

blast.ncbi.nlm.nih.gov/Blast.cgi

Job Title: **NG\_047085.2:21380-73566 Homo sapiens transmembrane**  
 RID: [H6KJ98AW016](#) Search expires on 07-20 03:32 am [Download All](#)  
 Program: BLASTN [Citation](#)  
 Database: pdbnt [See details](#)  
 Query ID: lcl|Query\_55779  
 Description: NG\_047085.2:21380-73566 Homo sapiens transmembrane se ...  
 Molecule type: dna  
 Query Length: 52187  
 Other reports: [Distance tree of results](#) [MSA viewer](#)

**Filter Results**  
 Organism: only top 20 will appear  exclude  
 Type common name, binomial, taxid or group name  
 + [Add organism](#)  
 Percent Identity:  to   
 E value:  to   
 Query Coverage:  to   
[Filter](#) [Reset](#)

Descriptions **Graphic Summary** Alignments Taxonomy

Alignment Scores:  <40  40 - 50  50 - 80  80 - 200  >= 200

4 sequences selected

**Distribution of the top 9 Blast Hits on 4 subject sequences**

Sequences producing significant alignments

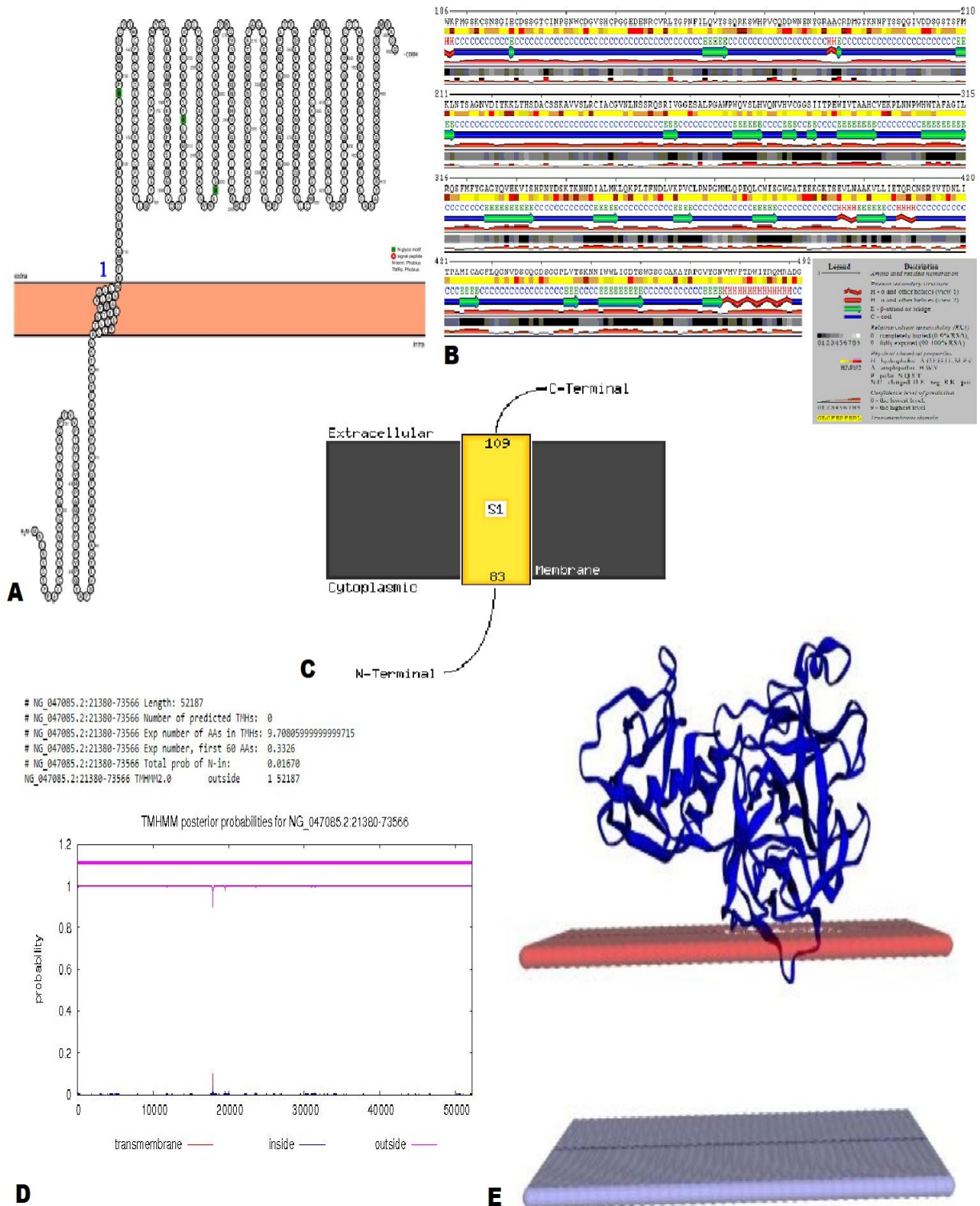
Download Manage Columns Show 100

select all 4 sequences selected

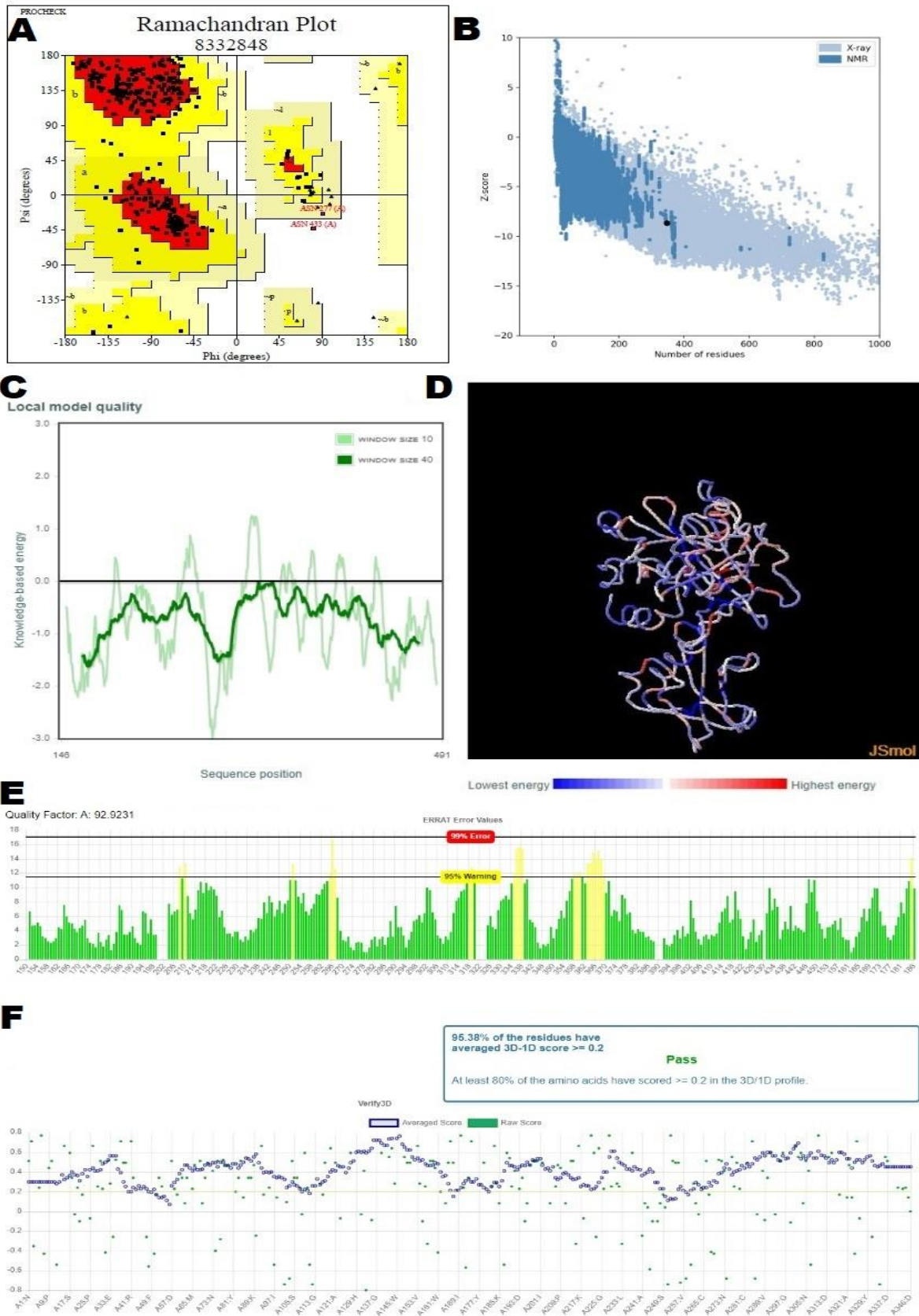
| Description   | Max Score | Total Score | Query Cover | E value | Per Ident | Accession              |
|---|-----------|-------------|-------------|---------|-----------|------------------------|
