Structure-based Discovery of Potential Inhibitors Targeting Sodium-bile Acid Co-transporter of Carcinogenic Liver Fluke Clonorchis Sinensis

Won Gi Yoo

Chung-Ang University

Fuhong Dai

Soochow University

Jhang Ho Pak

University of Ulsan College of Medicine

Sung-Jong Hong (✉ hongsj@cau.ac.kr)

Research

Keywords: Clonorchis sinensis, bile, transporter, structure, inhibitor, docking

Posted Date: July 24th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-47437/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background: *Clonorchis sinensis* requires bile acid transporters as this fluke inhabits bile juice-filled biliary ducts, which provide an extreme environment. *C. sinensis* sodium-bile acid co-transporter (CsSBAT) is indispensable for survival in the final host, as it circulates taurocholate and prevents bile toxicity in the fluke; hence, it is recognized as a useful drug target.

Results: In the present study, using structure-based drug discovery approach, we presented inhibitor candidates targeting a bile acid-binding pocket of CsSBAT. CsSBAT models were built using comparative tertiary structure modeling based on a bile acid transporter template (PDB ID: 3zuy and 4n7x), and were applied into AutoDock Vina for competitive docking simulation. First, potential compounds were identified from PubChem (holding more than 100,000 compounds) by applying three criteria: i) interacting more favorably with CsSBAT than with a human homolog, ii) intimate interaction to the inward- and outward-facing conformational states, iii) binding with CsSBAT more preferably than natural bile acids. Second, two compounds were identified following Lipinski's rule of five. Third, another two compounds of molecular weight higher than 500 Da ($M_r > 500$ Da) were presumed to efficiently block the transporter via a feasible rational screening strategy. These four inhibitor candidates exhibited least toxicity that may enhance drug-likeness properties.

Conclusion: It is proposed that four compounds act as potential inhibitors toward CsSBAT, and further studies are warranted for drug development process against clonorchiasis.

Background

Clonorchis sinensis is a trematode parasite commonly observed in humans, and is transmitted by eating raw or undercooked freshwater fish contaminated with its metacercariae [1]. The intensity and duration of the infection determine the disease complications caused by this trematode. World Health Organization has classified *C. sinensis* as a Group 1 biological carcinogen inducing cholangiocarcinoma in humans [2].

Praziquantel has been strongly recommended for treating trematode infections in humans including clonorchiasis [3]. Due to the extensive use of praziquantel, certain trematodes in the tropical countries have developed low sensitivity to this drug [4, 5]. Hence, new chemotherapeutic agents should be developed to circumvent the low sensitivity or drug resistance in these trematodes. To date, several drug candidates for trematodiasis have been tested on animals [6–10]. Among these, numerous drugs are under trials for clonorchiasis patients [11].

Structure-based drug discovery (SBDD) is a powerful approach for compound identification, and has recently gained immense attention in drug development [12]. Several studies have reported the effectiveness of computational approaches in discovering novel antiparasitic druggable compounds [13–15]. For *C. sinensis*, secretory phospholipase A_2 enzyme [16] and 20.6-kDa tegumental protein [17] were computationally evaluated as druggable targets; however, more compounds designed by SBDD are

necessary for expanding drug candidate repertoire because they can be optionally tested according to the worm burden or developmental stages of *C. sinensis*.

Notably, bile acids trigger physiological stimuli; however, these metabolites can have detrimental effects on *C. sinensis* survival, as evidence suggests that the accumulated bile acids can cause toxicity to the worm's tissues and cells [18, 19]. Recent studies reported that bile components can be toxic or detrimental to *C. sinensis* survival [10, 20, 21]. Under *in vitro* conditions, higher concentration (> 0.005%) of bile and lithocholic acid significantly decreased the worm survival [10, 20, 21]. Therefore, a defense system against the accumulation and toxicity of bile acids could be targeted to decrease *C. sinensis* survival.

Apical sodium-dependent bile acid transporter (ASBT) and Na⁺-taurocholate co-transporting polypeptide (NTCP) contribute to the enterohepatic circulation of bile salts in humans as homologs of *C. sinensis* sodium-bile acid cotransporter (CsSBAT) [22]. Since ASBT is a promising druggable target due to effectiveness in bile circulation, several inhibitors have been developed and evaluated [23–25]. Therefore, CsSBAT should be considered as a druggable target, whose taurocholate-binding pocket can be blocked and further inhibition can impede its transporter function.

In the present study, CsSBAT structures were reliably prepared using comparative tertiary structure modeling and refinements. For compound screening, the Lipinski's rule of five and the rational virtual screening strategy were applied. Here, we identified putative inhibitory compounds, which competitively targeted the taurocholate-binding pocket of CsSBAT.

Methods

Tertiary structure modeling and refinement

The full-length cDNA sequence of CsSBAT (Acc. No. KX756671) was retrieved from GenBank database [26]. To generate three-dimensional (3D) structure of CsSBAT, we compared the 3D structures built using different 3D modeling softwares such as Swiss-Model [27], I-TASSER [28], Phyre2 [29], RaptorX [30], HHpred [31], and I-TASSER [32]. The 3D models were then refined in two steps. First, low free-energy conformations of the 3D structure were refined by full-atomic simulations using either ModRefiner [33] or FG-MD [34]. Thereafter, backbone and side chains of the structure were refined using GalaxyRefine [35] using “both mild and aggressive relaxation” method based on repeated perturbation and overall conformational relaxation with short molecular dynamics simulations. Outward-facing (OF) and inward-facing (IF) conformations of CsSBAT were modeled using the templates *Yersinia frederiksenii* (YfASBT; PDB ID: 4n7x_A) [36] and *Neisseria meningitidis* ASBT (NmASBT; PDB ID: 3zuy_A) [37], respectively.

