

A hybrid ligand and structure-based virtual screening of NCI compound library identifies potential SAPT1 inhibitors

Suat SARI^{1*} , Nisha VALAND² , Umakhanth Venkatraman GIRIJA^{2*} 

¹ Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Hacettepe University, Sıhhiye 06100 Ankara, Turkey.

² Leicester School of Allied Health Sciences, Faculty of Health and Life Sciences, De Montfort University, Leicester LE1 9BH, United Kingdom.

* Corresponding Author. E-mail: suat.sari@hacettepe.edu.tr (S.S.); Tel. +90-312-305 18 72.

Received: 6 December 2021 / Accepted: 20 January 2022

ABSTRACT: Secretory aspartate proteases (SAPs) is a key virulence factor of *Candida* spp. enabling adherence to and invasion of host tissues through breakdown of host proteins related to immunological and structural defenses, making them potential drug targets for drug-resistant mycoses, especially where the available therapies fail. To date, no SAP inhibitors for *Candida tropicalis*, one of the top five most common species isolated in *Candida*-related mycoses, has been reported. In this study, we report first-time identification of a set of potential *C. tropicalis* SAP1 (SAPT1) inhibitors through a hybrid ligand-and structure-based virtual screening of National Cancer Institute (NCI) library compounds. The NCI library was refined by filtering off non-druglike molecules. Referring to a known *C. albicans* SAP inhibitor, a similarity search was performed for the refined library, in addition to a pharmacophore screen using a model of the ligand-receptor interactions between a peptide substrate and SAPT1 obtained from crystallographic data. The compounds selected from these screens were subject to molecular docking to the SAPT1 active site and the top-scoring ligand-receptor complexes were further included in MM-GBSA calculations to optimize the predicted binding affinities. Finally, the selected 16 compounds, which were confirmed to make key interactions with the catalytic residues, were *in silico* evaluated and found eligible for certain pharmacokinetic properties. As a future prospect, obtaining these virtual hits and testing them *in vitro* against SAPT1 could validate the virtual screening process and yield the first small-molecule inhibitors of SAPT.

KEYWORDS: secretory aspartate protease; *Candida tropicalis*; virtual screening; shape screening; pharmacophore modeling; molecular docking; MM-GBSA

1. INTRODUCTION

Candida spp. are opportunistic fungi and the leading cause of superficial, invasive, and systemic fungal infections. Emergence of intrinsically drug-resistant species, as well as drug-resistant isolates of some susceptible *Candida* spp., has caused a surge in severe and deadly fungal infections in recent years. *C. albicans* is known to be the most common species isolated in treatable and drug-resistant *Candida* infections. However, there is an increased emergence of drug resistance in other species like *C. parapsilosis*, *C. krusei*, *C. glabrata*, *C. auris*, and *C. tropicalis* [1].

Secretory aspartate proteases (SAPs) are catalytic proteases that break down peptide bonds. The *Candida* SAP catalytic site features an activated water molecule bound to two aspartate side chains [2]. SAPs of different *Candida* spp. share high sequence identity and a highly conserved active site [3]. *Candida* spp. express different SAP genes encoding isoenzymes which show diverse function and distribution in different infections. SAP is an important virulence factor of *Candida* spp. enabling adherence to and invasion of host tissues by the fungi through breakdown of proteins related to immunological and structural defenses, therefore it has drawn great attention as a potential drug target for drug-resistant mycoses, especially where the available therapies fail [4]. *C. tropicalis* is among the top five most common species isolated in *Candida* infections. The development of antifungal drug resistance in *C. tropicalis* with different mechanisms is alarming and needs diverse treatment

Sari S, Valand N, Girija UV. A hybrid ligand and structure-based virtual screening of NCI compound library identifies potential SAPT1 inhibitors. J Res Pharm. 2022; 26(3): 617-624.

strategies [5]. Although promising inhibitors have been identified against *C. albicans* SAPs [4, 6-11], there is no known small-molecule inhibitor of *C. tropicalis* SAP (SAPT).

In this study, we identified 16 potential inhibitors of SAPT1 through a ligand- and structure-based hybrid virtual screening of the National Cancer Institute (NCI) compound library, which is a collection of developmental compounds submitted to the NCI from all over the world for anticancer screening. The campaign featured ADMET (absorption, distribution, metabolism, excretion, toxicity) filtering, similarity-based screening, pharmacophore screening based on ligand-receptor interactions, molecular docking and MM-GBSA calculations (Figure 1). The identified virtual hits may hold promise against SAPT1 and be developed into leads against drug-resistant *Candida* infections.

