for further investigation due to its highest binding score against MPro in the molecular docking experiment. The allosteric inhibition potential of DHE toward MPro was demonstrated by applying some computational tools on the trajectory files which were obtained from the Molecular Dynamics Simulations. Results support that the hydrogen bonding interactions of DHE with GLU278 and THR280, located between Protomer A and Protomer B, affect the structure of the side chain of CYS145 at the catalytic site of MPro. Considering the role of CYS145 in the catalytic cycle, this structural change is likely to be a mechanism for inhibiting MPro.

Keywords: MPro, Main Protease, allosteric inhibition, Dihydroergotamine.

1. Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the causative agent of COVID-19, a single stranded RNA virus (+ssRNA) belonging to the Coronaviridae family, which has caused a global pandemic since the first quarter of 2020, threatens human health and public safety [1]. The size of coronavirus genome is in the range of 26 to 32 kb and comprise 6-11 open reading frames (ORFs) encoding 9680 amino acid polyproteins. The first ORF comprises approximately 67% of the genome, encoding 16 nonstructural proteins (nsps), while the remaining ORFs encode accessory and structural proteins [2,3]. Among the 16 encoded nsps, two viral cysteine proteases, main protease (nsp5) which is widely called as 3CLPro or MPro and papain-like protease (nsp3), were determined to play critical roles in the replication process of SARS-CoV-2. Besides nsps, ORFs also encode four major structural proteins: pointed surface spike (S), membrane (M), nucleocapsid protein (N), envelope (E), and accessory proteins. N-terminal glycosylated ectodomain is present at the N-terminal end of M protein that comprises of three transmembrane domains (TM) and a long C-terminal CT domain [4]. This first critical step of viral infection is catalyzed by its trimeric spike S glycoprotein which is a type I membrane protein and forms a trimer, anchored to the viral membrane by its transmembrane segment, while decorating the virion surface with it large ectodomain [5]. For the replication cycle, the mechanism of SARS-

[1] Corresponding Authors
e-mail: eeroglu@akdeniz.edu.tr

CoV-2 entry into a host cell starts with by anchoring to the ACE2 receptor of a host cell which is mediated by S protein, followed by its proteolytic activation, endocytosis, and membrane fusion (Jackson, Farzan, Chen, & Choe, 2022). Next to the anchoring to ACE2, S protein undergoes large structural rearrangements to promote membrane fusion [6]. Due to its important roles in viral infection and eliciting protective humoral and cell-mediated immune responses in hosts during infection [7], S protein is one of the primary target for existing vaccines as well as designing new antiviral therapeutics [8]. S protein of SARS-CoV-2 is cleaved into S1 and S2 subunits by proprotein convertases such as furin in the infected cells [9]. Divided subunits (S1 and S2) of S protein remain associated each other non-covalently with different functions: in the new target cell, while the S1 subunit binds the receptor and on the other hand the S2 subunit anchors the S protein to the virion membrane and mediates membrane fusion [10].

The main protease (so called MPro or 3CLPro) of SARS-CoV-2 virus is a cysteine enzyme critical for viral replication and transcription. Therefore, its inhibition can stall the production of infectious viral particles and thus alleviate disease symptoms of COVID-19 infection [11,12]. MPro contains a cysteine-histidine dyad at its catalytic center and cleaves viral polyproteins, pp1a and pp1ab, at 11 distinct sites with cleavage sequences LQ↓(S,A,G) generating 12 small functional proteins [4,13,15]. MPro has a length of 306 amino acids that form three different domains (domains I, II, and III). The substrate-binding site is located at a cleft between domain I (residues 8–101) and domain II (residues 102–184), which share a common fold consisting of antiparallel β-barrel structures [15]. Domain III (residues 201–303) is composed of an antiparallel globular cluster of five α-helices and takes a role in M-pro dimerization, which is essential for MPro activity. A QM/MM modeling study was conducted to explore the molecular mechanisms of proteolysis of MPro using the Ac-VAL-LYS-LEU-GLN-ACC polypeptide as a model substrate [16]. Results indicated that mechanism of action of MPro, taking place in four steps, slightly differs from that of other cysteine proteases. The first step of the catalytic cycle begins with the transfer of hydrogen which is bonded to the sulfur atom (SG) of CYS145 that becomes polarized when the substrate comes around the S1 pocket, to the nitrogen (NE2) of the imidazole ring of HIS41. Amongst the coronaviral targets that have been studied in the past, the MPro received major attention, particularly following the first SARS-CoV outbreak in the early 2000s [11,12]. The S protein, RNA-dependent RNA-polymerase (RdRp, nsp12), NTPase/helicase (nsp13) and papain-like protease (PLPro) are also recognized as alternative coronaviral drug targets [17].

When the literature is examined, there have been several studies involving the screening of chemical databases of small molecules using the virtual scanning method to rank them according to the predicted binding energy against MPro since the beginning of the COVID-19 outbreak. This strategy is quite useful in selecting the best candidates to be synthesized, acquired, or evaluated in wet experiments. In most of these virtual screening studies, the catalytic site of MPro was primarily targeted to find out appropriate covalent or non-covalent binders [18,19]. In some studies, the interface region between the protomers of MPro was chosen as the binding target of small molecules in order to prevent the dimerization process of MPro, because MPro can perform its function only when it is in the dimeric form [20-23]. Drug development studies based on MPro inhibition have been currently continued intensively using experimental and in silico methods by various research groups. The latest status in these studies was recently reviewed by [24].

Our main motivation in starting this research was to investigate allosteric inhibitors of this enzyme based on the assumption that MPro may have pocket of allosteric inhibitors. Because it was stated that allostery is an intrinsic property of all dynamic proteins and results from a redistribution of the conformational ensemble of a protein. Structural perturbation such as ligand binding at one site leads to a redistribution of the population [25]. Our main objective was to seek putative allosteric pockets of MPro and to find already approved drugs that bind to these pockets. In our approach, blind docking procedure was applied to drugs of the DrugBank database to filter out drugs that bind in pockets of MPro other than the catalytic site. Then, we tried to reveal whether the pocket to which drugs were bonded is a true allosteric pocket, using some validation tools.

2. Computational Method

2.1. Molecular docking and molecular dynamics simulation

Crystal structure coordinate of MPro (deposited in the PDB with accession number 6M03) was used for molecular docking and MD simulations. Its homodimer coordinates were downloaded from COVID-19 Proteins Library of CHARMM-GUI Archive [26]. Using MPro homodimer as a target, more than 2500 FDA-approved small molecule drugs from DrugBank database were subjected to virtual screening experiment using P_yR_x software [27,28]. P_yR_x is a virtual screening software in which AutoDock 4 docking simulation engine is implemented [28,29]. Our primary aim was to identify drugs with a high binding affinity without considering binding location within the enzyme, therefore we used the blind docking approach. Among the drugs with high docking scores, four drugs (namely Paliperidone, Dihydroergotamine (DHE), Balanol Analog 2, and Tadalafil) of which their binding affinity (K_d) was lower than -10 kcal/mol, were selected for further analysis. Each of these tightly bonded drugs was subjected to 20 ns MD simulations for a stability check. Among the four drugs, DHE exhibited the most stabile binding profile during the 20 ns MD simulation, therefore it was chosen for a further investigation to explore its allosteric inhibition potential toward MPro.

DHE-MPro dimer complex and apo-MPro dimer were subject to MD simulations using AMBER18 software package [30]. Cubic solvation boxes filled with TIP3P water model and protonation state of the titratable residues were prepared using the H++ webserver [31]. Appropriate number of Na+ ions was added to neutralize the system electrostatically. All simulations were performed by using the ff14SB force field [32]. The general amber force field (GAFF) parameters for DHE were obtained using the Antechamber and parmchk2 programs [33,34]. Partial atomic charges (AM1-BCC level of theory) needed for the parameterization was calculated using Gaussian 16 software [35,36]. For each of the simulation setups, minimizations were carried out with 20000 iterations in two parts. In the first part, minimization involving 10000 steepest descent steps was performed, while the positions of all heavy atoms were fixed by imposing harmonic restraints, and then followed by another 10000 steps of the conjugate gradient were run without any restraints. Minimized systems were gradually increased from 0 K to 310 K and equilibrated over 300 ps by an applying Langevin thermostat with a temperature coupling constant of 1.0 ps in the NVT ensemble. Density equilibrations (1 ns) and production runs (200 ns) were carried out using a constant pressure ensemble (NPT), the pressure was kept constant using a Berendsen barostat. All simulations were performed using periodic boundary conditions and a 2-fs time step. Long-range electrostatic interactions were calculated using the particle mesh Ewald method with a non-bonded cutoff of 9 Å, and the SHAKE algorithm was used to implement rigid constraints. RMSD, RMSF, hydrogen bonding and dihedral angle analysis of the obtained MD trajectory files were performed using CPPTRAJ module of the AMBER18 software package [37].

2.2. Dynamics cross-correlation (DCC)

The calculation of correlated motions between alpha carbons of the residues was performed using MD-TASK web server [38]. Initially, all atoms other than alpha carbons, C_α of the trajectory files were deleted using CPPTRAJ module. Next, resulting trajectory files were uploaded MD-TASK webserver to calculate the Pearson correlation coefficient from covariance matrix elements according to the following formula.

$$C_{i,j} = \frac{\langle (r_i - \langle r_i \rangle) \cdot (r_j - \langle r_j \rangle) \rangle}{\sqrt{(\langle r_i^2 \rangle - \langle r_i \rangle^2)(\langle r_j^2 \rangle - \langle r_j \rangle^2)}} \quad (1)$$

where bracket-enclosed quantities represent time-averaged values over the trajectory, and r_i and r_j are the positional vectors of C_α of residue i and j , respectively. Cross-correlation values span the range from -1 to +1. Values -1, 0, and 1 indicate perfect anti-correlation, no-correlation, and perfect correlation between fluctuations of Cα of residues i and j respectively.

2.3. Communication propensity (CP)

CPof a residue pair characterizes the efficiency of communication between residues i and j, which is based on the idea that signal transduction events in proteins are directly related to the distance fluctuation of the communicating atoms. This approach was developed by [39], firstly utilized to

the equilibrium fluctuations of residue pairs obtained by Elastic Network model calculations. Subsequently the approach was adapted for utilizing the equilibrium fluctuations of C_α of residue pairs obtained by classical MD simulation data [40]. CP of a residue pair is described as a function of the inter-residue distance fluctuations, where residues whose C_α-C_α distance fluctuates with low intensity, means that communications are more efficiently (faster) compared to residues whose distance fluctuations are large in which the amplitude of the fluctuations results in slower inter-residue communication. CP is calculated as the mean-square fluctuation of the inter-residue distance as fallow.

$$CP = \langle \left(d_{ij} - d_{ij,ave}\right)^2 \rangle \qquad (2)$$

where d_{ij} is the time dependent distance between the C_α atoms of residues i and j respectively and the bracket-enclosed quantities represent time-averaged values over the trajectory. In this study, CP values of residue pairs of DHE-MPro dimer complex and apo-MPro dimer systems were calculated using MD-TASK web server [38].

2.4. Dynamic residue network analysis (DRN)

To analyze communications between residues within a protein, a residue interaction network (RIN) is constructed, of which the C_β atoms of residues (C_α for glycine) are treated as nodes within the network, and edges between nodes defined within a distance cut of 6.7 Å [41]. Two graph theoretical metrics namely the shortest path length (L_{ij}) and the betweenness centrality (BC) were used for the dynamic residue network analysis of DHE-MPro dimer complex and apo-Mpro dimer systems. To determine impact of DHE binding on the properties of the dynamic residue network of MPro dimer, ΔBC (average BC of DHE-MPro complex minus average BC of apo-MPro) and ΔL_i (average L_i of DHE-MPro complex minus average L_i of apo-MPro) were calculated. In graph theory, the reachability of a residue is defined as the number of connections required to reach residue i from j using the shortest possible path. The average reachability of a residue L_i is thus defined as the average number of steps required to reach residue i from any other residue in the network. L_i of a residue gives an idea of how accessible the residue is within the protein. The metric BC is related to L_i in that it is a measure of how often on average a residue is utilized in shortest path navigation [42]. Constructions of DRN for each MD trajectory, in which RINs are constructed for every nth frame of the trajectory using a 200 ps time interval, to build a DRN matrix, were carried out using MD-TASK webserver [38]. The average L_{ij} is then calculated as the average number of steps that the residue may be reached from all other residues in the RIN as fallow:

$$L_i = \frac{1}{N-1}\sum_{j=1}^{N} L_{ij} \qquad (3)$$

BC is defined as the number of shortest paths running through a residue for a given RIN and provides a measure of usage frequency each residue during navigation of the network. Residues in a protein that have a high BC reveal locations that tend to be important for controlling inter-domain communication in a protein. MD-TASK webserver in which the calculation of BC is accomplished using the Dijkstra's algorithm [43], was also used to calculate BCs of the residues for DHE-MPro and apo-Mpro dimer systems as according to fallowing formula.

$$L_i = \frac{1}{N-1}\sum_{j=1}^{N} L_{ij} \qquad (4)$$

where, all s-t residue pairs that traverse a given residue V along their geodesic distance. For each residue, this value is averaged over each frame i from the total m frames over the trajectory.

