Structural and functional features of the transmembrane serine proteases – Implications in viral pathogenesis

Shan Aziz

Department of Natural Sciences and Mathematics, The University of Texas at Dallas

Dr. Mehmet Candas

Structural and functional features of the transmembrane serine proteases – Implications in viral pathogenesis

Abstract:

There are at least eight members of the transmembrane protease serine (TMPRSS) family whose protein sequences are known. These members, which include TMPRSS2, TMPRSS3, TMPRSS4, TMPRSS5, TMPRSS6, TMPRSS7, TMPRSS9, TMPRSS11, TMPRSS12, and TMPRSS13, have significant functions in human hemostasis and are linked to the onset of several ailments, such as certain cancers. TMPRSS proteins have also been shown to facilitate the entrance of respiratory viruses into human cells. Particularly, TMPRSS2 and TMPRSS4 stimulate the SARS-CoV-2 virus' spike protein. None of these proteins have been crystallized, despite the variety of roles they play in the human body. Identifying effective treatment options for diseases related to the TMPRSS protein family has been challenging due to the lack of structural data. However, this publication presents homology models for every member of the family, enabling identification of key amino acids and possible binding sites for drug targeting. Comparing the 3D structures of TMPRSS2 with similar enzymes has led to discoveries of distinct traits for each family member. This information can facilitate development of specific inhibitors or modulators for TMPRSS, potentially leading to effective treatments. These findings could significantly impact the field, paving the way for future research. TMPRSS2 is a protein that plays a critical role in the entry of SARS-CoV-2 virus into host cells. Its structure can be targeted for drug development to inhibit its activity and block viral entry, potentially leading to the development of antiviral drugs to treat COVID-19. Additionally, determining the structure of TMPRSS2 can help in the development of specific inhibitors that can target TMPRSS2 only while ignoring other related proteins, potentially minimizing the adverse effects of antiviral medications.

Introduction:

The Type II Transmembrane Serine Protease (TTSP) family is a group of enzymes that are critical for many physiological processes. They are divided into four subfamilies based on specific domains and require activation through cleavage of a conserved activation motif. Once activated, they are involved in processes such as blood coagulation, digestion, and tissue remodeling. The TTSP family is a target for drug development due to their involvement in various pathological conditions. Understanding their activation and regulation mechanisms is crucial for developing effective therapies. (Antalis et al., 2011; Bugge et al., 2009). After activation, the catalytic and transmembrane domains are kept connected by disulfide links (Antalis et al., 2011). Currently, there are 14 distinct TMPRSS proteins that have been identified. These include TMPRSS2, TMPRSS3, TMPRSS4, TMPRSS5, TMPRSS6, TMPRSS7, TMPRSS9, TMPRSS11A, TMPRSS11B, TMPRSS11D, TMPRSS11E, TMPRSS11F, TMPRSS12, and TMPRSS13 (Escalante et al., 2021). Although not all of this subfamily's physiological roles, including homeostasis and proteolytic cascades, have been discovered, their involvement in many pathogenicities is starting to become clear, as illustrated in Table 1 (Supplementary Information) (Webb et al., 2011a).

Table A:
TMPRSS proteins' recognized functions

Proteins	Function
TMPRSS 2	Vital function in the pathogenesis of several viruses, including SARS-CoV-2 and MERS-CoV.
TMPRSS 3 and 5	Cochlear hair cells, which are expressed in the inner ear and where mutations cause hearing loss.
TMPRSS 4	A pancreatic tumor that has an overexpression.
TMPRSS 6	Controls hepcidin levels to regulate iron homeostasis, and mutations result in iron-refractory iron deficiency anemia (IRDA).
TMPRSS 9	Enhances pancreatic cancer's ability to spread invasively.
TMPRSS 11A	Down-regulated in esophageal cancer.
TMPRSS 11B	Increases lactate export and glycolytic metabolism, which both serve to develop lung cancer.
TMPRSS 12	frequently shown in colorectal cancer.

Table A: This table shows the TMPRSS subfamily proteins and their functions. From (Cal et al., 2003; Chen et al., 2010; Vaarala et al., 2001; Wallrapp et al., 2000; Webb et al., 2011b, 2011a; Zang et al., 2020; Ziegler et al., 2020)

Due to TMPRSS2's critical function in controlling SARS-CoV-2 virulence and its potential contribution to the emergence of malignancies, the TMPRSS family has drawn a lot of attention. Lung cells primarily incorporate SARS-CoV-2 and other pathogens via the TMPRSS2 protein and ACE2 receptor. The virus's TMPRSS2 enzyme specifically targets the Spike protein at a location determined by Arg255 and Ile256 to initiate membrane insertion. ACE2, through a specific interaction with the Spike protein's receptor binding domain, attracts the Spike protein to the surface of the host cell. These mechanisms are critical for the virus's ability to infect and replicate within host cells. For COVID-19 and other respiratory diseases to be effectively treated and prevented, it is crucial to comprehend how TMPRSS enzymes, ACE2 receptors, and Spike proteins interact (Shen et al., 2017). The Spike protein of SARS-CoV-2 has two functional subunits, S1 and S2, responsible for binding to host cells and fusing viral and cellular membranes, respectively (Fung et al., 2019; S. Liu et al., 2004). In vivo investigations have revealed that TMPRSS2, the primary processing enzyme for virus penetration in lung cells, is a key target for developing medicines to treat SARS-CoV-2

infections. (Hatesuer et al., 2013; Iwata-Yoshikawa et al., 2019). As a result, TMPRSS2 has become a key target for the development of medicines to treat SARS-CoV-2 infections. Camostat, a recognized inhibitor of TMPRSS2, has shown promising results in keeping infected mice alive. However, it is not very selective for TMPRSS2 and lacks crystal structures for the TMPRSS protein subfamily. (Shen et al., 2017). Further research is needed to create more selective medicines. This study expands on earlier research on camostat bound to TMPRSS2 to find common features of recognition throughout the TMPRSS sub-family. Developing effective treatments for SARS-CoV-2 infections is crucial, and understanding the role of TMPRSS2 is an important step towards achieving this goal.

