

Research Paper

Inhibition Requirements of the Human Apical Sodium-Dependent Bile Acid Transporter (hASBT) Using Aminopiperidine Conjugates of glutamyl-Bile Acids

Pablo M. González,^{1,2} Chayan Acharya,¹ Alexander D. MacKerell Jr.,¹ and James E. Polli^{1,3}

Received December 15, 2008; accepted March 16, 2009; published online April 21, 2009

Purpose. Synthesize aminopiperidine conjugates of glutamyl-bile acids (glu-BAs) and develop a hASBT inhibition model using the conformationally sampled pharmacophore (CSP) approach.

Methods. glu-BAs aminopiperidine conjugates were synthesized. hASBT inhibition was measured as K_i . A CSP-SAR model was built using structural and physico-chemical descriptors and evaluated via cross-validation.

Results. Twenty-nine aminopiperidine conjugates were synthesized. All inhibited hASBT, with K_i ranging from 0.95 to 31.8 μM . Amidation of the piperidine nitrogen slightly decreased activity, while replacement by a carbon increased potency. Esterification of the glutamic acid linker had a minor impact, suggesting that a negative charge around C-24 is not required for binding. Three quantitative CSP-SAR models were developed. The best model ($r^2=0.813$, $Q^2=0.726$) included two descriptors: angle between 7-OH, α -substituent and centroid of rings B and C, and electrostatic contribution to the solvation free-energy. The model successfully distinguished between compounds with $K_i < 16\mu\text{M}$ and $K_i > 16\mu\text{M}$. Models indicated that hydrophobicity, α substituent orientation, and partially compacted side chain conformation promote inhibitory potency. Qualitative CSP-SAR analysis indicated that the presence of an internal salt bridge, resulting in a locked conformation of the side chain, yielded weaker inhibitors.

Conclusions. Aminopiperidine conjugates of glu-BAs were potent hASBT inhibitors. A predictive and robust CSP-SAR model was developed.

KEY WORDS: apical sodium-dependent bile acid transporter; bile acid; CHARMM; conformationally sampled pharmacophore; transporter.

INTRODUCTION

The human apical sodium-dependent bile acid transporter (hASBT) is a 348 amino acid intestinal transporter which

Electronic supplementary material The online version of this article (doi:10.1007/s11095-009-9877-3) contains supplementary material, which is available to authorized users.

Pablo Gonzalez and Chayan Acharya contributed equally to this work and should be considered joint First-Authors.

¹Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, 20 Penn Street, HSF2 room 623, Baltimore, Maryland 21201, USA.

²Departamento de Farmacia, Facultad de Química, Pontificia Universidad Católica de Chile, Santiago, Chile.

³To whom correspondence should be addressed. (e-mail: jpolli@rx.umaryland.edu)

ABBREVIATIONS: ABL, aqueous boundary layer; CDCA-glu, glutamyl-chenodeoxycholic acid; CSP-SAR, conformationally sampled pharmacophore structure activity relationship; DIPEA, diisopropylethyl amine; DMF, dimethyl formamide; EDC, 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide; hASBT, human apical sodium-dependent bile acid transporter; HBSS, Hanks balanced salt solution; HBTU, O-Benzotriazole-N,N,N',N'-tetramethyl-uronium-hexafluoro-phosphate; HOBT, N-hydroxybenzotriazole; J_{max} , maximum transport capacity; K_i , binding affinity; K_t , Michaelis-Menten transport constant; MD, molecular dynamics; MDCK, Madin-Darby canine kidney; P_p , passive permeability; TCA, taurocholic acid; UDCA-glu, glutamyl-ursodeoxycholic acid.

plays a key role in the enterohepatic recirculation of bile acids and homeostasis of cholesterol (1–6). hASBT shows high affinity and high capacity for conjugated primary bile acids such as taurocholate and taurochenodeoxycholate with K_i and K_t in the low micromolar range (7–10). Inhibitors of hASBT decrease plasma cholesterol, improve the global serum lipid profile, and increase the fecal excretion of bile acids (11–15). hASBT is a possible target for increasing drug absorption via a prodrug approach (16).

