In silico study on Spice-derived antiviral phytochemicals against SARS-CoV-2 TMPRSS2 target

Pradeep Kumar Yadav[γ]*, Amit Jaiswal[γ 1] and Rajiv Kumar Singh

Bioinformatics Centre, Faculty of Life Science, Rajiv Gandhi University, Rono Hills, Doimukh, Arunachal Pradesh, India 791112.

[γ] These authors contributed equally

* yadavpradeep755@gmail.com

Highlights

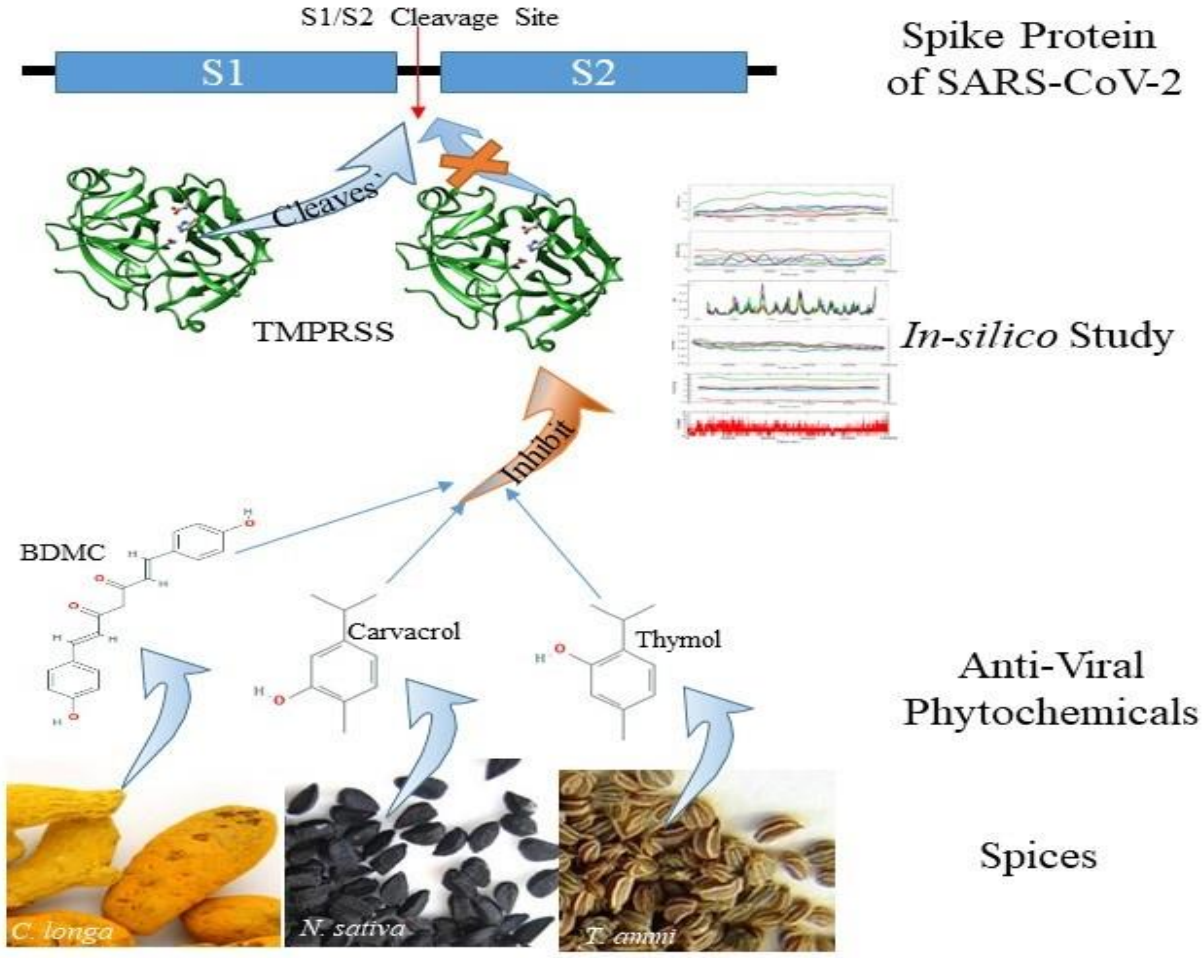
- TMPRSS2 facilitate the entry of SARS-CoV-2 in the host cell
- Spices have numerous potent anti-viral phytochemicals
- The study identified the phytochemicals (BDMC, carvacrol and thymol) as potent inhibitor candidates of TMPRSS2

Abstract

Corona Virus Disease (COVID-19) caused by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) is a pandemic that has claimed so far over half a million human life across the globe. Researchers all over the world are exploring various molecules including phytochemicals to get a potential anti-COVID-19 drug. Certain phytochemicals present in some spices are claimed to possess antiviral, anti-bacterial, and anti-fungal properties. Hence, an *in-silico* study was done by selecting eighteen well reported antiviral phytochemicals from some spices commonly used in Indian kitchen viz. *Curcuma longa* (Turmeric), *Nigella sativa* (Black cumin), *Piper nigrum* (Black pepper), *Trachyspermum ammi* (Carom) and *Zingiber officinale* (Ginger) to find out whether they can prevent SARS-CoV-2 infection. Firstly, we predicted the structure of TMPRSS2 (transmembrane protease serine 2), a host protein that truncates spike protein of SARS-CoV-2 thereby facilitating its endocytosis, and then docked against its catalytic domain the selected phytochemicals and camostat (a well-known synthetic inhibitor of TMPRSS2). Thereafter, stability of seven best docked phytochemicals and camostat were scrutinized by Molecular Dynamic Simulation (MDS). MDS analysis indicated bisdemethoxycurcumin (BDMC), carvacrol and thymol as better inhibitors than the camostat due to their stable binding with TMPRSS2 in its oxyanion hole and inducing subtle modification in the spatial arrangement

of the catalytic triad residues. Among these three phytochemicals, carvacrol appeared to be the best inhibitor, followed by BDMC, whereas thymol was least effective.

Graphical abstract



Keywords: SARS-CoV-2, COVID-19, TMPRSS2 inhibitor, Spices, Antiviral Phytochemical, Molecular Dynamics Simulation

Running Title: Phytochemicals from Spices as natural drug candidates for Covid-19

Introduction

The deadly pandemic pneumonia like Corona Virus Disease (COVID-19) caused by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) [1, 2] has spread worldwide with 10 million population affected including around half a million succumbing to death [3]. Presently, in absence of any known cure for the disease, tremendous efforts are on all over the world to find out an effective drug from different sources including phytochemicals [4-6]. Certain drugs such as Remdesivir, Favipiravir, etc. are currently under clinical trials in the fight against SARS-Cov-2 [7].

The SARS-CoV-2 belongs to Coronavirus family (Coronaviridae), a cluster of viruses mainly hosted by bats [8]. Three viruses of this family, Middle-East Respiratory Syndrome Coronavirus (MERS-CoV), Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV) and SARS-CoV-2) are reported to have crossed the species barrier and become deadly for humans since the dawn of the 21st century [9-11]. Entry of these viruses into the host cell is facilitated by the binding of their spike proteins (a highly glycosylated surface protein) to the cellular membrane receptor. The spike protein of SARS-CoV and the SARS-CoV-2 binds to the host receptor called Angiotensin-converting enzyme 2 (ACE2), whereas in case of MERS-CoV, it binds to dipeptidyl peptidase-4 [12]. It is also reported that infection gradient of SARS-CoV-2 in the respiratory tract is correlated with expression of ACE2 with occurs maximally in the nose and bronchus and decreases throughout the lower respiratory tract [13]. The spike protein of SARS-CoV-2 has two domains: S1 (receptor binding domain) and S2 (membrane fusion domain); and a cleavage between these domains by the host transmembrane protease serine 2 (TMPRSS2) is a pre-requisite for the entry of these viruses through endocytosis into the host cell [14-16].

Iwata-Yoshikawa et al. have reported a reduction in SARS-CoV and MERS-CoV infection in the absence of TMPRSS2 in mouse [16, 17]. Matsuyama et al. found enhanced infection rates while activating TMPRSS2 [18]. Hoffmann et al. has reported that inhibition of TMPRSS2 by camostat mesylate, a synthetic inhibitor approved for clinical use, blocks the entry of SARS-CoV-2 into the host cell [16, 17]. These studies highlight that TMPRSS2 is one of the prime targets whose inhibition can prevent spread of these viruses within the host.

Some traditional spices used routinely across the Indian sub-continent are well known for their medicinal values, antiviral properties and the least side effects [19-25]. Therefore, in the present study, we selected eighteen well-known antiviral phytochemicals (Table 1) present in some commonly used spices in Indian kitchens viz. *Curcuma longa* (Turmeric), *Nigella sativa* (Black cumin), *Piper nigrum* (Black pepper), *Trachyspermum ammi* (Carom), *Zingiber officinale* (Ginger) and camostat, and carried out Molecular Dynamic Simulations (MDS) analysis after docking them individually against our predicted three dimensional (3D) molecular model of TMPRSS2 for identifying potential phytochemicals that can alter the catalytic domain of TMPRSS2.

