



Computational Identification of a Putative Allosteric Binding Pocket in TMPRSS2

Jacopo Sgrignani^{1*} and Andrea Cavalli^{1,2*}

¹ Institute for Research in Biomedicine, Università della Svizzera Italiana, Bellinzona, Switzerland, ² Swiss Institute of Bioinformatics, Lausanne, Switzerland

OPEN ACCESS

Edited by:

Massimiliano Bonomi,
Institut Pasteur, France

Reviewed by:

Therese E. Mallavin,
Institut Pasteur, France
Matteo Masetti,
University of Bologna, Italy

*Correspondence:

Jacopo Sgrignani
jacopo.sgrignani@irb.usi.ch
Andrea Cavalli
andrea.cavalli@irb.usi.ch

Specialty section:

This article was submitted to
Biological Modelling and Simulation,
a section of the journal
Frontiers in Molecular Biosciences

Received: 10 February 2021

Accepted: 01 April 2021

Published: 30 April 2021

Citation:

Sgrignani J and Cavalli A (2021)
Computational Identification of a
Putative Allosteric Binding Pocket
in TMPRSS2.
Front. Mol. Biosci. 8:666626.
doi: 10.3389/fmolb.2021.666626

Camostat, nafamostat, and bromhexine are inhibitors of the transmembrane serine protease TMPRSS2. The inhibition of TMPRSS2 has been shown to prevent the viral infection of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and other viruses. However, while camostat and nafamostat inhibit TMPRSS2 by forming a covalent adduct, the mode of action of bromhexine remains unclear. TMPRSS2 is autocatalytically activated from its inactive form, zymogen, through a proteolytic cleavage that promotes the binding of Ile256 to a putative allosteric pocket (A-pocket). Computer simulations, reported here, indicate that Ile256 binding induces a conformational change in the catalytic site, thus providing the atomistic rationale to the activation process of the enzyme. Furthermore, computational docking and molecular dynamics simulations indicate that bromhexine competes with the N-terminal Ile256 for the same binding site, making it a potential allosteric inhibitor. Taken together, these findings provide the atomistic basis for the development of more selective and potent TMPRSS2 inhibitors.

Keywords: TMPRSS2 protein, molecular modeling, allosteric pocket, docking, MD simulation

INTRODUCTION

Since the early days of the pandemic coronavirus disease 2019 (COVID-19) started from the Chinese city of Wuhan, Hubei province, in December 2019, many reports highlighted the crucial role of transmembrane serine protease 2 (TMPRSS2) in the spread and progression of the viral infection (Hoffmann et al., 2020; Sungnak et al., 2020). TMPRSS2 has been identified as one of the proteases responsible for the proteolytic priming of SARS-CoV-2 spike protein which leads to the release of the fusion peptide. In addition to that, TMPRSS2 has been put in relation with the spread of other viruses, such as influenza A viruses, severe acute respiratory syndrome coronavirus 2 (SARS-CoV), and Middle East respiratory syndrome coronavirus (MERS-CoV), and it has been studied as a potential therapeutic target for prostate cancer therapy (Lucas et al., 2014; Shen et al., 2017). Finally, as TMPRSS2 expression is regulated by the androgen receptor, it has been hypothesized that its crucial role in the viral infection might help explain why males have more frequently severe complications and a worse clinical outcome than females and if androgen deprivation therapy (ADT) can have a protective effect against SARS-CoV-2 infection (Montopoli et al., 2020). These observations stimulated intense investigations, and the number of papers with the TMPRSS2 keyword in the title indexed in PubMed during 2020 raised from an average of 80–100/year to 601.

TMRSS2 is a membrane protein belonging to the type II transmembrane serine protease (TTSP) family. It is functionally classified as a trypsin-like protease (TLP). Like other serine proteases, TMPRSS2 cleaves peptide bonds that are present after positively charged residues (lysine or arginine), and its enzymatic activity depends on the presence of a catalytic triad formed by His296, Asp345, and Ser441. The catalytic selectivity is achieved with the presence of a negatively charged Asp residue at the bottom of a cavity usually indicated as “S1 specificity pocket” (Laporte and Naesens, 2017; Singh et al., 2020).

Structurally, TMPRSS2 is characterized by the presence of a cytoplasmic N-terminal domain, a transmembrane helical domain, and three extracellular domains: low-density lipoprotein (LDL)-receptor class A domain, scavenger receptor cysteine-rich (SRCR) domain, and the peptidase S1 domain, also called serine protease domain (SPD) (Figure 1A).

An autocatalytic cleavage between Arg255 and Ile256 activates the 492-residue long TMPRSS2 zymogen. This modification enables the binding of Ile256 into a putative allosteric pocket (A-pocket), which induces a conformational rearrangement of the catalytic site (Bertram et al., 2010). After the cleavage, membrane TLPs, such as TMPRSS2, remain bound to the transmembrane N-terminal domains by a conserved disulfide bond, although a small fraction of the protein can be detected into the extracellular milieu (Szabo et al., 2003; Pászti-Gere et al., 2016; Shen et al., 2017).

Two different species are reported in the literature, one with a mass of ~55 kDa that corresponds to the full-length protein and one of ~30 kDa which represents the SPD released in the extracellular space if the disulfide bond is not formed (Afar et al., 2001; Chen et al., 2010).

To date, no atomistic structure of the entire TMPRSS2, or the SPD, is available. However, important information can be derived from the structure of homologous proteins such as matriptase, DESC1, and several kallikreins.

Several inhibitors of TMPRSS2 have been identified in the last years. These include organic compounds such as camostat, nafamostat, and bromhexine (BH) (Figure 1B) and peptidomimetics (Meyer et al., 2013; Lucas et al., 2014; Shen et al., 2017; Bestle et al., 2020; Hoffmann et al., 2020; Zang et al., 2020). Of particular note is BH, a component of widely used medicaments against respiratory disorders characterized by viscid or excessive mucus. In fact, following the report of a selective TMPRSS2 inhibition by Lucas et al. (2014), the use of BH for the prevention and therapy of the SARS-CoV-2 infection has been hypothesized (Depfenhart et al., 2020; Habtemariam et al., 2020; Maggio and Corsini, 2020). However, to date, only a limited number of clinical trials have been carried out, and their results remain inconclusive (Ansarin et al., 2020; Li et al., 2020).

In this work, we used computational and experimental methods, such as homology modeling, molecular docking, molecular dynamics (MD), and microscale thermophoresis (MST), to investigate the structure and dynamics of TMPRSS2 and clarify its activation mechanism and the interaction with various inhibitors at an atomistic level of details. We focused, in particular, on the differences in the mode of action of camostat/nafamostat and BH. In fact, while camostat and

nafamostat inhibit TMPRSS2 by forming a covalent adduct, the mode of action of BH remains unclear.

Besides the generation of a reliable model of the TMPRSS2 catalytic domain, the results of our investigations confirmed that both camostat and nafamostat are competitive inhibitors efficiently binding the active site. Contrarily, they indicated that the binding of BH to the active site is unlikely, leading us to the identification of a putative allosteric binding pocket.

MATERIALS AND METHODS

Homology Modeling

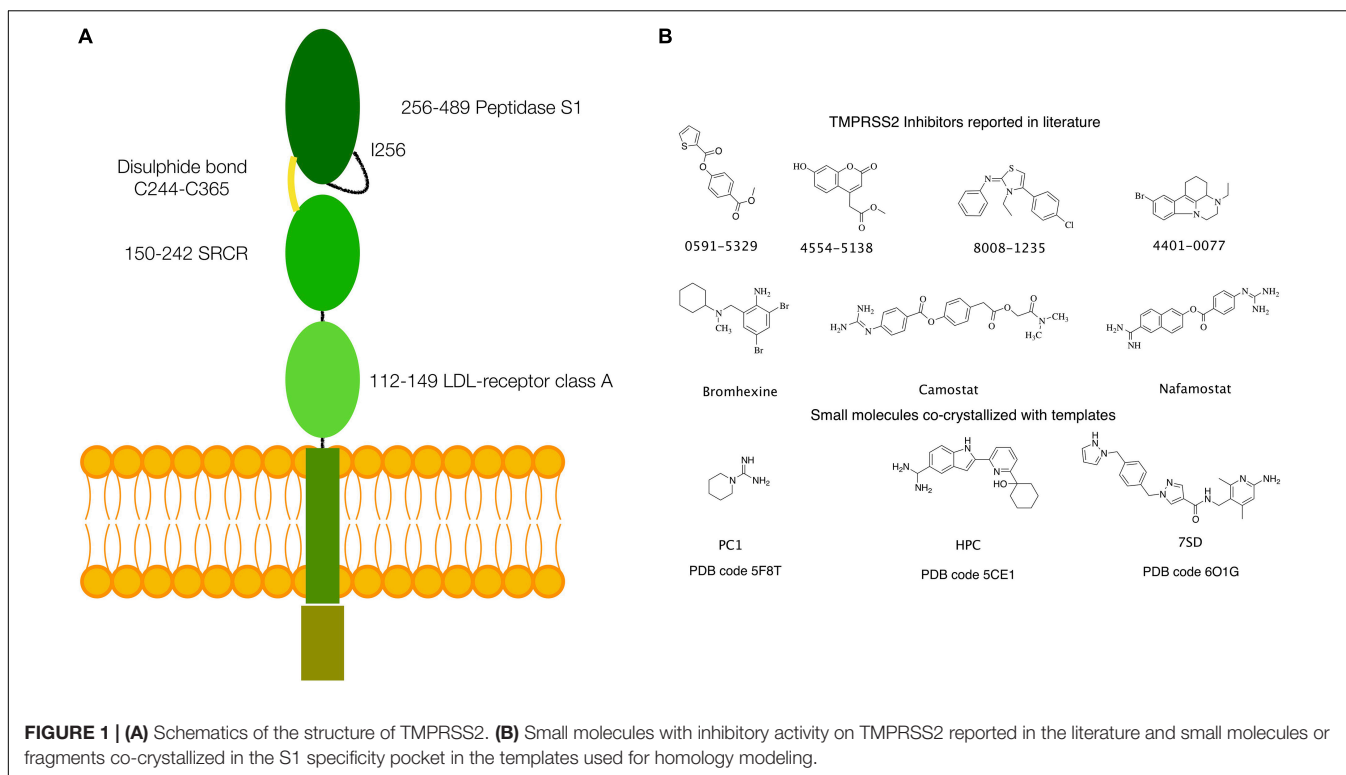
An atomistic model of the SPD of TMPRSS2 (UniProt code O15393), covering residues from Ile256 to Gly492, was generated by homology modeling. The most suitable templates were identified using the SWISS-MODEL webserver (Waterhouse et al., 2018). This search provided three templates [Protein Data Bank (PDB) codes: 5F8T, 5CE1, and 6O1G], having a sufficient degree of similarity (between 38 and 41%), thus well-suited for an accurate model generation (Xiang, 2006; Cavasotto and Phatak, 2009; Sgrignani et al., 2018). The alignment between the target sequence and the templates was performed using the Prime-STA algorithm, included in the Schrodinger suite for molecular modeling (Schrodinger Suite 2020-1). This algorithm, in addition to the sequence alignment, considers secondary structure matching, providing better alignments also in poorly conserved regions. Next, 10 models were generated for each of the three templates using PRIME, keeping small ligand molecules, such as piperidine-1-carboximidamide (PC1), 2-[6-(1-hydroxycyclohexyl)pyridin-2-yl]-1H-indole-5-carboximidamide (HCP), and N-[(6-amino-2,4-dimethylpyridin-3-yl)methyl]-1-({4-[(1H-pyrazol-1-yl)methyl]phenyl)methyl}-1H-pyrazole-4-carboxamide (7SD) (Figure 1), bound to the protein active site in the templates 5F8T, 5CE1, and 6O1G, to preserve their respective conformations.