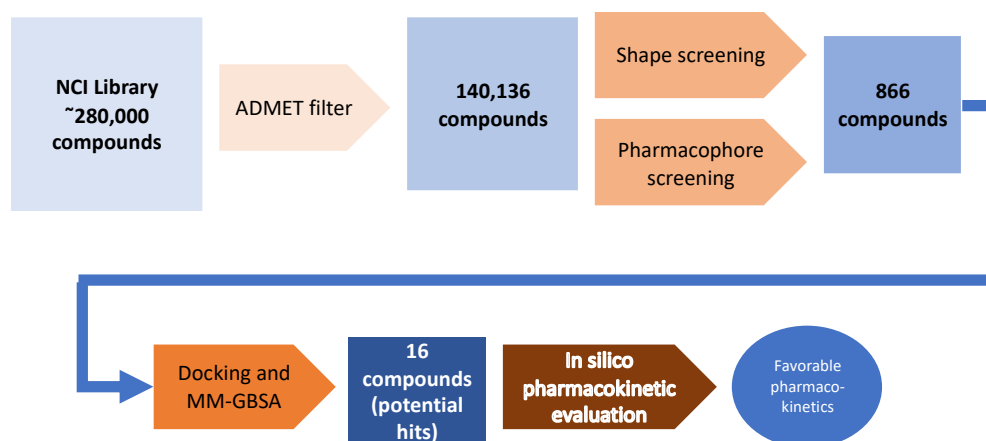


Figure 1. The virtual screening pipeline of the current study.

2. RESULTS

The NCI compound library, which comprises of more than 280,000 compounds, was filtered (ADMET filter) using QikProp (2021-2, Schrödinger LLC, New York, NY) software and the molecular descriptors relevant to druglikeness, namely MW (molecular weight), RB (number of rotatable bonds), HD (hydrogen bond donor count), HA (hydrogen bond acceptor count), logP, and PSA (polar surface area), to discard non-druglike molecules, leaving 140,136 eligible compounds. The cutoff values of the descriptors to identify outliers were defined by QikProp (Table 1).

Table 1. Molecular descriptors used in ADMET filter.

Descriptor	Description	Ideal values range
MW	Molecular weight (Da)	130-725
RB	Number of non-trivial and non-hindered rotatable bonds	0-15
HD	Estimated number of hydrogen bond donors	0-6
HA	Estimated number of hydrogen bond acceptors	2-20
logP	Predicted octanol/water partition coefficient	-2.0-6.5
PSA	Polar surface area: Van der Waals surface area (Å ²) of polar nitrogen and oxygen atoms and carbonyl carbon atoms	7-200

The similarity screen was performed with the eligible compounds by describing each structure as a collection of pharmacophore features. Since there is no known small molecule inhibitor of SAPT1 and the SAPs share high level of homology among *Candida* spp., especially for the active site, a highly potent SAP2 inhibitor (92% inhibition at 10 µM) [11] was used as a query compound (Figure 2). 783 compounds obtained similarity scores higher than 0.55, which were selected for molecular docking step. Seven of the 16 compounds identified at the end of this study came from the shape screening (Table 2).

Table 2. Virtual screening scores of the 16 compounds identified at the end of this study.

Compound	Similarity score ^a	Phase screen score ^b	Docking score (kcal/mol) ^c	ΔG bind score (kcal/mol) ^d
626866	0.57	-	-5.9	-42.0
666251	-	1.61	-5.4	-38.7
143489	0.55	-	-5.1	-37.4
41914	0.60	-	-5.4	-36.1
723978	0.62	-	-5.4	-35.0
290805	-	1.60	-5.8	-33.2
631814	0.55	-	-5.4	-32.7
46611	-	1.55	-5.6	-32.5
668036	-	1.57	-5.6	-32.3
187723	0.58	-	-5.5	-32.0
263627	-	1.52	-6.3	-29.9
203449	-	1.58	-5.8	-28.7
86286	-	1.59	-5.1	-28.2
86196	-	1.59	-6.0	-26.8
55841	-	1.58	-5.1	-23.5
666388	0.57	-	-5.4	-22.3

^a Obtained from the shape screen. Compounds with a similarity score ≥ 0.55 were selected (cells with a dash indicate that the regarding compound obtained a value lower than the cutoff).