Materials and Methods

Structure prediction and analysis of TMPRSS2

The protein sequence of TMPRSS2 (Uniprot ID:O15393) was collected from Uniprot [26] and NCBI blast search was performed against protein data bank (PDB) to find out a suitable structure for the study. The best available structure showed a sequence identity of 42.56% and a query coverage of only 48%. In absence of any deposited structure, the structure prediction for TMPRSS2 was done by an online web server Phyre2 based on multi-template and *ab-initio* [27].

The PDB model was then verified by getting Ramchandran plot from an online web server Procheck [28, 29] . Once the structure was validated, the catalytic domain was considered for molecular docking followed by MDS analysis.

Docking studies

Anti-viral phytochemicals reported from the spices (*Z. officinale*, *C. longa*, *T. ammi*, *N. sativa* and *P. nigrum*) were obtained through literature review (Table 1) and their three-dimensional structures were collected from the Pubchem database [30].

Table 1: Phytochemicals from spices used docking with TMPRSS2

Sl. No.	Spice Name	Antiviral Compound	PubChem ID	References
1.	*Trachyspermum ammi* (Carom)			
		Thymol	6989	[23, 31-33]
		p-cymene	7463	[34, 35]
		γ-terpinene	7461	[31]
2.	*Curcuma longa* (Turmeric)			
		Curcumin	969516	[36]
		Bisdemethoxycurcumin (BDMC)	5315472	[24, 36-38]
		Desmethoxycurcumin	5469424	[36]
		Tetrahydrocurcumin	124072	
3.	*Zingiber officinale* (Ginger)			
		[6]-Gingerol	442793	[25, 39-42]
		[8]-Gingerol	168114	

		[10]-Gingerol	168115	
		[6]-Shogaol	5281794	
		[8]-Shogaol	6442560	
		[10]-Shogaol	6442612	
4.	*Piper nigrum* (Black pepper)			
		β-caryophyllene	5281515	[43]
		Limonene	22311	[44]
5.	*Nigella sativa* (Black cumin)			
		β-pinene (also present in black pepper)	14896	[44]
		Carvacrol	10364	[45]
		Thymoquinone	10281	[46]

The modelled structure of TMPRSS2 was aligned with Thrombin (4UD9, a serine protease) to identify its active site and residues. The standard serine protease residue numbers, catalytic triad of active site residues, and oxyanion hole were also identified and located (Fig. 1A-C; Supplementary Fig. SF1). To know the binding site of the inhibitor, a 3D structure of Prostasin (3FVF, a serine protease) complexed with camostat (a well-known inhibitor of TMPRSS2) was analyzed [47]. The camostat structure was separated from 3FVF and included in the list of molecules to be docked with TMPRSS2 for comparative analysis. The co-ordinates of catalytic triad residues (His296, Asp345 and Ser441; residue number are as per TMPRSS2 sequence) and oxyanion hole of catalytic domain of TMPRSS2 were chosen as the binding site for docking

studies with a total of 18 phytochemicals along with camostat. For docking purpose, the preparation of ligand and protein molecules followed by docking search run and analysis were done by graphical user interface software "AutoDockTools1.5.7" [48]. Autogrid was used to attain grid box with dimension (56 x 32 x 62 $Å^3$) and center at Cα atom of Ser441. Further, autodock4.2 was used with lamarckian genetic algorithm to get the best docking conformations [49]. The complexes with the best conformations were put under MDS scrutiny.

Molecular Dynamics Simulation (MDS) studies

Gromacs molecular dynamics package [50] was employed to have an insight of the conformational changes in the catalytic domain of TMPRSS2 in the apo-form and the holo-forms. The simulations were performed using GROMOS96-54a7 force field while TIP3P water model with cubic box was used for solvating the models. The topology of the ligands were obtained from the PRODRG server by submitting their structures [51]. The system's total charge calculated was +1 which was neutralized using chloride counter-ions by replacing a water molecule. The steepest descent algorithm followed by the conjugate algorithm was utilized with 50,000 steps energy minimization of the system. All the bond angles were restrained with the LINCS algorithm. Equilibration of the solvated system was performed with NVT (constant number of particles, volume and temperature) followed by NPT (constant number of particles, pressure and temperature) with 300K and 1.0 atm respectively. Finally, the pre-equilibrated systems were put on production run for 100ns. Final molecular dynamics trajectories were analyzed by GROMACS analysis packages and the graphs/plots were visualized in qtGrace. For determining the variations in binding energy throughout the trajectory for each complex, the frames at every 100ps were extracted and submitted to the online webserver "PRODIGY"[52].

Dihedral PCA ($_d$PCA) analysis

Dihedral PCA (principal component analysis) was used to describe the high-amplitude concerted motion from the MD trajectories of protein based on eigenvectors calculated using covariance matrix [53, 54]. The dihedral angles of protein defined the atomic fluctuation throughout the MD simulation, and were described by the cosine values of the PC of covariance matrix. The cosine values checked whether the trajectory has ensembled enough to show the free energy landscape obtained from the $_d$PCA analysis [55, 56]. The range of cosine value from 0 to 1 in the total time of MD simulation (T) is given by

$$Ci = \frac{2}{T}\left(\int_0^T \cos\left(\frac{i\pi t}{T}\right) p_i(t)dt\right)^2 \left(\int_0^T p_i^2(t)dt\right)^{-1}$$

where $p_i(t)$ is the ith PC's value. Thus, absolute and sensitive parameters of trajectory was measured by getting numerous free-energy minima, which relates to conformations mapping with their respective energy basins as available in the free energy landscape of the selected PCs. Generally, the first few PCs contributions define nature of the protein. However, in most of the eigenvectors, the cosine values were close to one due to a large-scale motion in the protein dynamics, and, hence, not used [57, 58].

Results and Discussion

Structure prediction and characterization of TMPRSS2

The 3D model of TMPRSS2 obtained in the present study from Phyre2 webserver was manually analyzed. The domains of TMPRSS2 were identified as Low Density Lipoprotein

(LDL) domain (113-148), Scavenger Receptor Cysteine-Rich (SRCR) domain (149-242) and the Catalytic domain (255-492) (Fig. 1A) [59].

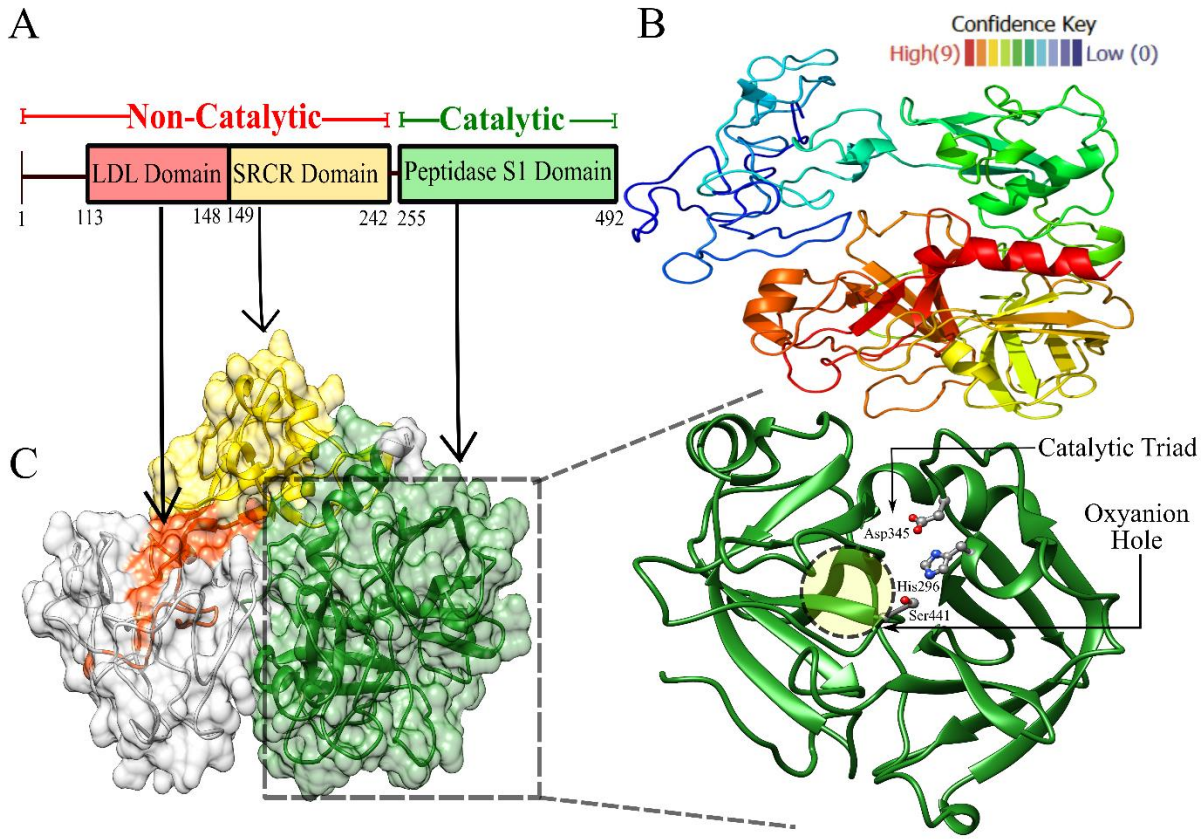


Fig. 1. **Structural characterization of TMPRSS2**. **(A)** Different domains (indicated by different colors and residue numbers). **(B)** Predicted 3D structure (from Phyre webserver). Colors in the cartoon representation indicate the confidence of prediction. **(C)** Surface representation of complete TMPRSS2 (The catalytic domain is further zoomed to show catalytic triad residues and oxyanion hole).