Finally, the models (subsequently indicated as M-5F8T, M-5CE1, and M-6O1G) with the lowest OPLS3e (Harder et al., 2016) potential energy after minimization were selected for the subsequent calculations.

Molecular Dynamics Simulations

Atomistic models were prepared for MD simulation with the following protocol: (1) the PROPKA program was used to assign the residue protonation state at a reference pH of 7.4 (Olsson et al., 2011) and (2) the structures were solvated in a box of water with a minimal distance from the protein surface of 10 Å. A proper number of counterions were added to the systems to ensure charge neutrality. All the non-solvent molecules were parametrized using the OPLS3 (Harder et al., 2016) force field, while TIP3P model (Jorgensen et al., 1983) was used for water molecules.

Before the MD production runs, the following simulation protocol was used to equilibrate the systems: (1) Brownian dynamics was run for 100 ps in an NVT ensemble ($T = 10$ K) applying harmonic restraints on solute heavy atoms (force constant 50 kcal/mol/Å²); (2) NVT ($T = 10$ K) MD simulation



of 12 ps in NVT ensemble conserving the same restraints applied in (1); (3) NPT ($T = 300$ K and $P = 1$ atm) MD simulation (12 ps) conserving the same restraints applied in (1); and (4) NPT ($T = 300$ K and $P = 1$ atm) MD simulation (24 ps) without restraints. The pressure and the temperature were fixed at 300 K and 1 atm by the Martyna–Tobias–Klein barostat (Martyna et al., 1994) and the Nosé–Hoover chain thermostat (Martyna et al., 1992), respectively. All the simulations were run using GPU accelerated DESMOND code. A summary of the simulations run in this work is reported in **Table 1**.

Root mean square deviation (RMSD), root mean square fluctuation (RMSF), and radius of gyration (Rg) analysis were computed using Maestro (Schrodinger Suite 2020-1). Cluster analysis was performed with the program TtClust (Tubiana et al., 2018), focusing on residues belonging to the catalytic site, namely, Cys281, Thr293, Ala294, Ala295, His296, Cys297, Val298, Glu299, Tyr337, Asp338, Ser339, Lys342, Asn343, Asn344, Asp345, Ile346, Ala347, Met424, Cys437, Gln438, Asp440, Ser441, Asp458, Thr459, Ser460, Trp461, and Phe480. Contrarily, the analysis of the loop that regulates the access to the S1 specificity pocket was performed considering all the residues between Gly462 and Val473. The optimal number of clusters was automatically determined using the “elbow” method with k-means (Tibshirani et al., 2001).

Computational Docking of TMPRSS2 Ligands

Computational docking was performed using the software GLIDE (Friesner et al., 2004). The analysis of the structural parameters and the analysis of MD simulations (see section

“Results and Discussion”) indicated M-5FT8 as having a higher quality and more stable among the generated models.

In analogy to the previous studies (Amaro et al., 2008, 2018; Sgrignani et al., 2009), to account for target flexibility, snapshots from MD simulations of M-5FT8 were selected using the previously described cluster analysis. In particular, four snapshots were selected from the simulations run with positively charged His296 and four from the simulations with His296 protonated on the ϵ nitrogen (see also section “Results and Discussion”).

The grids for docking were centered in the geometric center of all the atoms of the three residues forming the catalytic triads (His296, Asp345, and Ser441). A distinct grid file was generated for all selected snapshots.

Contrarily, for the docking of BH in the putative site predicted by Sitemap, the grid was centered using the corresponding sitepoints. In this context, sitepoints are points in a grid, contiguous, or bridged by short gaps in exposed regions, that define the shape of a putative binding site (Halgren, 2007).

All docking calculations were performed using the standard precision (SP) protocol and GlideScore. Furthermore, docking was performed on all selected snapshots, and, finally, the pose with the best GlideScore, together with the receptor, was saved for the analysis and MD simulations.

The structures of the small molecule ligands were prepared with LIGPREP. In the case of BH, the results indicated a protonation of the ternary amino group; therefore, both enantiomeric molecules (S and R) were considered in docking calculations, but only the complex with best GlideScore was used in MD simulations.

TABLE 1 | Summary of the performed molecular dynamics.

Description of the system	Number of independent simulations	Simulation length (ns)	Ligand
M-5F8T	3 + 3 (His296 protonated on the ϵ nitrogen)	250	
M-5CE1	3	250	
M-6O1G	3	250	
M-5F8T	1 + 1 (His296 protonated on the ϵ nitrogen) for both camostat and nafamostat	500	Camostat and nafamostat
M-5F8T	3	2 × 1,000 1 × 500 (the complex decomposed)	BH in Site_1
M-5F8T	1	100	BH in Site_2
Apo-C-M-5F8T	1	1,000	
Apo-M-5F8T	1	1,000	
C-M-5F8T	3	500	(R)-BH in the A-pocket
C-M-5F8T	3	500	(S)-BH in the A-pocket
C-M-5F8T	3	500	(R)-BH in the A-pocket (IFD docking)
C-M-5F8T	3	500	(S)-BH in the A-pocket (IFD docking)

The models from different templates were indicated as M-5F8T, M-5CE1, and M-6O1G. The suffix C indicates the models where the first two residues (Ile256 and Val257) were deleted.

Docking of BH in the A-pocket (see section “Results and Discussion” for a definition) was performed using a representative structure of the open and closed conformations (Figures 2B–E) sampled during the MD simulations of C-M-5F8T (see section “Results and Discussion”). In this case, the grid was centered in the COG of the residues Ile381, Ser382, Gly383, Gly385, Ala386, Thr387, Glu388, Asn398, Ala399, Ala400, Asn433, Val434, Asp435, Ser436, Cys437, Asp440, Cys465, and Ala466.

The results of these calculations showed a better GlideScore for the complex in the closed conformation (~ -3.0 kcal/mol vs. ~ -4.8 kcal/mol for open and closed conformations). However, also in this case, the complex with the best GlideScore dissociated during MD simulations.

Regarding this point, it is important to notice that in M-5F8T, the S-pocket is occupied by a small aminoacidic tail. It is, therefore, reasonable to assume that a side chain rearrangement is needed to accommodate different ligands.

Consequently, the docking was performed again using the induced-fit docking (IFD) protocol of GLIDE, with default input parameters. In particular, only the orientations of the side chains of the residues within a distance of 5 Å from the ligand were optimized. Finally, the complex with the lowest IFD score [a specific score that combines GlideScore, Glide_Ecoul energy, and Prime protein conformation energy (Sherman et al., 2006)] was selected as the best model and used in MD simulations.

Prediction of Putative Allosteric Binding Sites

Several algorithms to detect allosteric pockets in proteins have been developed in the last years (Halgren, 2007, 2009; Yu et al., 2010; Panjkovich and Daura, 2014; Kozakov et al., 2015; Jiménez et al., 2017; Xu et al., 2018; Guarnera and Berezovsky, 2019). Sitemap, proposed by Halgren in 2007 (Halgren, 2007, 2009) and implemented in the Schrodinger suite for molecular modeling, is among the most widely used. Furthermore, it provides a clear assessment of the druggability of the identified pockets.

Consequently, we used this algorithm to investigate the presence of allosteric pockets in both M-5F8T and C-M-5F8T. All the calculations were performed using default values provided by the Maestro interface (Schrodinger Suite 2020-1). In addition to that, to validate these results, the same structures were analyzed also with other algorithms (PARS, Deepsite, and FTMap), using the respective webservers^{1,2,3}.

MST Experiments for the TMPRSS2/BH Binding

The binding affinity between TMPRSS2 and BH was measured by MST. Recombinant human TMPRSS2 (106–492aa, 6xHisTag) was acquired from Cusabio (CSB-YP023924HU) and labeled using a His-tag-specific dye (Monolith His-Tag Labeling Kit RED-tris-NTA (MO-L018), NanoTemper® Technologies GmbH, München, Germany), according to manufacturer instructions. A fixed concentration of the labeled TMPRSS2 (5 nM) was mixed with 16 1:1 serial dilution of BH. MST measurements were performed using premium-coated capillary tubes on a NanoTemper instrument.

BH was first dissolved in DMSO at a 5 mM concentration. In all subsequent experiments, both protein and BH were dissolved in Dulbecco’s Phosphate-Buffered Saline (PBS; D8537, Sigma Aldrich, Saint Louis, MO, United States).

Two independent experiments were performed to compute the Kd values. Data were analyzed with the NanoTemper analysis software MO.Affinity Analysis (v. 2.3). Kd values were obtained fitting compound concentration-dependent changes in normalized fluorescence (Fnorm).