| <input checked="" type="checkbox"/> <a href="#">Chain F, ALU JO CONSENSUS RNA</a>   | 91.6      | 256         | 0%          | 4e-17   | 93.44%    | <a href="#">5A0X_F</a> |
| <input checked="" type="checkbox"/> <a href="#">Chain C, ALU JO CONSENSUS RNA</a>   | 91.6      | 256         | 0%          | 4e-17   | 93.44%    | <a href="#">5A0X_C</a> |
| <input checked="" type="checkbox"/> <a href="#">Chain C, Alu Domain Of The Mammalian Srp (Potential Alu Retroposition Intermediate)</a>                   | 84.2      | 84.2        | 0%          | 6e-15   | 91.67%    | <a href="#">1E8S_C</a> |
| <input checked="" type="checkbox"/> <a href="#">Chain A, Structural basis for targeting and elongation arrest of Bacillus signal recognition particle</a> | 80.5      | 161         | 0%          | 8e-14   | 88.06%    | <a href="#">4UE5_A</a> |

Feedback

**Supplementary fig.S1:** PSI BLAST analysis for selected TMPRSS2 Amino acid sequence with no Significant PDB structure.

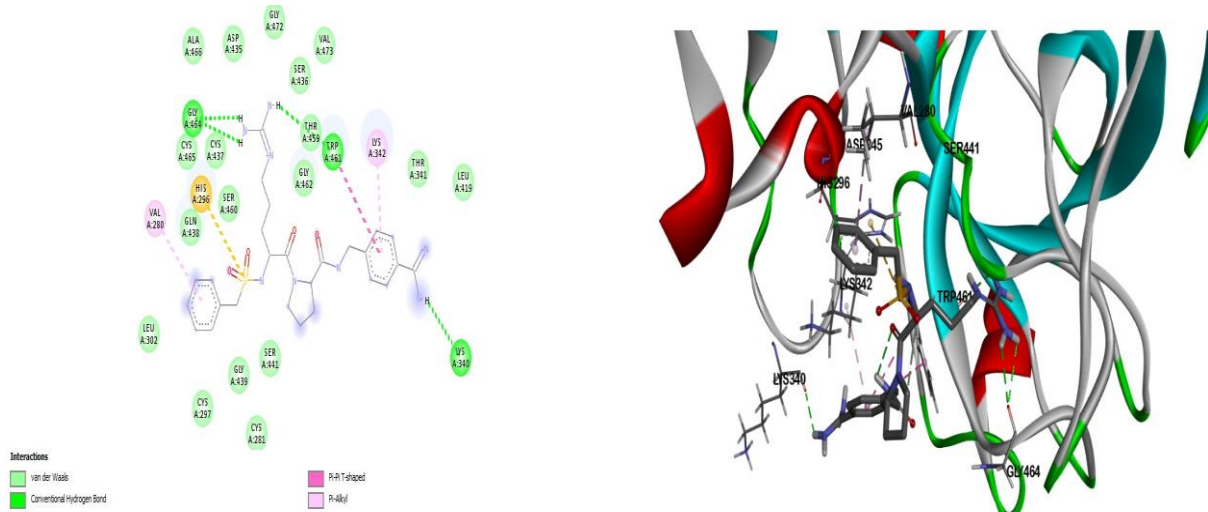


**Supplementary fig.S2: A.** Intracellular, transmembrane & extracellular domains of TMPRSS2 from PROTTER **B.**  $\alpha$ -helix,  $\beta$ -helix, coils and strands information from POLYVIEW **C.** C- & N- TERMINAL from PSIPRED **D.** TMHMM data showing no transmembrane helices and **E.** PSIPRED data for confirmation of TMPRSS2 extracellular domain.