The catalytic domain was well predicted with high confidence (Fig. 1B) whereas prediction for non-catalytic regions comprising of LDL and SRCR domains showed a low confidence. The structure alignment of TMPRSS2 catalytic domain with the Thrombin showed Root Mean Square Deviation (RMSD) of 0.59 Å. The Ramachandran plot for the truncated catalytic domain of TMPRSS2 (265-492) from Procheck server showed 98% of the residues in

the allowed or the most favored regions whereas only one residue (Gln276) lay in the disfavored region (Supplementary Fig. SF2). Thus, the modelled structure was found suitable for further docking and MDS studies.

Molecular docking studies

Autodock results were analyzed in terms of low binding energy, high number of hydrogen-bonds and ligands docked-poses. The best docked poses of the top seven phytochemicals were selected (Supplementary Table ST1). The binding energy of these phytochemicals were better (lower) than camostat (Fig. 2, Table 2) and docked in the proximity of oxyanion hole/catalytic triad of the active site of TMPRSS2. The stability of the selected docked-poses for the molecule was tested by running MDS.

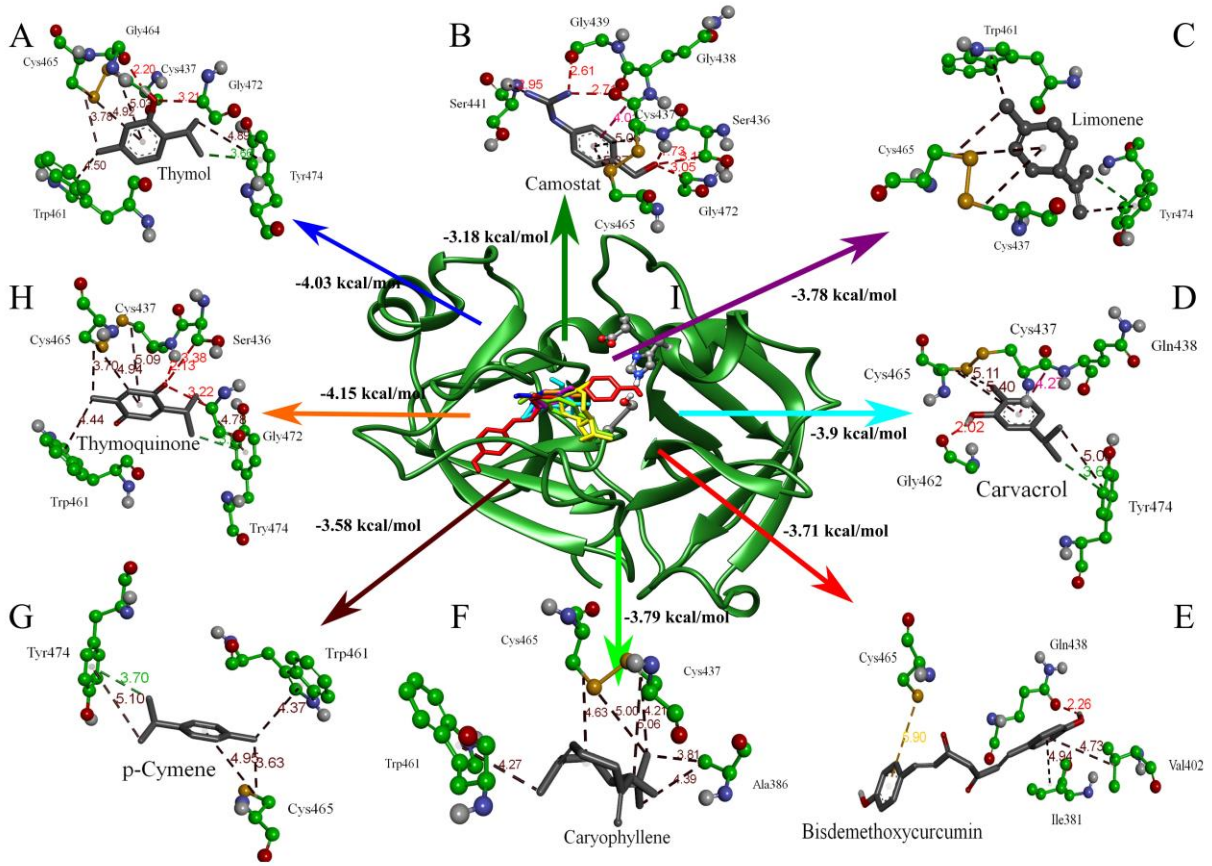


Fig. 2. **Docked poses of selected phytochemicals and interacting residues**. Picture in the center shows the docked poses of phytochemicals with catalytic domain of TMPRSS2. (A-H) Non-bonded interaction of each phytochemical. Dotted line in Red denotes hydrogen bond, Green denotes Pi-Sigma Bonds, Brown denotes Alkyl bonds, Orange denotes Pi-Sulphur bonds and Pink denotes Pi-Amide bonds. Binding energy scores are written beside the arrows.

Table 1: Seven best phytochemicals with their binding energy and H-bond for the best docked poses

Phytochemical compound	**CID**	**Binding Energy** (kcal/mol)	**H-bonds**	**Interacting Residues for H-bond**
Thymoquinone	10281	-4.15	3	Ser436, Cys437,

				Gly472
Thymol	6989	-4.03	2	Gly464, Gly472
Carvacrol	10364	-3.9	1	Gly462
Caryophyllene	5281515	-3.79	0	None
Limonene	22311	-3.78	0	None
Bisdemethoxycurcumin (BDMC)	5315472	-3.71	1	Gln438
p-Cymene	7463	-3.58	0	None
Camostat (Reference Molecule)	5284360	-3.18	6	Ser436, Cys437, Gly438, Gly439, Ser441, Gly472

MDS of Apo- and Holo- form of TMPRSS2

Root Mean Square Deviation (RMSD) and Root Mean Square Fluctuation (RMSF)

Protein-RMSD and Ligand-RMSD averaged over 500 frames for both the apo-form and the holo-forms of TMPRSS2 (Catalytic domain) were monitored and the graphs are presented in Fig. 3A and Fig. 3B respectively. The Protein-RMSD for TMPRSS2 apo-form showed a slight rise within initial 5ns and thereafter remained almost stable throughout the simulation showing only a slight fluctuation at 19ns and 60ns. The Protein-RMSD with thymol and carvacrol also showed stability except at 20ns for thymol and between 70-90ns for carvacrol. Further, except bisdemethoxycurcumin (BDMC), other ligands showed more RMSD fluctuation in comparison to the apo-form of TMPRSS2. High fluctuation in Protein-RMSD was observed in TMPRSS2 with limonene, camostat, p-cymene, thymoquinone and caryophyllene. The Protein-RMSD of TMPRSS2 with camostat showed a continuous rise from around 3.0Å to 5.3Å within first 40ns and remained at around 5.0 Å with slight fluctuations afterwards.