RESULTS AND DISCUSSION

Homology Modeling of the Serine Protease Domain of TMPRSS2

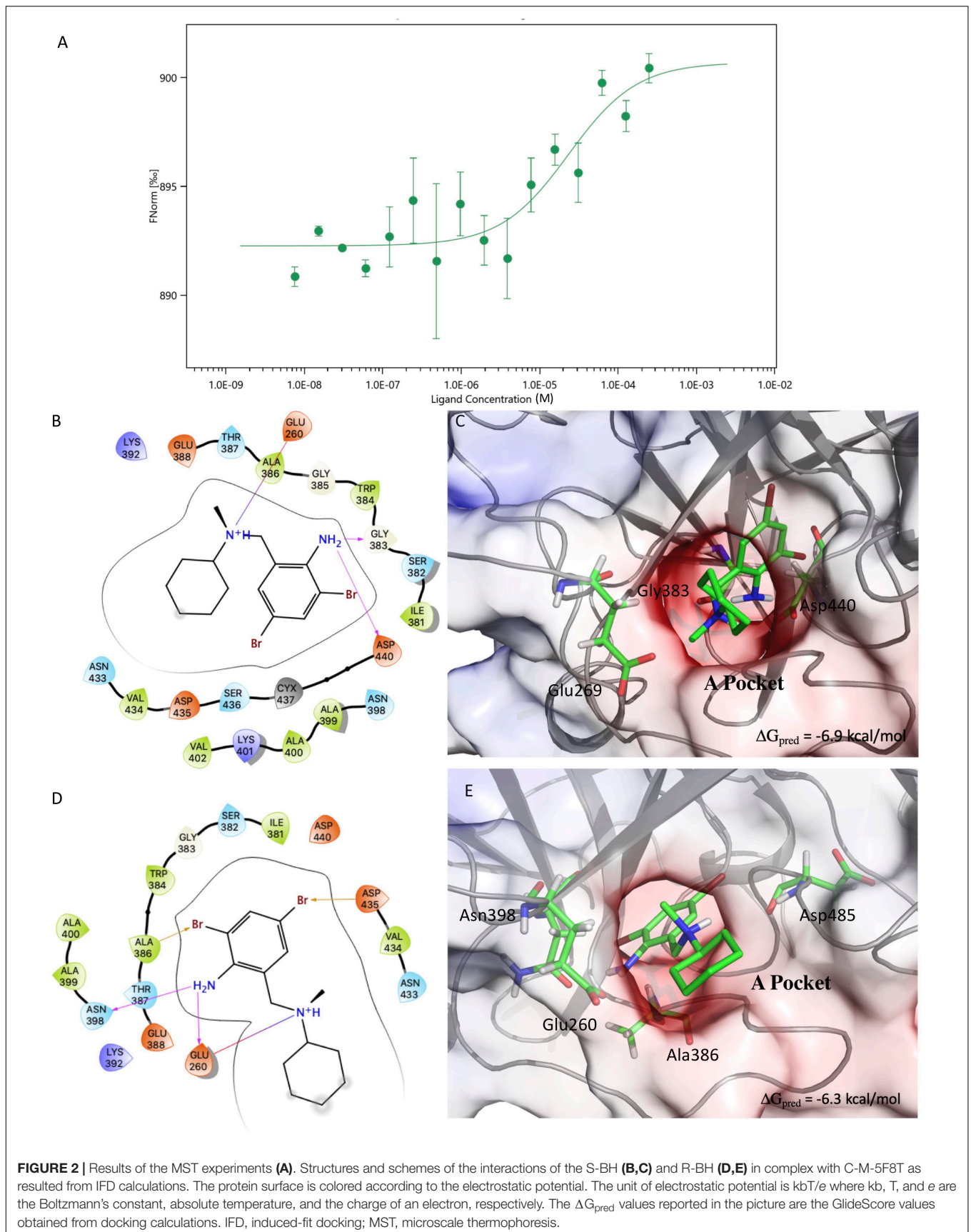
Considering its relevance for both the drug design and the enzymatic function, we focused our attention on the TMPRSS2 SPD (Ile256 to Gly492).

A search performed with the SWISS MODEL webserver identified three very similar structures (Figure 3 and Table 2) as suitable templates to generate TMPRSS2 models: (1) two structures of the human plasma kallikrein (PK), a serine

¹<http://bioinf.uab.cat/cgi-bin/pars-cgi/pars.pl>

²<https://playmolecule.org/deepsite>

³<http://ftmap.bu.edu/login.php>



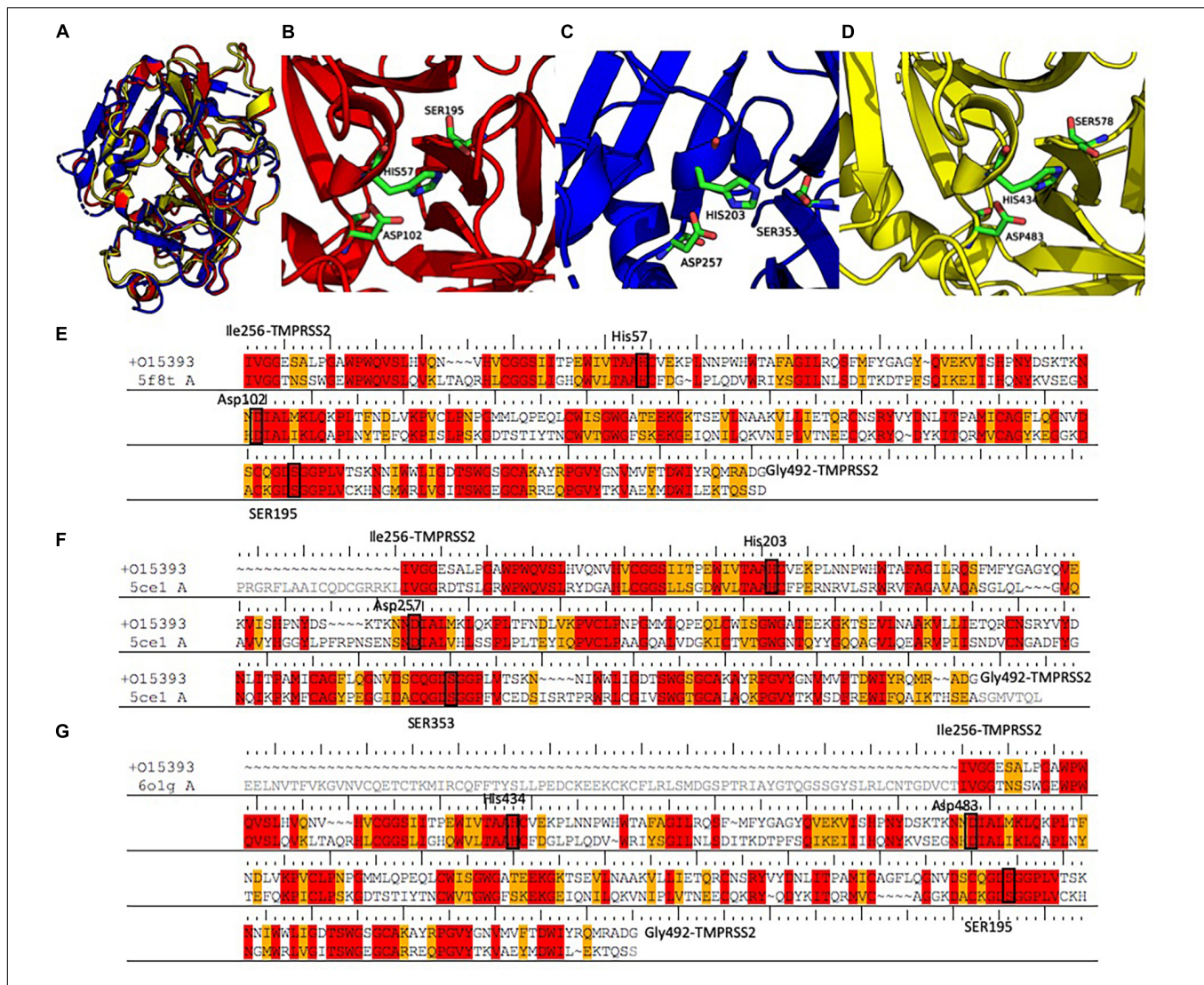


FIGURE 3 | (A) Structural alignment between the three selected templates 5F8T (red), 5CE1 (blue), and 6O1G (yellow). Details of the catalytic sites of the PK structures deposited with the PDB code 5F8T **(B)** and 6O1G **(C)** and of the HP structure deposited with the code 5CE1 **(D)**. **(E–G)** Sequence alignments between the three templates and the SPD of TMPRSS2. Identical residues are colored in red; conserved residues (according to the BLOSUM62 scoring matrix) are colored in orange. Abbreviations: PDB, Protein Data Bank; SPD, serine protease domain.

protease that cleaves high-molecular-weight kininogen (HMWK) to generate bradykinin (BK) (Schmaier, 2013) [PDB codes: 5F8T and 6O1G (Partridge et al., 2019), resolution 1.75 and 2.50 Å] and (2) the structure of hepsin (HP), a membrane-bound serine protease able to catalyze protein cleavage after basic amino-acid residues (PDB code: 5CE1, resolution 2.50 Å). In fact, the pairwise RMSD computed using the C α atoms and the program ALMOST (Fu et al., 2014) is smaller than 0.5 Å.

The sequences of the three selected templates were aligned to TMPRSS2 using the PRIME-STA procedure (Figures 3E–G), and 10 models were generated starting from each template. Finally, the model with the lowest potential energy was selected from the three different groups. As expected, all the three models were very similar, with a pairwise C α – RMSD below 0.5 Å. Furthermore, visual inspection of the three structures confirmed

the similarity between all models, with the exception of the region between Tyr322 and Ser333. In fact, while in the two models derived from PK structures (M-5F8T and M-6O1G), this region

TABLE 2 | Summary of the sequence–sequence alignment between the sequence of the serine protease domain of TMPRSS2 and the three selected templates.

Template PDB code	Score	Identities (%)	Positives (%)	Gaps (%)
5F8T	1,218	41	59	2
5CE1	1,152	39	55	5
6O1G	1,185	41	57	4

The score is the BLAST bit score. Identities is the percentage of residues that are identical between the sequences. Positives is the percentage of residues that are positive matches according to the similarity matrix (BLOSUM62). Gaps is the percentage of gaps in both the query and homolog as returned by BLAST.

is a β -sheet, in the model form HP structure (M-5CE1), it is modeled as a long loop. This is not surprising because in the sequence–sequence alignment between HP and TMPRSS2 used for model generation, this region is characterized by the insertion of three amino acids.

The quality of the models was evaluated with the Protein Structure Quality viewer implemented in Maestro, computing structural parameters widely used in the evaluation of homology models (Sgrignani et al., 2009) and by the PROSA-Web server (Wiederstein and Sippl, 2007; **Table 3**). This analysis did not show any critical points for all generated models. Nevertheless, the number of violations of the allowed regions in the Ramachandran plot, and other violations from the ideal structural parameters were higher for the models generated using 5CE1 and 6O1G.