^b Obtained from the pharmacophore screening. Compounds with a Phase screen score ≥ 1.5 were selected (cells with a dash indicate that the regarding compound obtained a value lower than the cutoff).

^c Obtained from the molecular docking

^d Free energy of binding obtained from the MM-GBSA calculations of the selected ligand-receptor complexes
Dash means no data

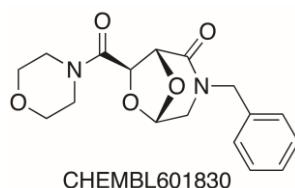


Figure 2. The query compound used in shape screening.

The only pdb structure for SAPT1 includes a peptide substrate in its active site [12], which makes a wide range of polar interactions with the receptor (Figure 3A). A pharmacophore model was created using the peptide structure and these interactions comprising a positive ionizable group (P15), two hydrogen donors (D7 and D11), two hydrogen acceptors (A3 and A4), and a negative ionizable group (N14), as well as excluded volumes (Figure 3B). 83 compounds scoring higher than 1.5 according to the Phase screen scores were selected for the next step. The compound with the best Phase screen score among the final 16 virtual hits, compound 666251 (Table 2), was observed to align well with the pharmacophore hypothesis (Figure 3C).

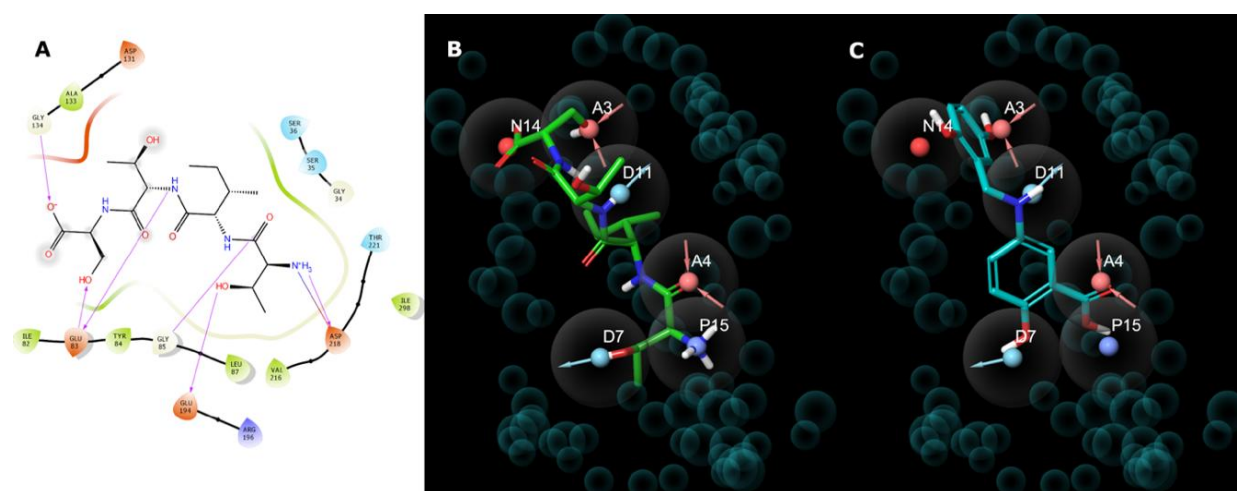


Figure 3. Interactions of the peptide in SAPT1 active site (A), the pharmacophore model overlaid with the peptide (B) and compound 666251, the highest scoring compound among the identified 16 virtual hits (C).

From the shape and pharmacophore screens a total of 866 compounds were selected for the molecular docking step, where the SAPT1 crystal structure was used. Molecular docking predicts the preferred binding mode of a ligand-receptor complex and the affinity of this complex via scoring functions. As molecular docking ignores solvation effects continuum solvation methods like MM-GBSA can predict better ligand-receptor affinity [13]. 35 compounds with a docking score lower than -5.0 kcal/mol were identified and MM-GBSA calculation were run for the selected ligand-receptor complex of each compound. Compounds with a free energy of binding, referred to as " ΔG_{bind} ", found lower than -22.0 kcal/mol from these calculations were visually evaluated for interactions with the receptor, resulting in 16 potential hits (Figure 4).