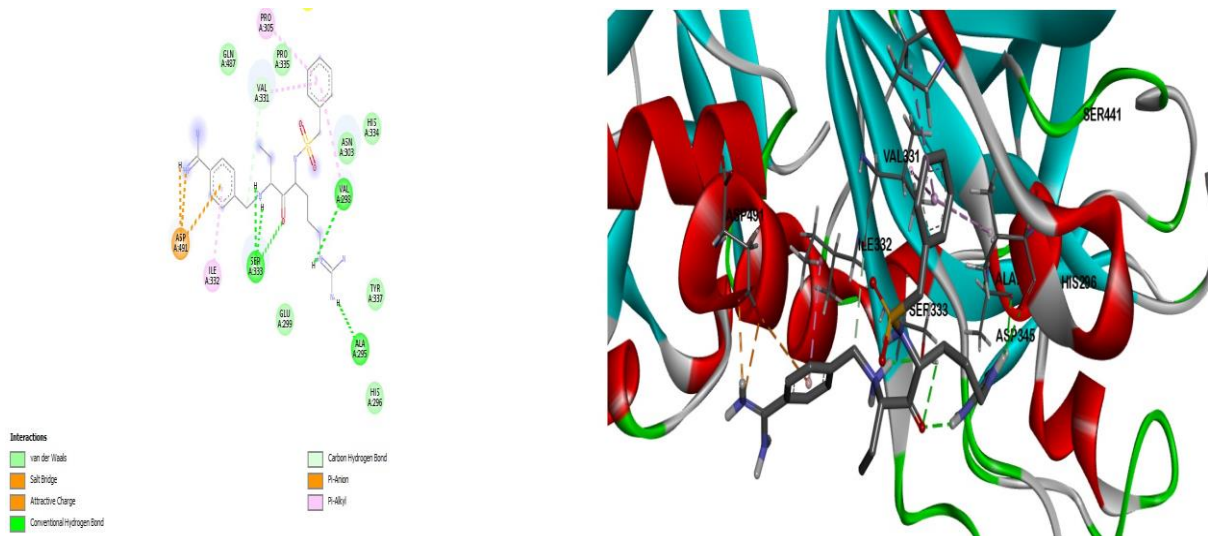


**Supplementary fig.S3:** Quality assessment and structural validation of Tmprss2 homology model. (A) PROCHECK Ramachandran plot (B) ProSA Z-score plot of the overall model

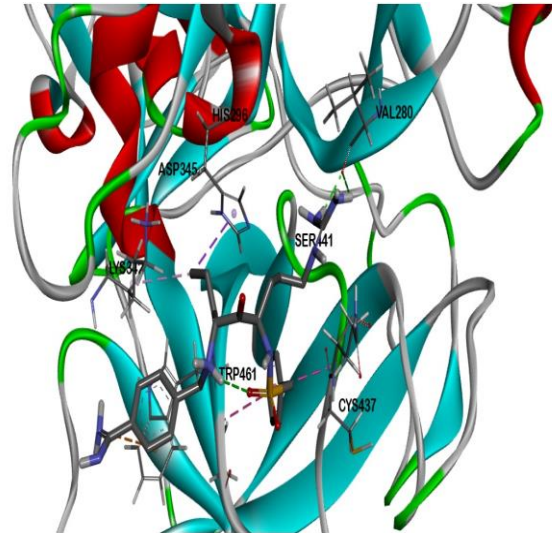
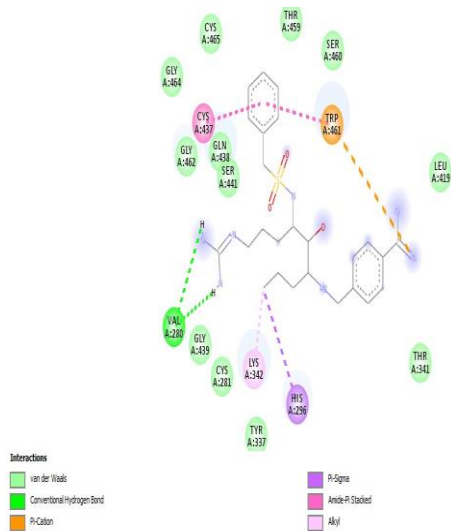
quality (C) ProSA amino acid local energy profile (D) Energy levels of TMPRSS2 by PSIPRED (E) ERRAT plot for TMPRSS2 model. Green bars: properly folded regions, yellow bars: less properly folded regions at 95% confidence level, red bars (none in this model): improperly folded regions at 99% confidence. (F) VERIFY 3D plot for TMPRSS2 model



**Supplementary fig.S4:** The representation of 2D and 3D plots of L1 docked with TMPRSS2 and interactive site residues along with hydrogen bonds as well as hydrophobic bond display.



**Supplementary fig. S5:** The representation of 2D and 3D plots of L2 docked with TMPRSS2 and interactive site residues along with hydrogen bonds as well as hydrophobic bond display.