Molecular Dynamics Simulations of the TMPRSS2 Models

Aimed to (1) understand the overall stability of the generated models, (2) to detect problematic or poorly modeled regions, and (3) to generate an ensemble of protein conformation for docking (Amaro et al., 2008; Sgrignani et al., 2009), we performed three 250 ns long MD simulations for each of the selected models. PROPKA calculations with the model from 5F8T indicated a positively charged catalytic histidine (His296) as the most probable state. However, considering that the same residue was predicted as His- ϵ in the other two models and that this specific protonation state would be required to start the enzymatic reaction (Ishida and Kato, 2003), we simulated this specific residue in both protonation states.

The simulation outputs were analyzed using consolidated observables such as RMSD, Rg, and the per-residue RMSF (**Figure 4**). This analysis highlighted a higher stability of M-5F8T with respect to M-5CE1 and M-6O1G. In particular, the simulations of M-5F8T always converged to a maximum RMSD of <3 Å from the starting model and Rg values similar to the starting one. Contrarily, M-5CE1 and M-6O1G showed continuously increasing RMSD and Rg profiles, suggesting that these models are less stable.

The RMSF profiles of M-5F8T and M-6O1G did not show anything relevant, substantially confirming the stability of M-5F8T. Contrarily, in M-5CE1, the protein region between positions 320 and 350, which contains the Tyr322 and Ser333 loop discussed before, was characterized by high RMSF values.

Small molecules in the catalytic site (PC1, HPC, 7SD, **Figure 1**) of the templates were preserved in the TMPRSS2 models, as the

behavior of these molecules during MD simulations provides important hints about their quality and suitability to bind drugs. In the case of M-5F8T, the PC1 molecule remained in the S1 specificity pocket through a salt-bridge with Asp435 (**Figure 5A**). A similar behavior was also observed in the MD simulations of M-5CE1 (**Figure 5B**) for HPC. Contrarily, P4C rapidly dissociated from M-6O1G in all the simulations, probably because of the lack of a positively charged group docking the molecule to the S1 specificity pocket.

The good structural parameters (**Table 3**), the higher stability with respect to the M-5CE1 and M-6O1G (**Figure 4**), and the stable binding observed for PC1 in all the performed simulations suggested M-5F8T as the most reliable TMPRSS2 model. We, therefore, analyzed this model more deeply, focusing on the geometry of the catalytic triad (Asp345, His296, and Ser441). The analysis of distances between the three residues (**Figures 5C–F**) showed that this region of the protein remained stable during all the performed simulations. However, the system with a charged His296 adopted a conformation more similar to the starting model in which the C γ @Asp345–C γ @His296 and C β @Ser441–C γ @His296 distances are 5.1 and 4.4 Å, respectively.

Docking of Camostat and Nafamostat to TMPRSS2

As in the MD simulations, docking of camostat and nafamostat in M-5F8T was performed with His296 in two protonation states, that is, positively charged and protonated on ϵ .

The outcomes of these calculations (**Figure 6**) indicated that camostat adopts a similar binding mode irrespectively to the His296 protonation state. In particular, camostat places its guanidine group in the S1 specificity pocket where it forms a salt bridge with Asp435 orienting the other part of the molecule in the same direction.

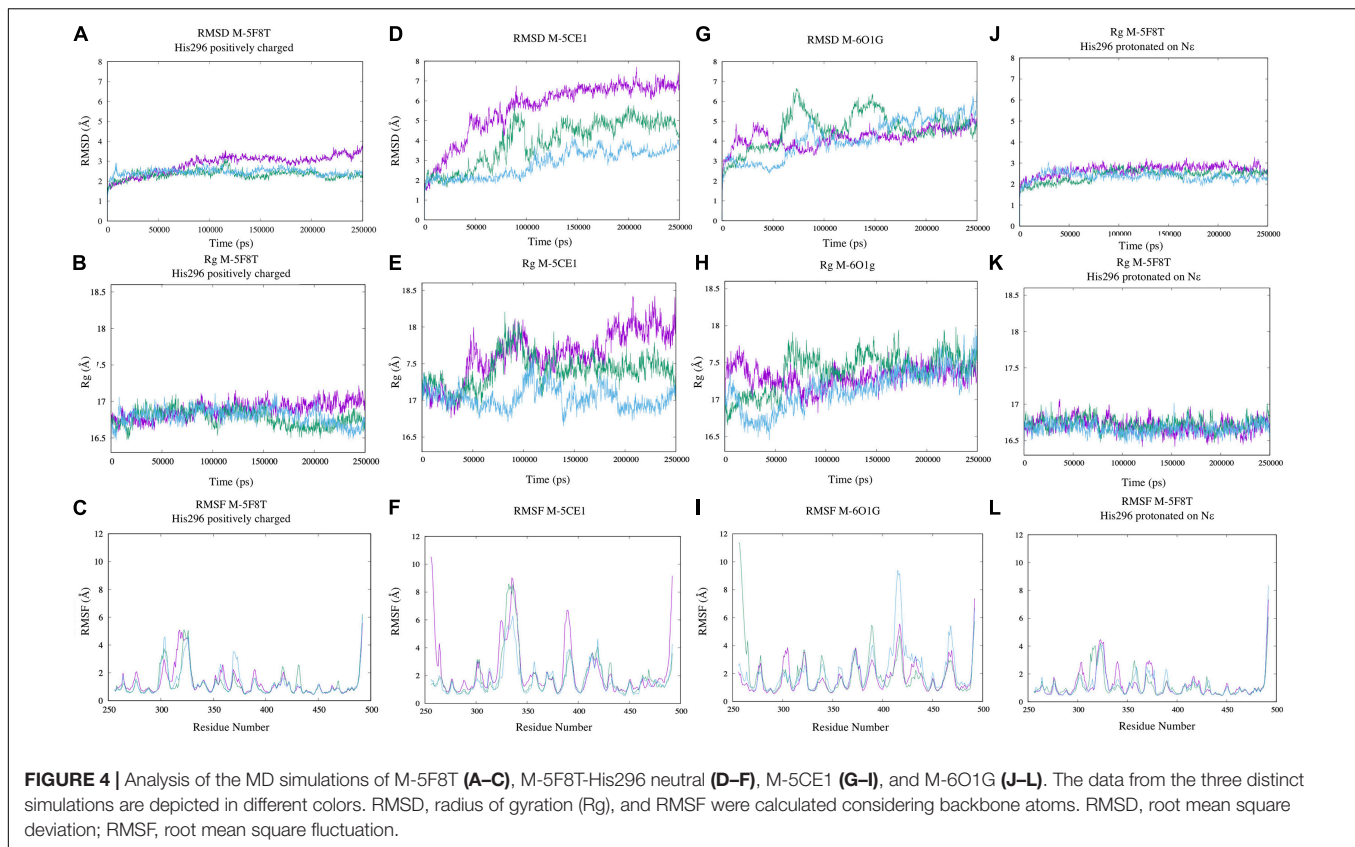
Nafamostat is characterized by the presence of a guanidine group and one its isoster. Therefore, while one of these is placed in the S1 specificity pocket, the other forms different interactions depending on the His296 protonation state. In fact, in the model with ϵ protonated His296, the guanidine moiety forms a salt bridge with Glu299. On the contrary, in the model with the positively charged His296, the isosteric group directly binds Asp345, that is one of the members of the catalytic triad (**Figure 6**).

Interestingly, the binding score is not affected by the His296 protonation state. However, the predicted score is lower for nafamostat than camostat, which is in a qualitative agreement with literature that reports an IC50 value for nafamostat 10 times lower than for camostat (Yamamoto et al., 2016).

TABLE 3 | Results of the structure quality evaluation.

Model name	Ramachandran violation	RMS bond dev.	RMS angle dev.	Backbone	Sidechains	Peptide planar dev.	Sidechains planar dev.	Torsion planar dev.	Z-score
M-5F8T	9	0.020	2.21	5	17	5.80	0.007	1.13	−6.82
M-5CE1	15	0.022	2.27	11	25	6.36	0.008	1.18	−7.12
M-6O1G	19	0.021	3.05	18	23	6.65	0.010	1.45	−6.47

Z-score is a measure of the overall model quality, and it was calculated by the Prosa-webserver (Wiederstein and Sippl, 2007). All the other parameters were calculated by Protein Structure Quality viewer implemented in the Schrodinger suite for molecular modeling.



Bromhexine Binding to TMPRSS2 Investigated by Microscale Thermophoresis

There have been discordant reports on the ability of BH to inhibit TMPRSS2. In fact, while the results of Lucas et al. (2014) appeared robust and convincing, a recent investigation by Hall and coworkers (Shrimp et al., 2020) concluded that BH is completely inactive as a TMPRSS2 inhibitor.

It is, however, important to consider that TMPRSS2 is a membrane protein with a peculiar and poorly understood activation mechanism. The purification of the active form of the enzyme, necessary for the inhibition tests, is thus extremely difficult. Furthermore, we noted that the protein quantity used in TMPRSS2 enzymatic assay is rarely reported (Meyer et al., 2013; Lucas et al., 2014), and that when reported (Shrimp et al., 2020), it extremely high (1 μ M) with respect to the 1–2 nM concentrations used for other similar proteases (Hammamy et al., 2013; Ivanova et al., 2017). This suggests that the active species could be only a small fraction of the total protein, making it more difficult to observe a non-covalent weaker inhibition as that of BH. Thus, to better understand the existence and the strength of a BH/TMPRSS2 complex, we performed MST experiments.

MST is a recently developed biophysical technique that enables the investigation of molecular complexes measuring changes, upon binding, of the migration of target proteins in a laser-induced thermal gradient (Jerabek-Willemsen et al., 2011, 2014; Fassi et al., 2019).

The results of the MST experiments confirmed the BH/TMPRSS2 interaction with a K_d of $24 \pm 13 \mu$ M (Figure 2A).