Unfortunately, in the absence of a crystal structure of the transporter, little is known about the structural requirements of binding and translocation by hASBT. In particular knowledge about the influence of chemical substituents beyond the C-24 region on hASBT binding affinity is lacking. Our laboratory has previously reported the synthesis and kinetic characterization of a series of glutamic acid and lysine conjugates of CDCA through the C-24 acid (17). In these series, the influence of bulkiness and charge state of the amino acid in C-24 was varied, and their ability to serve as inhibitors/substrates of hASBT evaluated. Results showed that neutral, monoanionic, and cationic CDCA conjugates were potent hASBT inhibitors. Dianionic conjugates did not inhibit hASBT. Baringhaus *et al.* developed a 3D QSAR inhibition model for rabbit ileal Na^+ /bile acid cotransporters; however, their set of inhibitors did not contain any positively charged compounds (8).

The conformationally sampled pharmacophore (CSP) method is a novel approach developed in our laboratory to

generate 3-D QSAR models for ligand-based drug design. This method is particularly suited for ligands with high structural flexibility, such as glutamyl-bile acid conjugates (18–20). The CSP method combined with the QSAR approach (denoted CSP-SAR) offers the advantage of including all possible conformations sampled by the training set compounds during model development while also accounting for physical characteristics. By considering all accessible conformers of the compounds, this method maximizes the probability of including the bound conformation of the ligands in the model.

The objective of the present study was to systematically explore the influence of piperidine probes beyond the γ -acid in the glutamic acid linker for their impact on binding to hASBT. The piperidine ring was chosen as the probe moiety since piperidine is a common scaffold in drug structures. Twenty-nine aminopiperidine conjugates of glu-BAs were synthesized. The glu-BAs were glutamyl-chenodeoxycholate (CDCA-glu) and glutamyl-ursodeoxycholate (UDCA-glu), either free or protected as a α -benzyl ester. A CSP-SAR model for hASBT inhibition was developed using a subset of 19 compounds from which a detailed model of the structural and physical characteristics that effect inhibition was obtained.

MATERIALS AND METHODS

Materials. [^3H]-Taurocholic acid (10 $\mu\text{Ci}/\text{mmol}$) was purchased from American Radiolabeled Chemicals, Inc (St. Louis, MO). Taurocholate (TCA) was from Sigma Aldrich (St. Louis, MO). Chenodeoxycholate (CDCA) and ursodeoxycholate (UDCA) were obtained from TCI America (Portland, OR). Protected glutamic acid analogs were from Novabiochem (Gibbstown, NJ). Chiral amino- and aminomethyl-1-N-boc-piperidine analogs were purchased from Astatech Inc. (Bristol, PA). 1-N-phenylpiperidin-4-one was from Atlantic SciTech Group (Linden, NJ). Geneticin, fetal bovine serum (FBS), trypsin, and DMEM were purchased from Invitrogen (Rockville, MD). All other reagents and chemicals were of the highest purity commercially available.