2.5. Transfer entropy analysis (TE)

In this study, the information theory measure of transfer entropy was used, is based on a concept developed by Schreiber [44]. The concept subsequently was adapted to classical MD simulations data by Kamberaj and van der Vaart [45]. Implementation of transfer entropy on MD simulation trajectory addresses the causal relationship between fluctuations of residue pairs. If the motion of residues *i* and *j* is correlated, does residue *i* drive the motion of residue *j*, or does residue *i* respond to the motion of *j*? Note that DCC contains only linear statistical dependencies and is sensitive to the relative spatial orientation of the motions. Assume that in equation 1, if fluctuation vectors of residues *i* and *j* are perpendicular each

other and perfectly correlated, but the dot product of the vectors in the numerator results in orthogonal motions always yielding a correlation of zero. Therefore, by DCC, even perfectly correlated motions may be unobserved if they are perpendicular to one another. Like mutual information, transfer entropy contains all pairwise linear and nonlinear statistical dependencies and is insensitive to the relative spatial orientation of the motions. Furthermore, an important feature that distinguishes transfer entropy from mutual information and DCC is that it is not symmetrical. So, the transfer entropy from residue i to j does not equal the transfer entropy from residue j to i. This allows us to determine which residue is the driver and which residue is the driven (responder). This driver-driven relation that maps the causality between correlated fluctuations of a pair of residues should be of great importance for allosteric mechanisms because it points out which residues to manipulate. In this regard, a driver residue is more critical than the driven residue and manipulating the driver will be perturb the existing correlations more efficiently.

In this study, we will present the directionality indexes of selected pairs of residues for apo-MPro and DHE-MPro complex. A complete derivation of formula for the directionality index (net transfer entropy of a pair of residues) is given elsewhere [45]. Here, we present a brief definition for the quantities to calculate the transfer entropy between a pair of residues. For a given MD simulation trajectory, when the positional fluctuations of atoms (C_α) i and j are considered, $r_i(t_n) \equiv |X_i(t_n) - \langle X_i \rangle|$where X is the atomic position and $\langle\ \rangle$ denotes a time average over the trajectory. Here, $t_n = n\Delta t$ is discrete time, with $\Delta t = M\delta t$ a multiple of the MD integration time step δt and M the frequency of saving coordinates. If the discrete time processes as stationary Markov processes of order m is described, the dynamics of the system can be represented by m-dimensional state vectors, which is denoted by I. Using the method of time delayed embedding; the state vectors are given by

$$\mathrm{I}_k^{\mu_\mathrm{i}} \equiv \left(r_\mathrm{i}(t_k), r_\mathrm{i}\left(t_{k+\tau_\mathrm{i}}\right), \cdots, r_\mathrm{i}\left(t_{k+(m_\mathrm{i}-1)\tau_\mathrm{i}}\right)\right)^\mathrm{T} \quad (5a)$$

$$\mathrm{J}_k^{\mu_\mathrm{j}} \equiv \left(r_\mathrm{j}(t_k), r_\mathrm{j}\left(t_{k+\tau_\mathrm{j}}\right), \cdots, r_\mathrm{j}\left(t_{k+(m_\mathrm{j}-1)\tau_\mathrm{j}}\right)\right)^\mathrm{T} \quad (5b)$$

where the time shift τ is a multiple of Δt, and the superscript T indicates the transpose operation. The m and τ embedding parameters can be different for atoms i and j; their specific values are indicated by the superscript $\mu \equiv (m, \tau)$. Taking time $t_{k+(m_\mathrm{i}-1)\tau_\mathrm{i}}$ as the present, $\mathrm{I}_k^{\mu_\mathrm{i}}$ and $\mathrm{J}_k^{\mu_\mathrm{j}}$ describe the histories of the processes, which consist of the present and past $(m-1)$ fluctuations. Similarly, the future fluctuations of atoms i and j are given by

$$\mathrm{I}_{k+1} \equiv r_\mathrm{i}\left(t_{k+m_\mathrm{i}\tau_\mathrm{i}}\right) \quad (6a)$$

$$\mathrm{J}_{k+1} \equiv r_\mathrm{j}\left(t_{k+m_\mathrm{j}\tau_\mathrm{j}}\right) \quad (6b)$$

In equation 6a and 6b, the correct choice of the m and τ embedding parameters is crucial for the proper characterization of the structure of the time series for a given dynamical system. Software so-called Symbolic Information Flow Measurement (SIFM) which was used in this study, is designed to calculate the transfer entropy and the mutual information between pairs of residues for a given MD simulation trajectory [46,47]. SIMF can optimize the embedding parameters τ and m individually. Algorithms implemented in SIFM firstly handle the optimization of τ for each atom by means of adjusting the number of bins that maximize the Shannon information entropy. Then m value for each atom is optimized by using the false nearest-neighbors method.

If the fluctuations of atoms i and j are independent processes, the conditional probability to observe the future fluctuation of atom i, given its history, is independent of the history of atom j,

$$p\left(\mathrm{I}_{k+1} \middle| \mathrm{I}_k^{\mu_\mathrm{i}}\right) = p\left(\mathrm{I}_{k+1} \middle| \mathrm{I}_k^{\mu_\mathrm{i}}, \mathrm{J}_k^{\mu_\mathrm{j}}\right) \quad (7)$$

where the left side of the equation 7 is the conditional probability to observe the future fluctuation of atom i, given its history, and the right side is the conditional probability to observe the future fluctuation of atom i, given the histories of atoms i and j. The transfer entropy quantifies the deviation from independence, and is given by the Kullback-Leibler distance between these probability distributions [44-48];

$$T_{\mathrm{j}\to\mathrm{i}} = \sum_{k=0}^{N_m} p\left(\mathrm{I}_{k+1}, \mathrm{I}_k^{\mu_\mathrm{i}}, J_k^{\mu_\mathrm{j}}\right) \log \frac{p\left(\mathrm{I}_{k+1} \middle| \mathrm{I}_k^{\mu_\mathrm{i}}, \mathrm{J}_k^{\mu_\mathrm{j}}\right)}{p\left(\mathrm{I}_{k+1} \middle| \mathrm{I}_k^{\mu_\mathrm{i}}\right)} \quad (8)$$

where $p\left(\mathrm{I}_{k+1}, \mathrm{I}_k^{\mu_\mathrm{i}}, J_k^{\mu_\mathrm{j}}\right)$ is the joint probability distribution of observing the future atomic fluctuation of atom i, and the histories of atoms i and j. All possible state vectors is summed: $N_m = N - (m-1)\tau$, with N the total number of the frames of the trajectory. The transfer entropy is a positive quantity, its value ranges from zero to maximum of the entropy rate at that point the fluctuations of atoms i and j are independent

$$h\left(\mathrm{I}_{k+1} \middle| \mathrm{I}_k^{\mu_\mathrm{i}}\right) = -\sum_{k=0}^{N_m} p\left(\mathrm{I}_{k+1}, \mathrm{I}_k^{\mu_\mathrm{i}}\right) \log p\left(\mathrm{I}_{k+1} \middle| \mathrm{I}_k^{\mu_\mathrm{i}}\right) \quad (9)$$

when the fluctuations of atoms i and j are completely coupled; this maximum is reached when i=j. It was shown that the transfer entropy can be reformulated as conditional mutual information (I) [49]:

$$T_{\mathrm{j}\to\mathrm{i}} = I\left(\mathrm{I}_{k+1}; \mathrm{J}_k^{\mu_\mathrm{j}} \middle| \mathrm{I}_k^{\mu_\mathrm{i}}\right) \quad (10)$$

Transfer entropy is based on conditional probabilities, thus $T_{\mathrm{j}\to\mathrm{i}} \neq T_{\mathrm{i}\to\mathrm{j}}$ in general. This inequality can be used to identify causal relationships. High values of $T_{\mathrm{j}\to\mathrm{i}}$ indicate that the fluctuations of atom j are strongly driving the fluctuations of atom i, in the opposite case dependence between atoms i and j are low. Normalized directional index, ($D_{\mathrm{j}\to\mathrm{i}}$) can be used to quantify, whether the fluctuation of atom j drives the fluctuation of atom i, or vice versa the fluctuation of atom j is driven by the fluctuation of atom i.

$$D_{\mathrm{j}\to\mathrm{i}} = \frac{T_{\mathrm{j}\to\mathrm{i}}}{h\left(\mathrm{I}_{k+1} \middle| I_k^{\mu_\mathrm{i}}\right)} - \frac{T_{\mathrm{i}\to\mathrm{j}}}{h\left(\mathrm{J}_{k+1} \middle| \mathrm{J}_k^{\mu_\mathrm{j}}\right)} \quad (11)$$

where $D_{\mathrm{i}\to\mathrm{j}} = -D_{\mathrm{j}\to\mathrm{i}}$, with values -1 and +1. In the case of $D_{\mathrm{j}\to\mathrm{i}} > 0$, the fluctuations of atom j drive the fluctuations of atom i, in the opposite case $D_{\mathrm{j}\to\mathrm{i}} < 0$, the motion of atom i drives the motion of atom j.

3. Results and discussion

3.1. *Dynamics of apo-MPro and DHE-MPro complex*

We performed two replicas of 200 ns md simulations for both DHE-MPro and apo-MPro dimer systems. Figure 1 (A) demonstrates DHE-MPro heterodimer system which consists of protomer A and protomer B (dark grey). Each protomer comprises three domains which are presented with different color for protomer A in Figure 1 (A). Locations of DHE binding site and catalytic dyad (HIS41/CYS145) which are about 40 Å apart from each other are shown as the zoomed views in the Figure 1 (B) and Figure 1 (C) respectively. Figure S1 (Supplementary Material) shows plots of RMSD for the heavy atoms of the DHE-MPro complex and apo-MPro dimer systems for both replicas. As can be seen, both replicas exhibit similar results, on that account, henceforth for remaining analyses of the trajectory were carried out using the data of replica 1. Plots of RMSF of the DHE-MPro complex and apo-MPro dimer systems are presented for protomer A and B in Figure 2 (A) and (B) respectively. Comparing Figure 2 (A) and (B), it is seen that DHE binding to MPro affects the fluctuations of residue in Protomer A and Protomer B differently. The most obvious difference arises in the fluctuations of some of linker residues located between domain II and domain III. While DHE binding appears to increase the fluctuations of linker residues in Protomer A, on the contrary, it stabilizes fluctuations of linker residues in Protomer B. In the following section, to explore whether ligand interactions in the interface region affect the structure of catalytic site or not, we focused to determine any structure changes come about in the structure of catalytic dyad residues due to these interactions.

3.2. Effect of DHE binging on the catalytic site structure: Hydrogen bonding analysis, distance, and dihedral measures

Hydrogen bonding interactions between DHE and the interface residues were monitored by CPPTRAJ program which is a part of Amber18 simulation package. Figure S2 demonstrates hydrogen bonding pattern over the trajectory. Perusal of Figure S2 indicates that in the first quarter part of the trajectory hydrogen bonds occur mostly between the hydrogen bound to oxygen DHE (O4) and THR280 (HG1), GLY278 (O). These interactions occur in 14637 frames which about 29 percent of total simulation time. We wondered if there was a DHE binding-related change in the structure of

CYS145 and HIS41, which are catalytic site residues during the simulation. To understand this, we tracked the sidechain torsion angel, χ1 of CYS145 and the distance between the SG atom of CYS145 and the NE2 atom of HIS41 over frames of the trajectories. Figure 3 (A) and (C) demonstrate χ1 angles of CYS145 in the trajectories of DHE-MPro complex and apo-MPro respectively. Figure 3 (B) and (D) demonstrate the SG-NE2 distances. In the trajectory of the DHE-MPro complex, it was observed that χ1 angle of CYS145 changes from 60 degree to 180 degree at the end of the first quarter, it remains at 180 degrees over the trajectory except for about 2500 frames (10 ns) in the last quarter where it is 60 degrees again. In the trajectory of the Apo-MPro system, the χ1 angle of CYS145 mostly remains at 180 degrees over the trajectory, except for a small portion in the first quarter where it turns 60 degrees for a while. It is seen that when the χ1 angle is at 180 degrees, the SG-NE2 distance is about 3.5 Å. If the χ1 angle turns from 180 degrees to 60 degrees, the SG-NE2 distance increase to 4.5 Å. A perusal of Figure 3 (A) and (B) indicate that the SG-NE2 distance is longer than 4.5 Å, even if where the χ1 angle remains at 180 degrees, especially in the second quarter of the trajectory. The reason for this large distance may be due to the movement of the loop where CYS145 is in domain II.