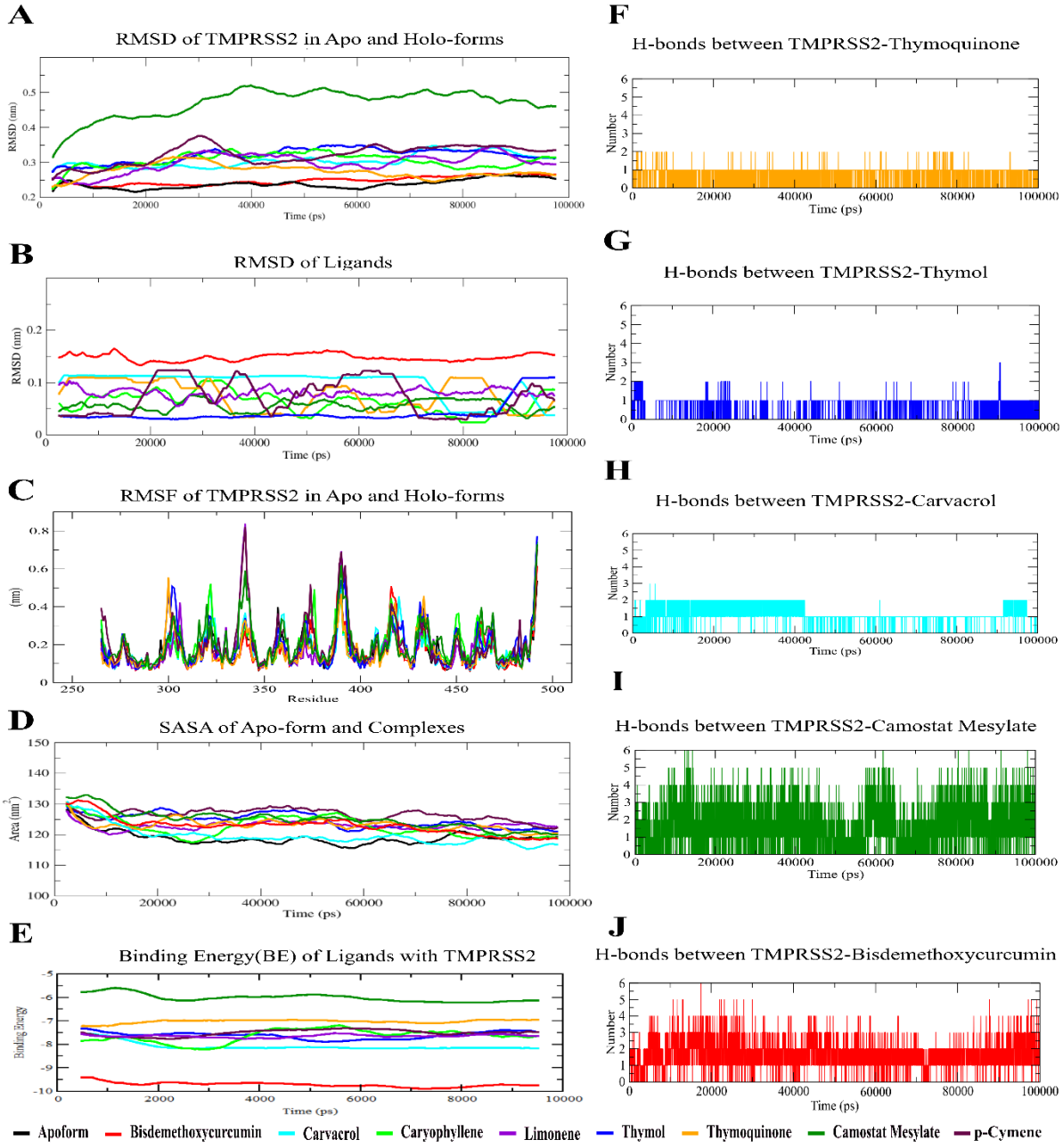


Fig. 3. **Molecular Dynamics Simulation analysis of TMPRSS2**. **(A)** RMSD of the apo-form and the holo-forms of TMPRSS2. **(B)** RMSD of the ligands. **(C)** RMSF of the apo-form and the holo-forms of TMPRSS2. **(D)** Solvent Accessible Surface Area (SASA) each complex. **(E)** Binding energy of each ligand throughout the trajectory. **(F-J)** Number of hydrogen bonds between TMPRSS2 and ligands. Black line represents the apo-form and the colored lines represent the holo-forms (assigned color for each ligand is given at the base).

High fluctuation in ligand-RMSD was observed in limonene, camostat, p-cymene, thymoquinone and caryophyllene against a very low fluctuation in thymol and carvacrol except towards the end of the simulation. On the other hand, BDMC-RMSD remained stable throughout the simulation with moderate fluctuations. Thus it can be concluded that the binding poses for BDMC, carvacrol and thymol are more stable in comparison to camostat and the rest four phytochemicals.

RMSF of TMPRSS2 for the apo-form and all holo-forms showed almost similar pattern with only a little difference. A major fluctuation was observed in loop regions in comparison to helix or sheets region (Fig. 3C).

Solvent accessible surface area (SASA), Hydrogen bond (H-Bonds) and Binding energy (BE)

SASA was calculated for each frame throughout the MD trajectory for the apo-form and all the complexes, and plotted averaging over 500 frames (Fig. 3D). All the SASA started at around 130±3nm^2 and ended at around 120±5nm^2. For TMPRSS2 complexed with caryophyllene, SASA showed the highest fluctuation with a decrement of about 10nm^2 during first 30ns, followed by an increment of 10nm^2 in next 30ns, and again a decrement at the end of the trajectory. The possibility of a high fluctuation can be correlated with the change in position of caryophyllene as it moved away from the active site (Supplementary Fig. SF4). Both for limonene and p-cymene, the complexes showed a consistency in SASA with slight fluctuations. BDMC, carvacrol, camostat and thymoquinone showed a decrement in SASA within 20ns and thereafter remained almost stable throughout the trajectory. TMPRSS2, in the apo-form, and the

complex form with carvacrol showed similar curves around 120nm^2. Its complexes with other ligands had the SASA curves within 120-130nm^2 which was more than the apo-form (Fig. 3D).

The presence of H-bond between ligand and protein was observed for each frame throughout the simulation. Five phytochemicals viz. thymoquinone, thymol, carvacrol, camostat and BDMC formed Protein-ligand H-bond (Fig. 3F-3J) whereas other ligands did not show H-bonds. Both camostat and BDMC showed high number of H-bonds, with a maximum of six H-bonds and an average of 2 to 3 H-bonds throughout the simulation. Thymoquinone and thymol generally showed one H-bond but sometimes two H-bonds. For carvacrol, a consistency of two H-bonds was observed from 5ns to 43ns, and also after 85ns, but in the rest of the trajectory one H-bond was observed.

For accounting variations in the BE of each ligand throughout the MDS, the coordinates were extracted at every 100ps from the trajectory and submitted to the Prodigy server. The results have been presented in Fig. 3E. All phytochemicals showed better BE than camostat (-6.0 kcal/mol), and among them, BDMC showed the lowest average BE (-9.7 kcal/mol). A stability in the BE curve was observed for camostat, thymoquinone, carvacrol and BDMC from 25ns onwards till the end of the trajectory whereas caryophyllene, p-cymene and limonene showed fluctuations throughout the trajectory. In initial 20ns, there was a decrement in the BE for BDMC, carvacrol, thymol, p-cymene, whereas there was an increment for camostat, thymoquinone and caryophyllene.

Dihedral principle component analysis ($_d$PCA)

$_d$PCA was performed to understand the structural behavior of TMPRSS2 in both the apo-form and the holo-forms. The free energy landscape (FEL) was drawn using the largest two principle components with the cosine value less than 0.2 [55]. Analysis of FEL showed that the apo-form

and the complex of TMPRSS2 with thymol, carvacrol and BDMC have converged into a big low energy cluster symbolizing inter-convertible low energy conformational population, thus concluding that the protein has attained a stable form (Supplementary Fig. SF3). Other complexes could not attain stable conformations within 100ns as reflected by formation of either several small clusters (p-cymene) or several medium size low free energy conformational clusters (caryophyllene, thymoquinone, carvacrol, limonene and camostat). Thus, the coordinate from minima of the largest cluster was extracted for TMPRSS2 complexed with thymol, carvacrol and BDMC for the analysis of binding pose/location (Fig. 4, Supplementary Table ST2, ST3 & ST4), which was subsequently zoomed to observe non-bonded interactions.

Carvacrol was found to bind deep in the oxyanion hole with one H-bond and ten hydrophobic interactions. Thymol remained at the entry point of the oxyanion hole forming one H-bond and four hydrophobic interactions. On the other hand, BDMC as a larger molecule occupied the entire oxyanion hole with two H-bonds and five hydrophobic interactions.