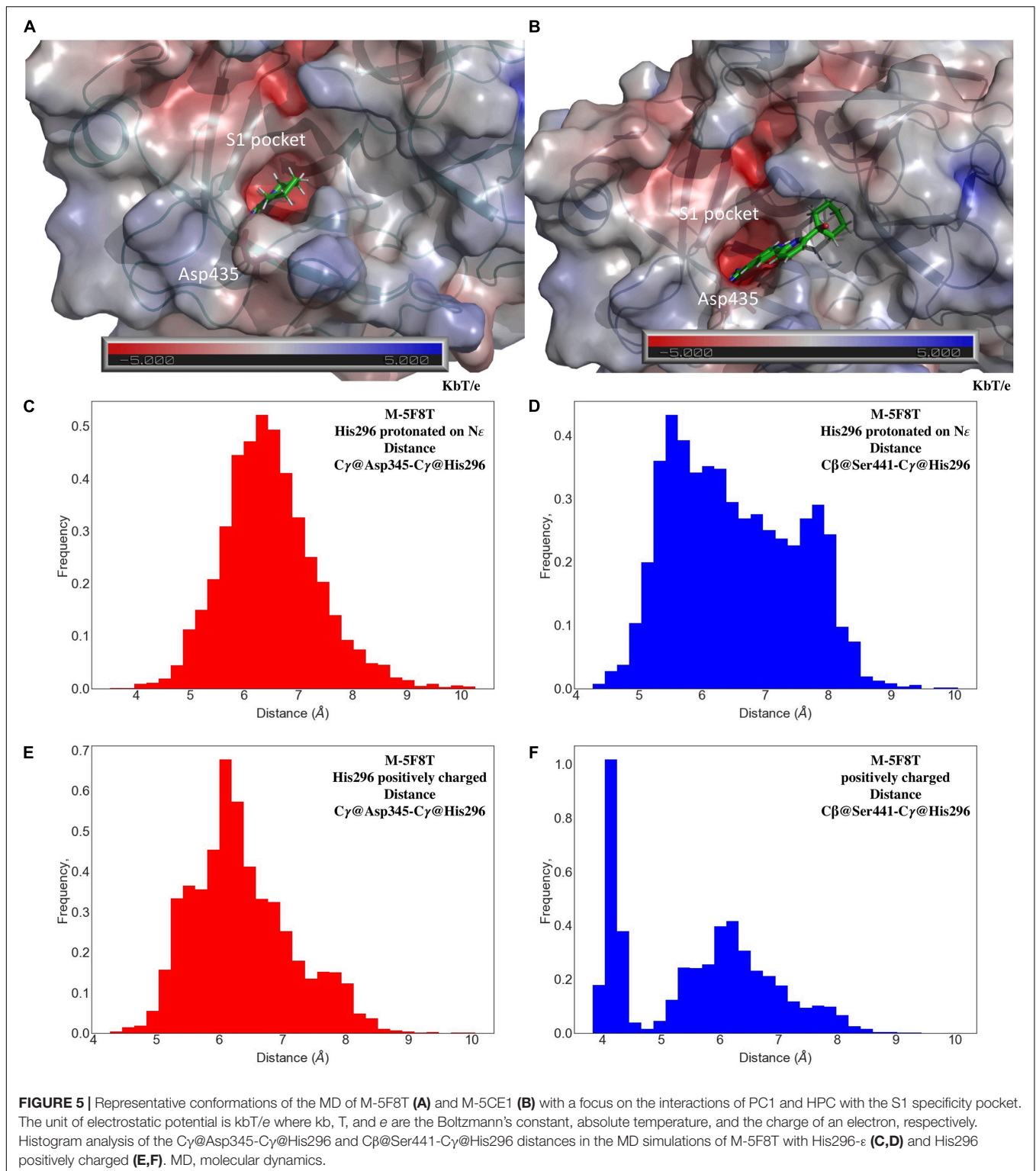
Modeling the Interaction Between BH and TMPRSS2

Motivated by literature data (Lucas et al., 2014) and by the results of our MST experiments, we used computational methods to investigate the BH/TMPRSS2 interaction at an atomistic level.

A closer look at the chemical structure of BH revealed that it cannot form a covalent bond with the protein and, therefore, should have a different inhibition mechanism compared to camostat and nafamostat. We, therefore, performed docking calculations considering the catalytic site as a putative BH binding site. However, when simulated by MD, the ligand–protein complexes dissociated after few nanoseconds suggesting a low reliability of the obtained structures. This observation was validated by performing several simulations starting from slightly different initial poses of BH in the catalytic site obtained using runs with different grids (data not shown): inevitably, the BH-TMPRSS2 complex dissociates in few nanoseconds.

In their 2007 review, Laporte and Naesens (Laporte and Naesens, 2017) suggested that, because of its selectivity for TMPRSS2 over matriptase, trypsin, or thrombin, BH could exert its inhibitory effect binding to an allosteric site.

To better explore this hypothesis, we analyzed our models with Sitemap (Halgren, 2007, 2009), a computational tool already



applied to the identification of allosteric sites (Sgrignani et al., 2014; Kots et al., 2017; Sanchez-Martin et al., 2020).

This analysis highlighted the existence of two putative drug binding sites (M-5F8T_site_1 and M-5F8T_site_2), for which a SiteScore value of >0.8 was obtained (Table 4). To note, while

Site_1 describes a zone quite far from the active site, Site_2 includes also a part of the active sites Ser441 and His296.

In recent years, several algorithms for the prediction of putative allosteric sites (see also section “Materials and Methods”) have been developed. Therefore, to obtain a more comprehensive

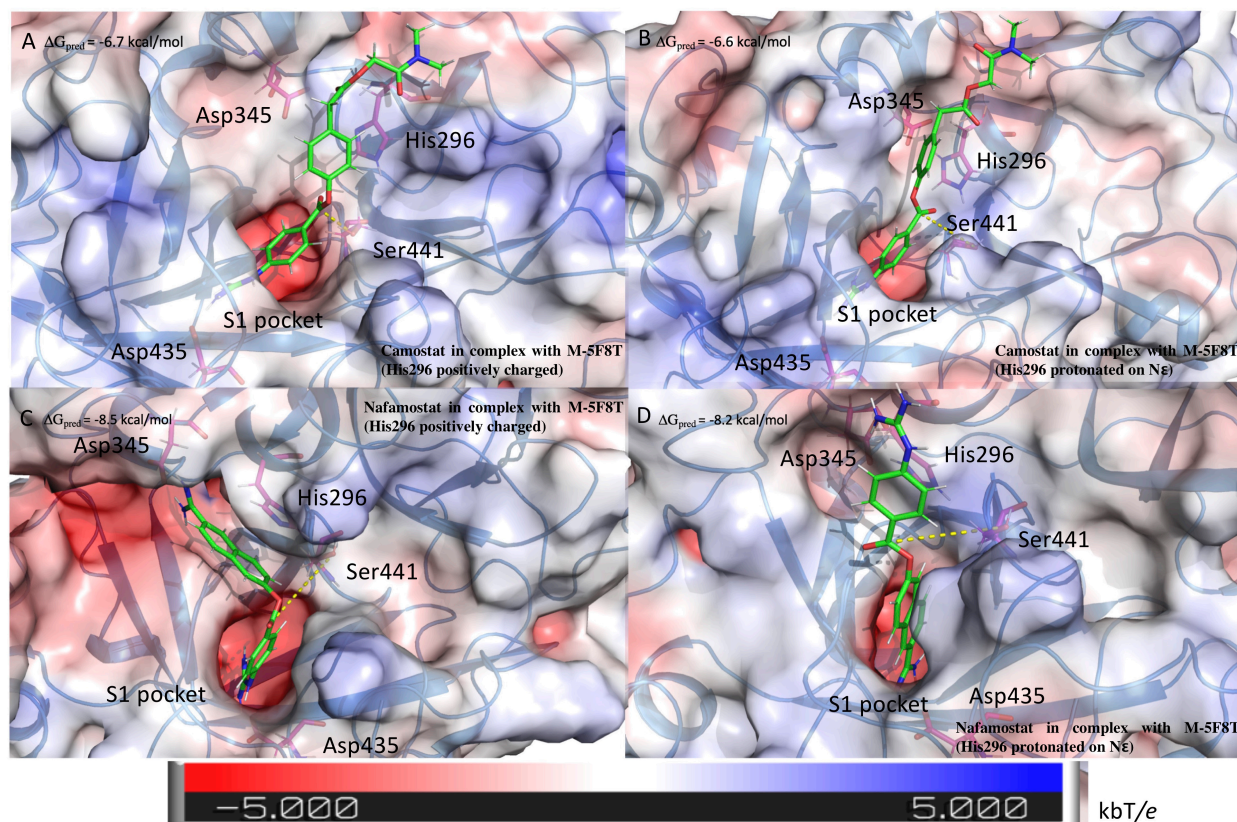


FIGURE 6 | Structure of the complexes between TMPRSS2 and camostat (A,B) or nafamostat (C,D). The pictures in the right and left columns refer to the docking calculation ran on M-5F8T considering His296 in its positively charged state or protonated on the ϵ nitrogen, respectively. The protein surface is colored in the function of the electrostatic potential according to the shown bar. The unit of electrostatic potential is $k_B T/e$ where k_B , T , and e are the Boltzmann's constant, absolute temperature, and the charge of an electron, respectively. The ΔG_{pred} values reported in the picture are the GlideScore values obtained from docking calculations.

analysis, we carried out the same calculation using three additional algorithms (PARS, DeepSite, and FTMap). All these calculations confirmed the existence of Site₁, while Site₂ was identified by PARS and FTMap but not DeepSite.

Next, we docked BH in M-5F8T_{site_1} and M-5F8T_{site_2} and performed MD simulations of the complexes. While the simulations with BH bound to M-5F8T_{site_2} resulted in a complex dissociating in the first 100 ns, the complex between BH and TMPRSS2 bound to M-5F8T_{site_1} and remained stable for $\sim 1 \mu\text{s}$. In fact, the drug remained close to the protein although it did not find a stable binding mode. Consequently, we performed two additional MD simulations to clarify this point. In the first control simulation, the ligand dissociated in the first 500 ns, while the second control simulation BH remained close to the protein surface without finding a stable binding mode, as in the first run. It should be also noted that in these simulations, while close to the protein surface, BH has a distance of $\sim 30 \text{ \AA}$ from the catalytic triad, making it difficult to imagine a direct effect on the catalytic activity from that position.

Summarizing the results of our simulations indicated that M-5F8T_{site_1} and M-5F8T_{site_2} are unsuitable to bind BH, leaving unsolved the question about the position of the BH allosteric site.

We, therefore, explored the possibility of the existence of a hidden allosteric site.

The Role in Protein Activity of Free Isoleucine at the N-Terminal Side

The essential role for the enzymatic activity of the free isoleucine at the N-terminal side of TLPs has been previously reported (Stubbs et al., 1998; Huber, 2013; Meyer et al., 2013).

From visual inspection of M-5F8T, it can be seen that the N-terminal fragment of the protein occupies a negatively charged cavity (subsequently A-pocket, **Figure 7**) where the positively charged amino group of the N-terminal Ile256 forms a salt bridge with Asp440. Interestingly, Asp440 is contiguous to Ser441, one of the members of the catalytic triads, but oriented in a different direction.

To investigate the importance of this structural feature (i.e., the presence of free isoleucine at N-terminal site) for TMPRSS2, we generated a model of the enzyme deleting the first two residues at N-terminal (Ile256 and Val257) from M-5F8T (this model is subsequently indicated as C-M-5F8T) and performed an MD simulation for $1 \mu\text{s}$.