Quality validation and binding pockets of 3D models

Potential errors in the 3D models were evaluated using PROCHECK [38], ProSA [39], and ERRAT [40]. Residue-by-residue stereochemical quality of 3D models was verified by Ramachandran plot [41] of PROCHECK. Overall quality score was analyzed by calculating atomic coordinates of the model using ProSA with a Z-score of experimentally determined structures deposited in PDB [42]. Statistics of

nonbonded atom–atom interactions were validated in comparison to a database of reliable high-resolution crystallographic structures using ERRAT [40]. All structures and protein–compound interactions were visualized using UCSF Chimera v1.14 [43]. Disordered region was predicted using ESpritz [44]. Substrate-binding sites in CsSBAT were predicted using COACH [45], referring to the analogs with similar binding sites.

In silico screening of putative inhibitors

Two strategies were applied to screen the putative inhibitors against IF-/OF-CsSBAT and IF-/OF-*Homo sapiens* ASBT (HsASBT) (Fig. 1): i) Compounds (99,288 ea) following the Lipinski's rule of five [46] were screened against “diverse-lib” compounds database using MTiOpenScreen [47]. A list of compounds from each dataset was compared using Venn diagrams generated via jvenn [48]; ii) Compounds (1,255 ea) with high molecular weight (Mr) ranging 500–1,200 Da were collected from PubChem database [49] as of June 2018 and were screened using AutoDock Vina v1.1.2 [50]. Both docking tools were run with following parameters specifying substrate-binding pocket. Grid center parameters were set with X:77.63, Y:76.50, and Z:63.58 for IF conformation, and X:65.26, Y:74.20, and Z:68.12 for OF conformation. These were calculated using Autoligand v1.10 [51, 52]. Grid sizes were set at X:40, Y:40, and Z:40. SDF and PDBQT formats were converted into MOL2 format using Open Babel [53]. Each protein–ligand interaction map was generated using ViewDock method of UCSF Chimera v1.14 [43] and visualized using LigPlot⁺ v1.4.5 [54]. Massive data were analyzed iteratively and parsed using in-house Python scripts. For positive controls, bile acids (7 ea) including taurocholic acid were collected from PubChem database [49].

Toxicity assessment

Toxicity risk and oral toxicity (LD₅₀) were predicted using ProTox [55]. Higher LD₅₀ dose led to less toxic compound. Toxicity class ranging 4–6 indicates that the compound is safe.

Results And Discussion

The best 3D model of CsSBAT was built from comparative homology modeling

Both N- and C-terminal regions of CsSBAT were predicted to be hypothetical as well as disordered. In particular, disordered regions can result in long simulation time and may lead to errors in the structural clustering process [56]. Therefore, these regions were excluded from building 3D models (Additional file 1: Fig. S1). Functional region (residues 185–492) of CsSBAT matched well with experimentally characterized ASBTs such as IF-NmASBT (PDB ID: 3zuy_A) [37] and OF-YfASBT (PDB ID: 4n7x_A) [36].

To obtain reliable homologous models of parasites, combined approach of 3D modeling methods and refinement were employed [17–19, 57–59]. All predicted 3D models of IF-CsSBAT were evaluated using homology modeling programs such as Swiss-Model [27], IntFOLD [28], Phyre2 [29], RaptorX [30], and

HHpred [31], and threading-based modeling program such as I-TASSER [32] (Table 1). Swiss-Model, IntFOLD, and HHpred revealed values greater than 91.0% in the most favored region of Ramachandran plot [41]. Except for IntFOLD presenting erroneous ERRAT value, Swiss-Model, I-TASSER, and HHpred were evaluated further in terms of refinement because I-TASSER is efficient in building structure of unaligned regions by employing *ab initio* modeling [60].

Table 1
Quality verifications of initial models of IF-CsSBAT according to different programs

3D modeling program	Verification value/score			
	Ramachandran ¹	ERRAT	ProSA	Cov ²
Swiss-Model	91.0% (0.8%)	46.0	−3.15	0.56
I-TASSER	75.8% (1.9%)	92.0	−1.07	1.00
IntFOLD	93.8% (0.2%)	n.a. ³	−1.81	1.00
Phyre2	73.2% (5.1%)	26.9	−3.29	1.00
RaptorX	89.3% (0.4%)	41.9	−2.87	1.00
HHpred	91.7% (0%)	50.2	−1.43	1.00
¹ Ramachandran plot, % of most favored region (disallowed region)				
² Cov: Sequence coverage, ratio of predicted region relative to whole region				
³ n.a.: not available				

Homology modeling of membrane proteins is challenging when the target protein shares low sequence identity (approximately 20%) with a template [61]. This issue was circumvented by refining the initial models indicating poor quality [17, 18]. Therefore, the initial models of IF-CsSBAT were refined with ModRefiner, FG-MD, and GalaxyRefine (Table 2). After comprehensive evaluation, we applied Swiss-Model for reliable 3D modeling, and thereafter, FG-MD and GalaxyRefine were used for effective refinement. In fact, Swiss-Model has been a powerful tool for transporter modeling [62, 63]. Refining enhanced the structural quality of the final model compared to that of the initial model, particularly in terms of ERRAT values, which were increased from 46.0–97.7% (Tables 1 and 2, Additional files 1 and 2: Figs. S1 and S2); however, I-TASSER and HHpred could not overcome poor values either in most favored regions of Ramachandran plot or in unacceptable ERRAT plot. Moreover, 3D models of OF-CsSBAT, OF-HsASBT and IF-HsASBT were prepared as aforementioned.