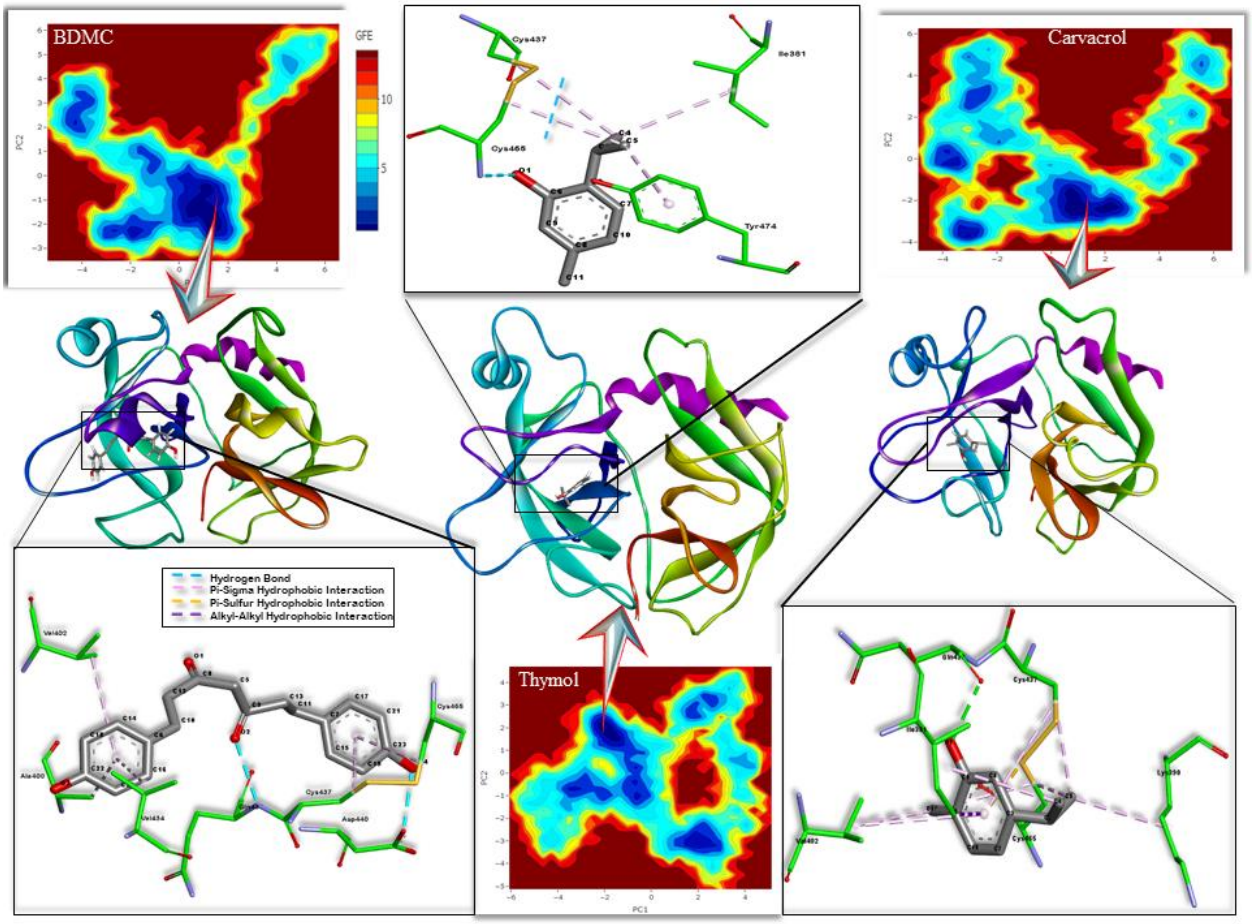


Fig. 4. **Free Energy Landscape (FEL) plot of BDMC, carvacrol and thymol**. Representative structures from the most populated low-energy cluster is shown to depict the binding pose. Zoomed image depict non-bonded interaction of ligand with TMPRSS2. (Zoomed image has been re-orientated for showing the best view).

Influence of carvacrol, thymol, BDMC on the catalytic triad

The distance between the Cα atoms of the catalytic triad residues of TMPRSS2 (Cα-His296, Cα-Asp345 and Cα-Ser441) were also monitored throughout the simulation in both the apo-form and the holo-form with carvacrol, thymol and BDMC (Fig. 5). A significant increment in the distance between Cα-His296 and Cα-Ser441 as well as Cα-Asp345 and Cα-Ser441 was observed due to binding of carvacrol. Thymol increased the distance between Cα-Asp345 and Cα-Ser441 whereas BDMC showed negligible change in Cα distance of the catalytic triad residues.

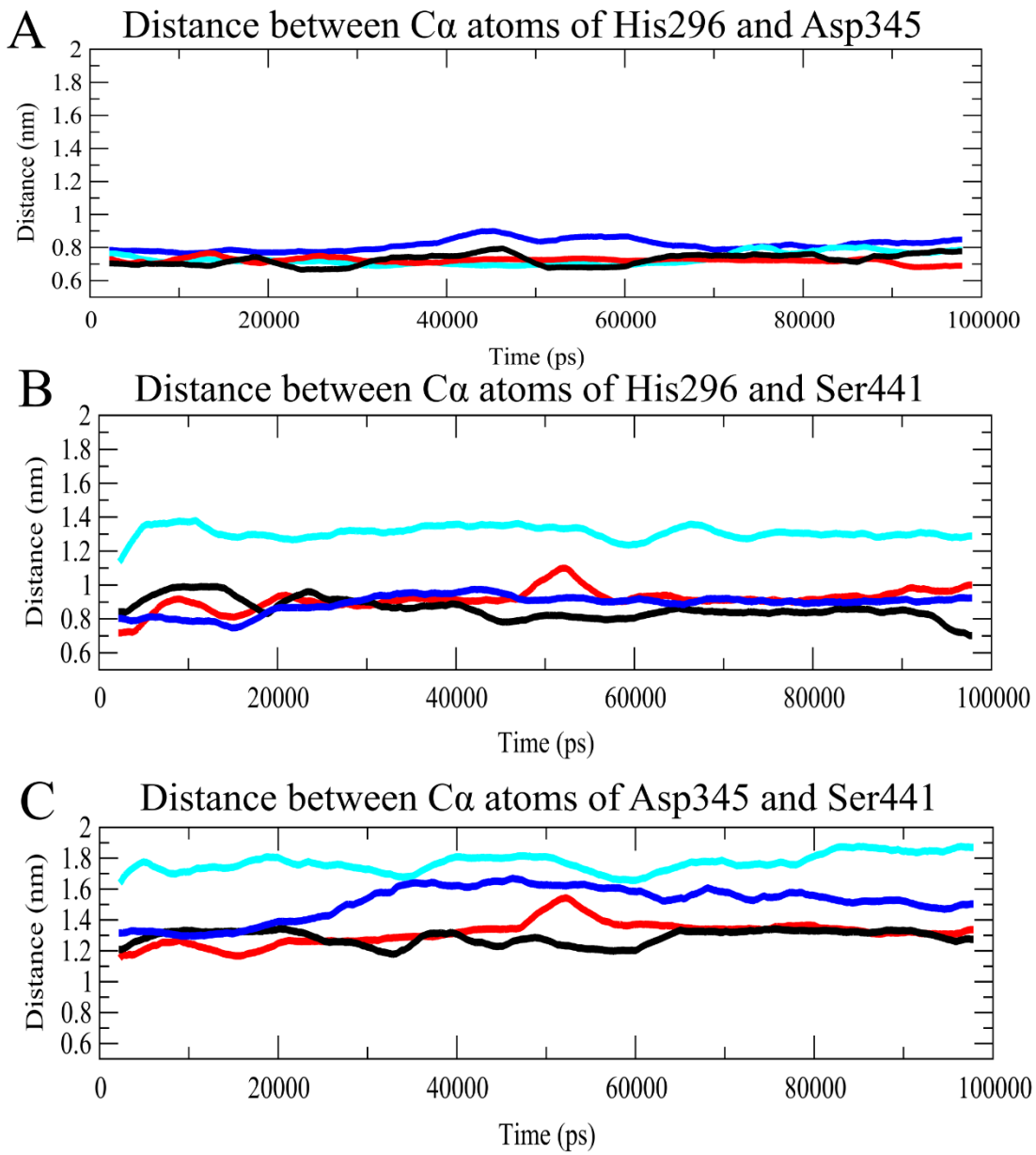


Fig. 5. **Distance between Cα atoms of catalytic triad residues throughout the trajectory. (A)** Cα-His296 and Cα-Asp345 **(B)** Cα-His296 and Cα-Ser441 **(C)** Cα-Asp345 and Cα-Ser441. Color depiction: Apo-form of TMPRSS2 (black line) and holo-form of TMPRSS2 with BDMC (red line), carvacrol (cyan line) and thymol (blue line). The distance between catalytic triad residues shows maximum deviation due to binding of carvacrol followed by thymol whereas the least with BDMC.

Picking up of the best TMPRSS2 inhibitor

Stable binding of ligands to oxyanion hole can block interaction of arginine of the substrate with the oxyanion hole which plays a crucial role in accommodating residue arginine of the substrate to ignite cleavage of its peptide bond.

On the basis of the results (stable RMSD, low binding energy, number of H-bonds throughout simulation and $_{d}$PCA analysis), we conclude that the BDMC, carvacrol and thymol form stable binding with TMPRSS2 in the oxyanion hole and modified the spatial arrangement of the catalytic triad residues. The change in the spatial arrangement of the catalytic triad was the highest in carvacrol, followed by thymol, and the least with BDMC. However, BDMC being a large molecule could effectively shield the oxyanion hole. Levels of inhibition among stable bound molecules would be graded as: best carvacrol, then BDMC, and least thymol. Therefore, we conclude that blocking of TMPRSS2 by these phytochemicals is expected to prevent ACE2- and TMPRSS2-mediated cell entry of SARS-CoV-2 and other viruses into the host cells.

Conclusion

The *in-silico* study conducted to designate some potential antiviral phytochemicals present in some common spices used in Indian kitchens to inhibit the activity of TMPRSS2 included docking of phytochemicals with the catalytic domain of TMPRSS2 for specifying the best binding pose in terms of docked binding energy followed by MD simulation scrutiny and identified three potential antiviral phytochemicals namely carvacrol, thymol and BDMC found in *N. sativa*, *T. ammi*, and *C. longa* respectively which might be studied further as potential drug candidates against SARS-CoV-2.

Acknowledgements

Authors acknowledge Department of Biotechnology (DBT), Govt. of India for providing financial support to the Bioinformatics Centre of the University where this work was carried out, and also to the DBT electronic library consortium for providing access to the e-resources. First author thanks the DBT for providing Post-doctoral Fellowship. Critical suggestions during the finalization of the manuscript from Dr. K. C. Mishra, a faculty of the University are acknowledged.