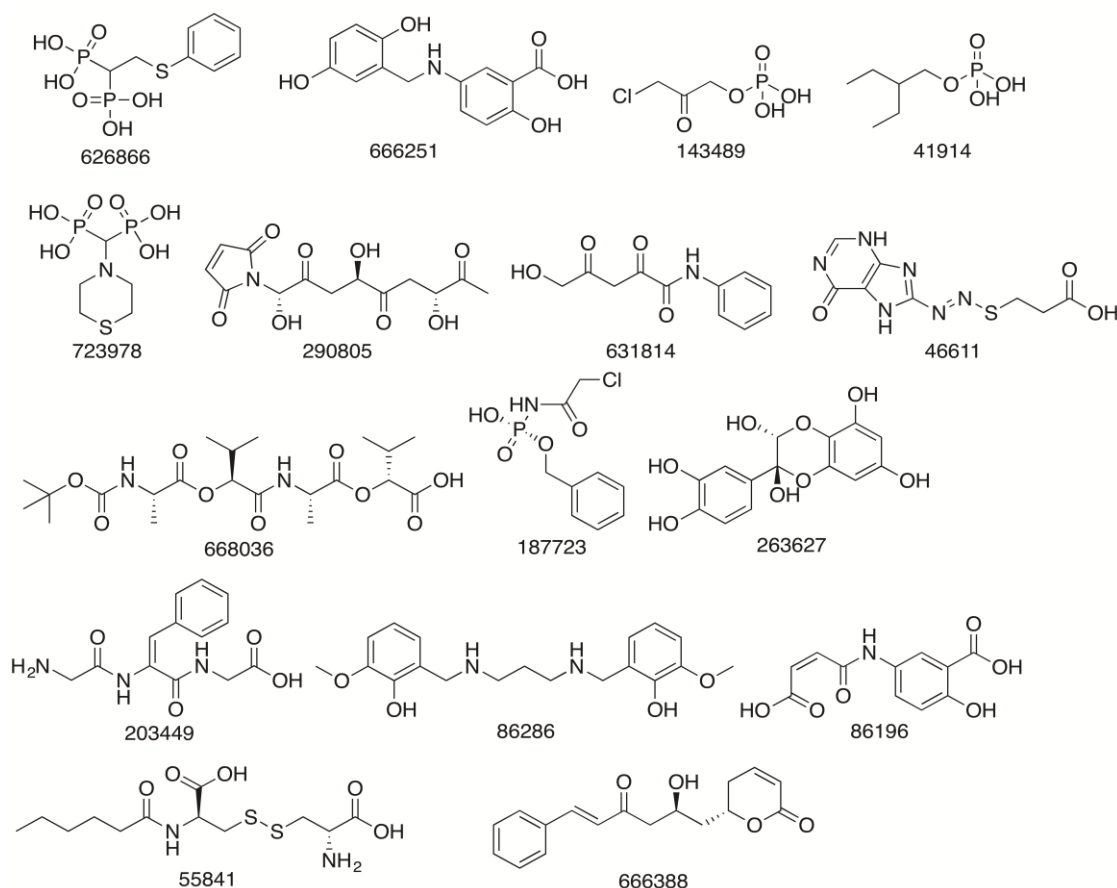


Figure 4. Potential SAPT1 inhibitors selected via virtual screening.

A number of pharmacokinetic parameters of the selected compounds were evaluated *in silico* (Table 3). The compounds were predicted water soluble according to the three solubility models (ESOL, Ali, and Silicos-IT) with the exception of 86286, which was found poorly soluble by Silicos-IT. High gastrointestinal (GI) absorption was predicted for ten of the compounds and only 666388 was not blood-brain barrier (BBB) permeant. Except 668036 and 86286, nine of the compounds were substrate of P-glycoprotein (Pgp), an efflux pump responsible for drug resistance due to its ability to pump drug molecules from the intestinal epithelium back into the intestinal lumen or from liver cells into bile ducts, etc. [14]. The compounds were predicted not to inhibit common CYPs responsible for drug-drug interactions.