Two superimposed representative snapshots of χ1 angles of CYS145 from two different frames shown in Figure 1 (C). Henceforth, remaining of this article, when χ1 angles of CYS145 are 60 and 180 degrees will be termed 'out conformation' and 'in conformation' respectively. Changing the χ1 angle of CYS145 should cause a change in the distance between HIS41 (NE2) and CYS145 (SG) atoms, which is vital for MPro catalytic activity. A computational study was carried out in which MPro catalyzed proteolysis reaction of the polypeptide Ac-Val-Lys-Leu-Gln-ACC (ACC is the 7-amino-4-carbamoylmethylcoumarin fluorescent tag) was modeled [16]. Authors showed that a proton (HG) transfer is necessary from CYS145 to nitrogen (NE2) of HIS41 for the initial step of

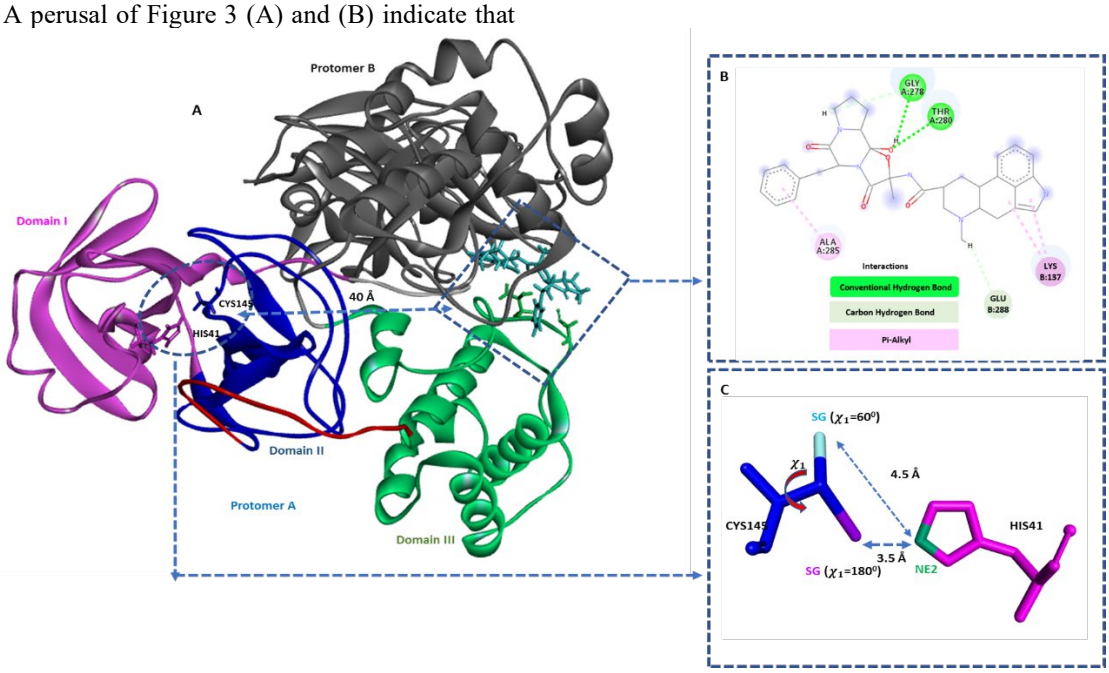

Figure 1. Illustration of MPro dimer. (**A**) Domains I, II, and III are colored in purple, navy blue, and green respectively for protomer A. Linker residues between domains II and III are colored in red. Protomer B is colored in gray. Catalytic dyad residues are framed by an ellipse. The proposed allosteric pocket (DHE binding site) is 40 Å away from the catalytic dyad and framed by a rectangle. (**B**) Inset-zoomed in view of interactions of DHE with the surrounding residues. (**C**) Inset-zoomed in view of two conformations (in and out) of the side chain of CYS145 and HIS41.

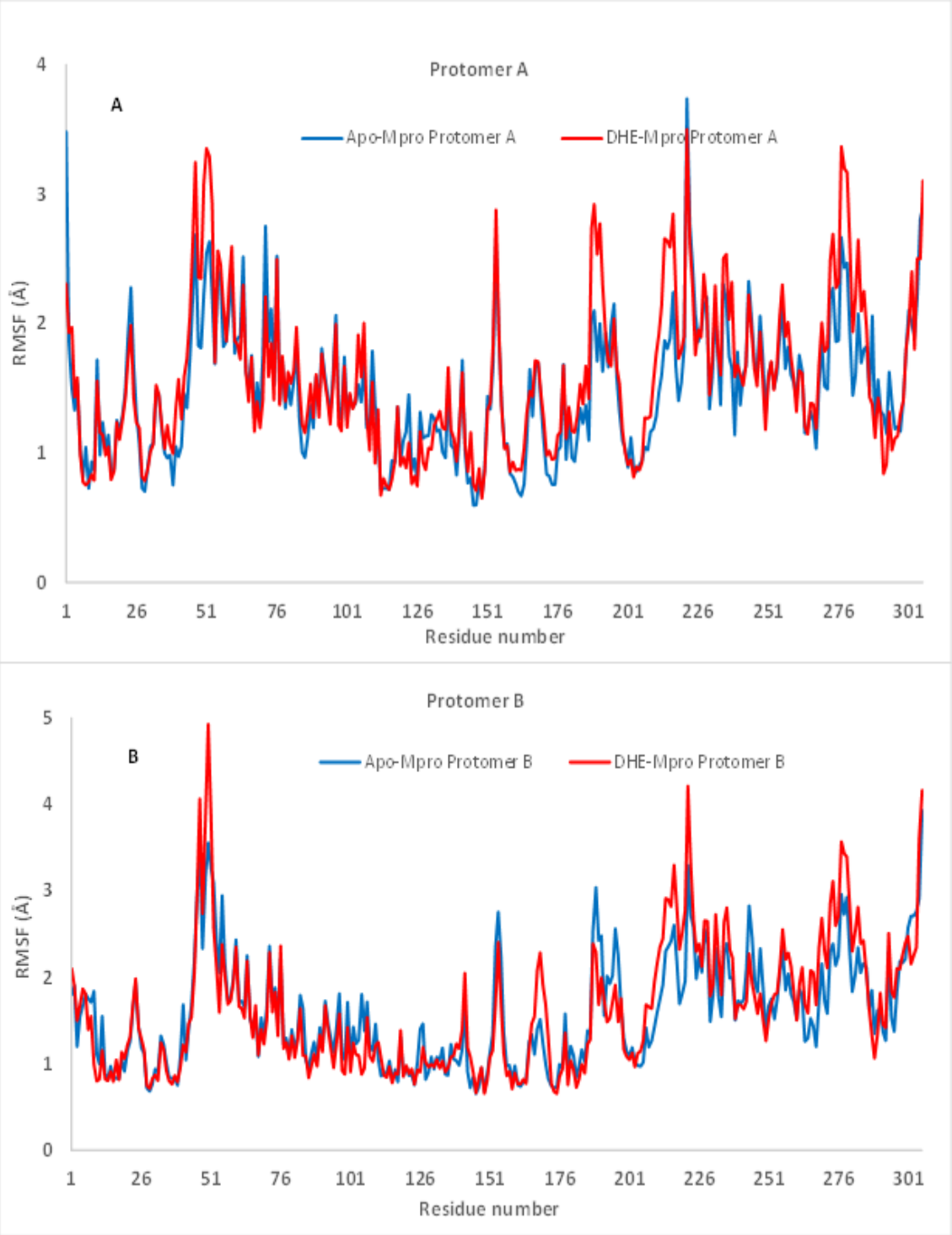

Figure 2. Fluctuations of residues in Protomer A and Protomer. (A) Fluctuations of the residues in Apo-MPro and DHE-MPro complex for Protomer A are shown in blue and red colors respectively. (B) Fluctuations of the residues in Apo-MPro and DHE-MPro complex for Protomer B are shown in blue and red colors respectively.

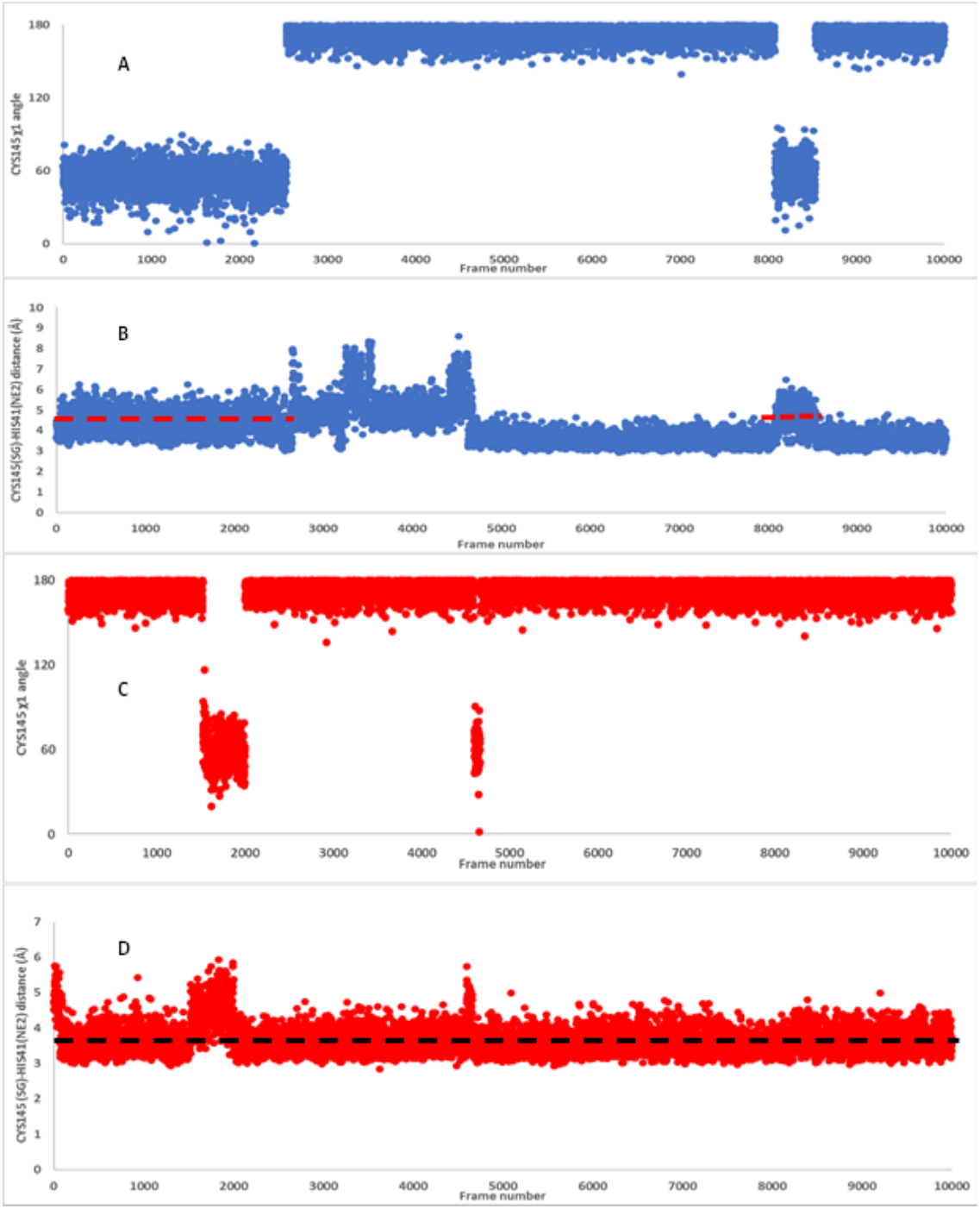

Figure 2. Fluctuations of residues in Protomer A and Protomer. (A) Fluctuations of the residues in Apo-MPro and DHE-MPro complex for Protomer A are shown in blue and red colors respectively. (B) Fluctuations of the residues in Apo-MPro and DHE-MPro complex for Protomer B are shown in blue and red colors respectively.