Abbreviations

COVID-19 = Corona Virus Disease

SARS-CoV-2 = Severe Acute Respiratory Syndrome Coronavirus 2

TMPRSS2 = transmembrane protease serine 2

BDMC = bisdemethoxycurcumin

MDS = Molecular Dynamic Simulation

ACE2 = Angiotensin-converting enzyme 2

$_{d}$PCA = Dihedral PCA

RMSD = Root Mean Square Deviation

References

1. Lu, R., et al., *Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding.* The Lancet, 2020. **395**(10224): p. 565-574.
2. Chan, J.F.-W., et al., *A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster.* The Lancet, 2020. **395**(10223): p. 514-523.
3. WHO, *Coronavirus disease (COVID-2019) situation reports - 162*. 2020: World Health Organization.
4. Prasad, A., M. Muthamilarasan, and M. Prasad, *Synergistic antiviral effects against SARS-CoV-2 by plant-based molecules.* Plant Cell Reports, 2020.
5. Mani, J.S., et al., *Natural product-derived phytochemicals as potential agents against coronaviruses: A review.* Virus Res, 2020. **284**: p. 197989.
6. Orhan, I.E. and F.S. Senol Deniz, *Natural Products as Potential Leads Against Coronaviruses: Could They be Encouraging Structural Models Against SARS-CoV-2?* Nat Prod Bioprospect, 2020.
7. Costanzo, M., M. De Giglio, and G. Roviello, *SARS CoV-2: Recent Reports on Antiviral Therapies Based on Lopinavir/Ritonavir, Darunavir/Umifenovir, Hydroxychloroquine, Remdesivir, Favipiravir and Other Drugs for the Treatment of the New Coronavirus.* Current medicinal chemistry, 2020.
8. Forni, D., et al., *Molecular Evolution of Human Coronavirus Genomes.* Trends Microbiol, 2017. **25**(1): p. 35-48.

9. Zaki, A.M., et al., *Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia.* N Engl J Med, 2012. **367**(19): p. 1814-20.

10. Drosten, C., et al., *Identification of a novel coronavirus in patients with severe acute respiratory syndrome.* N Engl J Med, 2003. **348**(20): p. 1967-76.

11. Huang, C., et al., *Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China.* Lancet, 2020. **395**(10223): p. 497-506.

12. Mou, H., et al., *The receptor binding domain of the new Middle East respiratory syndrome coronavirus maps to a 231-residue region in the spike protein that efficiently elicits neutralizing antibodies.* J Virol, 2013. **87**(16): p. 9379-83.

13. Hou, Y.J., et al., *SARS-CoV-2 Reverse Genetics Reveals a Variable Infection Gradient in the Respiratory Tract.* 2020.

14. Iwata-Yoshikawa, N., et al., *TMPRSS2 Contributes to Virus Spread and Immunopathology in the Airways of Murine Models after Coronavirus Infection.* J Virol, 2019. **93**(6).

15. Glowacka, I., et al., *Evidence that TMPRSS2 activates the severe acute respiratory syndrome coronavirus spike protein for membrane fusion and reduces viral control by the humoral immune response.* J Virol, 2011. **85**(9): p. 4122-34.

16. Hoffmann, M., et al., *SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor.* Cell, 2020.

17. Iwata-Yoshikawa, N., et al., *TMPRSS2 contributes to virus spread and immunopathology in the airways of murine models after coronavirus infection.* Journal of virology, 2019. **93**(6): p. e01815-18.

18. Matsuyama, S., et al., *Enhanced isolation of SARS-CoV-2 by TMPRSS2-expressing cells.* Proceedings of the National Academy of Sciences, 2020. **117**(13): p. 7001-7003.

368 19. Vasanthi, H.R. and R.P. Parameswari, *Indian spices for healthy heart - an overview.* Curr Cardiol
369 Rev, 2010. **6**(4): p. 274-9.
370 20. Chongtham, A. and N. Agrawal, *Curcumin modulates cell death and is protective in Huntington's*
371 *disease model.* Sci Rep, 2016. **6**: p. 18736.
372 21. Saleem, U., et al., *Pharmacological Screening of Trachyspermum ammi for Antihyperlipidemic*
373 *Activity in Triton X-100 Induced Hyperlipidemia Rat Model.* Pharmacognosy Res, 2017. **9**(Suppl
374 1): p. S34-S40.
375 22. Singh, S., *From exotic spice to modern drug?* Cell, 2007. **130**(5): p. 765-8.
376 23. Alburn, H.E. and G. Greenspan, *Thymol as an anti-influenza agent*. 1972, Google Patents.
377 24. Chen, T.Y., et al., *Inhibition of enveloped viruses infectivity by curcumin.* PLoS One, 2013. **8**(5): p.
378 e62482.
379 25. Rathinavel, T., et al., *Phytochemical 6-Gingerol–A promising Drug of choice for COVID-19.*
380 26. Consortium, U., *UniProt: a hub for protein information.* Nucleic acids research, 2015. **43**(D1): p.
381 D204-D212.
382 27. Kelley, L.A., et al., *The Phyre2 web portal for protein modeling, prediction and analysis.* Nature
383 protocols, 2015. **10**(6): p. 845.
384 28. Laskowski, R., M. MacArthur, and J. Thornton, *PROCHECK: validation of protein-structure*
385 *coordinates.* 2006.
386 29. Ramachandran, G.N., *Stereochemistry of polypeptide chain configurations.* J. Mol. Biol., 1963. **7**:
387 p. 95-99.
388 30. Kim, S., et al., *PubChem Substance and Compound databases.* Nucleic Acids Res, 2016. **44**(D1): p.
389 D1202-13.
390 31. Astani, A., J. Reichling, and P. Schnitzler, *Comparative study on the antiviral activity of selected*
391 *monoterpenes derived from essential oils.* Phytother Res, 2010. **24**(5): p. 673-9.

392 32. Knight, V., et al., *Thymol suppression of protein synthesis in influenza virus-infected and*
393 *uninfected chick embryo fibroblast cells.* Proc Soc Exp Biol Med, 1977. **155**(1): p. 128-36.
394 33. Wu, Q.F., et al., *Chemical compositions and anti-influenza activities of essential oils from Mosla*
395 *dianthera.* J Ethnopharmacol, 2012. **139**(2): p. 668-71.
396 34. Marchese, A., et al., *Update on Monoterpenes as Antimicrobial Agents: A Particular Focus on p-*
397 *Cymene.* Materials (Basel), 2017. **10**(8).
398 35. Sharifi-Rad, J., et al., *Susceptibility of herpes simplex virus type 1 to monoterpenes thymol,*
399 *carvacrol, p-cymene and essential oils of Sinapis arvensis L., Lallemantia royleana Benth. and*
400 *Pulicaria vulgaris Gaertn.* Cell Mol Biol (Noisy-le-grand), 2017. **63**(8): p. 42-47.
401 36. Anggakusuma, et al., *Turmeric curcumin inhibits entry of all hepatitis C virus genotypes into*
402 *human liver cells.* Gut, 2014. **63**(7): p. 1137-49.
403 37. Mounce, B.C., et al., *Curcumin inhibits Zika and chikungunya virus infection by inhibiting cell*
404 *binding.* Antiviral Res, 2017. **142**: p. 148-157.
405 38. Ting, D., et al., *Multisite inhibitors for enteric coronavirus: antiviral cationic carbon dots based on*
406 *curcumin.* 2018. **1**(10): p. 5451-5459.
407 39. Kaushik, S., et al., *Anti-viral activity of Zingiber officinale (Ginger) ingredients against the*
408 *Chikungunya virus.* Virusdisease, 2020: p. 1-7.
409 40. Chang, J.S., et al., *Fresh ginger (Zingiber officinale) has anti-viral activity against human*
410 *respiratory syncytial virus in human respiratory tract cell lines.* J Ethnopharmacol, 2013. **145**(1):
411 p. 146-51.
412 41. Imanishi, N., et al., *Macrophage-mediated inhibitory effect of Zingiber officinale Rosc, a*
413 *traditional oriental herbal medicine, on the growth of influenza A/Aichi/2/68 virus.* Am J Chin
414 Med, 2006. **34**(1): p. 157-69.

42. Rajagopal, K., et al., *Activity of phytochemical constituents of black pepper, ginger, and garlic against coronavirus (COVID-19): An in silico approach.* 2020. **9**(5): p. 43.

43. Astani, A., J. Reichling, and P. Schnitzler, *Screening for antiviral activities of isolated compounds from essential oils.* Evid Based Complement Alternat Med, 2011. **2011**: p. 253643.

44. Astani, A. and P. Schnitzler, *Antiviral activity of monoterpenes beta-pinene and limonene against herpes simplex virus in vitro.* Iran J Microbiol, 2014. **6**(3): p. 149-55.

45. Sanchez, C., R. Aznar, and G. Sanchez, *The effect of carvacrol on enteric viruses.* Int J Food Microbiol, 2015. **192**: p. 72-6.

46. Umar, S., et al., *Synergistic effects of thymoquinone and curcumin on immune response and anti-viral activity against avian influenza virus (H9N2) in turkeys.* Poult Sci, 2016. **95**(7): p. 1513-1520.