Table 2
Comparison on refined structural evaluation

3D modeling program	Refinement program			Verified score/value produced by			
	ModRefiner	FG-MD	GalaxyRefine	Ramachandran ¹	ERRAT	ProSA	Cov ²
Swiss-Model	□ ³		□	93.3% (0.8%)	89.7	−3.84	0.6
		□	□	92.5% (1.2%)	97.7	−3.73	0.6
I-TASSER	□		n.a. ⁴	86.8% (1.9%)	89.0	−1.07	1.0
		□	n.a.	74.7% (1.9%)	92.9	−1.24	1.0
HHpred	□		n.a.	94.1% (0%)	69.2	−1.89	1.0
		□	n.a.	76.0% (1.3%)	81.5	−1.89	1.0
¹ Ramachandran plot, % of most favored region (disallowed region)							
² Cov: Sequence coverage, ratio of predicted region relative to whole region							
³ Symbol “□”: corresponding software was applied.							
⁴ n.a.: not available							

Bile acid-binding cavity and sodium-binding sites

For the translocation of bile acids across the cell membrane, alternative conformational changes in IF and OF conformations of secondary active transporters were proposed [36, 64]. One of the two conformational states was observed only in a particular state because it is difficult to crystallize the structure under other states. Structural information of one state was applied to predict another conformation of transporters [65]. Here, we predicted the OF-CsSBAT and IF-CsSBAT models based on OF-YfASBT [36] and IF-NmASBT [37], respectively. In CsSBAT, the bile acid-binding pocket was presumed to be formed with 9 residues in an extracellular cavity with a volume of 908 Å³ in OF-CsSBAT (Fig. 2A) and with 11 residues in an intracellular cavity with a volume of 986 Å³ in IF-CsSBAT (Fig. 2B). Among the pocket forming residues, five residues (Phe₁₉₆, Phe₂₂₂, Ala₂₈₈, Ala₂₉₁, and Met₂₉₅) participated in both conformations.

ASBT is a symporter dependent on Na⁺ gradient that drives bile acid transport, which carries one bile acid with two Na⁺ ions [66]. Two Na⁺-binding sites of CsSBAT, Na1 and Na2, were predicted based on NmASBT (Fig. 3A). Na1 site comprised Ser₂₈₉, Asn₂₉₀, Ser₃₀₃, Thr₃₀₇, and Glu₄₄₁ residues; whereas, Na2 site comprised Gln₂₅₂, Glu₄₄₁, Thr₄₄₂, Ile₄₄₄, and Gln₄₄₅ residues (Fig. 3B and C). These two sites were highly conserved with ASBTs and NTCs [36]. Furthermore, Glu₄₄₁ was predicted to participate in Na⁺-binding of the two sites, implying that it might have a functionally significant role (Fig. 3C). In NmASBT, mutation at Glu₂₆₀ (corresponding to Glu₄₄₁ of CsSBAT) markedly alters the taurocholate transporting activity [37].

Recently, a putative third Na⁺-binding site was proposed, albeit rather speculative without experimental evidence [64]. The third Na⁺-binding site was reported from similar transporters such as glutamate [67] and leucine [68] transporters. In CsSBAT, residues Ile₂₈₀, Gly₂₈₁, Ser₂₈₃, and Gln₄₄₅ were predicted to act as binding sites of the third Na⁺ ion, which were superposed with corresponding residues on NmASBT (Fig. 3D). Residue Gln₄₄₅ could be a key residue carrying Na⁺-ion from Na₂ to Na₃ site (Fig. 3). Mutation of Gln₂₅₈ in YfASBT (corresponding to Gln₄₄₅ in CsSBAT) was reported to significantly reduce the Na⁺-binding capacity of Na₂ and Na₃ sites [36]. Therefore, Glu₄₄₁ and Gln₄₄₅ in CsSBAT might act as molecular arms and transport Na⁺ ion from one site to the next.

For accurate molecular docking of CsSBAT against compounds in the library, grid center and size were precisely specified in the extracellular and intracellular bile acid-binding pockets of OF-CsSBAT (Fig. 2A) and IF-CsSBAT (Fig. 2B), respectively, rather than sodium-binding sites (Fig. 3B–D), as described in “Putative inhibitors screening” section of Materials and Methods.

Putative inhibitors targeting at bile acid-binding pocket

Structure-based virtual screening was used to select putative inhibitors of CsSBAT, which satisfied the following criteria (Fig. 1). i) A compound should interact more favorably with CsSBAT than with HsASBT to ensure accurate targeting. ii) OF-ASBT conformation should be considered as a target although IF-ASBT binding with taurocholate was used as a template for virtual docking [25], because the ASBTs transfer bile acid inward *via* conformational change from OF- to IF-conformation [69]. iii) The compound to be identified as a competitive inhibitor of taurocholate should reveal higher affinity than natural bile acids, ranging from – 6.2 to – 9.0 kcal/mol [25]. Theoretically, the binding energies of several bile acids against IF-CsSBAT and OF-CsSBAT conformations ranged from – 6.1 to – 8.7 kcal/mol (Table 3) [19]; however, those of HsASBT (CsSBAT homolog) were – 9.0 kcal/mol and – 9.2 kcal/mol against natural bile acids and poly(acrylic acid)-tetraDOCA conjugate (PATD; lead compound blocking HsASBT), respectively [25]. Thus, a cut-off value was set at – 9.2 kcal/mol since the present study aimed to explore the most probable inhibitor candidates binding to CsSBAT.

Table 3
Binding energy for bile acids OF-/IF-SBAT of *Clonorchis sinensis* and OF-/IF-ASBT of *Homo sapiens* using AutoDock vina

		<i>C. sinensis</i> (kcal/mol)		<i>H. sapiens</i> (kcal/mol)	
Bile acids	PubChem ID	OF	IF	OF	IF
Chenodeoxycholic acid	24875071	-6.7	-7.6	-7.7	-7.0
Taurochenodeoxycholic acid	312642451	-6.4	-7.7	-8.0	-8.0
Glycochenodeoxycholic acid	177011774	-6.4	-7.6	-7.7	-7.3
Glycodeoxycholic acid	57309861	n.a. ¹	n.a.	n.a.	n.a.
Deoxycholic acid	222528 ²	-6.1	-7.7	-7.3	-6.9
Taurocholic acid	828139	-6.6	-8.7	-7.1	-7.0
Glycocholic acid	177011773	-6.8	-8.5	-8.0	-6.8
¹ n.a.: not available					
² PubChem compound ID but other bile acids are PubChem substance ID					