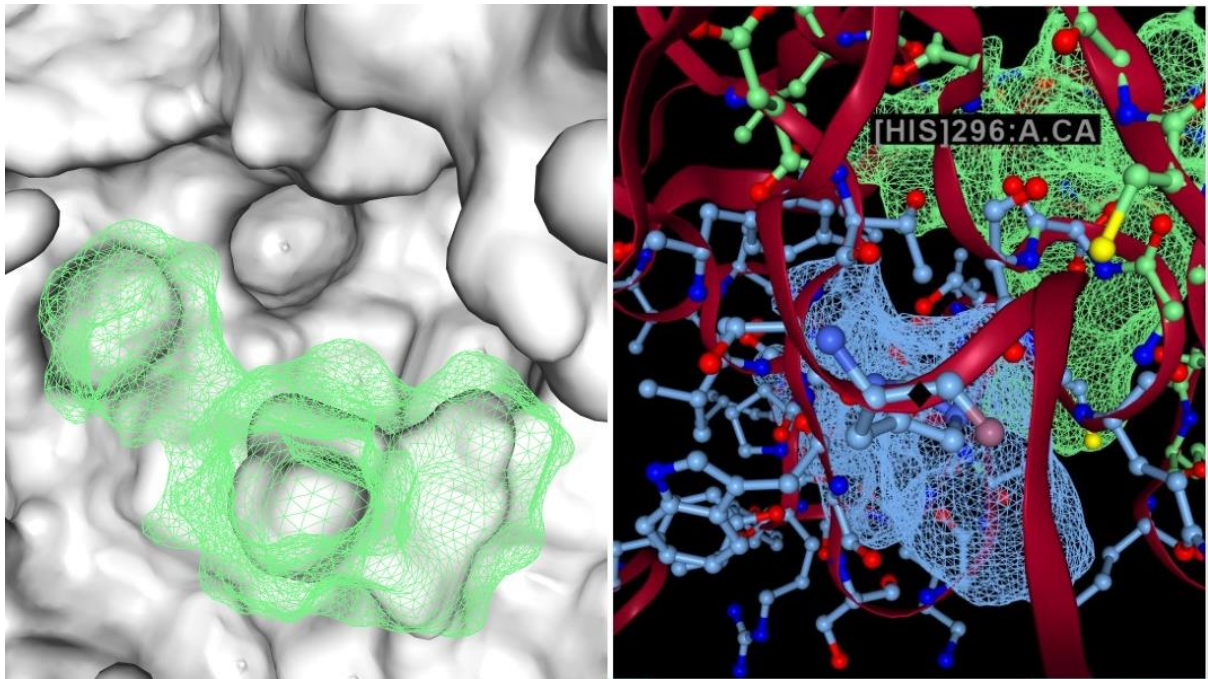


**Supplementary fig.S6:** The representation of 2D and 3D plots of L3 docked with TMPRSS2 and interactive site residues along with hydrogen bonds as well as hydrophobic bond display.

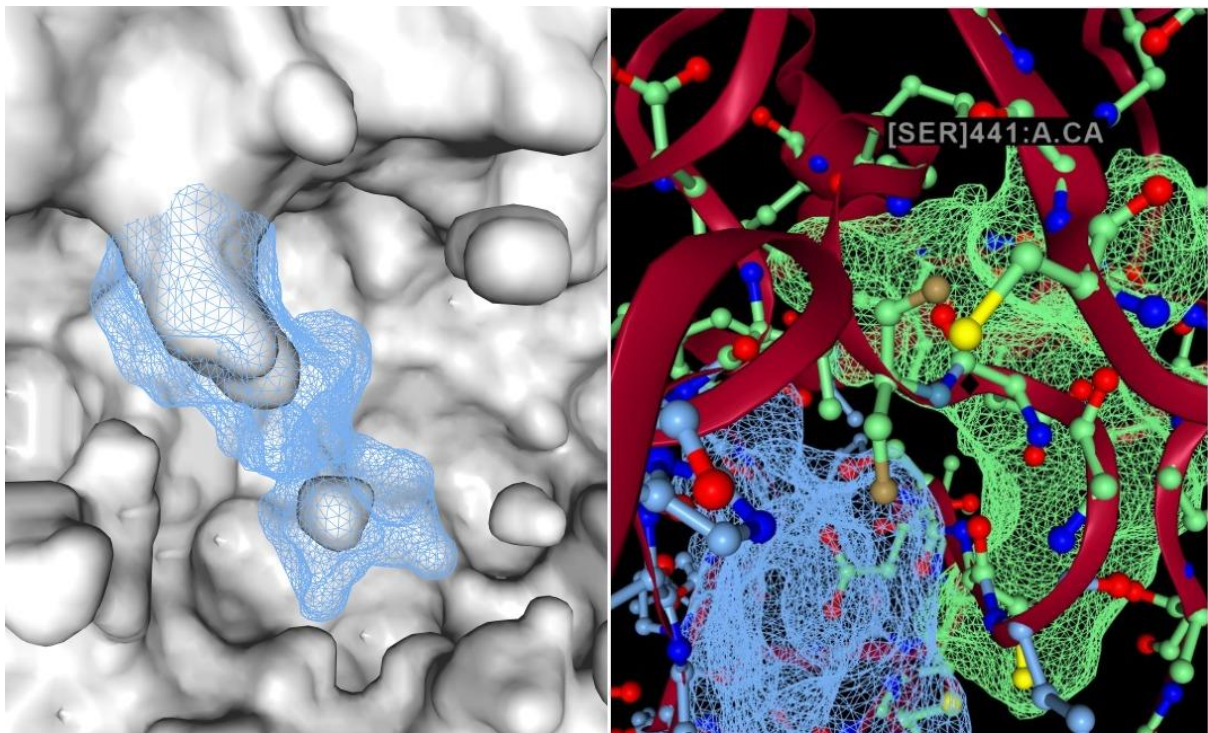
Alignment based on TMPRSS2\_HOMO.pdb (#0) chain A  
File Edit Structure Headers Numberings Tree Info Preferences

|   |                        |                   |                   |                   |                                  |
|---|------------------------|-------------------|-------------------|-------------------|----------------------------------|
| <b>TMPRSS2_HOMO.pdb (#0) chain A</b><br>added | 1<br>1 WKFMGSKCSN      | 11<br>SGIECDSSGT  | 21<br>CINPSNWCDG  | 31<br>VSHCPGGEDG  | 41<br>NRCVRLYGP<br>NRCVRLYGP     |
| <b>TMPRSS2_HOMO.pdb (#0) chain A</b><br>added | 51<br>51 FILLQVYSSQR   | 61<br>KSWHPVCQDD  | 71<br>WNEYGRAAC   | 81<br>RDMGYKNNFY  | 91<br>SSQGIIVDDSG<br>SSQGIIVDDSG |
| <b>TMPRSS2_HOMO.pdb (#0) chain A</b><br>added | 101<br>101 STSFMKLNITS | 111<br>AGNVDIYKKL | 121<br>YHSDACSSKA | 131<br>VVSLRCIACG | 141<br>VNLNSSRQSR<br>VNLNSSRQSR  |
| <b>TMPRSS2_HOMO.pdb (#0) chain A</b><br>added | 151<br>151 IVGGESALPG  | 161<br>AWPWQVSLHV | 171<br>QNVHVCGGSI | 181<br>ITPEWIVTAA | 191<br>HCVEKPLNNP<br>HCVEKPLNNP  |
| <b>TMPRSS2_HOMO.pdb (#0) chain A</b><br>added | 201<br>201 WHWTAFAGIL  | 211<br>RQSFMFYAG  | 221<br>YQVEKVISHP | 231<br>NYDSKTKNND | 241<br>IALMKLQKPL<br>IALMKLQKPL  |
| <b>TMPRSS2_HOMO.pdb (#0) chain A</b><br>added | 251<br>251 TFNDLVKPV   | 261<br>LPNPGMMLQP | 271<br>EQLCWISGWG | 281<br>ATEEKGKTSE | 291<br>VLNAAKVLII<br>VLNAAKVLII  |
| <b>TMPRSS2_HOMO.pdb (#0) chain A</b><br>added | 301<br>301 ETQRCNSRYV  | 311<br>YDNLITPAMI | 321<br>CAGFLQGNVD | 331<br>SCQGDSSGGL | 341<br>VTSKNNIWWL<br>VTSKNNIWWL  |
| <b>TMPRSS2_HOMO.pdb (#0) chain A</b><br>added | 351<br>351 TGDTSWGS    | 361<br>AKAYRPGVYG | 371<br>NVMVFTDWIY | 381<br>RQMRAD     |                                  |