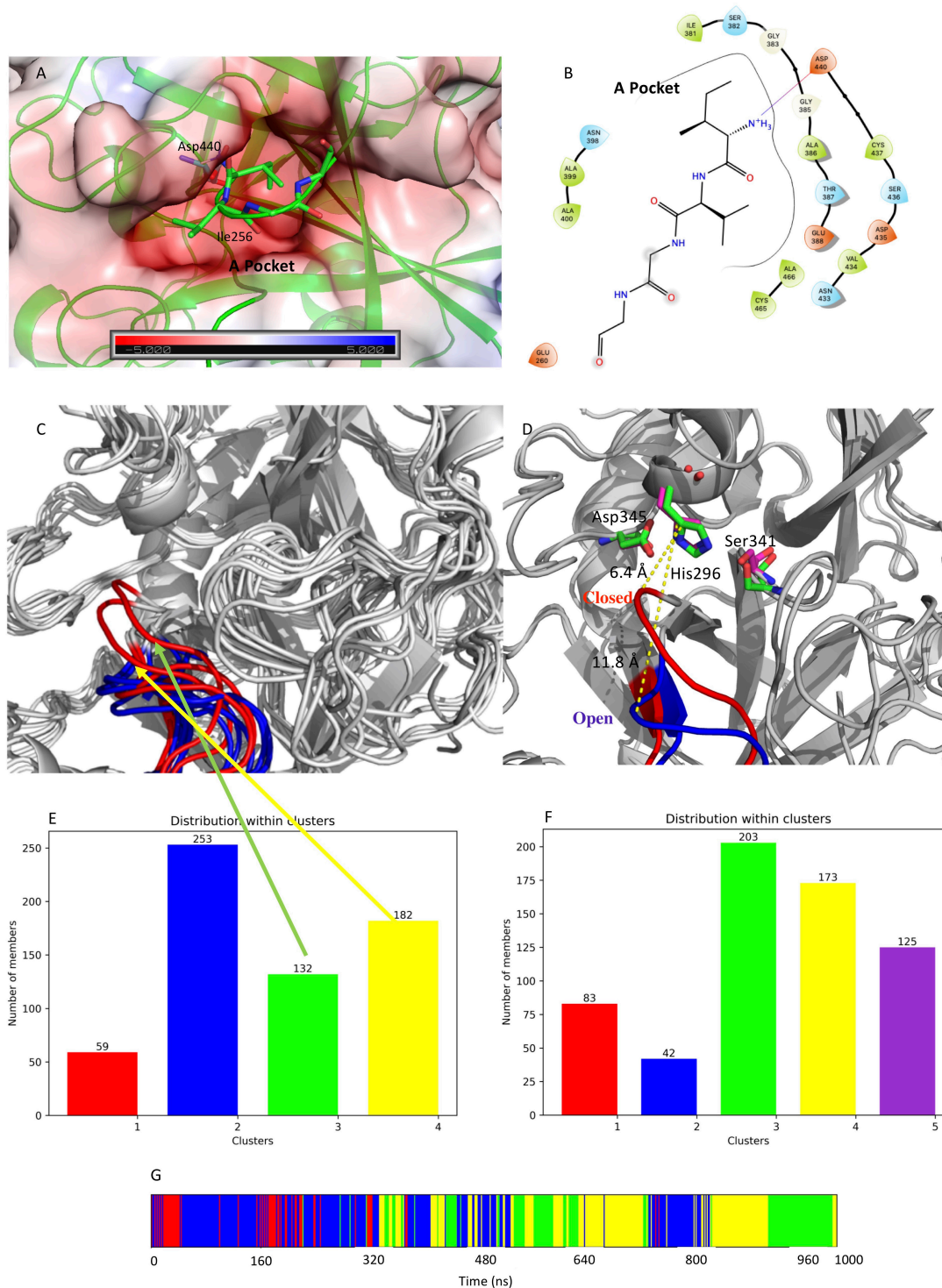
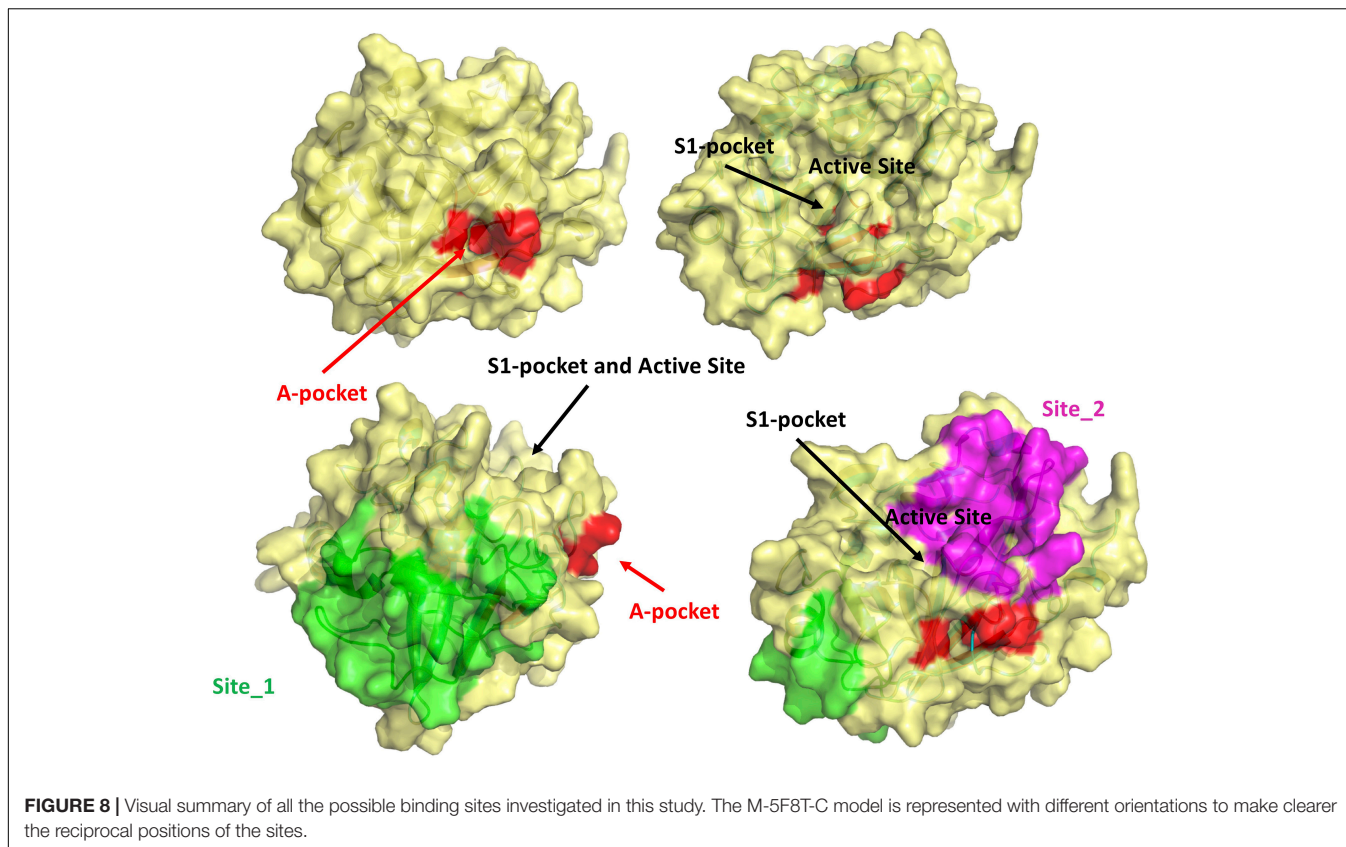


FIGURE 7 | (A) Structure of the negatively charged cavity that hosts the N-terminal tail of the catalytically active TMPRSS2. The protein surface is colored by the electrostatic potential value calculated by the APBS plugin implemented in Pymol-2.3.4. The unit of electrostatic potential is $k_B T/e$ where k_B , T , and e are the Boltzmann's constant, absolute temperature, and the charge of an electron, respectively. **(B)** Scheme of the interactions between the N-terminal end and its binding site on the TMPRSS2 structure. **(C)** Conformations of the loop Gly462-Val473 in the representative structures from the identified clusters. The conformations from the simulation of apo-M-5F8T are shown in blue while those from the simulation of C-M-5F8T in red. **(D)** Comparison between the loop conformation assumed in the cluster3 of the C-M-5F8T MD (in red) and that of cluster3 of the M-5F8T MD (in blue). **(E)** Distribution of the C-M-5F8T conformations over the identified clusters. **(F)** Distribution of the M-5F8T conformations over the identified clusters. **(G)** Time evolution of the clusters obtained from the C-M-5F8T MD simulation. MD, molecular dynamics.



Next, we compared the outputs of this simulation with an identical simulation of M-5F8T in its apo form.

Given its importance for the substrate recognition in TLPs and its structural proximity to the binding site, we first focused our analysis on the effect of the presence/absence of Ile256-Val257 on the structure of the S1 specificity pocket. Visual inspection of M-5F8T suggested that the Gly462-Val473 loop could regulate the access of the substrates to the S1 specificity pocket. We, therefore, analyzed the effects of the N-terminal truncation on the conformation of this protein region. Interestingly, we observed (**Figure 7C**) that, while in the simulation of the M-5F8T, the loop conserves a conformation similar to that adopted in the starting model; in the second part of the C-M-5F8T simulation, it moves closer to the catalytic triad (**Figure 7D**) occupying a position that should reduce the efficiency of both substrate recognition and catalysis.

Taken together, these observations strongly suggest that the binding of the N-terminal tail into the A-pocket (**Figures 7A,B, 8**) is important to stabilize the structure of the TMPRSS2 active site and, in particular, of the S1 specificity pocket.

Is the A-Pocket Relevant for Drug Design?

Considering the importance, highlighted by the previously discussed simulations, of the binding of the N-terminal tail into the A-pocket for the stability of the catalytic site, we performed some analysis to explore its druggability.

To this end, we first used Sitemap to analyze the C-M-5F8T model. This analysis showed that the cavity was made accessible by the deletion of the first two residues of M-5F8T and was highly suitable for drug binding, with a value of SiteScore of 0.93 over 1.00, where a druggable cavity should have SiteScore > 0.80.

TABLE 4 | Results of the Sitemap analysis carried out on M-5F8T.