MATERIALS AND METHODS

Homology Modeling:

In this study, we utilized the Uniprot database to retrieve amino acid sequences for 14 proteins, identified by gene IDs O15393, P57727, Q9NRS4, Q9H3S3, Q9DBI0, Q7RTY8, Q7Z410, Q6ZMR5, Q86T26, O60235, Q9UL52, Q6ZWK6, Q86WS5, and Q9BYE2. To determine the crystal structures with the highest sequence identity for these proteins, we employed the BLASTp methodology and searched the Protein Databank (PDB). We then utilized templates from 23 crystal structures to construct models for the TMPRSS proteins. Notably, three crystal structures stood out: (6O1G) a human plasma kallikrein, DESC1, a member of the TTSP family, and (1O5E) a urokinase-type plasminogen activator with a mutation at position 190 (alanine). These findings provide valuable insights into the structure and function of TMPRSS proteins and may have implications for their potential use as therapeutic targets.

Ligands are essential for understanding the crystal structures of different proteins. Recently, three ligands were identified and attached to crystal structures of TMPRSS proteins. These

ligands were sourced from compounds 1O5E, 2OQ5, and 6O1G and were named 6-Chloro-2-(2-Hydroxy-Biphenyl-3-yl)-1H-indole-5-carboxamidine, benzamide, and 7SD, respectively. All three ligands were successfully retained in the homology models that were developed. Using the InterPro, Pfam member database, the trypsin domain-corresponding amino acid region for each TMPRSS protein was determined. The crystal structures discovered through the Pfam database were then reduced to only their trypsin domains. This process has provided valuable insights into the structure and function of TMPRSS proteins, which could have significant implications for drug discovery and development. Each sequence was aligned using the T-Coffee Consistency-based MSA tool, which attempts to get beyond the drawbacks of progressive alignment methods. Any conserved secondary structure alignments were detected as shown in (Fig A). Each TMPRSS protein has at least five crystal structures chosen and homology models were predicted and refined using Modeller; these crystal structures are provided in Supplementary Table 2. Each homology model that was created particularly for a given protein was then evaluated using SAVES v6.0 and then used to create a final consensus model. Each TMPRSS protein's final model is designated as hmTMPRSSx, where x is the protein's particular name.

Molecular Docking:

The CB-Dock2 tool was used to conduct a molecular docking research to determine the binding affinity between the ligand molecules and the receptor protein (Y. Liu et al., 2019). The more advanced CB-Dock2 technique uses four key processes to dock proteins with ligands: data input processing, cavity identification, docking, and visualization and analysis of the docked complex (Y. Liu et al., 2022). Transmembrane protease serine 2 inhibitor was retrieved using the PubChem database (https://pubchem.ncbi.nlm.nih.gov/) (CID_2536) the ligands and proteins used in the molecular interaction are entered in the SDF/MOL2 formats

and the PDB format, respectively. The tool checks the submitted protein automatically for any side chain atoms and hydrogen atoms that are absent. The ligands that were submitted have been modified by partial charges and hydrogen addition. The program used by CB-dock to do docking predicts the binding sites for the receptor protein, determines the center and size using a curvature-based cavity identification method, and performs docking.

Results:

According to Fig. A, all fifteen sequences exhibit a sufficient level of identity to template models (>40%). As a result, we were able to create homology models for each TMPRSS family member. The Modeller Loop Refinement was used to perform loop refinement on all of the homology models. The trypsin-like serine protease domain of TMPRSS2 refers to a specific region of the protein that is structurally similar to trypsin, a digestive enzyme that breaks down proteins. The region of TMPRSS2 that is being discussed plays a crucial role in the proteolytic activity of the protein. This activity enables it to cleave and activate the spike protein on the surface of the SARS-CoV-2 virus, which is essential for the virus to enter human cells. The catalytic triad, consisting of histidine, aspartic acid, and serine, is conserved in all TMPRSS proteins, as demonstrated in Fig 2. Although the catalytic triad appears to be at different points across the amino acid sequences, the histidine, aspartic acid, and serine are all located at comparable places in structural space inside the trypsin area. Furthermore, the geometry of the catalytic triad is maintained based on the structural alignment, as shown in Fig 3. These findings provide valuable insights into the functioning of TMPRSS2 and its potential as a therapeutic target for COVID-19.