Compounds satisfying Lipinski's rule of five

Compounds following the Lipinski's rule of five [46] were screened using MTiOpenScreen [47]. Of the top 1,000 scoring compounds under docking simulation against OF- and IF-conformations of CsSBAT and HsASBT, 19 compounds that could interact with only OF-CsSBAT or IF-CsSBAT were selected (Fig. 4 and Additional file 3: Table S1). Of these, two compounds met our strict criteria. Compound 49734421 formed a hydrogen bond with Ala₂₉₁ of IF-CsSBAT and Asn₄₄₆ of OF-CsSBAT. Compound 124948115 formed two hydrogen bonds with OF-CsSBAT but not with IF-CsSBAT (Table 4 and Fig. 5). Majority of the residues of these two compounds were involved in hydrophobic interaction with residues on CsSBAT, implying that these interactions might play a crucial role in compound-protein interactivity. It has been reported that aromatic moieties with high hydrophobicity can enable beneficial interactions with nonpolar residues in the binding pocket [70].

Table 4

Inhibitory compounds subjected to virtual screening using MTiOpenScreen and selected by applying Lipinski's rule of five

PubChem ID	Binding energy (kcal/mol)		nRot	HBA	HBD	LogP	Mr	TPSA	Hydrogen bond		Toxicity class (LD ₅₀ : mg/kg)
	OF	IF							OF	IF	
49734421	-10.9	-9.3	4	7	1	2.3	396.5	94.8	1	1	4 (1,077)
124948115	-10.3	-9.3	5	6	1	2.8	397.5	67.6	2	0	4 (1,500)
Abbreviations: nRot, number of rotatable bonds; HBA, hydrogen bond acceptors; HBD, hydrogen bond donors; LogP, lipophilicity; Mr, molecular weight; TPSA, topological polar surface area.											

Since Lipinski's rule of five was introduced in 1997 [46], absorption or drug permeability is presumed to be more likely when there are less than 5 hydrogen bond donors, less than 10 hydrogen bond acceptors, a molecular weight with less than 500, and a calculated LogP smaller than 5. Recently, it was suggested that antiparasitic drugs should be exempted from this rule because several drug leads for infectious diseases do not follow Lipinski's rule of five [46, 71]. Less stringent criteria may allow to identify more lead compounds for further assays. The suggestion motivated us to find out more effective strategy for antiparasitic drugs.

Large compounds satisfying high affinities

PATD was recently synthesized and evaluated as a potent inhibitor against ASBT [25]. Molecular weight of PATD molecule is larger than 500 Da because it has several polyacrylic acids and tetraDOCAs. Surprisingly, PATD is a hydrophobic substrate, which violates ideal molecular weight of the Lipinski's rule of five. Nonetheless, it successfully inhibits ASBT by filling up the bile acid-binding cavity [46]. This finding motivated us to screen compounds with molecular weight higher than 500 Da, which are assumed to tightly dock CsSBAT.

Compounds of high molecular weight (500–1,200 Da) were retrieved from PubChem compound library [49] and screened using AutoDock Vina v1.1.2 [50]. Of the 1,255 compounds, 49 compounds satisfied the three given criteria. By strictly applying the third criterion (higher affinity than natural bile acids), we selected 25 compounds with high affinity for CsSBAT, but low affinity for HsASBT (Additional file 4: Table S2). Five compounds presented toxicity values of 4–6 with LD₅₀ ranging 500–5,000 mg/kg (Table 5). Eventually, two compounds 45375808 (Fig. 6A and B) and 9806452 (Fig. 6C and D) were selected as possible candidate inhibitors for CsSBAT because these compounds could form more than two hydrogen bonds each with OF-CsSBAT and IF-CsSBAT.

Table 5

Inhibitory compounds with high molecular weight ($M_r > 500$ Da) selected using AutoDock Vina v1.1.2

PubChem ID	Mr ¹	MF ²	Binding energy (kcal/mol)				Hydrogen		Toxicity
			<i>C. sinensis</i>		<i>H. sapiens</i>		bond		class
			OF	IF	OF	IF	OF	IF	(LD ₅₀ : mg/kg)
441243	670.8	C ₃₈ H ₅₀ N ₆ O ₅	-12.3	-10.0	-8.9	-7.9	5	1	4 (500)
4701	508.6	C ₃₁ H ₃₂ N ₄ O ₃	-11.2	-10.9	-8.9	-8.7	4	0	4 (800)
92727	628.8	C ₃₇ H ₄₈ N ₄ O ₅	-10.0	-10.0	-9.0	-8.7	4	0	5 (5,000)
45375808	529.5	C ₂₂ H ₂₉ FN ₃ O ₉ P	-9.7	-9.7	-8.4	-8.4	2	4	6 (12,000)
9806452	512.7	C ₃₀ H ₄₁ FN ₂ O ₄	-9.3	-9.3	-7.6	-8.3	4	2	5 (3,990)
¹ Mr, molecular weight									
² MF, molecular formula									

Notably, through docking simulation on compound–protein interactions, residues Glu₂₂₉ and Gly₂₈₇ participated in hydrogen bonding in taurocholate-IF-CsSBAT complex, whereas residues Gly₂₈₇, Gln₃₄₅, and Gln₃₄₈ participated in hydrogen bonding in taurocholate-OF-CsSBAT complex (Additional file 2: Fig. S2). Gly₂₈₇ was involved both in taurocholate-IF-CsSBAT and taurocholate-OF-CsSBAT complexes. In majority of the compound-OF-CsSBAT complexes, this conserved residue Gly₂₈₇ was involved in hydrogen bond interaction (Fig. 6A, C, E, and G), and Gln₃₄₅ was involved in compound 92727-OF-CsSBAT complex formation (Fig. 6I). Among compound-IF-CsSBAT complexes, only compound 441243 formed hydrogen bond with Glu₂₂₉ (Fig. 6F). Compared to taurocholate-CsSBAT complexes, Ala₂₈₈ acted as a key residue for either hydrogen bonds or hydrophobic interactions in three compound-CsSBAT complexes (Fig. 6A–F). Compounds 441243 and 3693566 presented 5 and 6 hydrogen bonds, respectively; however, these compounds were excluded owing to less hydrogen bonds in compound-IF-CsSBAT interactions, which could result in off-target binding with adverse drug reactions [72].