Comp.	Aqueous solubility			GI absorp.	BBB permeant	Pgp substrate	CYP inhibitor				
	ESOL	Ali	Silicos- IT				CYP1A2	CYP2C19	CYP2C9	CYP2D6	CYP3A4
626866	Very soluble	Very soluble	Soluble	Low	No	No	No	No	No	No	No
666251	Soluble	Moderately soluble	Soluble	High	No	No	No	No	No	No	No
143489	Very soluble	Very soluble	Soluble	High	No	No	No	No	No	No	No
41914	Very soluble	Very soluble	Soluble	High	No	No	No	No	No	No	No
723978	Highly soluble	Highly soluble	Soluble	Low	No	No	No	No	No	No	No
290805	Highly soluble	Highly soluble	Soluble	Low	No	No	No	No	No	No	No
631814	Very soluble	Very soluble	Soluble	High	No	No	No	No	No	No	No
46611	Very soluble	Soluble	Soluble	Low	No	No	No	No	No	No	No
668036	Soluble	Poorly soluble	Soluble	Low	No	Yes	No	No	No	No	No
187723	Very soluble	Very soluble	Soluble	High	No	No	No	No	No	No	No
263627	Soluble	Soluble	Soluble	High	No	No	No	No	No	No	No
203449	Highly soluble	Highly soluble	Soluble	High	No	No	No	No	No	No	No
86286	Soluble	Moderately soluble	Poorly soluble	High	No	Yes	No	No	No	Yes	No
86196	Soluble	Soluble	Soluble	High	No	No	No	No	No	No	No
55841	Highly soluble	Very soluble	Soluble	Low	No	No	No	No	No	No	No
666388	Soluble	Soluble	Soluble	High	Yes	No	Yes	No	No	No	No

Table 3. Calculated pharmacokinetic parameters for the selected compounds.