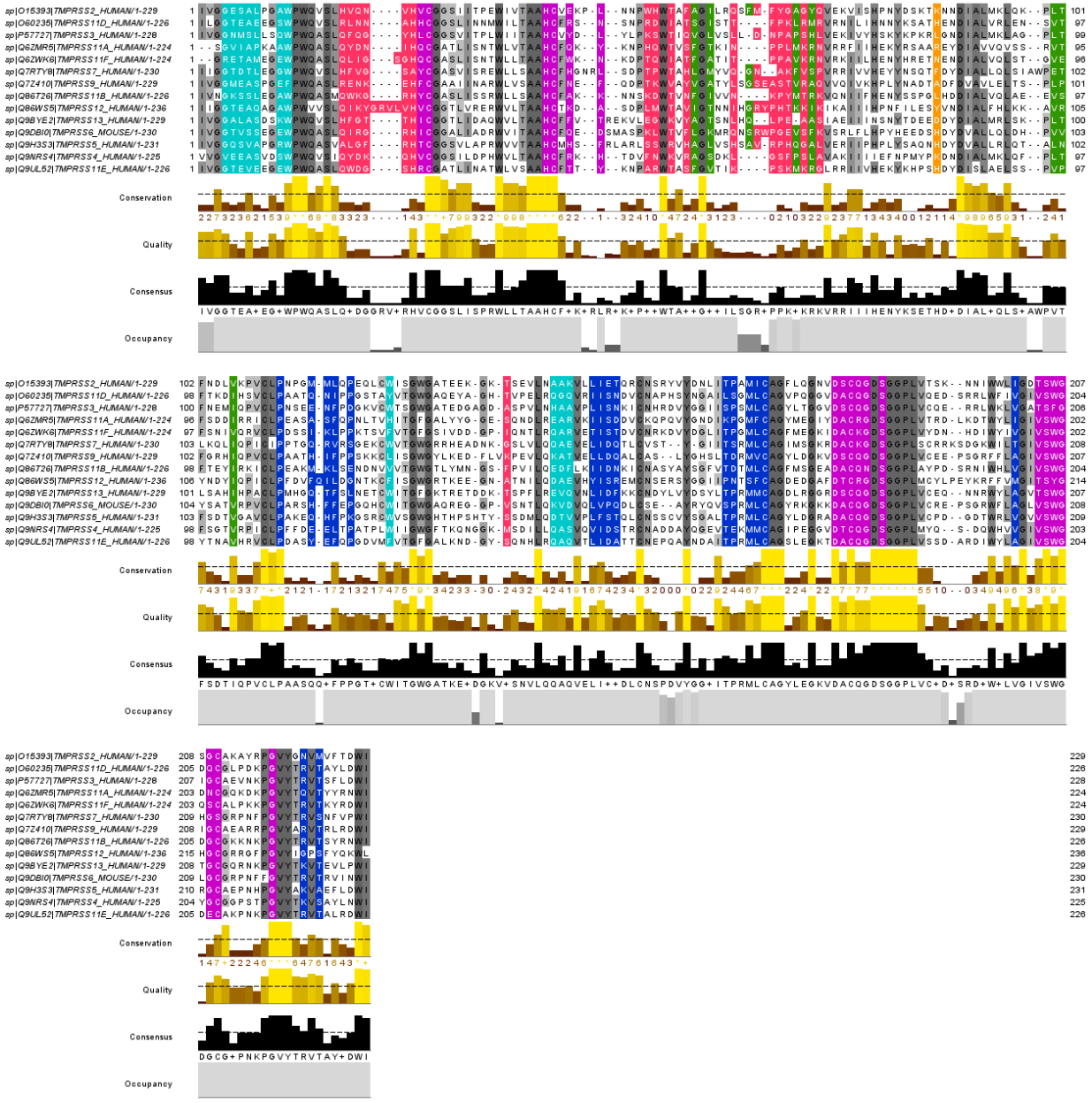


Fig 1: Sequence alignment for every member of the TMPRSS family using Jalview (2.11.2.6). Residues that are totally conserved by more than 50% of all TMPRSS sequences are highlighted in dark gray, whereas high-identity residues are highlighted in light gray. pockets (C1-C5) close to the TMPRSS2 active site that help ligands bind properly. Each of these components in Fig. 1 has been highlighted by color: Position 87 is colored yellow, pocket C1 is red, pocket C2 is blue, pocket C3 is purple, pocket C4 is green, and pocket C5 is sky-blue.

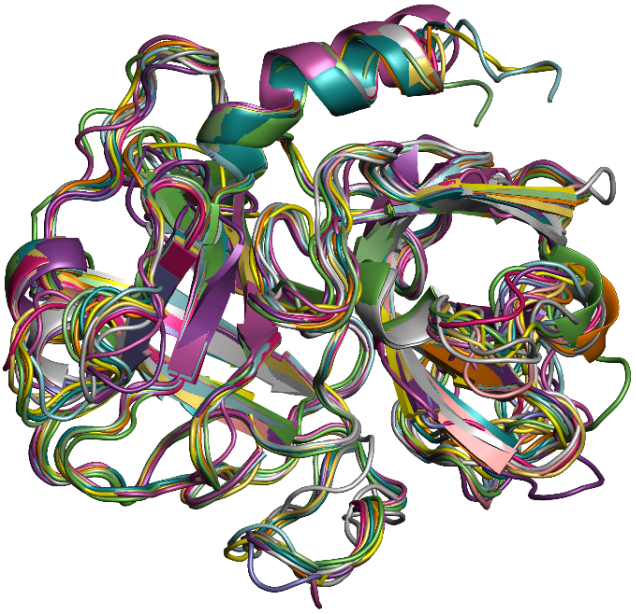


Fig 2. The alignment process involves superimposing the three-dimensional structures of the TMPRSS protein family models, taking into account structural similarities and differences. This allows us to visualize the overall architecture of the protein family and identify regions that may be important for function or regulation.