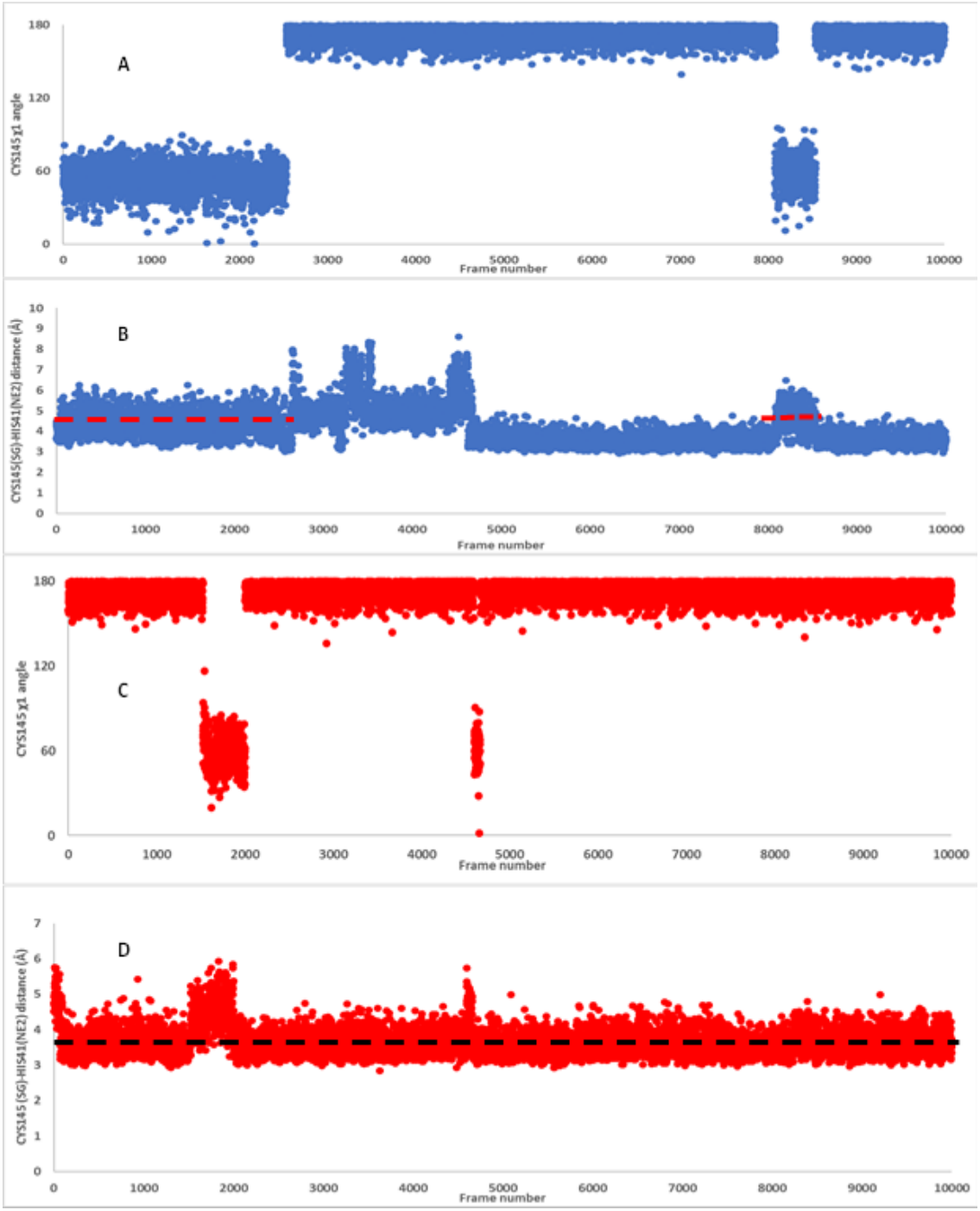

Figure 3. The χ1 angles of CYS145 and the SG-NE2 distances over the trajectories. (**A**) Change in value of the χ1 angle of CYS145 for the DHE-MPro complex. (**B**) Change in value of the SG-NE2 distance for the DHE-MPro complex. (**C**) Change in value of the χ1 angle of CYS145 for the Apo-MPro, and (**D**) Change in value of the SG-NE2 distance for the Apo-MPro.

the proteolysis reaction. Then, the reaction proceeds with a nucleophilic attack on the carbonyl carbon atom of the peptide bond by the sulfur (SG) of CYS145, thus leading to a thiohemiketal (THA) intermediate. Based on above result, we can argue that the distance between hydrogen donor SG and acceptor NE2 is a critical key to initialize the proteolysis reaction. We may hypnotize that if somehow this distance is kept longer than a critical value since the proton transfer reaction cannot occur so that the catalytic function of MPro diminishes. Based on the above findings, there may be a relationship between the hydrogen bond interactions occurring at the dimer interface and the

χ1 angle of CYS145 at the catalytic site. Because if Figure S2 and Figure 3 (B) are evaluated together, it is observed that the χ1 angle is 60 degrees when DHE makes hydrogen bonds with THR278 and GLU280 (in the first quarter of the trajectory).
Change in the χ1 angle of CYS145 from out conformation (60 degree) to in conformation (180 degree) results in shorter SG-NE2 distance. In the case of short SG-NE2 distance, proton transfer from SG to NE2 more likely happens, which is needed to initialize the proteolysis reaction [16]. Therefore, the remaining of this paper, 'in conformation' and 'out conformation' will be termed ‘active state’ and inactive state’ respectively. It seems that certain hydrogen bonding interactions between DHE and the surrounding amino acids (especially with GLY278 and THR280) likely affect the structure of the catalytic site allosterically. Therefore, GLY278 and THR280 residues together with their spatial neighbors GLY2, PHE3, ARG4, GLN127, ASP216, ARG217, SER284 and ALA285 henceforth will be termed allosteric residues. Throughout the following sections, we will try to rationalize the existence of such an allosteric effect using different trajectory analysis tools.

3.3. Dynamics cross-correlation (DCC) and Communication propensity (CP)

As mentioned in the previous sections, an examination of DHE-MPro complex trajectory over the course of 200 ns simulation revealed that DHE is quite mobile in the dimer interface region, and it exhibits different interactions with the surrounding residues. In the first quarter, part of the simulation, the catalytic site of MPro is in an inactive state, whereas the remaining of the simulation is an active state, except for about 2500 frames (10 ns) in the last quarter of the trajectory. Here, the differences in cross-correlations and communication propensity of the residues for the active state and inactive state as well as the apo-MPro state were explored. Figure 4 (A), (B), and (C) show the dynamics cross-correlation maps for the active and the inactive states of the DHE-MPro complex, as well as for the apo-MPro states respectively. An examination of the dynamics cross-correlation maps indicates that the overall inter-domain correlations between domain I and domain II are higher in the active state in comparison to that of the inactive state of DHE-MPro complex and apo-MPro. In the active state, considering the interactions of DHE with surrounding amino acids (especially hydrogen bonds with GLY278 and THR280), it may be interpreted as these interactions increase signal transduction between domain I and domain II. If the motions of neighboring amino acids are highly correlated with one another, this can give rise to a “domino effect” in which perturbations to one amino acid can produce long-range allosteric effects by propagating through networks of highly correlated neighbors [50]. The increased correlation between domain I and domain II due to DHE binding and the change in the χ1 angle of the catalytic site of CYS145 support the possibility that the DHE binding site is an allosteric site.
CP maps of the residue pairs for the active state and the inactive state of the DHE-MPro complex, the apo-MPro are shown in Figure 5 (A), (B), (C) respectively. Figure 5 (D) represents the difference in values of CP between the inactive state of DHE-MPro and the apo-MPro. It is worth mentioning that CP is directly related to communication time, hence a low CP value of a residue pair indicates that they communicate with each other more efficiently [39,40]. A perusal of figure 4 indicates that for all three states, the maps of CP exhibit a similar trend, in which a significant number of domain III residues communicate with domain I and domain II residues less efficiently in comparison with the remaining ones. Figure 5 (D), which was obtained by subtracting CP values of the residue pairs of apo-MPro from the inactive state of DHE-MPro ones, may help us to determine whether communication efficiency is increased upon DHE binding or not. In Figure 5(D), negative values of residue pairs (dark blue to yellow) indicate increased CP values upon DHE binding. This result points out that CP values between the domain III residues and the rest of MPro are elevated due to DHE binding. Our particular interest focused on change in CP values between the allosteric site and catalytic site residues upon DHE binding. To investigate this aspect more thoroughly, we outlined CP values of residue pairs between both sites in Table 1. For both apo-MPro and DHE-MPro complex, first lines and columns of data in Table 1 represent average CP values of the residues which were calculated over whole dimer system. Note that the average CP values of ASN277, GLY278, and ARG279 of the DHE-MPro complex and GLY215, ASP216, ASN277, GLY278, ARG279, and THR280 of the Apo-MPro are quite high in comparison to other residues in Table 1. In the Apo-MPro, the CP value between GLU278 and the catalytic site amino acids was the highest among the selected amino acid pairs, while it is observed that these values decreased dramatically in response to

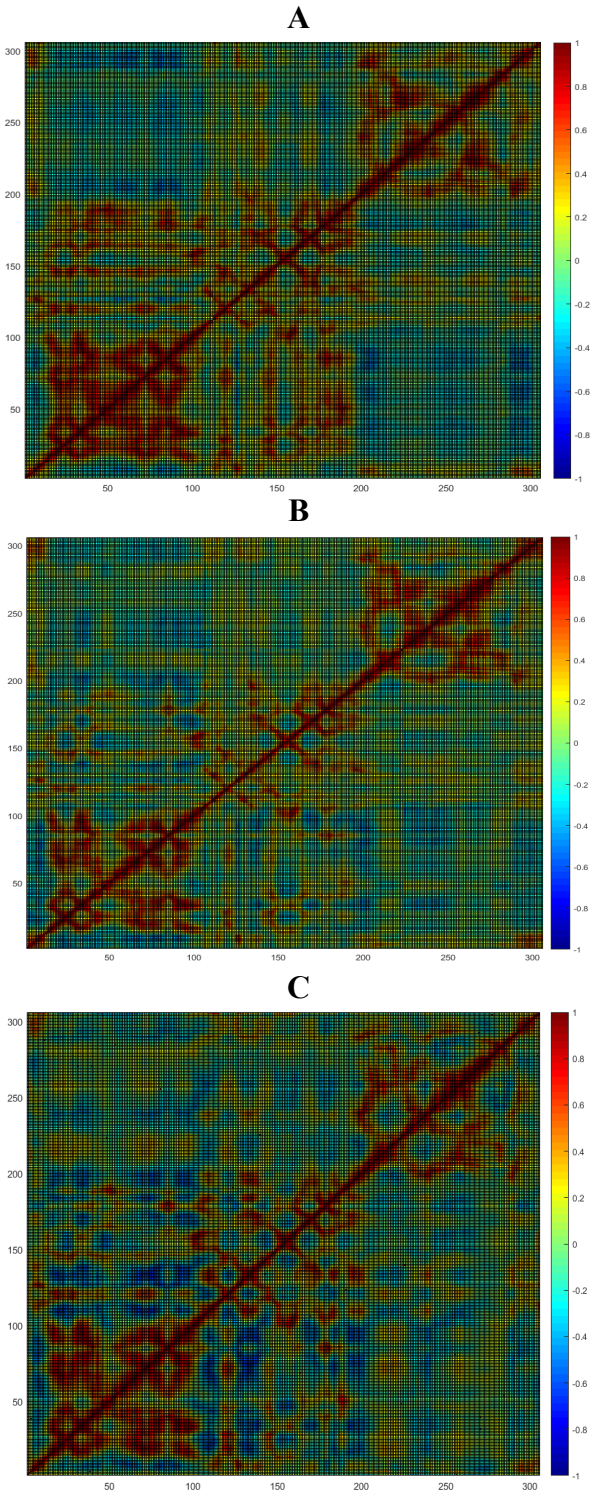

Figure 4. The dynamics cross-correlation maps for the active state (**A**), the inactive state (**B**) of the DHE-MPro complex, and for the apo-MPro (**C**).