Synthetic Procedures

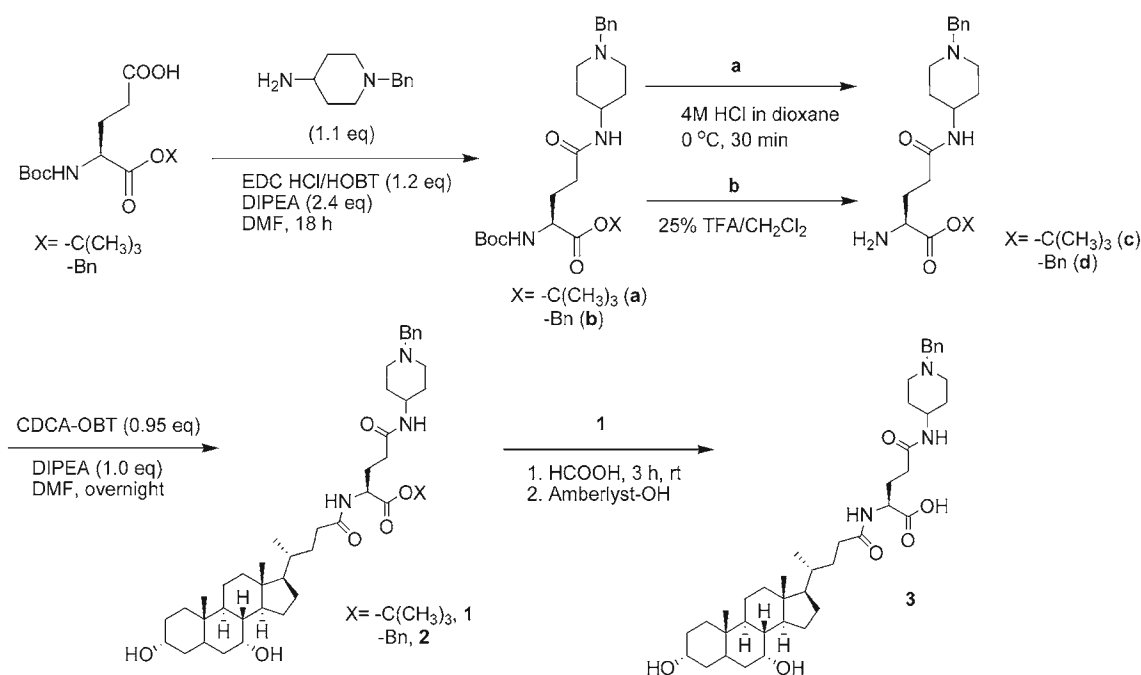
Synthesis of N-hydroxybenzotriazole (HOBT) Esters of CDCA and UDCA. Five grams (12.7 mmol) of CDCA were reacted with HOBT (6.35 mmol) and HBTU (12.7 mmol) in 20 mL of anhydrous dimethylformamide (DMF) at RT for 4 h. The reaction was quenched by adding 50 mL of cold water and extracting into 30 mL of ethyl acetate (3 \times). The combined organic extracts were washed thrice with 10% sodium bicarbonate and one time with water. The organic layer was dried with anhydrous sodium sulfate, filtered and evaporated under vacuum to yield a fluffy white powder (92% yield). An identical procedure was employed in the synthesis of UDCA-OBT ester. Mass spectrometry was performed on a LCQ ESI-MS (Thermo Scientific, Waltham, MD) and showed appropriate peaks: [M+1] 510.68, [M+Na] 532.68. Compounds were used in the next step without further purification.

Synthesis of α -benzyl Ester Glutamic Acid CDCA Amide (CDCA-glu). Three grams (5.89 mmol) of CDCA-OBT ester were reacted with α -benzyl-glutamic acid (6.45 mmol) along

with 1.0 eq of diisopropylethylamine (DIPEA) in anhydrous 20 mL DMF at 40°C overnight. The reaction was quenched by adding 50 mL of water and extracting into 30 mL of ethyl acetate (three times). The organic extract was washed with 15 mL of 1 N HCl (3 \times), and 15 mL of brine (1 \times). The ethyl acetate extract was evaporated under vacuum to yield a pale yellow oil. This oil was dissolved in the minimum amount of acetonitrile and poured into 250 mL of cold water with agitation. The suspension was then decanted and extracted into 50 mL of ethyl acetate. The organic layer was washed with brine, dried with anhydrous sodium sulfate, and evaporated under vacuum to yield α -benzyl ester glutamic acid CDCA amide (CDCA-glu) as an off-white solid (75% yield). An identical procedure was followed in the synthesis of α -benzyl ester glutamic acid UDCA amide (UDCA-glu). MS showed appropriate peaks: [M+1] 612.81, [M+Na] 634.81. ^1H and ^{13}C nuclear magnetic resonance spectra were recorded on a Varian Inova 500 MHz (Varian Inc., Palo Alto, CA). For product 1 and all other 29 products, see supplemental materials for NMR data.