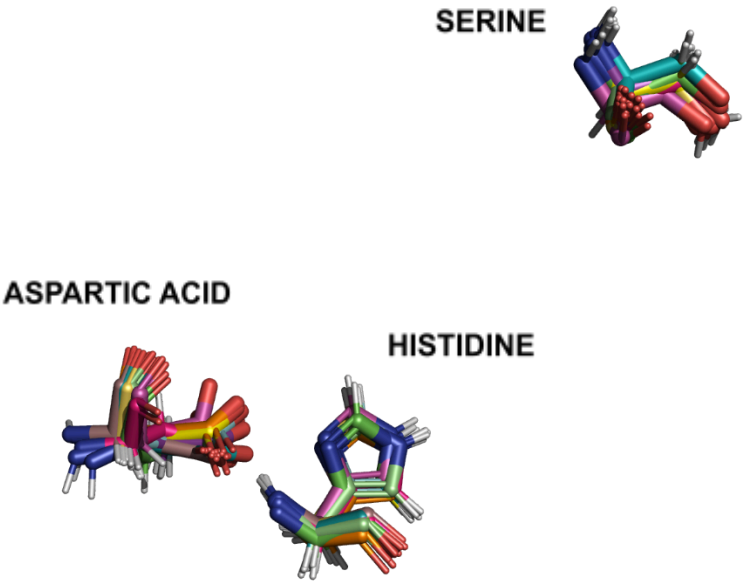
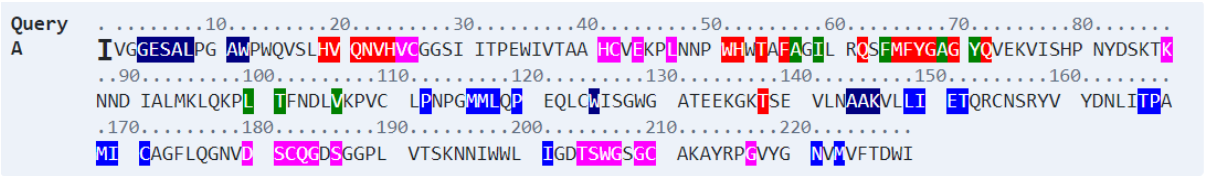


Fig 3. The structural alignment of catalytic triads is a crucial aspect of understanding the mechanisms of enzyme catalysis. Catalytic triads are composed of three amino acid residues, typically serine, histidine, and aspartate, that work together to facilitate enzymatic reactions.

Drug Binding Site:

Fig 4 shows how carboxylesterases hydrolyze camostat to create the active metabolite 4-(4-guanidinobenzoyloxy) phenylacetate (GBPA), which is consistent with observations in the literature. The negatively charged carboxylate group in the active metabolite GBPA is stabilized by forming a salt bridge with the lysine residue at position 87 of TMPRSS2 and is also attached by the guanidino group (Escalante et al., 2021). The drug is anchored by the salt bridge, which places the Camostat scissile bond in the best possible location for the catalytic serine to attack. The enzyme is acylated as a result of the catalytic serine's breakdown of the scissile link, which inhibits TMPRSS2. Since that K87 (K342(Escalante et al., 2021)) has been demonstrated to contribute to the stabilization of the camostat active metabolite(Escalante et al., 2021). It is essential to look for the amino acid at this position in the other TMPRSS proteins because doing so can reveal similar stabilizing mechanisms. According to Fig 1, the only other protein that conserves lysine is TMPRSS4. Histidine and arginine are used by TMPRSS5, TMPRSS6, and TMPRSS11A to maintain positively charged amino acids. On the other hand, aspartic acid allows TMPRSS13 to carry a negative charge. Asparagine, an uncharged polar amino acid, is present in TMPRSS11F. The aromatic hydrophobic side chains of TMPRSS7 and TMPRSS12 are phenylalanine and tyrosine, respectively. Chain hydrophobic amino acids are present in all the other TMPRSS proteins.



Sequence of the hmTMPRSS2 protein, that highlights the residues in each detected cavity. Pocket C1 – Red, Pocket C2 – Blue, Pocket C3 – Purple, Pocket C4 – Green, Pocket C5 - Dark Blue.

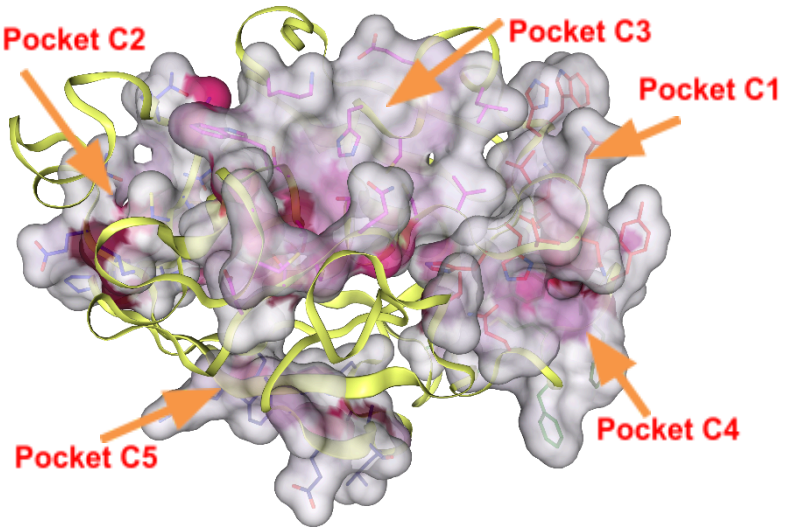


Fig 4: TMPRSS2 Binding Sites (Pockets) molecular docking based on AutoDock Vina.