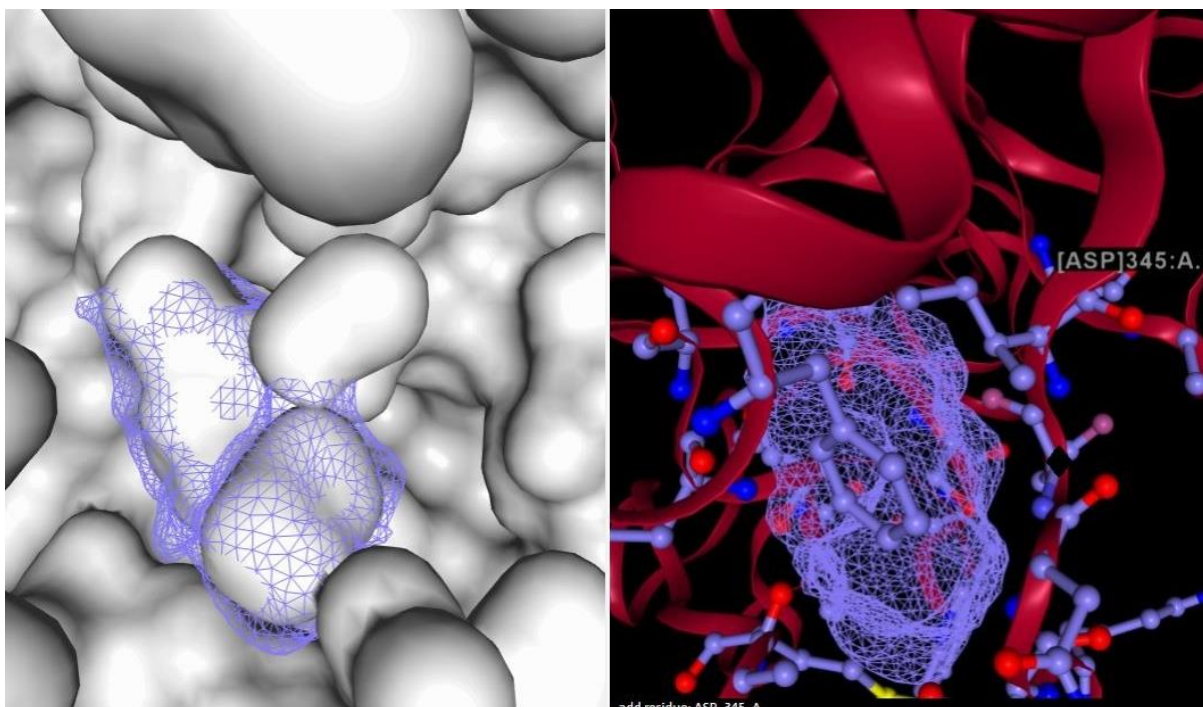
**Supplementary fig. S7:** Sequence alignment of original TMPRSS2 sequence obtained from UniProt with developed Homology protein structure using UCSF Chimera.



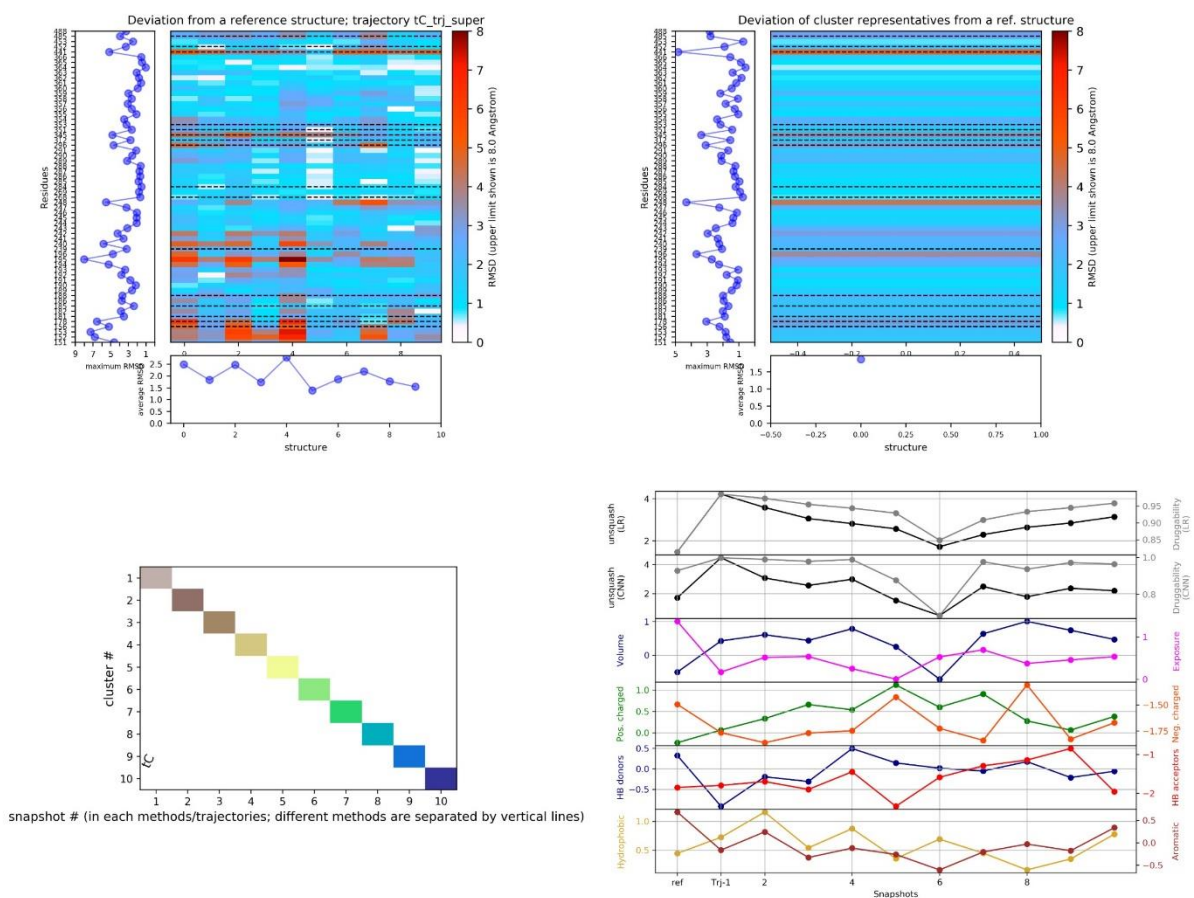
**Supplementary fig.S8.a:** Druggable Pocket 4 containing amino acid His 296 from Active Catalytic triad.