Synthesis of Aminopiperidine Conjugates of CDCA-Glutamic Acid or UDCA-Glutamic Acid. Three general different procedures (denoted A, B, and C) were employed, depending upon the N1-protection chemistry on piperidine probe. Procedure A was used when N1 was protected with a benzyl group. Procedure B was applied when piperidine nitrogen was boc protected. Procedure C was used when N1 was acylated or aromatic, or when isosterically replaced with carbon.

Procedure A. Procedure A was applied to synthesize products 1, 2, and 3, as shown in Scheme 1. First, 3.0 mmol of N-boc- α -benzyl-glutamic acid or N-boc- α -*t*-butyl-glutamic acid were coupled to 1-N-benzyl-4-aminopiperidine (1.1 eq) using EDC hydrochloride/HOBT (1.2 eq each) as coupling reagents in presence of 2.4 eq of DIPEA in 10 mL of anhydrous DMF at RT for 18 h. Reaction was quenched by adding 30 mL of water and extracting into 15 mL of ethyl acetate (three times). The organic extracts were washed with 15 mL of 1 N NaOH (3 \times), water (1 \times), and brine (1 \times). The ethyl acetate extract was dried using anhydrous sodium sulfate, filtered and evaporated under vacuum to provide a slightly orange oil. The crude product was purified by silica gel column chromatography using ethyl acetate as mobile phase. Fractions were collected and solvent evaporated to give intermediates a or b (85 and 79% yield, respectively). TLC showed a single spot. Intermediate a was then selectively N-deprotected using 4 M HCl in dioxane at 0°C for 30 min (21). Solvent was evaporated under vacuum at 4°C, and compound triturated with ethyl acetate to yield a white solid in almost quantitative yield (c). MS showed appropriate peaks: [M+1] 376.25, [M+Na] 398.25. Alternatively, intermediate b was N-deprotected using 25% trifluoroacetic acid (TFA) in anhydrous dichloromethane (DCM) until bubbling subsided (around 1 h). Solvent was evaporated under vacuum and compound triturated with hexanes to give an off-white solid in almost quantitative yield (d, [M+1] 410.52, [M+Na] 432.52). Second, 1.1 mmol of c or d were reacted with 0.95 eq of CDCA-OBT ester, along with 1 eq of DIPEA in 15 mL of anhydrous DMF at RT overnight. Reaction was terminated



Scheme 1. Synthetic approach to obtain compounds 1–3 from N1-benzyl-4-aminopiperidine.

by adding 50 mL of water and extracting into ethyl acetate (3 × 15 mL). The combined organic extracts were washed with 15 mL of 1 N NaOH (3 ×), water (1 ×), and brine (1 ×). Solvent was evaporated under vacuum and crude products purified by silica gel column chromatography using mixtures of ethyl acetate:methanol as mobile phase. TLC showed a single spot free of impurities. Products 1 and 2 were obtained as off-white solids in 75 and 81% yield, respectively. MS showed appropriate peaks: 1, [M+1] 750.53, [M+Na] 772.53; 2, [M+1] 784.52, [M+Na] 806.52. Product 1 was subsequently deprotected by stirring in 98% formic acid for 3 h at RT (22). Solvent was evaporated under vacuum and crude product was purified by ion-exchange chromatography in Amberlyst-OH (MS showed presence of impurity consistent with the methyl ester of 1) and slowly eluted with 10% water in methanol. Fractions were collected, and solvent evaporated under vacuum to give product 3 as a colorless oil in 35% yield. [M-1] 692.43.

Procedure B. Procedure B was applied to synthesize products 4–23, as shown in Scheme 2. These compounds were synthesized by reacting the appropriate amino or aminomethyl-1-N-boc-piperidine probe and CDCA-glu or UDCA-glu, as follows. In general, 250 mg of probe were reacted with 0.95 eq of CDCA-glu, along with 1.0 eq of HBTU and 1.0 eq of DIPEA in 15 mL of anhydrous DMF at RT overnight. Reaction was stopped by adding 50 mL of water and extracting into ethyl acetate (15 mL, 3 ×). The combined organic extracts were washed with 15 mL of 1 N NaOH (3 ×), 1 N HCl (3 ×), water (1 ×), and brine (1 ×). The ethyl acetate extract was dried over anhydrous sodium sulfate and evaporated under vacuum. Crude intermediates (e-n) were purified by column chromatography using mixtures ethyl acetate:methanol or DCM:methanol as required.