Compound 45375808, known as sofosbuvir, was proposed to inhibit nonstructural protein 5B polymerase in hepatitis C virus (HCV) [73]. Analogs of this compound exhibited inhibitory effect on HCV [74]. Compound 9806452 was reported as an inhibitor of matrix metalloproteinases [75] such as gelatinase A associated with tumor metastasis [76] and stromelysin-1 found in osteoarthritis [77]. Carboxyalkyl peptides

containing a biphenylethyl group inhibits adult *Schistosoma mansoni* [78]. Considering these reports, it is suggested that compounds 45375808 and 9806452 could be anthelmintic candidates for *C. sinensis*.

Conclusion

C. sinensis sodium-bile acid cotransporter (CsSBAT) should be indispensable for survival in the bile duct. Inhibition of the bile transporters may perturb bile acid transport, and prove detrimental to *C. sinensis*. We identified the inhibitory compounds targeting the bile acid-binding pockets in CsSBAT, based on the physiological essence of crucial importance using structure-based drug discovery approach. First, PubChem compounds 49734421 and 124948115 were selected by applying the Lipinski's rule of five. Furthermore, we devised a feasible rational screening strategy to search inhibitor candidates with molecular weight greater than 500 Da. The large inhibitory compounds 45375808 and 9806452 were selected and expected to tightly bind to CsSBAT. These four inhibitor candidates revealed least toxicity that may enhance their druggability. Collectively, four compounds were proposed as putative inhibitors of CsSBAT, deserving further *in vitro* and *in vivo* evaluation toward anthelmintic development.

Declarations

Acknowledgments

This research was supported by the National Research Foundation of Korea (NRF; (<http://www.nrf.re.kr>) funded by the Ministry of Science, ICT & Future Planning (grant No. 2014R1A2A1A11051870).

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and material

The data supporting the conclusions of this article are included in the article.

Competing interests

The authors declare that they have no competing interests.

Funding

This research was supported by the National Research Foundation of Korea (NRF; (<http://www.nrf.re.kr>) funded by the Ministry of Science, ICT & Future Planning (grant No. 2014R1A2A1A11051870).

Authors contributions

SJH and WGY conceived of and designed the experiments. WGY and FD analyzed the data. WGY contributed to utilizing the analysis tools. WGY and JHP wrote the paper. WGY and SJH reviewed the paper. All authors read and approved the final version of the manuscript.

Author details

¹ Department of Medical Environmental Biology, Chung-Ang University College of Medicine, Seoul 06974, Korea

²Department of Parasitology, School of Biology and Basic Medical Sciences, Medical College, Soochow University, Suzhou, Jiangsu, PR China

³ Department of Convergence Medicine, University of Ulsan College of Medicine and Asan Institute for Life Sciences, Asan Medical Center, Seoul 05505, Korea

References

1. Keiser J, Utzinger J. Emerging foodborne trematodiasis. *Emerg Infect Dis.* 2005;11:1507–14.
2. Bouvard V, Baan R, Straif K, Grosse Y, Secretan B, El Ghissassi F, et al. A review of human carcinogens—Part B: biological agents. *Lancet Oncol.* 2009;10:321–2.
3. Choi MH, Park SK, Li Z, Ji Z, Yu G, Feng Z, et al. Effect of control strategies on prevalence, incidence and re-infection of clonorchiasis in endemic areas of China. *PLoS Negl Trop Dis.* 2010;4:e601.
4. Tinga N, De N, Vien HV, Chau L, Toan ND, Kager PA, et al. Little effect of praziquantel or artemisinin on clonorchiasis in Northern Vietnam. A pilot study. *Trop Med Int Health.* 1999;4:814–8.
5. Fallon PG. Schistosome resistance to praziquantel. *Drug Resist Updat.* 1998;1:236–41.
6. Kim TI, Yoo WG, Li S, Hong ST, Keiser J, Hong SJ. Efficacy of artesunate and artemether against *Clonorchis sinensis* in rabbits. *Parasitol Res.* 2009;106:153–6.
7. Liu YH, Wang XG, Gao P, Qian MX. Experimental and clinical trial of albendazole in the treatment of *Clonorchiasis sinensis*. *Chin Med J (Engl).* 1991;104:27–31.
8. Xiao SH, Xue J, Li-li X, Zhang YN, Qiang HQ. Effectiveness of mefloquine against *Clonorchis sinensis* in rats and *Paragonimus westermani* in dogs. *Parasitol Res.* 2010;107:1391–7.