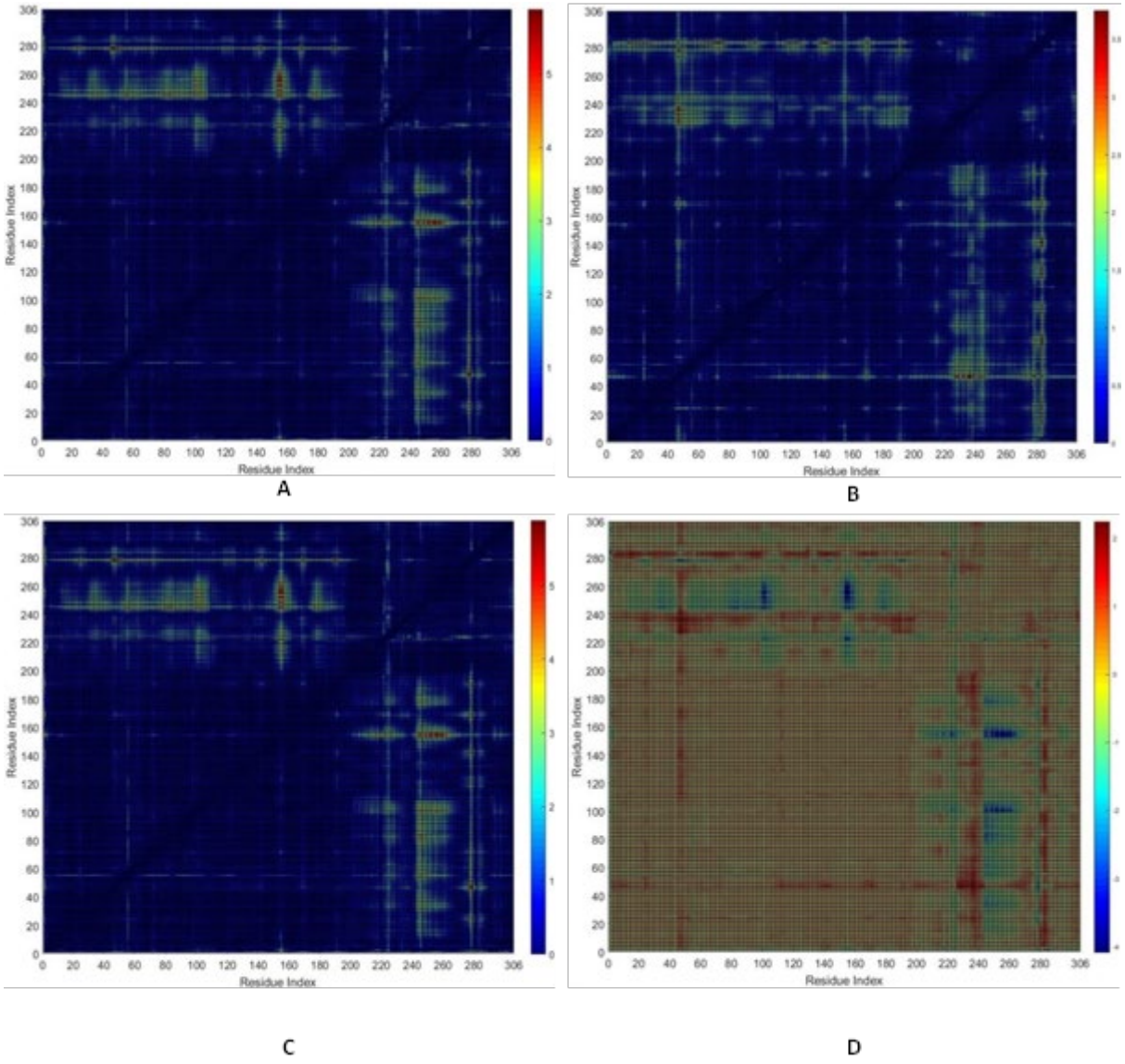

Figure 5. CP maps of the residue pairs for the active state (A), the inactive state (B) for the DHE-MPro complex, and the apo-MPro (C). (D) represents the difference in values of CP between the inactive state of DHE-MPro complex and the apo-MPro.

Table 1. Communication propensity values of residue pairs between the allosteric pocket and the catalytic pocket in protomer A.

		Apo-MPro state					
	Average CP	**ARG40**	**HIS41**	**VAL42**	**SER144**	**CYS145**	**GLY146**
Average CP		0.701	0.703	0.746	0.637	0.610	0.600
GLY2	0.939	0.437	0.409	0.437	0.477	0.420	0.475
PHE3	0.600	0.354	0.330	0.271	0.247	0.255	0.294
ARG4	0.550	0.175	0.182	0.171	0.197	0.164	0.145
ASN214	0.895	0.519	0.414	0.382	0.343	0.321	0.406
GLY215	1.003	0.455	0.367	0.348	0.364	0.305	0.357
ASP216	1.072	0.552	0.508	0.500	0.603	0.474	0.463
ASN277	1.618	1.575	1.862	1.995	2.145	1.856	1.468
GLY278	2.356	3.118	3.396	3.482	3.635	3.357	2.999
ARG279	1.227	0.900	1.088	1.178	1.376	1.121	0.864

THR280	1.025	0.603	0.750	0.878	1.106	0.852	0.642
ILE281	0.824	0.405	0.445	0.502	0.671	0.493	0.391
	DHE-MPro complex in active state						
	Average CP	**ARG40**	**HIS41**	**VAL42**	**SER144**	**CYS145**	**GLY146**
Average CP		0.728	0.807	0.828	0.698	0.686	0.614
GLY2	0.445	0.206	0.299	0.271	0.360	0.290	0.194
PHE3	0.387	0.201	0.279	0.256	0.336	0.271	0.193
ARG4	0.424	0.241	0.329	0.289	0.339	0.294	0.206
ASN214	0.626	0.421	0.476	0.516	0.671	0.546	0.504
GLY215	0.718	0.294	0.351	0.395	0.567	0.430	0.391
ASP216	0.629	0.214	0.280	0.295	0.401	0.301	0.254
ASN277	1.281	0.918	1.106	1.092	1.176	1.069	0.880
GLY278	1.462	1.47	1.640	1.656	1.862	1.718	1.547
ARG279	1.075	0.741	0.866	0.918	1.200	1.014	0.895
THR280	0.986	0.777	0.906	0.988	1.299	1.085	0.965
ILE281	0.799	0.634	0.730	0.795	1.080	0.890	0.809

DHE binding. When the CP values of apo-MPro and DHE-MPro are compared in general, it can be said that the CP values between residue pairs are lower in DHE-MPro complex. This means that the interaction of DHE with GLU278 and THR280 increases the communication between this site and the catalytic site. This result supports the possibility that the site where DHE binds is an allosteric site.

3.4. Dynamic residue network analysis (DRN)

Two graph theory-based network measures namely betweenness centrality (BC) and average shortest path (L) were used to explore the change in signal propagation pattern between the amino acid residues of MPro due to DHE binding. In this analysis (see Methods section), each MPro dimer (apo-MPro and DHE-MPro systems) is considered as a network of amino acids, where the beta carbons, C_β (except glycine, for which alpha carbon, C_α is considered) of each amino acids are regarded as to be nodes in the network and, and the edges between notes are built based on a separation cutoff distance of 6.7 Å. Each DRN was calculated using 5000 frames which were picked out from apo-MPro and DHE-MPro trajectories. Selecting the 5000 frames from DHE-MPro's trajectory was particularly picked up from the first quarter part of the whole trajectory, in where DHE is assumed to interact with the allosteric site residues. To determine the change in residues connectivity and consequently signal propagation in response to DHE binding to MPro, ΔBC (average BC of DHE-MPro complex minus average BC of apo-MPro) and ΔL (average L of DHE-MPro complex minus average L of apo-MPro) were calculated. Residues with significant change (more than ± 2 standard deviations) are presented in Table 2 and Figure 6. Both measures, BC and L are associated with each other. L between two residues, i and j, in a network is equal to minimum number of the edges that must be traversed reach from j to i [51]. In other words, L accents the mean topological spread of all residues from a given residue by considering the shortest paths to every other residue constituting the network [52]. Residues with large ΔL values may be interpreted as important ones for intra-protein communication. BC is a measure of the number of the shortest path passing through a given residue. Usage frequency of residues in cross network communication in where high usage residues are assumed to portray a taking part in controlling intra-protein signal transmission [42]. In other words, BC of a residue reflects the amount of control that this residue exerts over the interactions of other residues in the protein [53].

Residues with a positive ΔBC value mean that their role in intra-protein communication is ascended in response to DHE binding. As can be seen in Table 2 and Figure 6 (A), the residues in domain I and II regions in where they are mostly structured as alpha helix or anti-parallel beta-barrel have positive ΔBC values whereas the residues in the loop regions lean to have negative

ΔBC values. Spatial position of residues with significant ΔBC values (more than ± 2 σ) can be seen in Figure 6 (A). It can be noted that spatially the residues on one side of domain II (especially residues of linker from ASN184 to THR198) have positive ΔBC values however other side ones have negative ΔBC values. Many of residues of the linker which spatially lie down between domain I and domain III have positive ΔBC values. It seems that residues of this linker function as a signaling pathway between domain I and domain III regions. This signal transmission pathway favors the possibility of the existence of allosteric communication between the dimer interface region and catalytic site, since the proposed allosteric site (GLY278 and THR280) is in domain III, whereas the catalytic site is 40 Å distant from that of the

allosteric site and flanked by residues from both domains I and II.

Residues with a significant positive ΔL value are mostly accumulated in domain I, as can be seen from Figure 6 (B) and Table 2. ΔL values of the linker (from ASN180 to THR198) residues are also positive, this result may be interpreted as their reachability were increased in response to DHE binding, thus they play a role in residue communication between proposed allosteric site and catalytic site. This is in accord with the results of ΔBC values mentioned above**.**

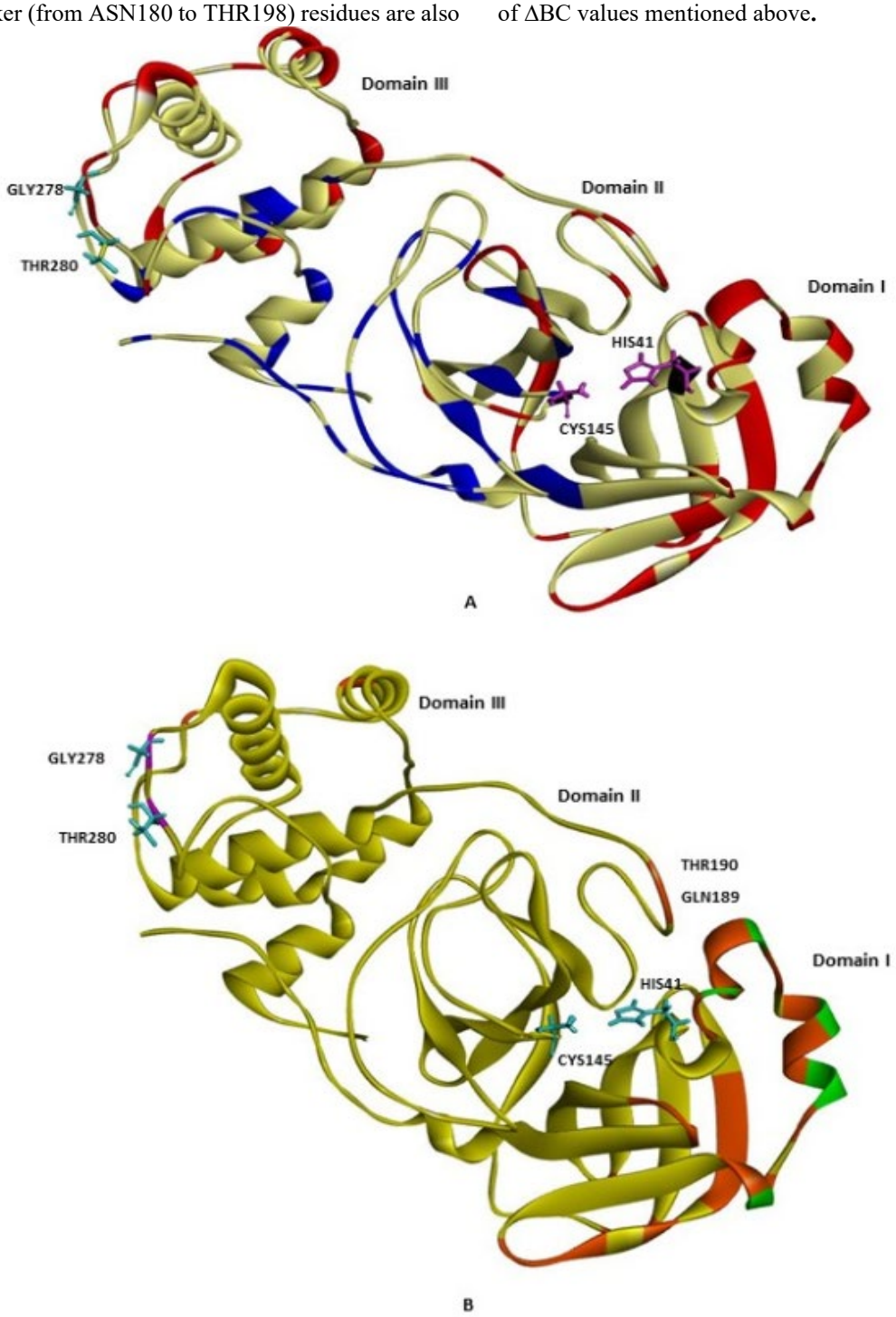

Figure 6. Indicates residues with significant changes in ΔBC and ΔL values in response to DHE binding. (**A**) Residues with an increase in BC value greater than 3 standard deviations are shown in red, and residues with a decrease greater than 3 standard deviations are shown in blue. (**B**) Amino acids with an increase in

L value greater than 3 standard deviations are shown in orange, amino acids with an increase greater than 2 standard deviations are shown in green.