47. Spraggon, G., et al., *Active site conformational changes of prostasin provide a new mechanism of protease regulation by divalent cations.* Protein Sci, 2009. **18**(5): p. 1081-94.

48. Goodsell, D.S., G.M. Morris, and A.J. Olson, *Automated docking of flexible ligands: applications of AutoDock.* Journal of molecular recognition, 1996. **9**(1): p. 1-5.

49. Morris, G.M., et al., *Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function.* 1998. **19**(14): p. 1639-1662.

50. Van Der Spoel, D., et al., *GROMACS: fast, flexible, and free.* Journal of computational chemistry, 2005. **26**(16): p. 1701-1718.

51. Schuttelkopf, A.W. and D.M. van Aalten, *PRODRG: a tool for high-throughput crystallography of protein-ligand complexes.* Acta Crystallogr D Biol Crystallogr, 2004. **60**(Pt 8): p. 1355-63.

52. Xue, L.C., et al., *PRODIGY: a web server for predicting the binding affinity of protein-protein complexes.* Bioinformatics, 2016. **32**(23): p. 3676-3678.

53. David, C.C. and D.J. Jacobs, *Principal component analysis: a method for determining the essential dynamics of proteins.* Methods Mol Biol, 2014. **1084**: p. 193-226.

54. Amadei, A., A.B. Linssen, and H.J. Berendsen, *Essential dynamics of proteins.* Proteins, 1993. **17**(4): p. 412-25.

55. Maisuradze, G.G. and D.M. Leitner, *Free energy landscape of a biomolecule in dihedral principal component space: sampling convergence and correspondence between structures and minima.* Proteins, 2007. **67**(3): p. 569-78.

56. Altis, A., et al., *Construction of the free energy landscape of biomolecules via dihedral angle principal component analysis.* J Chem Phys, 2008. **128**(24): p. 245102.

57. Miszkiel, A., M. Wojciechowski, and S. Milewski, *Long range molecular dynamics study of regulation of eukaryotic glucosamine-6-phosphate synthase activity by UDP-GlcNAc.* Journal of Molecular Modeling, 2011. **17**(12): p. 3103-3115.

58. Maisuradze, G.G. and D.M. Leitner, *Free energy landscape of a biomolecule in dihedral principal component space: Sampling convergence and correspondence between structures and minima.* Proteins-Structure Function and Bioinformatics, 2007. **67**(3): p. 569-578.

59. Paoloni-Giacobino, A., et al., *Cloning of the TMPRSS2 gene, which encodes a novel serine protease with transmembrane, LDLRA, and SRCR domains and maps to 21q22.3.* Genomics, 1997. **44**(3): p. 309-20.

60. Neitzel, J.J.J.N.E., *Enzyme catalysis: the serine proteases.* 2010. **3**(9): p. 21.

Supplementary Information

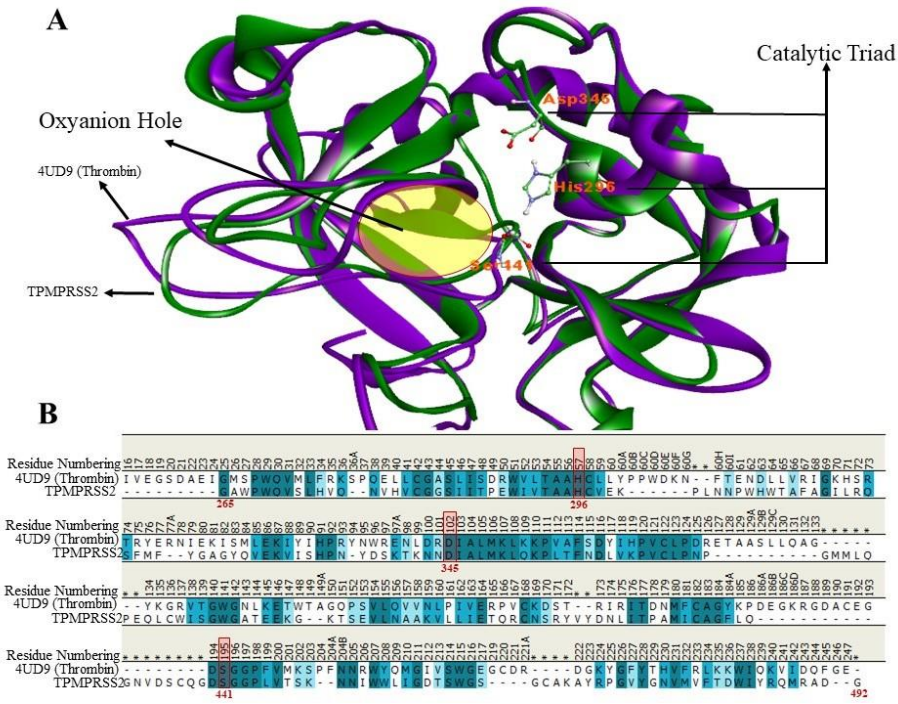


Fig. SF1: Structural alignment of TMPRSS2 with Thrombin (4UD9). **(A)** Cartoon representation for the structural alignment of Thrombin (Purple) with TMPRSS2 (Green). Oxyanion hole (yellow circle) and catalytic triad residues (ball and stick representation) are indicated. **(B)** The corresponding sequences of the structure alignment are reflected here. The first and second row shows residue-numbers and residues respectively of thrombin (4UD9). The third row show residues of TMPRSS2. The catalytic triad residues are marked with rectangular red boxes and the residue numbers of these residues in TMPRSS2 are shown just below the box.