Title	SiteScore	Dscore	Volume (Å ³)	Residues
M-5F8T_site_3	0.924	0.656	68.9	262, 263, 264, 265, 266, 267, 268, 270, 271, 272, 311, 312, 313, 314, 315, 316, 360, 384, 397
M-5F8T_site_2	0.892	0.907	201.3	274, 275, 277, 278, 279, 280, 281, 296, 300, 301, 302, 307, 308, 317, 384, 385, 386, 390, 391, 392, 393, 438, 439, 441
M-5F8T_site_1	0.863	0.872	279.2	369, 370, 371, 372, 373, 374, 376, 377, 403, 404, 405, 406, 407, 409, 413, 421, 422, 424, 425, 428, 429, 430, 469, 471, 476, 478, 479
M-5F8T_site_4	0.738	0.395	50.7	367, 368, 371, 372, 373, 375, 376, 447, 449, 454, 456, 478
M-5F8T_site_5	0.655	0.608	96.7	271, 291, 310, 311, 312, 325, 326, 327, 351, 355

Next, we investigated if this pocket could be a suitable site for the BH binding.

Preliminary calculations (see section “Materials and Methods”) showed that the cavity was optimized for the binding of the N-terminal tail and not for the binding of a small molecule. We, therefore, first computed the optimal BH/TMPRSS2 binding pose using the IFD protocol implemented in Glide, followed by three MD independent simulations of 500 ns each.

The results of the docking calculations indicated the both (S)- and (R)-BH bind the A-pocket with a similar predicted affinity (−6.9 and −6.3 kcal/mol, respectively, **Figures 2B–E**).

From the structural point of view, both the molecules place the ring bearing the two bromine atoms in a cavity delimited by Gly363, Ile381, Ser382, Trp384, and Asp440. Moreover, the amino group in position 5 of the same ring establishes h-bond interactions with Asp440 and Gly383 in the case of (S)-BH and with Asn390 and Glu260 for (R)-BH. In both the structures, the positively charged amino group of BH electrostatically interacts with Glu260.

All MD simulations confirmed the stability of the complexes obtained from docking, with BH bound into the A-pocket for the entire simulation time.

Finally, we also performed a cluster analysis to verify the conformation of the Gly462-Val473 loop, which regulates the access to the S1 specificity pocket. This analysis (**Supplementary Figure 1**) clearly showed that the loop conserves a closed conformation in all the representative structures extracted from the simulations of (S)- and (R)-BH inside the A-pocket.

CONCLUSION

TMPRSS2 is an exceptional and intriguing protein (Thunders and Delahunt, 2020), whose precise physiological function remains unknown. Despite this, it has been linked with several human diseases, such as prostate cancer, and has been shown to play a key role in viral infections.

In particular, the SPD of TMPRSS2 is critical for the priming of SARS-Cov-2 spike protein. This prompted us to investigate the interaction between TMPRSS2 various known drugs, using both computational and experimental methods.

While in the case of camostat and nafamostat, our computational studies confirmed that these two molecules bind to the active site of TMPRSS2 and form molecular adducts competent for the formation of covalent complexes;

REFERENCES

- Afar, D. E., Vivanco, I., Hubert, R. S., Kuo, J., Chen, E., Saffran, D. C., et al. (2001). Catalytic cleavage of the androgen-regulated TMPRSS2 protease results in its secretion by prostate and prostate cancer epithelia. *Cancer Res.* 61, 1686–1692.
- Amaro, R. E., Baron, R., and Mccammon, J. A. (2008). An improved relaxed complex scheme for receptor flexibility in computer-aided drug design. *J. Comput. Aided Mol. Des.* 22, 693–705. doi: 10.1007/s10822-007-9159-2
- Amaro, R. E., Baudry, J., Chodera, J., Demir, Ö, Mccammon, J. A., Miao, Y., et al. (2018). Ensemble docking in drug discovery. *Biophys. J.* 114, 2271–2278.

in the case of BH, our studies indicated that a competitive inhibition was unlikely.

On the other side, MST experiments confirmed a BH/TMPRSS2 interaction, leading us to ponder the hypothesis of an allosteric binding. We, therefore, used computer simulations to validate this hypothesis. The MD simulation confirmed that similar to other TLPs, the binding of a free isoleucine residue in the A-pocket is crucial to stabilize the catalytically competent active site conformation. Moreover, our calculations indicated that this cavity (**Figure 8**), fully accessible in the TMPRSS2 zymogen, is suitable to host BH or other more potent drugs that could be identified by virtual screening.

The study presented here provides further understanding of how the catalytic activity of TMPRSS2 can be modulated and new ways to develop more selective and potent antiviral treatments for current and future pandemics.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation, to any qualified researcher.

AUTHOR CONTRIBUTIONS

JS designed the study, performed and analyzed simulations and experiments, and wrote and revised the manuscript. AC designed the study, analyzed the results of simulations and experiments, and wrote and revised the manuscript. Both authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

We thank Torsten Steinmetzer for useful discussions and Flora Gruner for the generous support. We also thank Fondazione CARIPARO for the financial support.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2021.666626/full#supplementary-material>

- Ansarin, K., Tolouian, R., Ardalan, M., Taghizadieh, A., Varshochi, M., Teimouri, S., et al. (2020). Effect of bromhexine on clinical outcomes and mortality in COVID-19 patients: a randomized clinical trial. *Bioimpacts* 10, 209–215. doi: 10.34172/bi.2020.27
- Bertram, S., Glowacka, I., Blazejewski, P., Soilleux, E., Allen, P., Danisch, S., et al. (2010). TMPRSS2 and TMPRSS4 facilitate trypsin-independent spread of influenza virus in Caco-2 cells. *J. Virol.* 84, 10016–10025. doi: 10.1128/jvi.00239-10
- Bestle, D., Heindl, M. R., Limburg, H., Van Lam Van, T., Pilgram, O., Moulton, H., et al. (2020). TMPRSS2 and furin are both essential for proteolytic activation of SARS-CoV-2 in human airway cells. *Life Sci. Alliance* 3:e202000786. doi: 10.26508/lsa.202000786