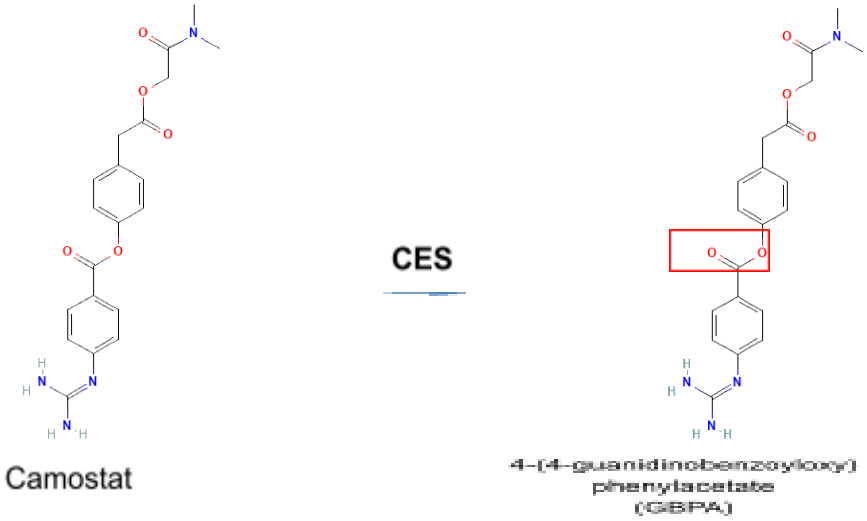


Fig: 5 The process of carboxylesterases hydrolyzing the drug camostat to create the active metabolite GBPA. The catalytic serine of TMPRSS2 is shown attacking the scissile ester bond in red rectangle.

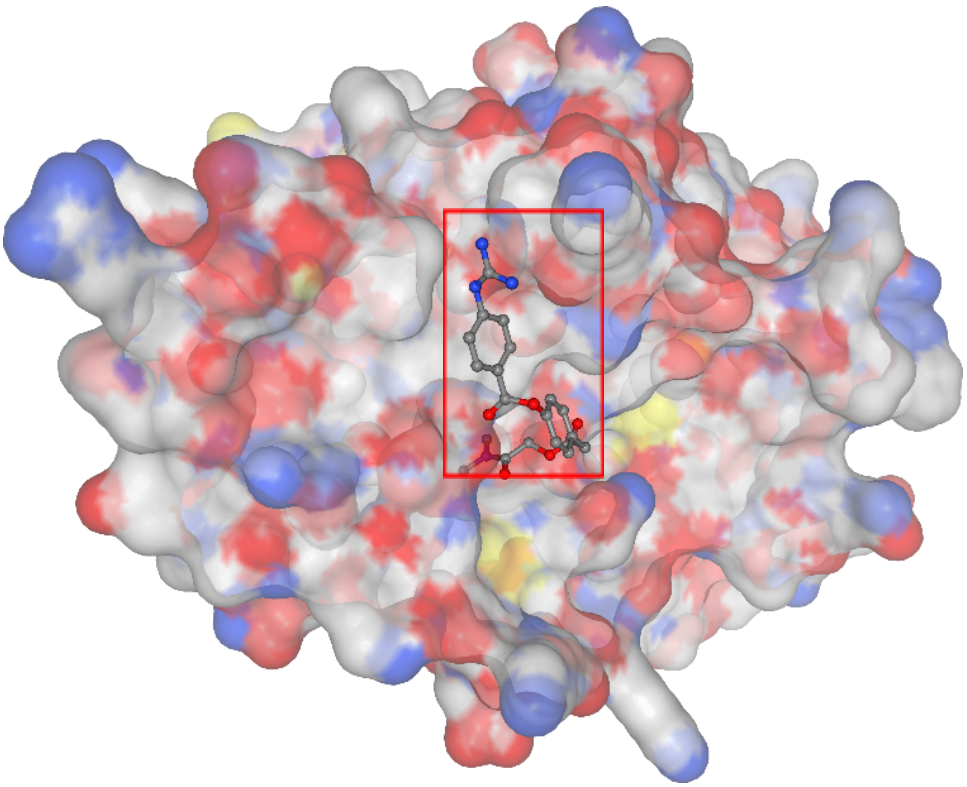


Fig 6: The red rectangle indicates the position of the active site catalytic triad in the Ligand and Receptor Docking.