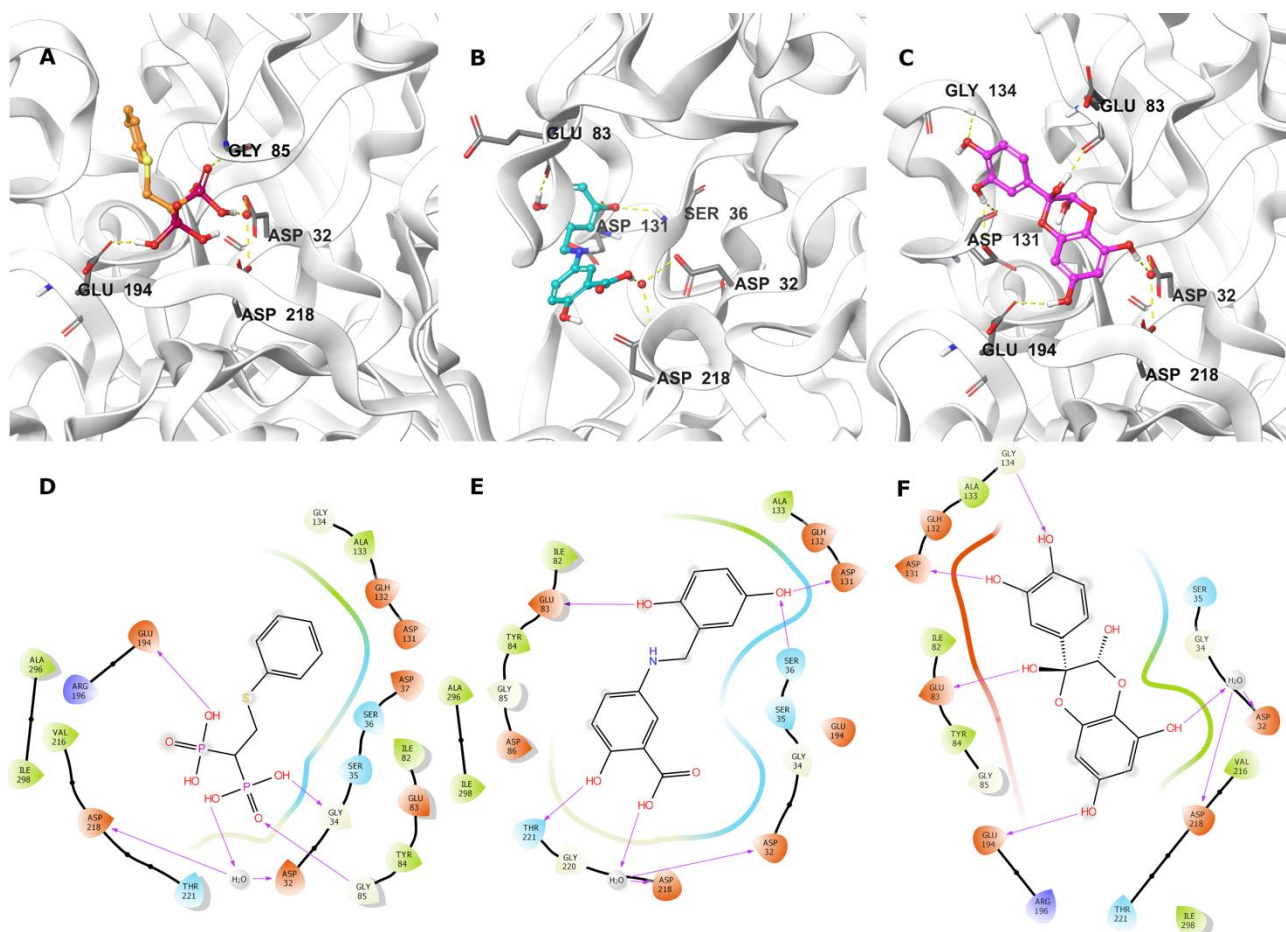
3. DISCUSSION

The ADMET filter eliminated almost half of the library. The druglikeness concept has a widespread use in virtual screening since it is helpful to eliminate such compounds with little or no prospect of being drug due to pharmacokinetic and toxicity issues in the early stages of drug discovery so that a great deal of time and cost is spared.

Shape similarity is a fast and efficient method to identify "similar, which will show similar if not better effects". The performance of shape similarity methods is usually praised although a consensus that combination of different ligand- and structure-based approaches are usually deemed more successful is present [15, 16]. Thus, the similarity and pharmacophore screens in our study is expected to extract potential hits from the library for the following steps.

The active site of SAPT1, like other SAPs, is dominated by polar residues and features two catalytic aspartates (Asp32 and Asp218) and a water molecule making triangle-shaped H bonds. The co-crystallized peptide makes polar contacts with these aspartates through its ammonium of the N-terminal threonine, which was included as a positively ionizable feature in the pharmacophore model (Figure 3), although the water-aspartates triangle can be contacted by non-ionized polar groups (Figure 5). Indeed, this feature is represented by hydroxyls of different kinds as seen in the case of the top scoring compounds in the pharmacophore screen and molecular docking (Figure 4). The pharmacophore model and the interactions of the peptide-SAPT1 complex indicate importance of contacts with other residues such as Glu83, Gly85, Gly134, and Glu194. Among the selected 16 hits, 626866 and 666251 were the best scoring compounds regarding ΔG bind score,

The *in silico* pharmacokinetic evaluation of the selected 16 compounds suggested favorable attributes for certain properties like aqueous solubility, GI absorption, avoiding Pgp and CYPs. Some of the compounds were predicted to pass BBB, which is a crucial property in the case of fungal infections of the CNS but could otherwise have implications for the CNS related conditions.



4. CONCLUSION

622

5. MATERIALS AND METHODS

The NCI compound library was downloaded (access date: 13.01.2021) and prepared using MacroModel (2021-2, Schrödinger LLC, New York, NY) and LigPrep (2021-2, Schrödinger LLC, New York, NY) to generate optimized 3D model of each compound according to the OPLS4 forcefield parameters [17]. Molecular descriptors were calculated using QikProp, ADMET and pharmacokinetic parameters were calculated using SwissADME server (www.swissadme.ch) [18]. The query compound for shape screening was modelled as defined for the NCI library above and shape screening was performed using the Shape Screening panel of Maestro (2021-2, Schrödinger LLC, New York, NY) at typed pharmacophore mode [19]. Crystal structure of SAPT1 in complex with a peptide substrate (PDB ID: 1J71 [12], resolution: 1.80 Å) was downloaded from the RCSB Protein Data Bank (www.rcsb.org) and a pharmacophore model based on the interactions of this complex was generated automatically and the selected library was screened against the model using Phase (2021-2, Schrödinger LLC, New York, NY) [20]. SAPT1 structure was prepared for docking using the Protein preparation Wizard of Maestro (2021-2, Schrödinger LLC, New York, NY) [21]. At this step redundant molecules were removed, hydrogen atoms were added, partial charges and bond orders were assigned, ionization states of the residues and H bonds were set. The receptor active site grids map was generated using the Receptor Grid Generator panel of Maestro by setting the central coordinates as 23.78 27.55 20.79 and the volume 15,625 Å³. Molecular docking was performed using Glide (2021-2, Schrödinger LLC, New York, NY) at extra precision mode with 50 runs per ligand [22]. The MM-GBSA calculations for the selected ligand-receptor complex were run using Prime (2021-2, Schrödinger LLC, New York, NY) with VSGB solvation model according to OPLS4 forcefield parameters [23].