9. Xiao SH, Xue J, Xu LL, Zhang YN, Qiang HQ. Comparative effect of mebendazole, albendazole, tribendimidine, and praziquantel in treatment of rats infected with *Clonorchis sinensis*. Parasitol Res. 2011;108:723–30.
10. Li S, Song JH, Kim TI, Yoo WG, Won MH, Dai F, et al. Chemotactic migration of newly excysted juvenile *Clonorchis sinensis* is suppressed by neuro-antagonists. PLoS Negl Trop Dis. 2019;13:e0007573.
11. Xu LL, Jiang B, Duan JH, Zhuang SF, Liu YC, Zhu SQ, et al. Efficacy and safety of praziquantel, tribendimidine and mebendazole in patients with co-infection of *Clonorchis sinensis* and other helminths. PLoS Negl Trop Dis. 2014;8:e3046.
12. Shoichet BK. Virtual screening of chemical libraries. Nature. 2004;432:862–5.
13. Cavasotto CN, Orry AJ, Murgolo NJ, Czarniecki MF, Kocsi SA, Hawes BE, et al. Discovery of novel chemotypes to a G-protein-coupled receptor through ligand-steered homology modeling and structure-based virtual screening. J Med Chem. 2008;51:581–8.
14. Cohen NC. Structure-based drug design and the discovery of aliskiren (Tekturna): perseverance and creativity to overcome a R&D pipeline challenge. Chem Biol Drug Des. 2007;70:557–65.
15. Kohring K, Wiesner J, Altenkamper M, Sakowski J, Silber K, Hillebrecht A, et al. Development of benzophenone-based farnesyltransferase inhibitors as novel antimalarials. ChemMedChem. 2008;3:1217–31.
16. Hariprasad G, Kaur P, Srinivasan A, Singh TP, Kumar M. Structural analysis of secretory phospholipase A2 from *Clonorchis sinensis*: therapeutic implications for hepatic fibrosis. J Mol Model. 2012;18:3139–45.
17. Kim YJ, Yoo WG, Lee MR, Kang JM, Na BK, Cho SH, et al. Molecular and Structural Characterization of the Tegumental 20.6-kDa Protein in *Clonorchis sinensis* as a Potential Druggable Target. Int J Mol Sci. 2017;18.
18. Lu Y, Yoo WG, Dai F, Lee JY, Pak JH, Sohn WM, et al. Characterization of a novel organic solute transporter homologue from *Clonorchis sinensis*. PLoS Negl Trop Dis. 2018;12:e0006459.
19. Dai F, Yoo WG, Lee JY, Lu Y, Pak JH, Sohn WM, et al. Multidrug resistance-associated protein 4 is a bile transporter of *Clonorchis sinensis* simulated by *in silico* docking. Parasit Vectors. 2017;10:578.
20. Li S, Yoo WG, Song JH, Kim TI, Hong SJ. Bile acids drive chemotaxis of *Clonorchis sinensis* juveniles to the bile duct. PLoS Negl Trop Dis. 2018;12:e0006818.
21. Uddin MH, Li S, Bae YM, Choi MH, Hong ST. *In vitro* maintenance of *Clonorchis sinensis* adult worms. Korean J Parasitol. 2012;50:309 – 15.
22. Claro da Silva T, Polli JE, Swaan PW. The solute carrier family 10 (SLC10): beyond bile acid transport. Mol Aspects Med. 2013;34:252–69.
23. Baghdasaryan A, Fuchs CD, Osterreicher CH, Lemberger UJ, Halilbasic E, Pahlman I, et al. Inhibition of intestinal bile acid absorption improves cholestatic liver and bile duct injury in a mouse model of sclerosing cholangitis. J Hepatol. 2016;64:674–81.
24. West KL, Zern TL, Butteiger DN, Keller BT, Fernandez ML. SC-435, an ileal apical sodium co-dependent bile acid transporter (ASBT) inhibitor lowers plasma cholesterol and reduces atherosclerosis in guinea

- pigs. *Atherosclerosis*. 2003;171:201–10.
25. Park J, Al-Hilal TA, Jeong JH, Choi J, Byun Y. Design, Synthesis, and Therapeutic Evaluation of Poly(acrylic acid)-tetraDOCA Conjugate as a Bile Acid Transporter Inhibitor. *Bioconjug Chem*. 2015;26:1597–605.
 26. Jenuth JP. The NCBI. Publicly available tools and resources on the Web. *Methods Mol Biol*. 2000;132:301–12.
 27. Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, et al. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res*. 2014;42::W252-8.
 28. McGuffin LJ, Atkins JD, Salehe BR, Shuid AN, Roche DB. IntFOLD: an integrated server for modelling protein structures and functions from amino acid sequences. *Nucleic Acids Res*. 2015;43::W169-73.
 29. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc*. 2015;10:845–58.
 30. Kallberg M, Margaryan G, Wang S, Ma J, Xu J. RaptorX server: a resource for template-based protein structure modeling. *Methods Mol Biol*. 2014;1137:17–27.
 31. Meier A, Soding J. Automatic Prediction of Protein 3D Structures by Probabilistic Multi-template Homology Modeling. *PLoS Comput Biol*. 2015;11:e1004343.
 32. Yang J, Zhang Y. I-TASSER server: new development for protein structure and function predictions. *Nucleic Acids Res*. 2015;43::W174-81.
 33. Xu D, Zhang Y. Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization. *Biophys J*. 2011;101:2525–34.
 34. Zhang J, Liang Y, Zhang Y. Atomic-level protein structure refinement using fragment-guided molecular dynamics conformation sampling. *Structure*. 2011;19:1784–95.
 35. Heo L, Park H, Seok C. GalaxyRefine. Protein structure refinement driven by side-chain repacking. *Nucleic Acids Res*. 2013;41::W384-8.
 36. Zhou X, Levin EJ, Pan Y, McCoy JG, Sharma R, Kloss B, et al. Structural basis of the alternating-access mechanism in a bile acid transporter. *Nature*. 2014;505:569–73.
 37. Hu NJ, Iwata S, Cameron AD, Drew D. Crystal structure of a bacterial homologue of the bile acid sodium symporter ASBT. *Nature*. 2011;478:408–11.
 38. Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R, Thornton JM. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR*. 1996;8:477–86.
 39. Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res*. 2007;35::W407-10.
 40. Colovos C, Yeates TO. Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci*. 1993;2:1511–9.
 41. Lovell SC, Davis IW, Arendall WB 3rd, de Bakker PI, Word JM, Prisant MG, et al. Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins*. 2003;50:437 – 50.