Table 2. Residues with significant change in BC and average L upon DHE binding.

	Residues with significant increase in BC	Residues with significant increase in average L	Residues with significant decrease in BC
Domain I	**GLY23, ASP34, SER46, GLU47, ASP48, LEU50, ASN51, ASN53, GLU55, ASP56, ILE59, ARG60, SER62, HIS64, LEU67, ASN72, VAL73, GLN74, ARG76, ILE78, GLY79, HIS80, SER81, ASN84, LYS90, ASP92, THR93, ALA94, PRO96, LYS100,** LYS12, GLY15, THR21, THR24, ASP33, THR45, LEU57, LEU58, LYS61, ASN63, ASN65, PHE66, GLY71, LEU75, MET82, GLN83, CYS85, LEU87, LYS88, ASN95, LYS97, THR98	**GLU47, ASN51, ASP56, ILE59, ARG60, HIS64,** GLY23, THR24, THR25, ASP34, CYS45, SER46, ASP48, MET49, LEU50, PRO52, ASN53, GLU55, LEU57, LEU58, LYS61, ASN63, ASN65, ASN72, VAL73, GLN74, ARG76, ILE78, GLY79, HIS80, SER81, ASP92, THR93	**GLY2, LYS5, MET6, ALA7, PRO9, GLY11, VAL13, GLU14, MET17, VAL18, LEU27, PRO39**
Domain II	**LYS102, PFE103, VAL104, ASN142, TYR154, ASP155, GLU166, PRO168, THR169, GLY170, ASN180, PRO184, GLN189, ALA191, GLY195** ARG105, GLN107, PRO108, LEU141, GLY143, ASP153, CYS156, MET165, GLU178, GLY179, TYR182, GLY183, VAL186, THR190, ALA193, THR196, ASP197	GLN189, THR190	**GLY109, THR111, PHE112, LEU115, ALA116, CYS117, PRO122, SER123, GLY124, VAL125, TYR126, CYS128, THR135, SER139, GLY146, VAL148, GLY149, HIS164, HIS172, ALA173**
Domain III	**TRP218, ARG222, PHE223, THR224, THR226, ASN228, ASP229, LEU232, MET235, LYS236, THR243, GLN244, ASP245, ASP248, PRO252, ALA255, GLN256, GLY258, ALA260, GLN273, ASN274, ASN277, GLY278, ARG279, GLY283,** VAL212, ILE213, GLY215, ARG217, LEU220, ASN221, THR225, LEU227, ASN231, VAL233, PRO241, HIS246, GLY251, THR257, ILE259, LEU262, ASP263, ALA266, GLU270, GLY275, THR280, CYS300	ARG222, ASP229	**ILE200, VAL204, ALA206, ASN214, LEU282, LEU286, LEU287, THR292, PRO293, PHE294, ARG298, VAL303**

Note: For residues in bold, changes in BC and L are higher than ± 3 standard deviations.

3.5. Transfer entropy analysis (TE)

To understand the effect of the change in uncertainty of fluctuations of residues in the proposed allosteric site, in response to DHE binding to MPro, on the uncertainty of fluctuations of residues at the catalytic site, we calculated the entropy transfer that takes place between residues at these sites. Allosteric site residues (GLY278 and THR280) together with their 15 spatial neighbor residues and catalytic site residues (HIS41 and CYS145) together with their 8 spatial neighbor residues, total of 27 residues were considered for the transfer entropy calculation. The normalized directional indexes ($D_{j\to i}$) between C_α atoms of

selected residues for Apo-MPro and DHE-MPro complex are presented in Table S1 (A) and (B) respectively, in which atom j is on the vertical and i on the horizontal axis. Table S1 (C) shows the change in the directionality indexes in response to DHE binding to MPro, $\Delta D_{j\rightarrow i}$ ($D_{j\rightarrow i}$ of DHE-MPro complex minus average $D_{j\rightarrow i}$ of apo-MPro). In Table S1(A) and Table S1 (B), the points with positive signs indicate that entropy is transferred from residue j to residue i, in the case of negative sign transfer of entropy is vice versa. Transfer entropy from j to i with a positive sign means that the fluctuations of residue j drive the fluctuations of residue i. In other words, residue j acts as a source, whereas i behaves as a sink in the correlation between residues j and i. Likewise, transfer entropy from j to i with a negative sign means that residue j is the sink of the correlations between residues j and i, and that residue j responds to the motion of atom i.

According to Table S1 (A), in the Apo-MPro state, proposed allosteric site residues (GLY278 and THR280) together with their 15 spatial neighbor residues drive catalytic site residues as evident from positive $D_{j\rightarrow i}$ values. As it can be seen from Table S1 (B), in response the DHE binding, the loop (ASN214, GLY215, ASP216) residues which are spatial neighbor of GLY278 and THR280 are no more act as entropy source for domain I residues ARG40, HIS41, VAL42 and ILE43. Mostly, the strength of transfer entropy values from allosteric region residues to catalytic region residues is increased in response to DHE binding as evident from their positive signs in Table S1 (C). When catalytic dyad residues (HIS41 and CYS145) are concerned, the strengths of transfer entropy value ($D_{j\rightarrow i}$) from THR280 to HIS41 and THR280 to CYS145 are increased drastically from 0.146 to 0.258 and from 0.308 to 0.426 respectively in response to DHE binding. Interestingly, although GLY278 drives some of the neighboring residues of HIS41 and CYS145, such as ARG40 and SER144, when the strengths of transfer entropy value ($D_{j\rightarrow i}$) from GLY278 to HIS41 and CYS145 are examined, there is no increment in strength, even though there is a small drop-off, -0.063 for both cases.

From drug design view, identification of driver-driven relations is important. By using entropy transfer analysis procedure between residue pairs, if one determines residues which behave as drivers (source) of the fluctuations of other residues, thereby determining causality that is inherent in the correlations. In this regard, driver residues are more critical than the driven residues (sink) and manipulating the driver will be perturbing the existing correlations more efficiently [54]. If an allosteric enzyme comprises such residues, some of which constitute an allosteric site, and act as drivers and these that constitute a catalytic site, and are driven, perturbation of the driver residues (such as by ligand binding) may cause a change in the fluctuations of driven residues, thereby may lead to modulate the function of catalytic site residues [45]. In the light of above information, analyses of the transfer entropy results for Apo-MPro and DHE-MPro complex support the idea that the proposed allosteric site is a true allosteric site. Because results of the analysis indicate that GLY278 and THR280 drive catalytic dyad residues HIS41 and CYS145. Perturbation of GLY278 and THR280 (by DHE binding through establishing hydrogen bonds) resulted in a drastic increase in the amount of information flow from THR280 to HIS41 and CYS145. A change in χ1 angle of CYS145, which is relevant for the MPro catalytic function, may be resulted from this drastic increase in the amount of information flow.

It was hypothesized that allostery is an intrinsic property of all dynamic proteins. Authors stated that allostery results from a redistribution of the conformational ensemble of a protein[25]. Structural perturbation such as ligand binding at one site leads to a redistribution of the population. Even if the enzymes may exhibit nonallosteric kinetics, they are likely to be allosteric if proper ligands are introduced or modified by a few mutations. In allosteric enzymes, the structures of orthosteric site residues are affected by altering protein dynamics in response to the ligand binding to the allosteric site, either through large-scale conformational changes or through more subtle changes in correlated motions between some of the residues from both sites [55,56].

In several recent studies, the researchers attempted to unclose potential allosteric sites of MPro dimer by using different computational and experimental approaches [57-70]. Here, we will discuss the results of this study considering those previous literatures.

Experimental x-ray and mass spectroscopy techniques were applied to screen small molecules, that bind distant from the active site, for their ability to modulate MPro activity [69]. Authors revealed three compounds among identified 71 hits, which bind to the MPro dimer interface region. The interface residues, which surround three compounds, include GLY2, PHE3, ARG4, LYS5, MET6, PHE8, SER123, GLN127, LEU282, ARG298, and VAL303. It should be noted that five of the listed residues here coincide with that of the residues listed as the spatial neighbors of DHE-MPro complex in the previous section this study. Authors commented that these compounds were that compounds that interfere with dimerization might serve as quasi-allosteric inhibitors of

protease activity. Furthermore, it was stated that three compounds could be exploited to design new compounds which might diminish the catalytic activity by disrupting the MPro dimerization process. Mass spectroscopy method was used to measure substrate turnover rates by following temporal changes in the enzyme-substrate complexes, for four small molecules, which bind distant from the active site, for their ability to modulate proteolytic activity of MPro [70]. Crystallographic X-ray structures of MPro complexed with these four small molecules were known and published with PDB codes, 5RFA, 5REC, 5REE and 5RGJ [69]. Two of them, namely x0390 and x0464 bind hydrophobic pocket, which is located end of the linker (from ASP187 to ILE200) and interact with THR196, THR198 and GLU240. Remaining two, x0425 and x1187 bind to the dimer interface (interacting with MET6, ASP295 and ARG298) and a solvent-exposed pocket (interacting with PRO96 and LYS97) respectively. The authors reported that all the four molecules inhibit the proteolytic activity of MPro. These results can be attributed as experimental evidence that the locations where these molecules interact with are allosteric sites. Note that PRO96, LYS87, THR196, THR198 and ARG298 were listed in Table 2 and they have positive ΔBC value mean that their roles in intra-protein communication are ascended in response to DHE binding. Our further interpretation about these residues was that ΔBC and ΔL values of the linker (from ASN185 to THR198) residues are also positive, this result may be interpreted as their reachability were increased in response to DHE binding, thus they play a role in residue communications between allosteric site and catalytic dyad. Above experimentally confirmed allosteric site are in accord with our findings given in Table 2, which were obtained from graph theoretical betweenness centrality and average shortest path measures. Günther et al., (2021) performed a high-throughput x-ray crystallographic screening using two repurposing drug libraries against MPro [67]. The authors experimentally revealed 37 hit molecules that bind to MPro. Besides the substrate-binding site, two allosteric binding sites were also identified. First one of these allosteric sites is a hydrophobic pocket, which is formed by residues ILE213, LEU253, GLN256, VAL297, CYS300, and SER301, hosting five of these 37 molecules. The second allosteric site is in the deep groove between the catalytic domains and the dimerization domain, formed by residues PRO108, GLN110, ILE249, ASP153, THR154 and PHE294, hosting one molecule. These six molecules, which bonded to allosteric sites, were subjected for testing to reveal their antiviral activity against SARS-CoV-2 in cell assays. It was shown that these molecules reduced viral RNA (vRNA) replication by at least two orders of magnitude in Vero E6 cells. By using NMR spectroscopy, a fragment-based screening was performed that led to the identification of 38 fragment hits[68]. Besides the substrate-binding pocket, one allosteric site was determined at the dimer interface of the MPro, based on a combined analysis of chemical shift perturbations (CSP) and the signal intensity changes upon fragment binding. This allosteric site comprises residues MET6, PHE8, ARG298, GLN299, and VAL303, its location coincides with the hydrophobic pocket that was determined by x-ray crystallography mass spectroscopy techniques by [69,70].