Cationic derivatives 4–13 were obtained after removal of boc protecting group on piperidine nitrogen from e-n by dissolving the appropriate intermediate in 1 part of anhydrous dioxane in presence of thioanisole as scavenger. Next, 4

parts of 2 N HCl in ether were rapidly added via syringe and the resulting suspension was vigorously stirred for 30 min (at this time the hydrochloride salt precipitated). Solvent was evaporated under vacuum and product was triturated several times with ethyl acetate. The Hydrochloride salts of products 4–13 were obtained in yields ranging from 69–82%. MS: [M+1] for 4 through 13 were 694.47, 694.47, 708.49, 708.49, 694.47, 694.47, 708.49, 708.49, 708.49, and 708.49, respectively. See supplemental materials for NMR data.

Zwitterionic products 14–23 were obtained by catalytic hydrogenolysis of benzyl esters of intermediates (e-n) using 10% Pd/charcoal in ethanol at 50 psi for 1 h. Suspension was filtered through Celite® and solvent evaporated under vacuum. Acidic intermediates were then subjected to removal of boc protecting group as described for products 4–13, obtaining the hydrochloride salts of derivatives 14–23 in yields ranging from 54–79%. MS: [M+1] for 14 through 23 were 604.42, 604.42, 618.44, 618.44, 604.42, 604.42, 618.44, 618.44, 618.44, and 618.44, respectively. See supplemental materials for NMR data.

Procedure C. Anionic conjugates 24–29 were synthesized from either N-acyl-4-aminopiperidine (24–27), N-phenyl-4-aminopiperidine (28), or cyclohexylamine (29) as follows. N-acyl-4-aminopiperidine probes were synthesized by reacting 4-Z-aminopiperidine and corresponding acylating agent (acyl chloride 24, benzoyl chloride 25, and phenylacetyl chloride 26), along with DIPEA and DMAP in THF at RT overnight (23). This procedure failed when pivaloyl chloride (27) was used as acylating agent. The 1-N-pivaloylamide was then obtained by reacting 4-Z-aminopiperidine and pivaloyl chloride along with EDC in presence of DMAP as previously described (24). Next, 4-Z-aminopiperidine-1-N-acyl amide was deprotected by catalytic hydrogenation with 10% Pd/C in ethanol. 1-N-phenyl-4-aminopiperidine was synthesized by reductive amination of 1-N-phenylpiperidin-4-one as previously described (25). Nitrogenated probes were then reacted with CDCA-glu, purified by column chromatography, and debenzylated by catalytic hydro-

genolysis, as described above in procedure B. Derivatives 24–29 were obtained in yields ranging from 45–86%. MS: [M-1] for 24 through 27 were 706.45, 720.47, 644.44, and 687.48, respectively. MS: [M+1] for 28 was 680.46. MS: [M-1] for 29 was 601.43. See supplemental materials for NMR data.