42. Rose PW, Bi C, Bluhm WF, Christie CH, Dimitropoulos D, Dutta S, et al. The RCSB Protein Data Bank: new resources for research and education. *Nucleic Acids Res.* 2013;41::D475-82.
43. Huang CC, Meng EC, Morris JH, Pettersen EF, Ferrin TE. Enhancing UCSF Chimera through web services. *Nucleic Acids Res.* 2014;42::W478-84.
44. Walsh I, Martin AJ, Di Domenico T, Vullo A, Pollastri G, Tosatto SC. CSpritz: accurate prediction of protein disorder segments with annotation for homology, secondary structure and linear motifs. *Nucleic Acids Res.* 2011;39::W190-6.
45. Yang J, Roy A, Zhang Y. Protein-ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment. *Bioinformatics.* 2013;29:2588–95.
46. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev.* 2001;46:3–26.
47. Labbe CM, Rey J, Lagorce D, Vavrusa M, Becot J, Sperandio O, et al. MTiOpenScreen: a web server for structure-based virtual screening. *Nucleic Acids Res.* 2015;43::W448-54.
48. Bardou P, Mariette J, Escudie F, Djemiel C, Klopp C. jvenn: an interactive Venn diagram viewer. *BMC Bioinformatics.* 2014;15:293.
49. Kim S, Thiessen PA, Bolton EE, Chen J, Fu G, Gindulyte A, et al. PubChem Substance and Compound databases. *Nucleic Acids Res.* 2016;44::D1202-13.
50. Forli S, Huey R, Pique ME, Sanner MF, Goodsell DS, Olson AJ. Computational protein-ligand docking and virtual drug screening with the AutoDock suite. *Nat Protoc.* 2016;11:905–19.
51. Harris R, Olson AJ, Goodsell DS. Automated prediction of ligand-binding sites in proteins. *Proteins.* 2008;70:1506–17.
52. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, et al. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem.* 2009;30:2785–91.
53. O'Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR. Open Babel: An open chemical toolbox. *J Cheminform.* 2011;3::33.
54. Laskowski RA, Swindells MB. LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J Chem Inf Model.* 2011;51:2778–86.
55. Drwal MN, Banerjee P, Dunkel M, Wettig MR, Preissner R. ProTox: a web server for the *in silico* prediction of rodent oral toxicity. *Nucleic Acids Res.* 2014;42::W53-8.
56. Yang J, Zhang Y. Protein Structure and Function Prediction Using I-TASSER. *Curr Protoc Bioinformatics.* 2015;52:5.8.1–15.
57. Lee J, Kim TI, Le HG, Yoo WG, Kang JM, Ahn SK, et al. Genetic diversity of *Plasmodium vivax* and *Plasmodium falciparum* lactate dehydrogenases in Myanmar isolates. *Malar J.* 2020;19:60.
58. Kang JM, Yoo WG, Le HG, Lee J, Sohn WM, Na BK. *Clonorchis sinensis* MF6p/HDM (CsMF6p/HDM) induces pro-inflammatory immune response in RAW 264.7 macrophage cells via NF-kappaB-dependent MAPK pathways. *Parasit Vectors.* 2020;13:20.

59. Dai F, Yoo WG, Lee JY, Lu Y, Pak JH, Sohn WM, et al. Molecular and structural characteristics of multidrug resistance-associated protein 7 in Chinese liver fluke *Clonorchis sinensis*. *Parasitol Res*. 2017;116:953–62.
60. Yang Y, Faraggi E, Zhao H, Zhou Y. Improving protein fold recognition and template-based modeling by employing probabilistic-based matching between predicted one-dimensional structural properties of query and corresponding native properties of templates. *Bioinformatics*. 2011;27::2076–82.
61. Franca TC. Homology modeling: an important tool for the drug discovery. *J Biomol Struct Dyn*. 2015;33:1780–93.
62. Ward AB, Szewczyk P, Grimard V, Lee CW, Martinez L, Doshi R, et al. Structures of P-glycoprotein reveal its conformational flexibility and an epitope on the nucleotide-binding domain. *Proc Natl Acad Sci U S A*. 2013;110:13386–91.
63. Murashita K, Yoshiura Y, Chisada S, Furuita H, Sugita T, Matsunari H, et al. Homologue gene of bile acid transporters ntcp, asbt, and ost-alpha in rainbow trout *Oncorhynchus mykiss*: tissue expression, effect of fasting, and response to bile acid administration. *Fish Physiol Biochem*. 2014;40:511–25.
64. Alhadeff R, Ganoth A, Arkin IT. Mechanistic studies of the apical sodium-dependent bile acid transporter. *Proteins*. 2015;83:1107–17.
65. Vergara-Jaque A, Fenollar-Ferrer C, Kaufmann D, Forrest LR. Repeat-swap homology modeling of secondary active transporters: updated protocol and prediction of elevator-type mechanisms. *Front Pharmacol*. 2015;6:183.
66. Weinman SA, Carruth MW, Dawson PA. Bile acid uptake via the human apical sodium-bile acid cotransporter is electrogenic. *J Biol Chem*. 1998;273:34691–5.
67. Bastug T, Heinzelmann G, Kuyucak S, Salim M, Vandenberg RJ, Ryan RM. Position of the third Na⁺ site in the aspartate transporter GltPh and the human glutamate transporter, EAAT1. *PLoS One*. 2012;7:e33058.
68. Zomot E, Gur M, Bahar I. Microseconds simulations reveal a new sodium-binding site and the mechanism of sodium-coupled substrate uptake by LeuT. *J Biol Chem*. 2015;290:544–55.
69. Balakrishnan A, Sussman DJ, Polli JE. Development of stably transfected monolayer overexpressing the human apical sodium-dependent bile acid transporter (hASBT). *Pharm Res*. 2005;22:1269–80.
70. Ferraro F, Merlino A, Dell Oca N, Gil J, Tort JF, Gonzalez M, et al. Identification of Chalcones as *Fasciola hepatica* Cathepsin L Inhibitors Using a Comprehensive Experimental and Computational Approach. *PLoS Negl Trop Dis*. 2016;10:e0004834.
71. McKerrow JH, Lipinski CA. The rule of five should not impede anti-parasitic drug development. *Int J Parasitol Drugs Drug Resist*. 2017;7:248–9.
72. Lounkine E, Keiser MJ, Whitebread S, Mikhailov D, Hamon J, Jenkins JL, et al. Large-scale prediction and testing of drug activity on side-effect targets. *Nature*. 2012;486:361–7.
73. Sofia MJ, Bao D, Chang W, Du J, Nagarathnam D, Rachakonda S, et al. Discovery of a beta-d-2'-deoxy-2'-alpha-fluoro-2'-beta-C-methyluridine nucleotide prodrug (PSI-7977) for the treatment of hepatitis C virus. *J Med Chem*. 2010;53:7202–18.