Besides the above experimental studies concerning allosteric inhibition of MPro, a few molecular modeling studies also tackled exploring such allosteric sites. Two groups used similar computational methods which are based on normal mode analysis of vibrational frequencies of the alfa carbons of the residues of MPro dimer [59,60]. These vibrational frequencies were obtained from coarse grained Gaussian network model (GNM) or Elastic network model (ENM). Two putative allosteric sites were propounded [59]. One is in near the interface region of the dimer, which comprises the residues namely GLY11, LYS12, GLY15, PRO96, and LYS97. Another one is at the interface region comprises the residues namely TRP207, LEU282, GLY283, LEU286, GLU288, ASP289 and GLU290. This second one is the spatially adjacent to our propounded allosteric site. Based on their calculation, Dubanevics et al., (2021) listed residues which may allosterically affect the structure of the catalytic dyad residues [60]. These residues occupy a place which was indicated as a binding site by [59] and include LYS5, PRO9, GLU14, GLY109, THR111, GLN127, ASP197, THR198, ILE200, ASN214, MET264, LEU268, LYS269, LEU272, MET276, ASN277, ILE281, LEU282, SER28685, AEU282 and THR292. Authors commented that the listed residues do not create a particular allosteric pocket, but rather they are swoon over a wide area and active in communication towards catalytic the dyad residues. Note that, results of our study indicate that the residues ASP197, THR198, and ILE200 are positioned in the linker which lay down over domain II, have positive ΔL values. This result means that their reachability was increased in response to DHE binding, thus they play a role in residue communications between the proposed allosteric site and catalytic site. It was concluded that the linker residues are in accord with our findings which point out that they are in communication with catalytic dyad residues

allosterically [60]. In the another modelling study, three possible allosteric binding sites were proposed based on an analysis of a 100 μs long MD simulation trajectory of MPro dimer using a non-parametric data analysis approach [61]. One of their proposed allosteric sites is localized at the end of the linker (THR196, ASP197, THR198, and THR199), which extends up to ARG131 and PRO132 in domain II and ASP289 in domain III. This result also supports our assumption (based on interpretation of the ΔBC and ΔL in the previous section) that the residues in the linker play a role in communicating with catalytic dyad residues. Four putative allosteric binding sites were proposed based on graph theoretical Bond-to-bond propensity and Markov Transient analysis method by using a X-ray crystal structures of MPro [58]. One of their proposed allosteric sites which was determined based on Bond-to-bond propensity analysis, spatially coincides with one which is mentioned in the previous literature, comprises that of the linker residues ASP197, THR199 and ARG131 and ASP289 as well. Two flavonoid compounds were fished out by applying a virtual screening procedure on a flavonoid database [63]. Authors defined two putative allosteric binding sites in which these two flavonoid compounds bind. The first allosteric site (called Dimerization Site, DS) includes residues ARG4, MET6, SER10, GLU11, GLU14, ASN28, SER139, PHE140, SER147, GLU166, GLU290 and ARG298. The second allosteric site (called Cryptic Site, CS) includes residues LYS5, MET6, PRO108, GLY109, ARG131, TRP218, PHE219, TYR239, GLU240, LEU271, LEU272, LEU287, GLU288, ASP289, GLU290, ARG298, GLN299 and VAL303. Note that these two sites are spatially overlapping each other as well as the allosteric site proposed in this study. By using molecular docking and MD simulations approaches, Verma and Pandey (2021) also studied the interactions between MPro and Quercetin which is a flavonoid dietary supplement [66]. It was reported that Quercetin binds a site that is indwelled in the region between the II and III domains, and comprises the residues LYS5, GLY127, CYS128, LYS137, ASP289, and GLU290. Based on their results, the authors concluded that the binding of Quercetin induces conformational changes in the structure of the catalytic cavity and causes a significant reduction in the binding affinity between MPro and a modeled substrate peptide (ALA-VAL-LEU-GLN-SER-GLY-PHE-ARG). Interactions between Ebselen and MPro were studied using the MD simulation approach [64]. Ebselen is an FDA-approved drug and its strong antiviral activity against SARS-CoV-2 was shown using cell-based assays [71]. The authors reported that there are two highly probable interaction sites between MPro and Ebselen. One of these interaction sites is indwelled within the catalytic cavity. The second one with a higher binding affinity is indwelled in the region between the II and III domains comprise residues GLY107, PRO108, GLN110, ILE200, VAL202, HIS246, ILE249, THR292, and PHE294. Based on their strain analysis results, the authors concluded that Ebselen bound between the II and III domains exert a pronounced allosteric effect that affects the loops regulating access to the catalytic dyad. Ensemble docking and MD simulation approaches was applied to find out potential allosteric sites of MPro [65]. Based on their results, the authors concluded that active site (the region around catalytic dyad) is not the best target for non-covalent inhibitors due to its shallowness and wideness. They propose an allosteric binding site which is in the groove between domains II and III. This groove is partially made of amino acids in the N-finger region (residues 1–7) and the crucial ARG298 residue, which are involved in the MPro dimerization and activation processes. Overall, one can make inferences from the above pieces of literature that the hydrophobic pocket (residues 290-302 of domain III, residues 108-110, 127-131 of domain II and the linker residues 185-200) accommodates different types of ligands. The antiviral activity of some of these ligands was experimentally validated using cell-based assays [67-71].

Although, according to our results, DHE binds in the interface region of MPro dimer as some ligands mentioned in above literatures. But our interpretation fundamentally differs from them, in a way that, they suppose that these ligands exhibit their functions trough obstructing the dimerization process of MPro. Whereas we put forward some computational results which support the idea that DHE binding causes a local structural change of the side chain of CYS145.

In summary, the results of a certain part of our MD simulation indicated the binding of DHE to the MPro dimer interface in where DHE established two hydrogen bonds with GLU278 and THR28. These bonding interactions seem to be able to quantitatively modify the χ1 angle of CYS145 of the catalytic dyad. We hypnotize that the modification of the χ1 angle results in a diminishing of catalytic activity of MPro due to the reason that the modified angle causes an enlargement of the distance between sulfur (SG) atom of CYS145 and nitrogen (NE2) atom of HIS41. This enlarged distance is unfavorable to initiate the catalytic cycle. By applying some computational tools which are convenient to use exploring an allosteric effect for a protein, we tried to demonstrate the reasoning behind the χ1 angle modification upon DHE binding to the distal place in MPro. As seen in the

results section, the analyses of the results of Dynamic cross-correlation, Communication propensity, Dynamic residue network, and Transfer entropy indicated that communication exists between catalytic dyad and DHE binding residues. The analyses indicated that upon DHE binding to MPro, communication increases to some degree, which may be inducing a change in the χ1 angle.

4. Conclusions

In conclusion, there are number of experimental and in silico studies which propose the potential allosteric sites of MPro in the literature. Locations of the proposed sites mainly are resided in the MPro dimer region or the groove between domains II and III. Analysis of trajectories of MD simulations of Apo-MPro and DHE-MPro complex let us to propose a new allosteric site. We applied some computational bioinformatics tools which are useful to evaluate an allosteric pathway for a given enzyme, on the obtained trajectories. Some pieces of evidence which were drawn from the DCC, CP, DRN, and TE analysis, support that structural change occurs in sidechain χ1 angle of CYS145 in response to the DHE binding to the proposed allosteric site (GLY278 and THR280). To increase the targetable space of the MPro and allow a broader approach to inhibitor discovery, here we gave a full computational analysis procedure that gives insights into the allosteric signaling pathway between the catalytic dyad residues and proposed allosteric site. Of course, experimental testing is necessary to confirm the predictions presented in this study; we hope that experimental testing will be performed by groups with the appropriate competencies.

Acknowledgments

This project was supported by Akdeniz University Scientific Research Projects Coordination Unit. Project ID:5408. The authors acknowledge Harran University High-Performance Computing Center for making computing resources available to us.

References

[1] B. Hu, H. Guo, P. Zhou, Z.L. Shi, Characteristics of SARS-CoV-2 and COVID-19, Nature Reviews Microbiology 19 (2021) 141-154.

[2] S. Kumar, R. Nyodu, V.K. Maurya, S.K. Saxena, Morphology, genome organization, replication, and pathogenesis of severe acute respiratory syndrome coronavirus 2 (SARS-COV-2), Coronavirus Disease 2019 (COVID-19): Epidemiology, Pathogenesis, Diagnosis, and Therapeutics (2020) 23-31.

[3] Y.R. Guo, Q.D. Cao, Z.S. Hong, Y.Y. Tan, S.D. Chen, H.J. Jin, K.S. Tan, D.Y. Wang, Y. Yan, The origin, transmission and clinical therapies on coronavirus disease 2019 (COVID-19) outbreak–an update on the status, Military Medical Research 7 (2020) 1-10.

[4] J.F.W. Chan, K.H. Kok, Z. Zhu, H. Chu, K. K.W. To, S. Yuan, K.Y. Yuen, Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan, Emerging Microbes & Infections 9 (2020) 221-236.

[5] J. Zhang, T. Xiao, Y. Cai, B. Chen, Structure of SARS-CoV-2 spike protein, Current Opinion In Virology 50 (2021) 173-182.

[6] P. Zhou, X.L. Yang, X.G. Wang, B. Hu, L. Zhang, W. Zhang, H.R. Si, Y. Zhu, B. Li, C.L. Huang, A pneumonia outbreak associated with a new coronavirus of probable bat origin, Nature 579 (2020) 270-273.

[7] L. Duan, Q. Zheng, H. Zhang, Y. Niu, Y. Lou, H. Wang, The SARS-CoV-2 spike glycoprotein biosynthesis, structure, function, and antigenicity: implications for the design of spike-based vaccine immunogens, Frontiers In Immunology 11 (2020) 2593.

[8] F. Amanat, F. Krammer, SARS-CoV-2 vaccines: status report, Immunity 52 (2020) 583-589.

[9] M. Hoffmann, H. Kleine-Weber, S. Pöhlmann, A multibasic cleavage site in the spike protein of SARS-CoV-2 is essential for infection of human lung cells, Molecular Cell 78 (2020) 779-784.

[10] C.B. Jackson, M. Farzan, B. Chen, H. Choe, Mechanisms of SARS-CoV-2 entry into cells, Nature Reviews Molecular Cell Biology 23 (2022) 3-20.

[11] K. Anand, J. Ziebuhr, P. Wadhwani, J.R. Mesters, R. Hilgenfeld, Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs, Science 300 (2003) 1763-1767.

[12] S. Ullrich, C. Nitsche, The SARS-CoV-2 main protease as drug target, Bioorganic & Medicinal Chemistry Letters 30 (2020) 127377.

[13] H.J. Lee, C.K. Shieh, A.E. Gorbalenya, E.V. Koonin, N. La Monica, J. Tuler, A. Bagdzhadzhyan, M.M. Lai, The complete sequence (22 kilobases) of murine coronavirus gene 1 encoding the putative proteases and

RNA polymerase, Virology 180 (1991) 567-582.

[14] J. Ziebuhr, E.J. Snijder, A.E. Gorbalenya, Virus-encoded proteinases and proteolytic processing in the Nidovirales, Journal of General Virology 81 (2000) 853-879.

[15] L. Z Zhang, D. Lin, X. Sun, U. Curth, C. Drosten, L. Sauerhering, S. Becker, K. Rox, R. Hilgenfeld, Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors, Science 368 (2020) 409-412.

[16] K. Świderek, V. Moliner, Revealing the molecular mechanisms of proteolysis of SARS-CoV-2 M pro by QM/MM computational methods, Chemical Science 11 (2020) 10626-10630.

[17] C. Wu, Y. Liu, Y. Yang, P. Zhang, W. Zhong, Y. Wang, Q. Wang, Y. Xu, M. Li, X. Li, Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods, Acta Pharmaceutica Sinica B 10 (2020) 766-788.

[18] R. Banerjee, L. Perera, L.V. Tillekeratne, Potential SARS-CoV-2 main protease inhibitors, Drug Discovery Today 26 (2021) 804-816.

[19] G. Macip, P. Garcia-Segura, J. Mestres-Truyol, B. Saldivar-Espinoza, M.J. Ojeda-Montes, A. Gimeno, A. Cereto-Massagué, S. Garcia-Vallvé, G. Pujadas, Haste makes waste: a critical review of docking-based virtual screening in drug repurposing for SARS-CoV-2 main protease (M-pro) inhibition, Medicinal Research Reviews 42 (2022) 744-769.

[20] B. Goyal, D. Goyal, Targeting the dimerization of the main protease of coronaviruses: a potential broad-spectrum therapeutic strategy, ACS Combinatorial Science 22 (2020) 297-305.

[21] S. Gupta, A.K. Singh, P.P. Kushwaha, K.S. Prajapati, M. Shuaib, S. Senapati, S. Kumar, Identification of potential natural inhibitors of SARS-CoV2 main protease by molecular docking and simulation studies, Journal of Biomolecular Structure and Dynamics 39 4334-4345.