Cell Culture and Assay Methods

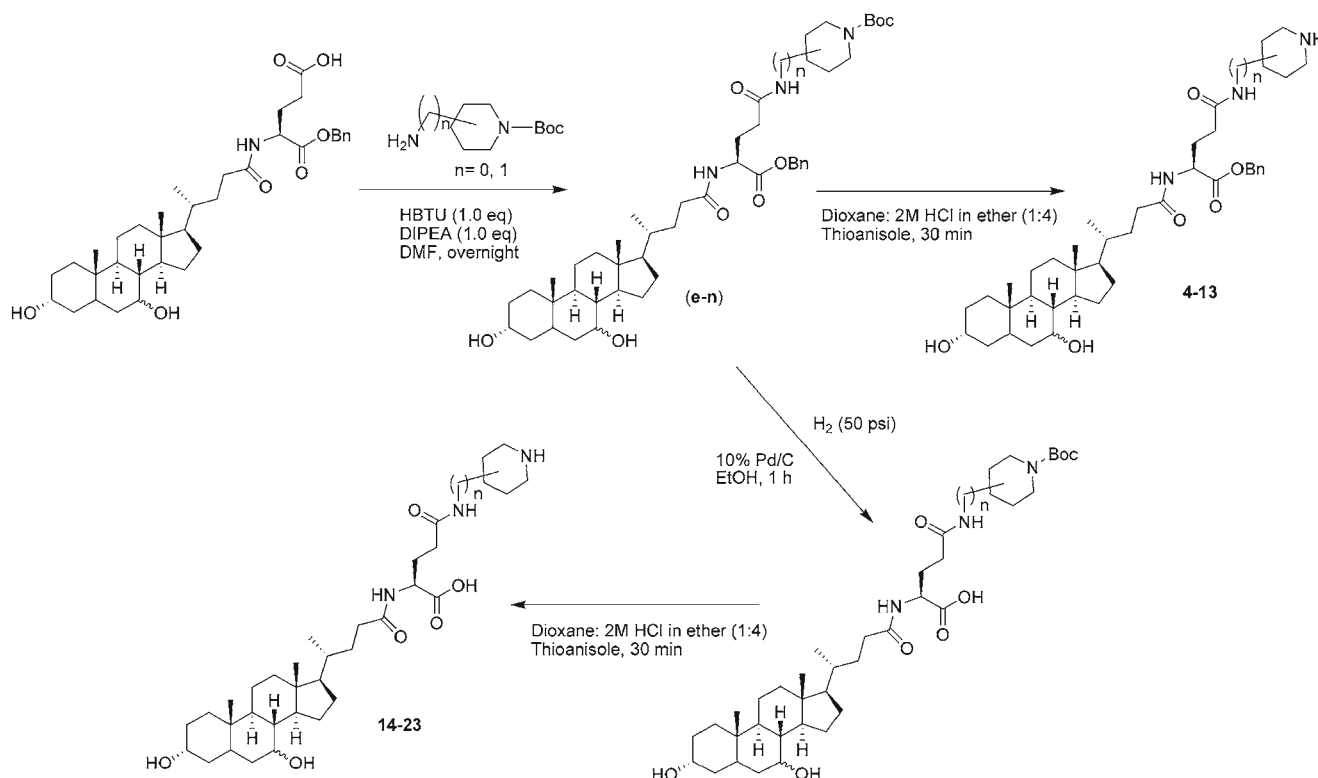
Cell Culture. Stably-transfected hASBT-MDCK cells were cultured as previously described (26). Briefly, cells were grown at 37°C, 90% relative humidity, 5% CO₂ atmosphere and fed every 2 days. Culture media consisted on DMEM supplemented with 10% FBS, 50 units/mL penicillin, and 50 µg/mL streptomycin. Geneticin was added at 1 mg/mL to maintain selection pressure. Cells were passaged after 4 days or after reaching 90% confluency.

Inhibition Assay. To characterize hASBT binding affinities, *cis*-inhibition studies of TCA uptake were conducted as described. Stably-transfected hASBT-MDCK cells were grown on 12-well plates (3.8 cm², Corning, Corning, NY) and grown under conditions described above. Briefly, cells were seeded at a density of 1.5 million/well and induced with 10 mM sodium butyrate 12–15 h at 37°C prior to study on day 4. Cells were washed thrice with Hank's balanced salt solution (HBSS) prior to assay. Studies were conducted at 37°C, 50 rpm for 10 min in an orbital shaker, as 10 min has been shown to be a suitable sample time (26). Uptake buffer consisted of HBSS, which contained 137 mM NaCl (pH 6.8) and was supplemented with 1 g/L of glucose. Cells were exposed to donor solutions containing substrate (2.5 µM

TCA+0.5 µCi/ml [³H]-TCA) and inhibitor (1–200 µM, *n*=3) for 10 min. After this time, donor solution was removed and cells were washed three times with chilled sodium-free buffer (where NaCl was replaced by 137 mM tetraethylammonium chloride). Cells were lysed using 250 µL of 1N NaOH and allowed to stand for at least 2 h. After that time cell lysate was neutralized with 250 µL of 1N HCl. Lysate was then counted for associated radioactivity (i.e [³H]-TCA) using an LS6500 liquid scintillation counter (Beckmann Instruments, Inc., Fullerton, CA). Inhibition data were analyzed in terms of inhibition constant *K_i* as described below.