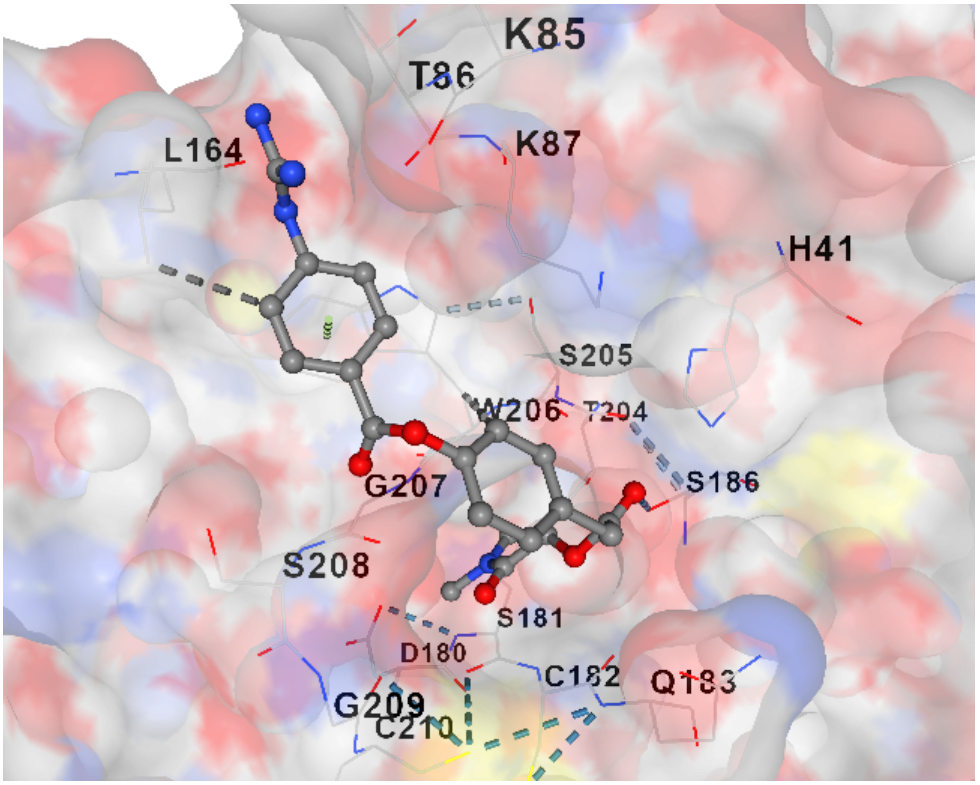


Fig 6: Showing interactions between Camostat and receptor molecule.

Type	Color	Style	Description
Hydrogen Bond			Hydrogen-bond between strong donor and acceptor atoms.
Weak Hydrogen Bond			Hydrogen-bond between a carbon donor atom and an acceptor, or a Pi group and a donor atom.
Hydrophobic Interaction			Interactions between alkyl groups, or a alkyl group and a Pi group.
Halogen Bond			Interactions with fluorine, chlorine, bromine or iodine atoms.
Ionic Interaction			Interactions between pairs of oppositely charged groups.
Cation-Pi Interaction			Interactions between a positively charged atom and the electrons of a delocalized Pi system.
Pi-Pi Stacking			Interactions between delocalized Pi systems.

Source: https://cadd.labshare.cn/cb-dock2/php/manual.php

Table B Comparison of Potential Pockets showing the Vina Score and Contact Residues of TMPRSS2(Receptor) and Camostat (Ligand). Selected pockets for each TMPRSS protein is provided in table 3 (Supplementary Information).

Pocket ID	Vina score	Cavity volume ($Å^3$)	Center (x, y, z)	Docking size (x, y, z)	Contact residues
C1	-6.8	248	-9, -4, 15	28, 28, 28	Chain A: HIS41 LYS85 THR86 LYS87 LEU164 ASP180 SER181 CYS182 GLN183 SER186 THR204 SER205 TRP206 GLY207 SER208 GLY209 CYS210
C2	-6.7	310	-2, 13, 14	28, 28, 28	Chain A: HIS19 ASN22 VAL23 HIS24 VAL25 GLN62 SER63 PHE64 MET65 TRP129 GLY130 ALA131 GLY136 LYS137 THR138 GLN183 GLY184
C3	-6.5	227	10, 12, 16	28, 28, 28	Chain A: PRO9 ALA57 GLY58 ILE59 LEU60 ARG61 PHE64 PHE66 ALA69 TYR71 LEU100 THR101 ASN103 VAL106
C4	-6.3	267	2, -19, 18	28, 28, 28	Chain A: MET116 MET117 LEU118 GLN119 PRO120 LEU149 ILE150 GLU151 THR152 MET169 ILE170 CYS171 ILE201 ASN221 MET223 VAL224
C5	-6.1	140	4, -3, 32	28, 28, 28	Chain A: GLU5 SER6 ALA7 LEU8 GLY10 ALA11 TRP12 PRO13 TRP14 LYS107 TRP125 ALA144 TRP198