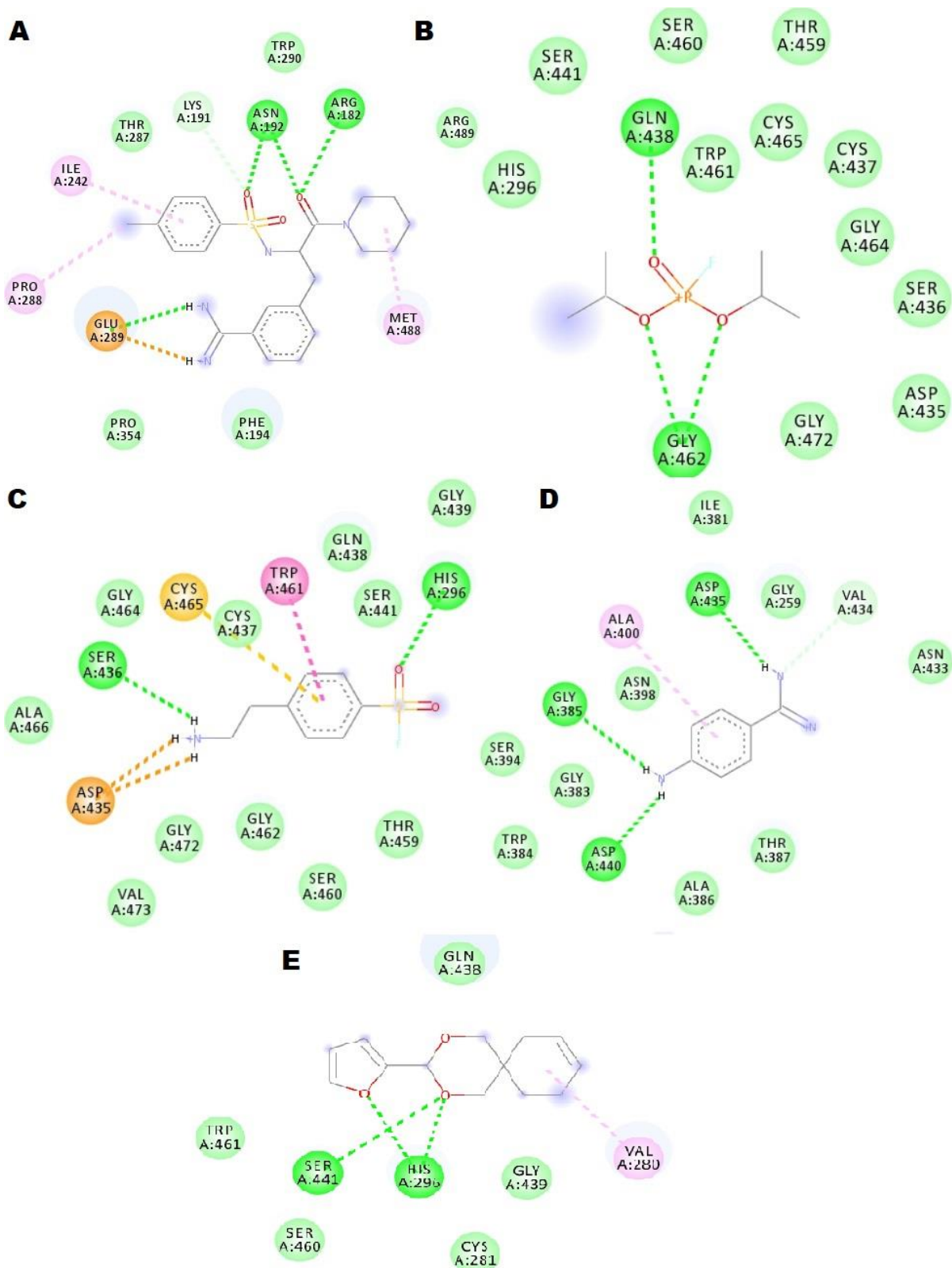
**Supplementary fig.S8.b:** Druggable Pocket 2 and pocket 4 containing Amino acid Ser 441 from Active Catalytic triad.