Kinetic Analysis. TCA uptake inhibition data were fitted to Eq. 1 in WinNonlin 5.2 (Pharsight, Mountain View, CA). Equation 1 is a modified version of the classical competitive inhibition model, but accounts for the presence of an aqueous boundary layer (7). Only *K_i* was estimated from Eq. 1, while other parameters (*J_{max}*, *P_p*, and *K_i*) were obtained from parallel TCA uptake studies performed on the same occasion (27). Across occasions, *J_{max}* varied from 3.91 × 10⁻⁴ nmol/cm²/s to 4.32 × 10⁻⁴ nmol/cm²/s, while *P_p* varied from 0.890 × 10⁻⁷ cm/s to 3.00 × 10⁻⁷ cm/s. TCA *K_t* was set to 5.03 mM, as obtained from pooled kinetic analysis of historical TCA uptake studies. *P_{ABL}* was set to 1.5 × 10⁻⁴ cm/s (28).

$$J = \frac{P_{ABL} \cdot \left(\frac{J_{max}}{K_i \left(1 + \frac{I}{K_i} \right) + S} + P_p \right)}{P_{ABL} + \frac{J_{max}}{K_i \left(1 + \frac{I}{K_i} \right) + S}} \cdot S \quad (1)$$



Scheme 2. Synthetic approach to obtain compounds 4–23 from N1-boc protected aminopiperidine probes.

Computational Methods

Molecular Dynamics. Molecules were built using the program CHARMM (29) with the all-atom CHARMM general force field (CGenFF, K.V. and A.D.M., Work in progress). Each molecule was subjected to 1,000 steps of steepest descent (SD) and 500 steps of Newton–Raphson (NRAP) energy minimization in the gas phase to a gradient of 10^{-4} kcal/mol/Å. CHARMM-minimized structures were subjected to replica exchange molecular dynamics (MD) simulations to obtain the conformational distribution of each molecule (19,30–32). Replica exchange MD simulations involved 20 ns simulations with four replicas of each molecule between 300 K and 400 K using an exponential scale (300 K, 330 K, 363 K, and 400 K). Exchange of replicas was attempted after every 250 MD steps. MD simulations were performed using Langevin dynamics (33) with an integration time step of 0.002 ps and the aqueous solvation was modeled implicitly via the Generalized Born Continuum Solvent Model (GBMV) (34,35). SHAKE was applied to all covalent bonds involving hydrogens (36). Conformations saved every 20 ps and obtained from all 4 replicas were used for the analysis. The protonation states of the ionizable groups present in the molecules were determined based on the experimental pH of 6.8. All free α -acids at R₁ (Table I) were assumed to be deprotonated and the basic nitrogens on piperidine rings were assumed to be protonated.

Model Development. Several physico-chemical descriptors were calculated for the molecules. The average electrostatic contribution to the solvation free energy (GBener) was calculated using the GBMV implicit solvent model. Solvent accessible surface area (SASA) and polar surface area (PSA) were calculated using the Lee and Richards method as implemented in CHARMM (37). The radius of the solvent molecule was taken as 1.4 Å, which approximates the radius of a water molecule (38). PSA for each compound was calculated by adding the contribution of polar atoms (N, O and their covalently bound hydrogen atoms). SASA and PSA were also calculated exclusively for the region beyond C-24 of the bile acid conjugates (SASA-side and PSA-side) to access the contribution of the substituents alone towards the interaction with hASBT. GBener and the surface area terms were obtained as averages over all the conformations saved from the MD simulations. Molecular weight (MW), molar refractivity (MR), number of rotatable bonds (brotN), partition coefficient (logP(o/w)) and flexibility parameter (KierFlex) were calculated using the program MOE (39).