C1-C5 Pocket Analysis

Pocket C1	There are two crucial locations in Pocket C1 that differ between different TMPRSS proteins. TMPRSS2 has an uncharged, polar amino acid residue at position 181, which is conserved in TMPRSS3, TMPRSS4, TMPRSS9, TMPRSS12, and TMPRSS13. Serine or threonine are the two specific amino acids that are utilized. There is an alanine in this position in other TMPRSS proteins. In Pocket C1, position 204 is a crucial site that contains an uncharged, polar amino acid residue in TMPRSS2, TMPRSS3, TMPRSS12, and TMPRSS13, whereas other TMPRSS proteins have a hydrophobic side chain amino acid (e.g., valine) at that position.
Pocket C2	Several potential discriminant residues for ligand recognition may be seen in Pocket C2, including variations at position 25. At position 25 in the genes TMPRSS2, TMPRSS3, TMPRSS4, TMPRSS6, TMPRSS12, and TMPRSS13, either valine or isoleucine is present. Among the TMPRSSs, TMPRSS7, TMPRSS9, and TMPRSS11B have aromatic hydrophobic side chains at this position. Conversely, TMPRSS5, TMPRSS11A, and TMPRSS11F have uncharged, polar amino acids such as threonine and glutamine at this site, while histidine and arginine are present in TMPRSS11D and TMPRSS11E, respectively.
Pocket C3	The family as a whole has substantial variation in Pocket C3. Both TMPRSS2 and TMPRSS3 have a hydrophobic amino acid (leucine) at position 47. Amino acids at this location are different in the two TMPRSS9 isoforms for this pocket. TMPRSS5, an isoform of TMPRSS9, TMPRSS11A, and TMPRSS11E, have aromatic hydrophobic amino acids such as phenylalanine and tyrosine. Histidine and lysine, which are positively charged amino acids, are present in TMPRSS4 and TMPRSS11B, respectively. The amino acid aspartate, which has a negative charge, is found in TMPRSS6. The uncharged, polar amino acids asparagine and threonine are present in TMPRSS7, the second isoform of TMPRSS9, TMPRSS11D, TMPRSS11F, and TMPRSS13. Finally, there is no amino acid in this corresponding position for TMPRSS12.
Pocket C4	Pocket C4 is significant due to the likely interaction between GBPA and lysine residue at position 87. The carboxylate of the active metabolite and the amide of the prodrug interact in this region. Only two proteins in this family, TMPRSS2 and TMPRSS4, have the lysine residue conserved in this pocket.
Pocket C5	While the interactions in pocket C5 do not seem to be directly involved in ligand binding based on their positions in relation to the binding pocket, changes across the amino acids are evident in this domain.

Discussion:

According to alignments and structural analysis, the TMPRSS family of proteins has considerable variation that can be used to create selective inhibitors. The variations in the main sequence of pockets C1, C2, C3, and C4 can significantly impact ligand binding and selectivity. Of particular interest is the position 87 in pocket C4, which plays a crucial role in

ligand binding. Lysine 87 is a key anchor site for the GBPA carboxylate group, the active metabolite of camostat. It is noteworthy that only TMPRSS2 and TMPRSS4 share this conserved lysine at position 87. This is significant because both TMPRSS2 and TMPRSS4 are known to cleave the Spike protein, which is essential for cell entry. The findings suggest that the differences in position 87 may influence the binding of ligands and selectivity of pockets, which could have implications for developing therapeutics targeting TMPRSS2 and TMPRSS4. Further research is needed to better understand the role of these variations in ligand binding and selectivity. Hence, focusing on this location might offer a strategy to limit camostat binding off target and selectively inhibit these two enzymes. Through its basic sequence, Pocket C1 also features a region of intense negative charges. This makes it possible for ligands with positively charged amino acids to engage strongly with TMPRSS proteins. The presence of large, bulky positively charged regions in current ligands, such as nafamostat, can lead to the formation of strong ionic connections with pocket C1. This interaction is a result of the specific core sequence of Pocket C1, which includes a densely negative charged region. The strength of these ionic connections can have significant implications for the binding and efficacy of the ligand. As such, understanding the interaction between the charge distribution of the ligand and the core sequence of Pocket C1 is crucial for the development of effective therapeutic agents. Further research in this area may lead to the identification of novel ligands with improved binding properties and therapeutic potential.

Conclusion:

The TMPRSS subfamily of the TTSP family has been linked to the development of cancer and infectious diseases like COVID-19. As a result, there has been a lot of interest in repurposing medications like camostat and nafamostat for COVID-19 pharmacological therapy due to the relationship between Spike protein processing, illness progression, and TMPRSS2 activity. Through presented alignments and analyses, noticeable changes in the amino acid residues inside the ligand binding pockets of the TMPRSS family of enzymes have been identified. The data provided is intended to aid in the identification of distinctive TMPRSS2 epitopes that can be utilized in the development of more selective medicines. This is important because drug selectivity and off target binding can have a significant impact on severe drug side effects. One of the more noteworthy findings of the study is the discovery of lys87 as a potential selectivity component designed to suppress TMPRSS2 and TMPRSS4. These are the only two members of the family enzymes that have been demonstrated to both break down the SARS-CoV Spike protein and conserve this lysine. If it were possible to selectively block TMPRSS2,4 and avoid off target binding to other TMPRSS proteins, better therapeutic treatments that prevent protease-mediated cleavage of the Spike protein could be developed. The identification of distinctive TMPRSS2 epitopes and the discovery of lys87 as a potential selectivity component are important steps towards the development of more selective medicines for COVID-19 pharmacological therapy. This will ultimately help reduce severe drug side effects and improve therapeutic treatments for infectious diseases like COVID-19.

References:

Antalis, T. M., Bugge, T. H., & Wu, Q. (2011). Membrane-anchored serine proteases in health and disease. *Progress in Molecular Biology and Translational Science*, *99*, 1–50. doi: 10.1016/B978-0-12-385504-6.00001-4

Bugge, T. H., Antalis, T. M., & Wu, Q. (2009). Type II transmembrane serine proteases. *The Journal of Biological Chemistry*, *284*(35), 23177–23181. doi: 10.1074/JBC.R109.021006

Cal, S., Quesada, V., Garabaya, C., & López-Otín, C. (2003). Polyserase-1, a human polyprotease with the ability to generate independent serine protease domains from a single translation product. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(16), 9185–9190. doi: 10.1073/PNAS.1633392100

Chen, Y.-W., Lee, M.-S., Lucht, A., Chou, F.-P., Huang, W., Havighurst, T. C., Kim, K., Wang, J.-K., Antalis, T. M., Johnson, M. D., & Lin, C.-Y. (2010). TMPRSS2, a serine protease expressed in the prostate on the apical surface of luminal epithelial cells and released into semen in prostasomes, is misregulated in prostate cancer cells. *The American Journal of Pathology*, *176*(6), 2986–2996. doi: 10.2353/ajpath.2010.090665