Author contributions: Concept – S.S., U.V.G.; Design – S.S., U.V.G.; Supervision – S.S., U.V.G.; Resources – S.S., U.V.G.; Materials – S.S., U.V.G.; Data Collection and/or Processing – S.S., N.V., U.V.G.; Analysis and/or Interpretation – S.S., N.V., U.V.G.; Literature Search – S.S., N.V.; Writing – S.S., N.V., U.V.G.; Critical Reviews – S.S., N.V., U.V.G.

Conflict of interest statement: The authors declared no conflict of interest.

REFERENCES

- [1] Sanglard D. Emerging Threats in Antifungal-Resistant Fungal Pathogens. *Front Med (Lausanne)*. 2016; 3: 11. [\[CrossRef\]](#)
- [2] Chen KC, Tang J. Amino acid sequence around the epoxide-reactive residues in pepsin. *J Biol Chem*. 1972; 247(8): 2566-2574. [\[CrossRef\]](#)
- [3] Mandujano-Gonzalez V, Villa-Tanaca L, Anducho-Reyes MA, Mercado-Flores Y. Secreted fungal aspartic proteases: A review. *Rev Iberoam Micol*. 2016; 33(2): 76-82. [\[CrossRef\]](#)
- [4] Pichova I, Pavlickova L, Dostal J, Dolejsi E, Hruskova-Heidingsfeldova O, Weber J, Ruml T, Soucek M. Secreted aspartic proteases of *Candida albicans*, *Candida tropicalis*, *Candida parapsilosis* and *Candida lusitanae*. Inhibition with peptidomimetic inhibitors. *Eur J Biochem*. 2001; 268(9): 2669-2677. [\[CrossRef\]](#)
- [5] Forastiero A, Mesa-Arango AC, Alastruey-Izquierdo A, Alcazar-Fuoli L, Bernal-Martinez L, Pelaez T, Lopez JF, Grimalt JO, Gomez-Lopez A, Cuesta I, Zaragoza O, Mellado E. *Candida tropicalis* antifungal cross-resistance is related to different azole target (Erg11p) modifications. *Antimicrob Agents Chemother*. 2013; 57(10): 4769-4781. [\[CrossRef\]](#)
- [6] Buchold C, Hemberger Y, Heindl C, Welker A, Degel B, Pfeuffer T, Staib P, Schneider S, Rosenthal PJ, Gut J, Morschhäuser J, Bringmann G, Schirmeister T. New cis-configured aziridine-2-carboxylates as aspartic acid protease inhibitors. *ChemMedChem*. 2011; 6(1): 141-152. [\[CrossRef\]](#)
- [7] Degel B, Staib P, Rohrer S, Scheiber J, Martina E, Buchold C, Baumann K, Morschhäuser J, Schirmeister T. Cis-Configured aziridines are new pseudo-irreversible dual-mode inhibitors of *Candida albicans* secreted aspartic protease 2. *ChemMedChem*. 2008; 3(2): 302-315. [\[CrossRef\]](#)