**Supplementary fig.S8.c:** Druggable Pocket 9 containing Amino acid Asp 345 from Active Catalytic triad.



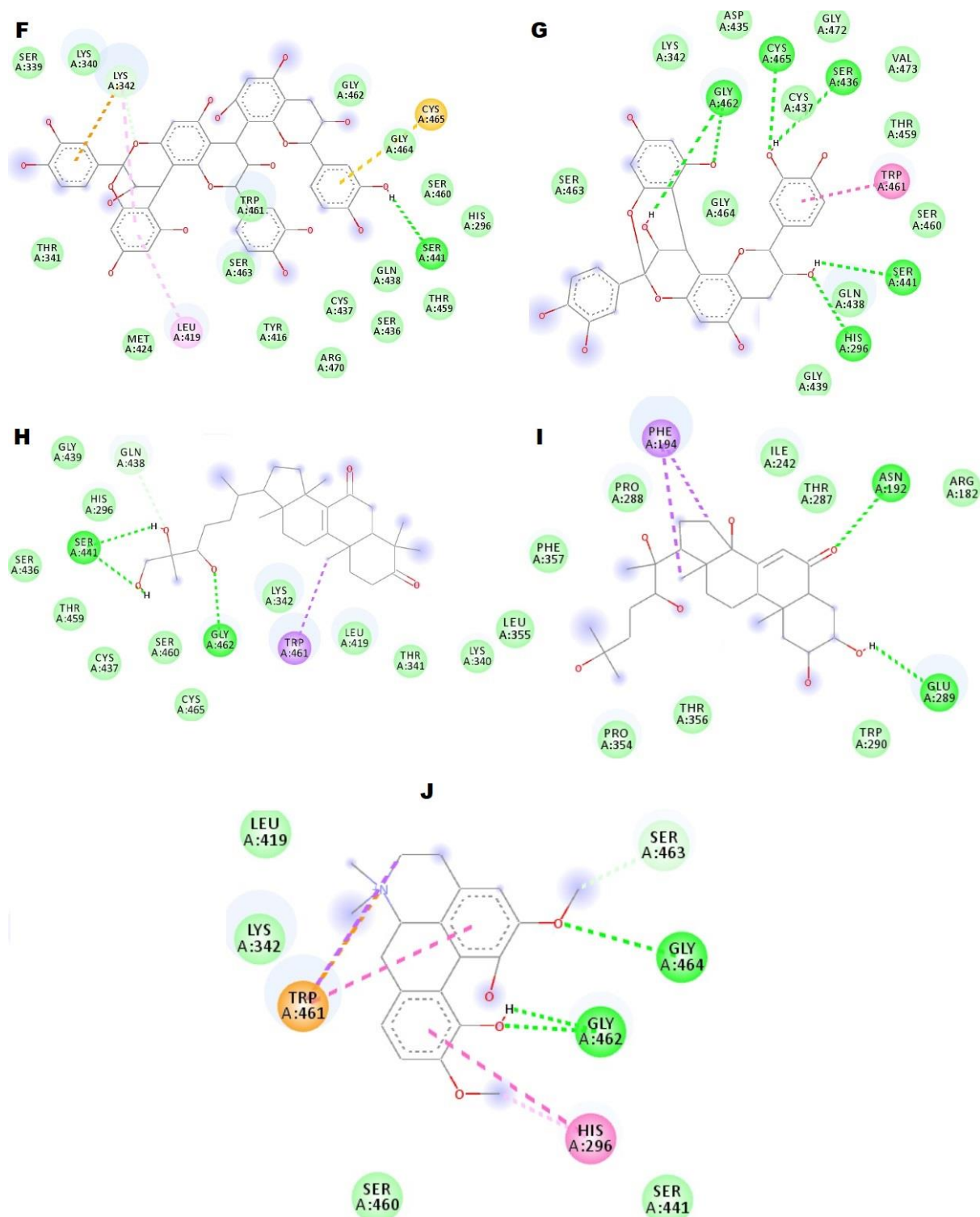
**Supplementary fig.S8.d:** Druggability & other properties of 10 most Druggable Pockets.



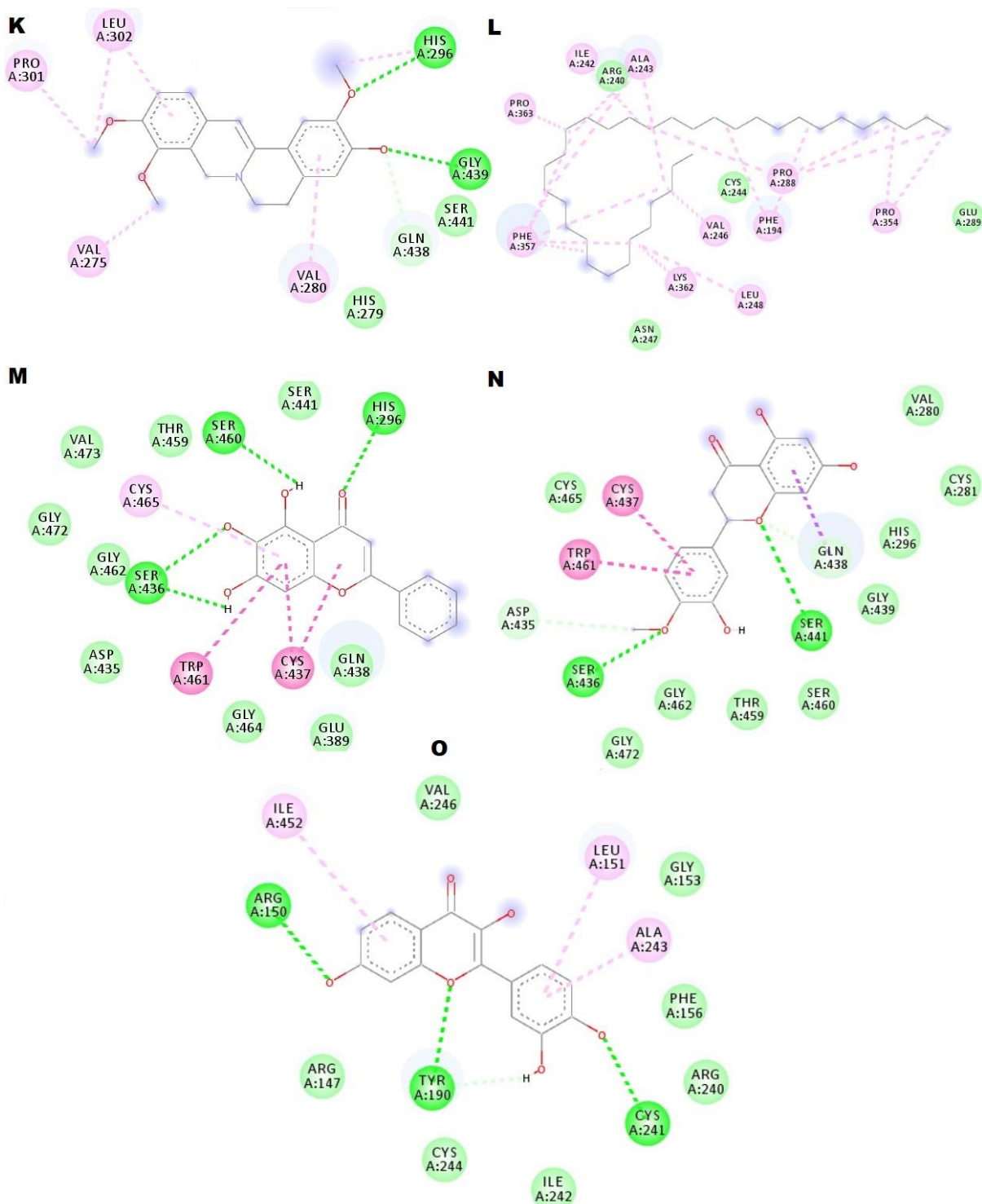
**Supplementary fig.S9.a:** The representation of 2D and 3D plots of (A) 3-TAPAP (B) DIF (C) AEBSF (D) PAB (E) ULINASTATIN docked with TMPRSS2 and interactive site residues along with hydrogen bonds as well as hydrophobic bond display.