Escalante, D. E., Wang, A., & Ferguson, D. M. (2021). Structural Modeling of the TMPRSS Subfamily of Host Cell Proteases Reveals Potential Binding Sites. *BioRxiv*, 2021.06.15.448583. doi: 10.1101/2021.06.15.448583

Fung, T. S., & Liu, D. X. (2019). Human Coronavirus: Host-Pathogen Interaction. *Annual Review of Microbiology*, *73*, 529–557. doi: 10.1146/ANNUREV-MICRO-020518-115759

Hatesuer, B., Bertram, S., Mehnert, N., Bahgat, M. M., Nelson, P. S., Pöhlman, S., & Schughart, K. (2013). Tmprss2 is essential for influenza H1N1 virus pathogenesis in mice. *PLoS Pathogens*, *9*(12), 1–10. doi: 10.1371/JOURNAL.PPAT.1003774

Iwata-Yoshikawa, N., Okamura, T., Shimizu, Y., Hasegawa, H., Takeda, M., & Nagata, N. (2019). TMPRSS2 Contributes to Virus Spread and Immunopathology in the Airways of Murine Models after Coronavirus Infection. *Journal of Virology*, *93*(6). doi: 10.1128/JVI.01815-18

Liu, S., Xiao, G., Chen, Y., He, Y., Niu, J., Escalante, C. R., Xiong, H., Farmar, J., Debnath, A. K., Tien, P., & Jiang, S. (2004). Interaction between heptad repeat 1 and 2 regions in spike protein of SARS-associated coronavirus: implications for virus fusogenic mechanism and identification of fusion inhibitors. *Lancet (London, England)*, *363*(9413), 938–947. doi: 10.1016/S0140-6736(04)15788-7

Liu, Y., Grimm, M., Dai, W. tao, Hou, M. chun, Xiao, Z. X., & Cao, Y. (2019). CB-Dock: a web server for cavity detection-guided protein–ligand blind docking. *Acta Pharmacologica Sinica 2019 41:1*, *41*(1), 138–144. doi: 10.1038/s41401-019-0228-6

Liu, Y., Yang, X., Gan, J., Chen, S., Xiao, Z. X., & Cao, Y. (2022). CB-Dock2: improved protein–ligand blind docking by integrating cavity detection, docking and homologous template fitting. *Nucleic Acids Research*, *50*(W1), W159–W164. doi: 10.1093/NAR/GKAC394

Shen, L. W., Mao, H. J., Wu, Y. L., Tanaka, Y., & Zhang, W. (2017). TMPRSS2: A potential target for treatment of influenza virus and coronavirus infections. *Biochimie*, *142*, 1–10. doi: 10.1016/J.BIOCHI.2017.07.016

Vaarala, M. H., Porvari, K. S., Kellokumpu, S., Kyllönen, A. P., & Vihko, P. T. (2001). Expression of transmembrane serine protease TMPRSS2 in mouse and human tissues. *The Journal of Pathology*, *193*(1), 134–140. doi: 10.1002/1096-9896(2000)9999:9999<::AID-PATH743>3.0.CO;2-T

Wallrapp, C., Hähnel, S., Müller-Pillasch, F., Burghardt, B., Iwamura, T., Ruthenbürger, M., Lerch, M. M., Adler, G., & Gress, T. M. (2000). A novel transmembrane serine protease (TMPRSS3) overexpressed in pancreatic cancer. *Cancer Research*, *60*(10), 2602–2606.

Webb, S. L., Sanders, A. J., Mason, M. D., & Jiang, W. G. (2011a). Type II transmembrane serine protease (TTSP) deregulation in cancer. *Frontiers in Bioscience (Landmark Edition)*, *16*(2), 539–552. doi: 10.2741/3704

Webb, S. L., Sanders, A. J., Mason, M. D., & Jiang, W. G. (2011b). Type II transmembrane serine protease (TTSP) deregulation in cancer. *Frontiers in Bioscience (Landmark Edition)*, *16*(2), 539–552. doi: 10.2741/3704

Zang, R., Gomez Castro, M. F., McCune, B. T., Zeng, Q., Rothlauf, P. W., Sonnek, N. M., Liu, Z., Brulois, K. F., Wang, X., Greenberg, H. B., Diamond, M. S., Ciorba, M. A., Whelan, S. P. J., & Ding, S. (2020). TMPRSS2 and TMPRSS4 promote SARS-CoV-2 infection of human small intestinal enterocytes. *Science Immunology*, *5*(47). doi: 10.1126/sciimmunol.abc3582

Ziegler, C. G. K., Allon, S. J., Nyquist, S. K., Mbano, I. M., Miao, V. N., Tzouanas, C. N., Cao, Y., Yousif, A. S., Bals, J., Hauser, B. M., Feldman, J., Muus, C., Wadsworth, M. H., Kazer, S. W., Hughes, T. K., Doran, B., Gatter, G. J., Vukovic, M., Taliaferro, F., … HCA Lung Biological Network. (2020). SARS-CoV-2 Receptor ACE2 Is an Interferon-Stimulated Gene in Human Airway Epithelial Cells and Is Detected in Specific Cell Subsets across Tissues. *Cell*, *181*(5), 1016-1035.e19. doi: 10.1016/j.cell.2020.04.035