[22] A. Ton, F. Gentile, M. Hsing, F. Ban, A. Cherkasov, Rapid identification of potential inhibitors of SARS-CoV-2 main protease by deep docking of 1.3 billion compounds, Mol Inform 39 (2020) 8 e2000028.

[23] J. Liang, C. Karagiannis, E. Pitsillou, K.K. Darmawan, K. Ng, A. Hung, T.C. Karagiannis, Site mapping and small molecule blind docking reveal a possible target site on the SARS-CoV-2 main protease dimer interface, Computational Biology and Chemistry 89 (2020) 107372.

[24] Z. Lv, K.E. Cano, L. Jia, M. Drag, T.T. Huang, S.K. Olsen, Targeting SARS-CoV-2 proteases for COVID-19 antiviral development, Frontiers in Chemistry (2022) 1221.

[25] K. Gunasekaran, B. Ma, R. Nussinov, Is allostery an intrinsic property of all dynamic proteins? Proteins: Structure, Function, and Bioinformatics 57 (2004) 433-443.

[26] CHARMM-GUI. MPro Dimer Structure 6M03 2021, December.

[27] D.S. Wishart, Y.D. Feunang, A.C. Guo, E.J. Lo, A. Marcu, J.R. Grant, T. Sajed, D. Johnson, C. Li, Z. Sayeeda, DrugBank 5.0: a major update to the DrugBank database for 2018, Nucleic Acids Research 46 (2018) D1074-D1082.

[28] S. Dallakyan, A.J. Olson, Small-molecule library screening by docking with PyRx, In Chemical Biology (2015) 243-250 Springer.

[29] G.M. Morris, R. Huey, W. Lindstrom, M.F. Sanner, R.K. Belew, D.S. Goodsell, A.J. Olson, AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility, Journal of Computational Chemistry 30 (2009) 2785-2791.

[30] Salomon-Ferrer, R., D. A. Case & R. C. Walker (2013) An overview of the Amber biomolecular simulation package. Wiley Interdisciplinary Reviews: Computational Molecular Science, 3, 198-210.

[31] VirginiaTech. 2022, January. H++.

[32] J.A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K.E. Hauser, C. Simmerling, ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB, Journal of Chemical Theory and Computation 11 (2015) 3696-3713.

[33] J. Wang, W. Wang, P.A. Kollman, D.A. Case, Automatic atom type and bond type perception in molecular mechanical calculations, Journal of Molecular Graphics and Modelling 25 (2006) 247-260.

[34] J. Wang, R.M. Wolf, J.W. Caldwell, P.A. Kollman, D.A. Case, Development and testing of a general amber force field, Journal of

Computational Chemistry 25 (2004) 1157-1174.

[35] A. Jakalian, B.L. Bush, D.B. Jack, C.I. Bayly, Fast, efficient generation of high-quality atomic charges. AM1-BCC model: I. Method, Journal Of Computational Chemistry 21 (2000) 132-146.

[36] M.J. Frisch, G. W. Trucks, H. B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, G.A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A.V. Marenich, J. Bloino, B.G. Janesko, R. Gomperts, B. Mennucci, H.P. Hratchian, J.V. Ortiz, A.F. Izmaylov, J.L. Sonnenberg, F.Williams Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V.G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J.A. Montgomery Jr., J.E. Peralta, F. Ogliaro, M. J. Bearpark, J.J. Heyd, E.N. Brothers, K.N. Kudin, V.N. Staroverov, T.A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A.P. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, J.M. Millam, M. Klene, C. Adamo, R. Cammi, J.W. Ochterski, R.L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, D. J.Fox, Gaussian 16 Rev. C.01, Wallingford, CT (2016).

[37] D. R. Roe, T. E. Cheatham III, PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data, Journal of Chemical Theory and Computation 9 (2013) 3084-3095.

[38] O. S. Amamuddy, M. Glenister, T. Tshabalala, Ö. T. Bishop, MDM-TASK-web: MD-TASK and MODE-TASK web server for analyzing protein dynamics, Computational and Structural Biotechnology Journal 19 (2021) 5059-5071.

[39] C. Chennubhotla, I. Bahar, Signal propagation in proteins and relation to equilibrium fluctuations, PLoS Computational Biology 3 (2007) e172.

[40] G. Morra, G. Verkhivker, G. Colombo, Modeling signal propagation mechanisms and ligand-based conformational dynamics of the Hsp90 molecular chaperone full-length dimer, PLoS Computational Biology 5 (2009) e1000323.

[41] A.R. Atilgan, P. Akan, C. Baysal, Small-world communication of residues and significance for protein dynamics, Biophysical Journal 86 (2004) 85-91.

[42] D.L. Penkler, C. Atilgan, O.Z. Tastan Bishop, Allosteric modulation of human Hsp90α conformational dynamics, Journal of Chemical Information and Modeling 58 (2018) 383-404.

[43] E. W. Dijkstra, A note on two problems in connexion with graphs, Numerische Mathematik 1 (1959) 269-271.

[44] T. Schreiber, Measuring information transfer, Physical Review Letters 85 (2000) 461.

[45] H. Kamberaj, A. van der Vaart, Extracting the causality of correlated motions from molecular dynamics simulations, Biophysical Journal 97 (2009) 1747-1755.

[46] D. Nebiu, H. Kamberaj, Symbolic Information Flow Measurement (SIFM): A software for measurement of information flow using symbolic analysis, SoftwareX 11 (2020) 100470.

[47] Kamberaj, H. 2020, February, 29. SifmV1.

[48] S. Kullback, R.A. Leibler, On information and sufficiency, The Annals of Mathematical Statistics 22 (1951) 79-86.

[49] B. Gourévitch, J.J. Eggermont, Evaluating information transfer between auditory cortical neurons, Journal of Neurophysiology 97 (2007) 2533-2543.

[50] T.C. McLeish, M.J. Cann, T.L. Rodgers, Dynamic transmission of protein allostery without structural change: spatial pathways or global modes? Biophysical Journal 109 (2015) 1240-1250.

[51] D.K. Brown, O.S. Amamuddy, Ö.T. Bishop, Structure-based analysis of single nucleotide variants in the renin-angiotensinogen complex, Global Heart 12 (2017) 121-132.

[52] A. Amusengeri, Ö. Tastan Bishop, Discorhabdin N, a South African natural compound, for Hsp72 and Hsc70 allosteric modulation: combined study of molecular modeling and dynamic residue network analysis, Molecules 24 (2019) 188.

[53] J. Yoon, A. Blumer, K. Lee, An algorithm for modularity analysis of directed and weighted biological networks based on edge-betweenness centrality, Bioinformatics 22 (2006) 3106-3108.

[54] A. Hacisuleyman, B. Erman, Entropy transfer between residue pairs and allostery in proteins: quantifying allosteric communication in

ubiquitin, PLoS Computational Biology 13 (2017) e1005319.

[55] R. Nussinov, Introduction to Protein Ensembles and Allostery, ACS Publications, 2016, 6263-6266.

[56] J.R. Wagner, C.T. Lee, J.D. Durrant, R.D. Malmstrom, V.A. Feher, R.E. Amaro, Emerging computational methods for the rational discovery of allosteric drugs, Chemical Reviews 116 (2016) 6370-6390.

[57] X. Tao, L. Zhang, L. Du, R. Liao, H. Cai, K. Lu, Z. Zhao, Y. Xie, P.H. Wang, J.A. Pan, Y. Zhang, G. Li, J. Dai, Z.W. Mao, W. Xia, Allosteric inhibition of SARS-CoV-2 3CL protease by colloidal bismuth subcitrate, Chem Sci 12 (2021) 14098-14102.

[58] L. Strömich, N. Wu, M. Barahona, S.N. Yaliraki Allosteric hotspots in the main protease of SARS-CoV-2, BioRxiv (2020).

[59] M. Yuce, E. Cicek, T. Inan, A.B. Dag, O. Kurkcuoglu, F.A. Sungur, Repurposing of FDA-approved drugs against active site and potential allosteric drug-binding sites of COVID-19 main protease, Proteins 89 (2021) 1425-1441.

[60] I. Dubanevics, T.C.B. McLeish, Computational analysis of dynamic allostery and control in the SARS-CoV-2 main protease, J R Soc Interface 18 (2021) 20200591.

[61] M. Carli, G. Sormani, A. Rodriguez, A. Laio, Candidate Binding Sites for Allosteric Inhibition of the SARS-CoV-2 Main Protease from the Analysis of Large-Scale Molecular Dynamics Simulations, J Phys Chem Lett 12 (2021) 65-72.

[62] T. Sztain, R. Amaro, J.A. McCammon, Elucidation of Cryptic and Allosteric Pockets within the SARS-CoV-2 Main Protease, J Chem Inf Model 61 (2021) 3495-3501.

[63] G. Jimenez-Avalos, A.P. Vargas-Ruiz, N.E. Delgado-Pease, G.E. Olivos-Ramirez, P. Sheen, M. Fernandez-Diaz, M. Quiliano, M. Zimic, C.W.G.İ. Peru, Comprehensive virtual screening of 4.8 k flavonoids reveals novel insights into allosteric inhibition of SARS-CoV-2 M(PRO), Sci Rep 11 (2021) 15452.

[64] C.A. Menéndez, F. Byléhn, G.R. Perez-Lemus, W. Alvarado, J.J. de Pablo, Molecular characterization of ebselen binding activity to SARS-CoV-2 main protease, Science Advances 6 (2020) eabd0345.

[65] J. Novak, H. Rimac, S. Kandagalla, P. Pathak, V. Naumovich, M. Grishina, V. Potemkin, Proposition of a new allosteric binding site for potential SARS-CoV-2 3CL protease inhibitors by utilizing molecular dynamics simulations and ensemble docking, J Biomol Struct Dyn (2021)1-14.

[66] S. Verma, A.K. Pandey, Factual insights of the allosteric inhibition mechanism of SARS-CoV-2 main protease by quercetin: an in silico analysis, 3 Biotech 11 (2021) 67.

[67] S. Günther, P.Y. Reinke, Y. Fernández-García, J. Lieske, T.J. Lane, H.M. Ginn, F.H. Koua, C. Ehrt, W. Ewert, D. Oberthuer, X-ray screening identifies active site and allosteric inhibitors of SARS-CoV-2 main protease, Science 372 (2021) 642-646.

[68] F.X. Cantrelle, E. Boll, L. Brier, D. Moschidi, S. Belouzard, V. Landry, F. Leroux, F. Dewitte, I. Landrieu, J. Dubuisson, B. Deprez, J. Charton, X. Hanoulle, NMR Spectroscopy of the Main Protease of SARS-CoV-2 and Fragment-Based Screening Identify Three Protein Hotspots and an Antiviral Fragment, Angew Chem Int Ed Engl 60 (2021) 25428-25435.

[69] A. Douangamath, D. Fearon, P. Gehrtz, T. Krojer, P. Lukacik, C.D. Owen, E. Resnick, C. Strain-Damerell, A. Aimon, P. Abranyi-Balogh, J. Brandao-Neto, A. Carbery, G. Davison, A. Dias, T.D. Downes, L. Dunnett, M. Fairhead, J.D. Firth, S.P. Jones, A. Keeley, G.M. Keseru, H.F. Klein, M.P. Martin, M.E.M. Noble, P. O'Brien, A. Powell, R.N. Reddi, R. Skyner, M. Snee, M.J. Waring, C. Wild, N. London, F. von Delft, M.A. Walsh, Crystallographic and electrophilic fragment screening of the SARS-CoV-2 main protease, Nat Commun 11 (2020) 5047.

[70] T.J. El-Baba, C.A. Lutomski, A.L. Kantsadi, T.R. Malla, T. John, V. Mikhailov, J.R. Bolla, C.J. Schofield, N. Zitzmann, I. Vakonakis, C.V. Robinson, Allosteric inhibition of the SARS-CoV-2 main protease: insights from mass spectrometry based assays, Angew Chem Int Ed Engl 59 (2020) 23544-23548.

[71] Z. Jin, X. Du, Y. Xu, Y. Deng, M. Liu, Y. Zhao, B. Zhang, X. Li, L. Zhang, C. Peng, Y. Duan, J. Yu, L. Wang, K. Yang, F. Liu, R. Jiang, X. Yang, T. You, X. Liu, X. Yang, F. Bai, H. Liu, X. Liu, L.W. Guddat, W. Xu, G. Xiao, C. Qin, Z. Shi, H. Jiang, Z. Rao, H. Yang, Structure of M(pro) from SARS-CoV-2 and discovery of its inhibitors, Nature 582 (2020) 289-293.