74. Jonckers TH, Vandyck K, Vandekerckhove L, Hu L, Tahri A, Van Hoof S, et al. Nucleotide prodrugs of 2'-deoxy-2'-spirooxetane ribonucleosides as novel inhibitors of the HCV NS5B polymerase. *J Med Chem.* 2014;57:1836–44.
75. Esser CK, Bugianesi RL, Caldwell CG, Chapman KT, Durette PL, Girotra NN, et al. Inhibition of stromelysin-1 (MMP-3) by P1'-biphenylethyl carboxyalkyl dipeptides. *J Med Chem.* 1997;40:1026–40.
76. Stetler-Stevenson WG, Liotta LA, Brown PD. Role of type IV collagenases in human breast cancer. *Cancer Treat Res.* 1992;61:21–41.
77. Okada Y, Shinmei M, Tanaka O, Naka K, Kimura A, Nakanishi I, et al. Localization of matrix metalloproteinase 3 (stromelysin) in osteoarthritic cartilage and synovium. *Lab Invest.* 1992;66:680–90.
78. Gorvin JH, Raison CG, Solomon W, Standen OD, Walls LP. The action of substances analogous to diaminodiphenoxyalkanes against *Schistosoma mansoni*. *Br J Pharmacol Chemother.* 1957;12:329–35.

Abbreviations

ASBT apical sodium-dependent bile acid transporter

BESP bile salt export pump

IF inward-facing

LD₅₀ median lethal dose

LogP lipophilicity

Mr molecular weight

MRP multidrug resistance-associated protein

NTCP Na⁺-taurocholate cotransporting polypeptide

OF outward-facing

PATD polyacrylic acid–tetraDOCA conjugate

SBAT sodium-bile acid co-transporter

SBDD structure-based drug discovery

TCH taurocholate

Figures

Figure 1

Strategy for structure-based virtual screening

Figure 2

Conformation change of CsSBAT accompanied with transport of Na⁺ and taurocholate. Both OF-CsSBAT (A) and IF-CsSBAT (B) are colored in tan. OF-YfASBT (A) is colored in pink and IF-NmASBT (B) in blue. ASBTs presenting identical conformation are superposed. Tunnel leading to the TCH-binding site is visualized in yellow eclipse. Residues forming TCH-binding cavity are depicted as green spheres in CsSBAT. Other residues participating in TCH-binding site are provided at each side. Na⁺ ions are in purple circle (Na1 and Na2). Taurocholate (TCH) is depicted as maroon pentagon

Figure 3

Na⁺- and taurocholate-binding sites of CsSBAT. All ligands and interaction were predicted using COACH [45] except for Na3 site (A). Na⁺-binding sites (B, C), Na1 site (B), and Na2 site (C). Putative Na3 site (D) as suggested by Alhadeff et al. [64]. The consensus residues forming the binding site are presented in stick mode and labeled. Side-chain oxygen and hydrogen atoms are indicated in red and white, respectively. Na⁺ ion is depicted as a purple ball and taurocholate substrate as stick mode

Figure 4

A Venn diagram presents compounds interacting with two SBATs and two ASBTs. Ellipse represents compounds screened against each of the four ASBTs. Gray area indicates compounds, which interact with either OF-CsSBAT or IF-CsSBAT. Black area depicts compounds that interact with both CsSBATs, but not with other ASBTs

Figure 5

Docked pose of two compounds with IF-CsSBAT and OF-CsSBAT following Lipinski's rule of five. Binding modes of compounds were obtained from the MTiOpenScreen [47]. In each panel, left box ribbon and stick indicate a compound placed in the active site of OF-CsSBAT (A and C) and IF-CsSBAT (B and D). Right panel generated by LigPlot+ v1.4.3 [54] indicates the schematic interactions of CsSBAT with 124948115 (A and B) and 49734421 (C and D). The compound and residues with hydrophobic interactions are visualized with 2D diagrams. The residues contributing to hydrogen bonds are colored green. Hydrogen bond is indicated as a lengthy broken line. Black arrow indicates compound movement

Figure 7

Docked poses of large compounds with OF-CsSBAT and IF-CsSBATs. The binding modes of compounds were obtained from the AutoDock Vina v1.1.2 [50]. Left (A, C, E, G, and I) and right (B, D, F, H, and J) panels represent OF-CsSBAT and IF-CsSBAT, respectively. The schematic interactions of CsSBAT with 4537808 (A and B), 9806452 (C and D), 441243 (E and F), 4701 (G and H), and 92727 (I and J) were generated using LigPlot+ v1.4.3 [54]. The compound and residues with hydrophobic interactions are visualized with 2D diagrams. The residues contributing to hydrogen bonds are colored green. Hydrogen bond is indicated as a broken line with length. Amino acid residues involved in the hydrophobic interactions are marked with red spoked arcs. Black arrow indicates compound movement

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additional4TableS2.docx](#)

- [Additional3TableS1.docx](#)
- [Additional2FigS2.pptx](#)
- [Additional1FigS1.pptx](#)