- [8] Korting HC, Schaller M, Eder G, Hamm G, Bohmer U, Hube B. Effects of the human immunodeficiency virus (HIV) proteinase inhibitors saquinavir and indinavir on in vitro activities of secreted aspartyl proteinases of *Candida albicans* isolates from HIV-infected patients. *Antimicrob Agents Chemother*. 1999; 43(8): 2038-2042. [CrossRef]
- [9] Tossi A, Benedetti F, Norbedo S, Skrbec D, Berti F, Romeo D. Small hydroxyethylene-based peptidomimetics inhibiting both HIV-1 and *C. albicans* aspartic proteases. *Bioorg Med Chem*. 2003; 11(22): 4719-4727. [CrossRef]
- [10] Calugi C, Trabocchi A, De Bernardis F, Arancia S, Navarra P, Cauda R, Cassone A, Guarna A. Bicyclic peptidomimetics targeting secreted aspartic protease 2 (SAP2) from *Candida albicans* reveal a constrained inhibitory chemotype. *Bioorg Med Chem*. 2012; 20(24): 7206-7213. [CrossRef]
- [11] Trabocchi A, Mannino C, Machetti F, De Bernardis F, Arancia S, Cauda R, Cassone A, Guarna A. Identification of inhibitors of drug-resistant *Candida albicans* strains from a library of bicyclic peptidomimetic compounds. *J Med Chem*. 2010; 53(6): 2502-2509. [CrossRef]
- [12] Symersky J, Monod M, Foundling SI. High-resolution structure of the extracellular aspartic proteinase from *Candida tropicalis* yeast. *Biochemistry*. 1997; 36(42): 12700-12710. [CrossRef]
- [13] Genheden S, Ryde U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin Drug Discov*. 2015; 10(5): 449-461. [CrossRef]
- [14] Schubert S, Barker KS, Znaidi S, Schneider S, Dierolf F, Dunkel N, Aid M, Boucher G, Rogers PD, Raymond M, Morschhäuser J. Regulation of efflux pump expression and drug resistance by the transcription factors Mrr1, Upc2, and Cap1 in *Candida albicans*. *Antimicrob Agents Chemother*. 2011; 55(5): 2212-2223. [CrossRef]
- [15] Kruger DM, Evers A. Comparison of structure- and ligand-based virtual screening protocols considering hit list complementarity and enrichment factors. *ChemMedChem*. 2010; 5(1): 148-158. [CrossRef]
- [16] Sastry GM, Inakollu VS, Sherman W. Boosting virtual screening enrichments with data fusion: coalescing hits from two-dimensional fingerprints, shape, and docking. *J Chem Inf Model*. 2013; 53(7): 1531-1542. [CrossRef]
- [17] Sánchez-Portocarrero J, Pérez-Cecilia E, Corral O, Romero-Vivas J, Picazo JJ. The central nervous system and infection by *Candida* species. *Diagn Microbiol Infect Dis*. 2000; 37(3): 169-179. [CrossRef]
- [18] Lu C, Wu C, Ghoreishi D, Chen W, Wang L, Damm W, Ross GA, Dahlgren MK, Russell E, Von Bargen CD, Abel R, Friesner RA, Harder ED. OPLS4: Improving Force Field Accuracy on Challenging Regimes of Chemical Space. *J Chem Theory Comput*. 2021; 17(7): 4291-300. [CrossRef]
- [19] Daina A, Michielin O, Zoete V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci Rep*. 2017; 7:42717. [CrossRef]
- [20] Sastry GM, Dixon SL, Sherman W. Rapid shape-based ligand alignment and virtual screening method based on atom/feature-pair similarities and volume overlap scoring. *J Chem Inf Model*. 2011; 51(10): 2455-2466. [CrossRef]
- [21] Dixon SL, Smondyrev AM, Knoll EH, Rao SN, Shaw DE, Friesner RA. PHASE: a new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology and preliminary results. *J Comput Aided Mol Des*. 2006; 20(10-11): 647-671. [CrossRef]
- [22] Sastry GM, Adzhigirey M, Day T, Annabhimoju R, Sherman W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J Comput Aided Mol Des*. 2013; 27(3): 221-234. [CrossRef]
- [23] Friesner RA, Murphy RB, Repasky MP, Frye LL, Greenwood JR, Halgren TA, Sanschagrin PC, Mainz DT. Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J Med Chem*. 2006; 49(21): 6177-6196. [CrossRef]
- [24] Jacobson MP, Friesner RA, Xiang Z, Honig B. On the role of the crystal environment in determining protein side-chain conformations. *J Mol Biol*. 2002; 320(3): 597-608. [CrossRef]

This is an open access article which is publicly available on our journal's website under Institutional Repository at <http://dspace.marmara.edu.tr>