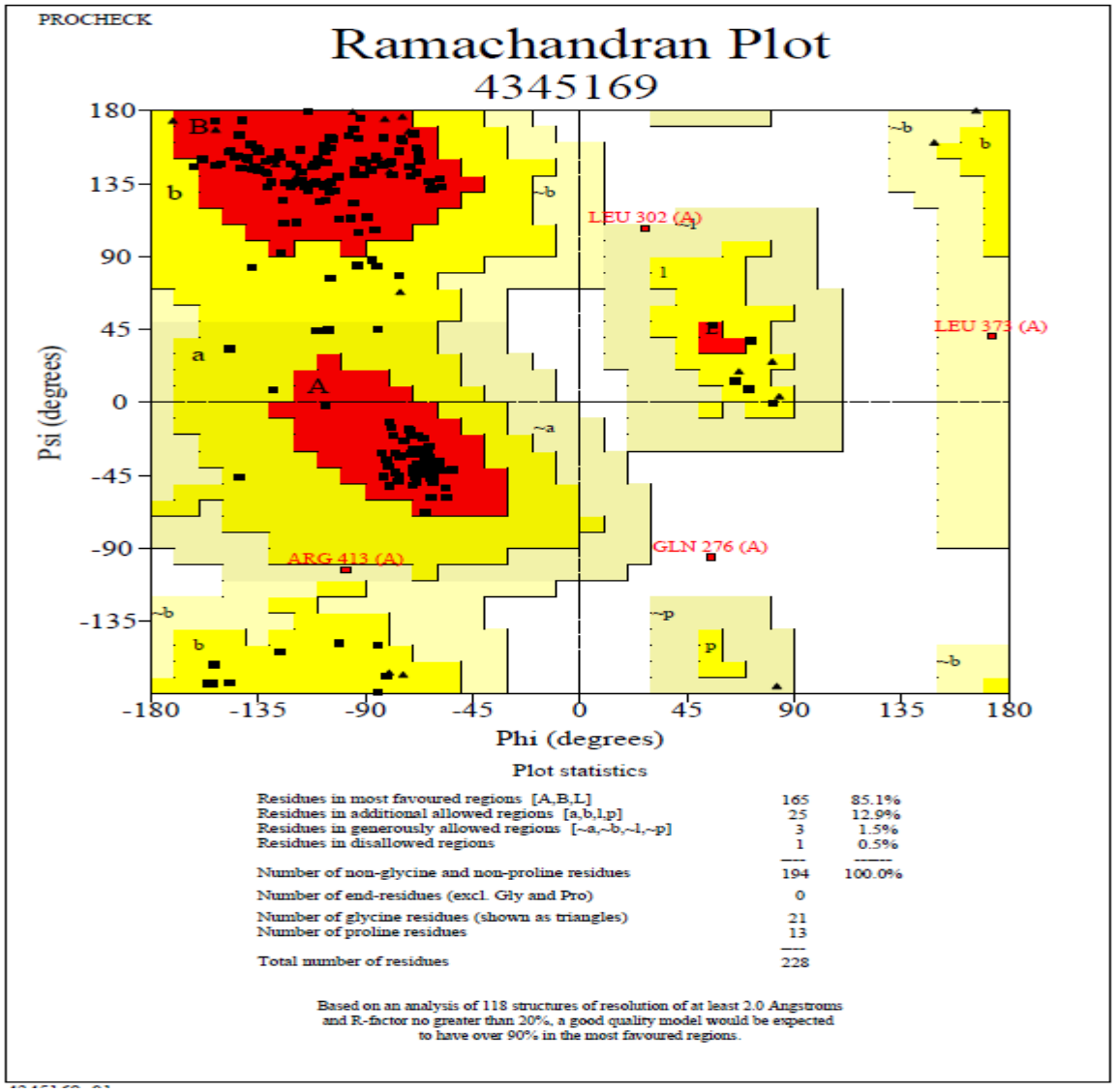


Fig. SF2: Ramachandran plot of modeled TMPRSS2

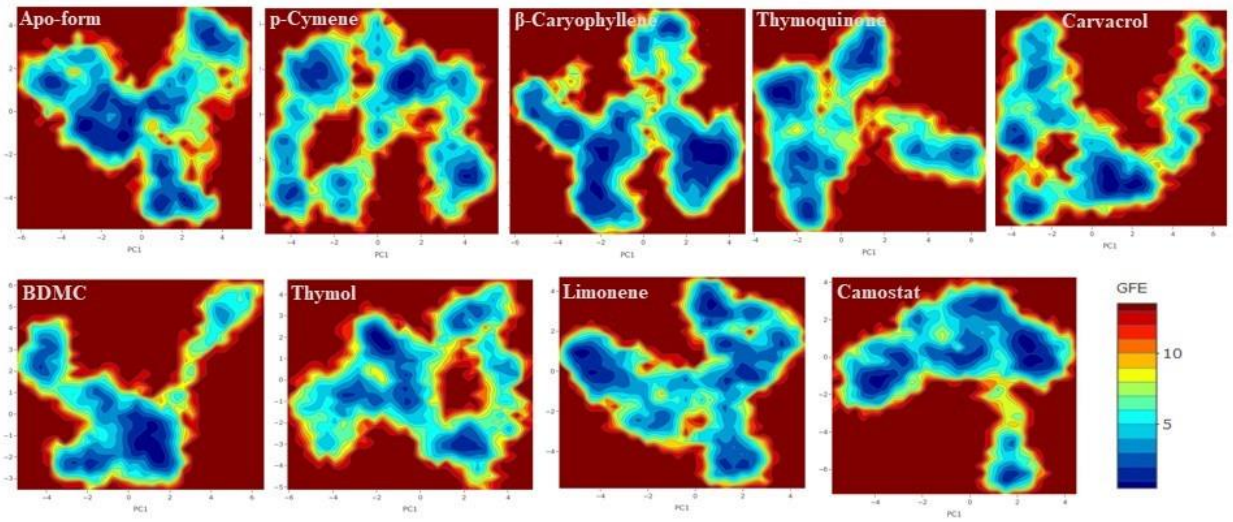


Fig. SF3: PCA analysis of TMPRSS2-Ligand complexes. Here we can observe that TMPRSS2 in apo-form, and with ligands like Carvacrol, BDMC and Thymol showed a big free energy cluster with several small clusters. TMPRSS2 with p-Cymene, Thymoquinone and limonene showed several small free energy clusters with no distinct big cluster whereas TMPRSS2 with β-Caryophyllene and Camostat exhibit three medium sized cluster with few small clusters.

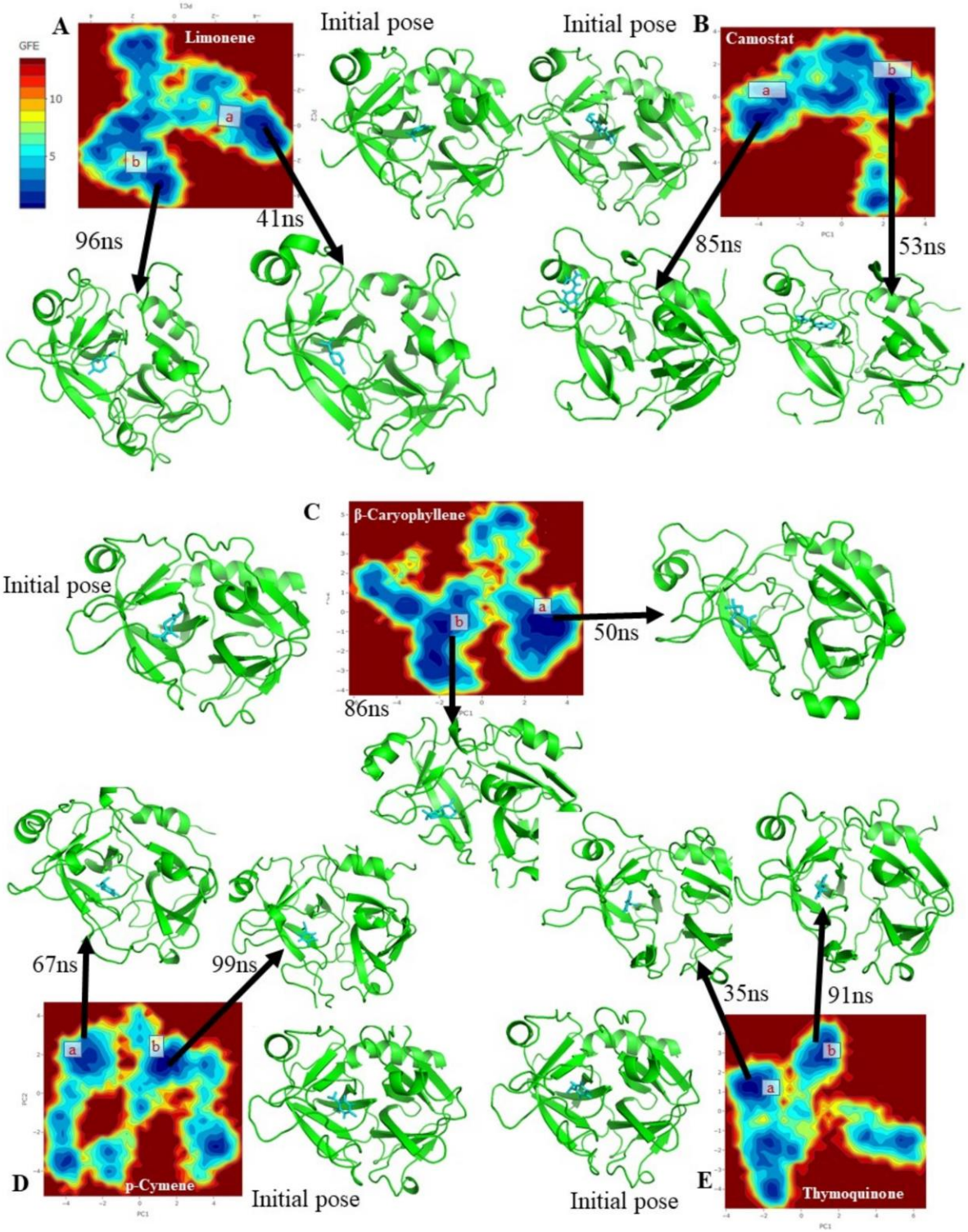


Fig. SF4: FEL for TMPRSS2 with rejected phytochemicals/ligands. Extracted frames from the larger two clusters to show the changes in the interaction pose of the ligands with TMPRSS2.

Supplementary Tables

Supplementary Table ST2: Non-bonded interaction of carvacrol with TMPRSS2 (shown in Fig. 4)

Sl. No.	TMPRSS2	Carvacrol	Distance (Å)	Type of Interaction
1	GLN438:O	O1	2.48	H-Bond
2	ILE381:CD	Pi-Orbitals	3.82	Hydrophobic
3	CYS465:SG	Pi-Orbitals	3.63	Hydrophobic
4	VAL402	C11	4.69	Hydrophobic
5	CYS465	C11	4.29	Hydrophobic
6	LYS390	C5	4.77	Hydrophobic
7	CYS437	C5	4.29	Hydrophobic
8	CYS465	C5	3.66	Hydrophobic
9	ILE381	C4	4.89	Hydrophobic
10	VAL402	Pi-Orbitals	5.33	Hydrophobic
11	CYS437	Pi-Orbitals	5.43	Hydrophobic

Supplementary Table ST3: Non-bonded interaction of thymol with TMPRSS2 (shown in Fig. 4)

Sl. No.	TMPRSS2	Thymol	Distance (Å)	Type of Interaction
1	CYS465:N	O1	3.18	H-Bond
2	ILE381	C4	4.39	Hydrophobic
3	CYS437	C4	4.99	Hydrophobic
4	CYS465	C5	4.47	Hydrophobic
5	TYR474	C5	4.13	Hydrophobic

Supplementary Table ST4: Non-bonded interaction of BDMC with TMPRSS2 (shown in Fig. 4)

Sl. No.	TMPRSS2	BDMC	Distance (Å)	Type of Interaction
1	GLN438:N	O2	2.91	H-Bond
2	ASP440:OD2	O4	2.61	H-Bond
3	ALA400	Pi-Orbitals	4.30	Hydrophobic
4	VAL402	Pi-Orbitals	5.42	Hydrophobic
5	VAL434	Pi-Orbitals	4.97	Hydrophobic
6	CYS437	Pi-Orbitals	4.84	Hydrophobic
7	CYS465	Pi-Orbitals	4.76	Hydrophobic