- Cavasotto, C. N., and Phatak, S. S. (2009). Homology modeling in drug discovery: current trends and applications. *Drug Discov. Today* 14, 676–683. doi: 10.1016/j.drudis.2009.04.006
- Chen, Y.-W., Lee, M.-S., Lucht, A., Chou, F.-P., Huang, W., Havighurst, T. C., et al. (2010). TMPRSS2, a serine protease expressed in the prostate on the apical surface of luminal epithelial cells and released into semen in prostatesomes, is misregulated in prostate cancer cells. *Am. J. Pathol.* 176, 2986–2996. doi: 10.2353/ajpath.2010.090665
- Depenhart, M., De Villiers, D., Lemperle, G., Meyer, M., and Di Somma, S. (2020). Potential new treatment strategies for COVID-19: is there a role for bromhexine as add-on therapy? *Intern. Emerg. Med.* 15, 801–812. doi: 10.1007/s11739-020-02383-3
- Fassi, E. M. A., Sgrignani, J., D'agostino, G., Cecchinato, V., Garofalo, M., Grazioso, G., et al. (2019). Oxidation state dependent conformational changes of HMGB1 regulate the formation of the CXCL12/HMGB1 Heterocomplex. *Comput. Struct. Biotechnol. J.* 17, 886–894. doi: 10.1016/j.csbj.2019.06.020
- Friesner, R. A., Banks, J. L., Murphy, R. B., Halgren, T. A., Klicic, J. J., Mainz, D. T., et al. (2004). Glide: a new approach for rapid, accurate docking and scoring. 1. method and assessment of docking accuracy. *J. Med. Chem.* 47, 1739–1749. doi: 10.1021/jm0306430
- Fu, B., Sahakyan, A. B., Camilloni, C., Tartaglia, G. G., Paci, E., Caflisch, A., et al. (2014). ALMOST: an all atom molecular simulation toolkit for protein structure determination. *J. Comput. Chem.* 35, 1101–1105. doi: 10.1002/jcc.23588
- Guarnera, E., and Berezovsky, I. N. (2019). Toward comprehensive allosteric control over protein activity. *Structure* 27, 866–878.e861.
- Habtemariam, S., Nabavi, S. F., Ghavami, S., Cismaru, C. A., Berindan-Neagoie, I., and Nabavi, S. M. (2020). Possible use of the mucolytic drug, bromhexine hydrochloride, as a prophylactic agent against SARS-CoV-2 infection based on its action on the Transmembrane Serine Protease 2. *Pharmacol. Res.* 157:104853. doi: 10.1016/j.phrs.2020.104853
- Halgren, T. (2007). New method for fast and accurate binding-site identification and analysis. *Chem. Biol. Drug Des.* 69, 146–148. doi: 10.1111/j.1747-0285.2007.00483.x
- Halgren, T. A. (2009). Identifying and characterizing binding sites and assessing druggability. *J. Chem. Inform. Modell.* 49, 377–389. doi: 10.1021/ci800324m
- Hammamy, M. Z., Haase, C., Hammami, M., Hilgenfeld, R., and Steinmetzer, T. (2013). Development and characterization of new peptidomimetic inhibitors of the West Nile virus NS2B-NS3 protease. *ChemMedChem* 8, 231–241. doi: 10.1002/cmdc.201200497
- Harder, E., Damm, W., Maple, J., Wu, C., Reboul, M., Xiang, J. Y., et al. (2016). OPLS3: a force field providing broad coverage of drug-like small molecules and proteins. *J. Chem. Theory Comput.* 12, 281–296. doi: 10.1021/acs.jctc.5b00864
- Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., et al. (2020). SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* 181, 271–280.e278.
- Huber, R. (2013). How I chose research on proteases or, more correctly, how it chose me. *Angew. Chem. Int. Ed. Engl.* 52, 68–73. doi: 10.1002/anie.201205629
- Ishida, T., and Kato, S. (2003). Theoretical perspectives on the reaction mechanism of serine proteases: the reaction free energy profiles of the acylation process. *J. Am. Chem. Soc.* 125, 12035–12048. doi: 10.1021/ja021369m
- Ivanova, T., Hards, K., Kallis, S., Dahms, S. O., Than, M. E., Künzel, S., et al. (2017). Optimization of substrate-analogue furin inhibitors. *ChemMedChem* 12, 1953–1968. doi: 10.1002/cmdc.201700596
- Jerabek-Willemsen, M., André, T., Wanner, R., Roth, H. M., Duhr, S., Baaske, P., et al. (2014). MicroScale thermophoresis: interaction analysis and beyond. *J. Mol. Struct.* 1077, 101–113. doi: 10.1016/j.molstruc.2014.03.009
- Jerabek-Willemsen, M., Wienken, C. J., Braun, D., Baaske, P., and Duhr, S. (2011). Molecular interaction studies using microscale thermophoresis. *Assay Drug. Dev. Technol.* 9, 342–353. doi: 10.1089/adt.2011.0380
- Jiménez, J., Doerr, S., Martínez-Rosell, G., Rose, A. S., and De Fabritiis, G. (2017). DeepSite: protein-binding site predictor using 3D-convolutional neural networks. *Bioinformatics* 33, 3036–3042. doi: 10.1093/bioinformatics/btx350
- Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, L. M. (1983). Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79, 926–935. doi: 10.1063/1.445869
- Kots, E. D., Lushchekina, S. V., Varfolomeev, S. D., and Nemukhin, A. V. (2017). Role of protein dimeric interface in allosteric inhibition of N-Acetyl-aspartate hydrolysis by human aspartoacylase. *J. Chem. Inform. Modell.* 57, 1999–2008. doi: 10.1021/acs.jcim.7b00133
- Kozakov, D., Grove, L. E., Hall, D. R., Bohnuud, T., Mottarella, S. E., Luo, L., et al. (2015). The FTMap family of web servers for determining and characterizing ligand-binding hot spots of proteins. *Nat. Protoc.* 10, 733–755. doi: 10.1038/nprot.2015.043
- Laporte, M., and Naesens, L. (2017). Airway proteases: an emerging drug target for influenza and other respiratory virus infections. *Curr. Opin. Virol.* 24, 16–24. doi: 10.1016/j.coviro.2017.03.018
- Li, T., Sun, L., Zhang, W., Zheng, C., Jiang, C., Chen, M., et al. (2020). Bromhexine hydrochloride tablets for the treatment of moderate COVID-19: an open-label randomized controlled pilot study. *Clin. Transl. Sci.* 13, 1096–1102. doi: 10.1111/cts.12881
- Lucas, J. M., Heinlein, C., Kim, T., Hernandez, S. A., Malik, M. S., True, L. D., et al. (2014). The androgen-regulated protease TMPRSS2 activates a proteolytic cascade involving components of the tumor microenvironment and promotes prostate cancer metastasis. *Cancer Discov.* 4, 1310–1325. doi: 10.1158/2159-8290.cd-13-1010
- Maggio, R., and Corsini, G. U. (2020). Repurposing the mucolytic cough suppressant and TMPRSS2 protease inhibitor bromhexine for the prevention and management of SARS-CoV-2 infection. *Pharmacol. Res.* 157:104837. doi: 10.1016/j.phrs.2020.104837
- Martyna, G. J., Klein, M. L., and Tuckerman, M. (1992). Nosé–hoover chains: the canonical ensemble via continuous dynamics. *J. Chem. Phys.* 97, 2635–2643. doi: 10.1063/1.463940
- Martyna, G., Tobias, D., and Klein, M. (1994). Constant pressure molecular dynamics algorithms. *J. Chem. Phys.* 101, 4177–4189. doi: 10.1063/1.467468
- Meyer, D., Sielaff, F., Hammami, M., Böttcher-Friebertshäuser, E., Garten, W., and Steinmetzer, T. (2013). Identification of the first synthetic inhibitors of the type II transmembrane serine protease TMPRSS2 suitable for inhibition of influenza virus activation. *Biochem. J.* 452, 331–343. doi: 10.1042/bj20130101
- Montopoli, M., Zumerle, S., Vettor, R., Ruggie, M., Zorzi, M., Catapano, C. V., et al. (2020). Androgen-deprivation therapies for prostate cancer and risk of infection by SARS-CoV-2: a population-based study (N = 4532). *Ann. Oncol.* 31, 1040–1045. doi: 10.1016/j.annonc.2020.04.479
- Olsson, M. H., Sondergaard, C. R., Rostkowski, M., and Jensen, J. H. (2011). PROPKA3: consistent treatment of internal and surface residues in empirical pKa predictions. *J. Chem. Theory Comput.* 7, 525–537. doi: 10.1021/ct100578z
- Panjikovich, A., and Daura, X. (2014). PARS: a web server for the prediction of protein allosteric and regulatory sites. *Bioinformatics* 30, 1314–1315. doi: 10.1093/bioinformatics/btu002
- Partridge, J. R., Choy, R. M., Silva-Garcia, A., Yu, C., Li, Z., Sham, H., et al. (2019). Structures of full-length plasma kallikrein bound to highly specific inhibitors describe a new mode of targeted inhibition. *J. Struct. Biol.* 206, 170–182. doi: 10.1016/j.jsb.2019.03.001
- Pásztí-Gere, E., Czimmermann, E., Ujhelyi, G., Balla, P., Maiwald, A., and Steinmetzer, T. (2016). In vitro characterization of TMPRSS2 inhibition in IPEC-J2 cells. *J. Enzyme Inhib. Med. Chem.* 31, 123–129. doi: 10.1080/14756366.2016.1193732
- Sanchez-Martin, C., Moroni, E., Ferraro, M., Laquatra, C., Cannino, G., Masgras, I., et al. (2020). Rational design of allosteric and selective inhibitors of the molecular chaperone TRAP1. *Cell Rep.* 31:107531. doi: 10.1016/j.celrep.2020.107531
- Schmaier, A. H. (2013). “Chapter 638 - prekallikrein and plasma kallikrein,” in *Handbook of Proteolytic Enzymes*, Third Edition. eds N. D. Rawlings and G. Salvesen (Cambridge, MA: Academic Press), 2885–2892. doi: 10.1016/b978-0-12-382219-2.00638-4
- Sgrignani, J., Bon, M., Colombo, G., and Magistrato, A. (2014). Computational approaches elucidate the allosteric mechanism of human aromatase inhibition: a novel possible route to small-molecule regulation of CYP450s activities? *J. Chem. Inf. Mod.* 54, 2856–2868. doi: 10.1021/ci500425y
- Sgrignani, J., Bonaccini, C., Grazioso, G., Chioccioli, M., Cavalli, A., and Gratteri, P. (2009). Insights into docking and scoring neuronal alpha4beta2 nicotinic receptor agonists using molecular dynamics simulations and QM/MM calculations. *J. Comput. Chem.* 30, 2443–2454. doi: 10.1002/jcc.21251

- Sgrignani, J., Garofalo, M., Matkovic, M., Merulla, J., Catapano, C. V., and Cavalli, A. (2018). Structural biology of STAT3 and its implications for anticancer therapies development. *Int. J. Mol. Sci.* 19:1591. doi: 10.3390/ijms19061591
- Shen, L. W., Mao, H. J., Wu, Y. L., Tanaka, Y., and 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
- Sherman, W., Day, T., Jacobson, M. P., Friesner, R. A., and Farid, R. (2006). Novel procedure for modeling ligand/receptor induced fit effects. *J. Med. Chem.* 49, 534–553. doi: 10.1021/jm050540c
- Shrimp, J. H., Kales, S. C., Sanderson, P. E., Simeonov, A., Shen, M., and Hall, M. D. (2020). An enzymatic TMPRSS2 assay for assessment of clinical candidates and discovery of inhibitors as potential treatment of COVID-19. *ACS Pharmacol. Transl. Sci.* 3, 997–1007. doi: 10.1021/acspstci.0c00106
- Singh, N., Decroly, E., Khatib, A.-M., and Villoutreix, B. O. (2020). Structure-based drug repositioning over the human TMPRSS2 protease domain: search for chemical probes able to repress SARS-CoV-2 Spike protein cleavages. *Eur. J. Pharm. Sci.* 153, 105495–105495. doi: 10.1016/j.ejps.2020.105495
- Stubbs, M. T., Renatus, M., and Bode, W. (1998). An active zymogen: unravelling the mystery of tissue-type plasminogen activator. *Biol. Chem.* 379, 95–103.
- Sungnak, W., Huang, N., Becavin, C., Berg, M., Queen, R., Litvinukova, M., et al. (2020). SARS-CoV-2 entry factors are highly expressed in nasal epithelial cells together with innate immune genes. *Nat. Med.* 26, 681–687. doi: 10.1038/s41591-020-0868-6
- Szabo, R., Wu, Q., Dickson, R. B., Netzel-Arnett, S., Antalis, T. M., and Bugge, T. H. (2003). Type II transmembrane serine proteases. *Thromb. Haemost.* 90, 185–193. doi: 10.1160/th03-02-0071
- Thunders, M., and Delahunt, B. (2020). Gene of the month: TMPRSS2 (transmembrane serine protease 2). *J. Clin. Pathol.* 73, 773–776. doi: 10.1136/jclinpath-2020-206987
- Tibshirani, R., Walther, G., and Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *R. Stat. Soc.* 63, 411–423. doi: 10.1111/1467-9868.00293
- Tubiana, T., Carvaillo, J.-C., Boulard, Y., and Bressanelli, S. (2018). TTClust: a versatile molecular simulation trajectory clustering program with graphical summaries. *J. Chem. Inform. Modell.* 58, 2178–2182. doi: 10.1021/acs.jcim.8b00512
- Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., et al. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* 46, W296–W303.
- Wiederstein, M., and Sippl, M. J. (2007). ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucl. Acids Res.* 35, W407–W410.
- Xiang, Z. (2006). Advances in homology protein structure modeling. *Curr. Protein Pept. Sci.* 7, 217–227. doi: 10.2174/13892030677452312
- Xu, Y., Wang, S., Hu, Q., Gao, S., Ma, X., Zhang, W., et al. (2018). CavityPlus: a web server for protein cavity detection with pharmacophore modelling, allosteric site identification and covalent ligand binding ability prediction. *Nucleic Acids Res.* 46, W374–W379.
- Yamamoto, M., Matsuyama, S., Li, X., Takeda, M., Kawaguchi, Y., Inoue, J. I., et al. (2016). Identification of nafamostat as a potent inhibitor of middle east respiratory syndrome coronavirus S protein-mediated membrane fusion using the split-protein-based cell-cell fusion assay. *Antimicrob. Agents Chemother.* 60, 6532–6539. doi: 10.1128/aac.01043-16
- Yu, J., Zhou, Y., Tanaka, I., and Yao, M. (2010). Roll: a new algorithm for the detection of protein pockets and cavities with a rolling probe sphere. *Bioinformatics* 26, 46–52. doi: 10.1093/bioinformatics/btp599
- Zang, R., Gomez Castro, M. F., Mccune, B. T., Zeng, Q., Rothlauf, P. W., and Sonnek, N. M. (2020). TMPRSS2 and TMPRSS4 promote SARS-CoV-2 infection of human small intestinal enterocytes. *Sci. Immunol.* 5:eabc3582. doi: 10.1126/sciimmunol.abc3582

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Sgrignani and Cavalli. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.