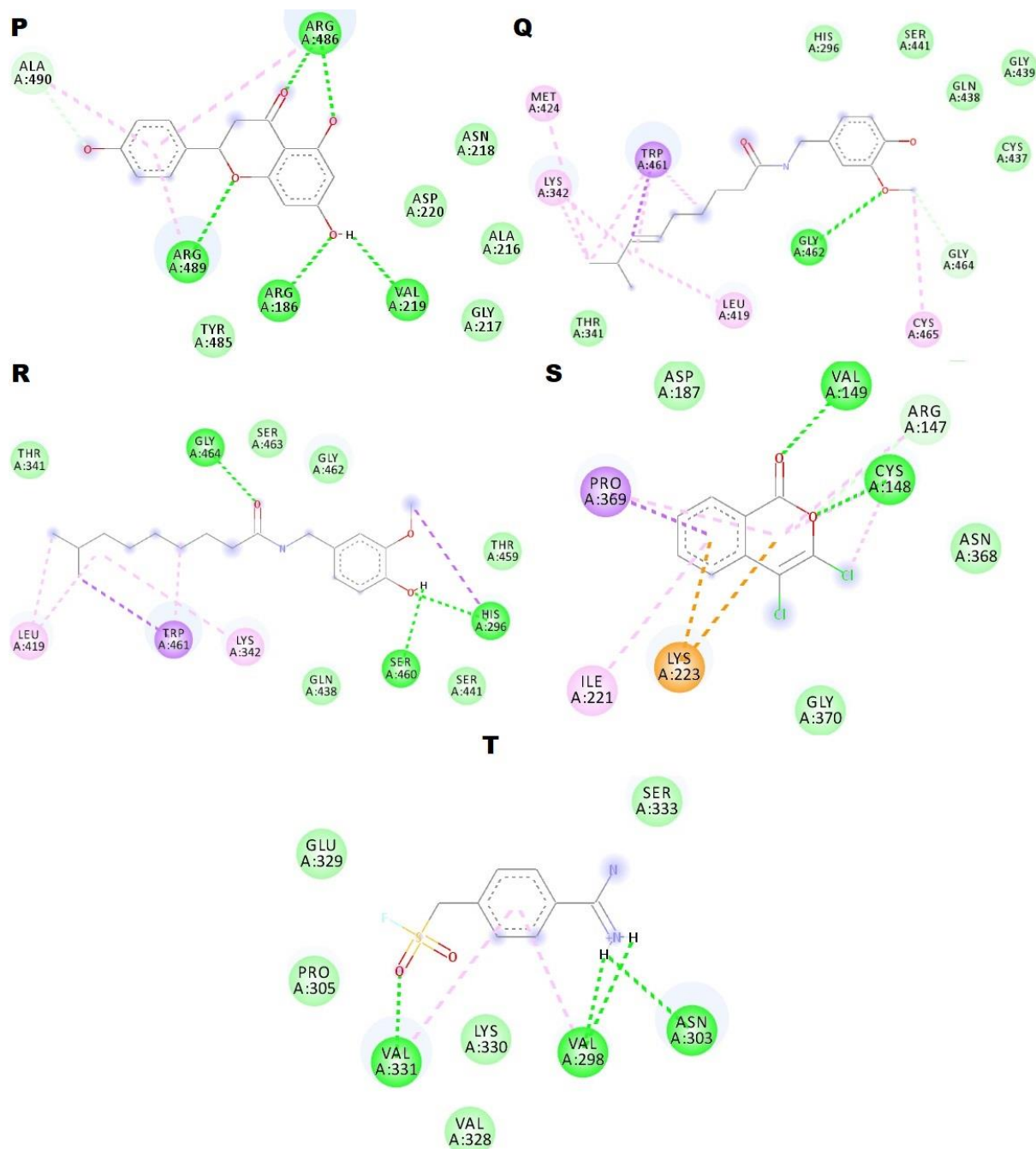
## NATURAL COMPOUNDS



**Supplementary fig. S9.b:** The representation of 2D and 3D plots of (F) AESCULITANNIN B (G) PROANTHOCYANIDIN A2 (H) GANODERIOL D (I) ECDYSTERONE (J) MAGNOFLORINE docked with TMPRSS2 and interactive site residues along with hydrogen bonds as well as hydrophobic bond display.



**Supplementary fig. S9.c:** The representation of 2D and 3D plots of (K) JATRORRHIZINE (L) TETRATRIACONTANE (M) BAICALEIN (N) HESPERETIN (O) FISSETIN docked with TMPRSS2 and interactive site residues along with hydrogen bonds as well as hydrophobic bond display.



**Supplementary fig. S9.d:** The representation of 2D and 3D plots of (P) NARINGENIN (Q) CIS-CAPSAICIN (R) DIHYDROCAPSAICIN (S) 3,4-DICHLOROISOCOUMARIN (T) APMSF docked with TMPRSS2 and interactive site residues along with hydrogen bonds as well as hydrophobic bond display.