In order to elucidate the effect of the aminopiperidine group on the rest of the molecule 10 distances (O3-AS, O7-AS, AS-BC, AC-C20, O3-OA, O7-OA, N1-OA, N1-CG, N1-O3 AND N1-O7), 18 virtual valence angles (O3-AS-BC, O3-BC-AS, BC-O3-AS, O3-AS-C20, O7-AS-BC, O7-BC-AS, BC-O7-AS, O7-AS-C20, O7-C20-AS, C20-O7-AS, O3-OA-N1, O3-C20-N1, O7-OA-N1, O7-C20-N1, C20-N1-OA, OA-N1-CG, OA-CG-N1, CG-OA-N1) and two virtual dihedral angles (OA-C24-CG-N1 and OG-NG-CG-N1) were calculated involving the pharmacophoric feature points shown in Fig. 1 for all conformations obtained from the MD trajectories. 1-D probability distributions of all structural descriptor were derived with a bin size of 0.1 Å, 1° and 5° for distances,

valence angles and dihedral angles, respectively. In order to complement the quantitative model, several 2D probability distributions of the structural descriptors in combinations of two were also derived. 2D probability distributions were obtained using a bin size of 0.2 Å, 3° and 5° for distances, angles and dihedrals, respectively.

Overlap coefficient (OC) of each structural descriptor were used to quantify the extent of similarity of the regions of conformational space as defined by the various distances, angles and dihedrals sampled by each compound. OC for continuous probability density functions is given by Reiser and Faraggi (40),

$$OC = \sum_{i=1}^N \min \{f_1(x), f_2(x)\} dx \quad (2)$$

where $f_1(x)$ and $f_2(x)$ are probability density functions of the two distributions. For discrete probability functions Eq. 2 can be written as,

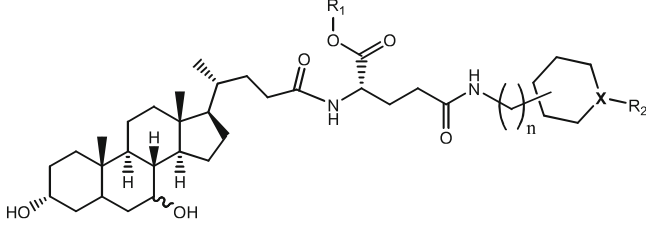
$$OC = \sum_{i=1}^N \min(P_i^A, P_i^B) \quad (3)$$

where P_i^A and P_i^B are the probability in bin i for compounds A and B and N is the total number of bins. OC values for each compound were calculated with respect to the most potent inhibitor in the data set (Compound 9, $K_i=0.953 \mu\text{M}$). The overlap coefficients of all molecules were regressed with respect to their inhibition coefficients. Similarly physico-chemical descriptors were also subjected to linear regression against the respective K_i values of the molecules. Microsoft Excel 2000 was used for the regression analysis. The structural and physico-chemical descriptors were then combined to make a complete set of molecular descriptors for the model development. All descriptors having linear $r^2 > 0.01$ were selected for multivariable regression analysis.

Molecular descriptors were initially subjected to multivariable regression in all possible combinations of two. However, any pair of descriptors having an internal correlation (i.e. correlation between those two descriptors) greater than 0.8 was discarded from multivariable regression. The correlation matrix used to determine the internal correlation between descriptors is presented in Table S1 of the Supplemental Material. All pairs of descriptors with p -values of the independent variables, as well as the intercept, less than 0.05 were selected for further analysis. Next, additional descriptors were added one at a time to each selected pair of descriptors. This approach was applied to all selected groups until all possible combinations of descriptors yielded p -values greater than 0.05. All the combinations of descriptors with p -values less than 0.05 formed the set of candidate models.

Model Evaluation. AIC_C (modified Akaike Information Criterion) (41,42) analysis was performed on the candidate models to rank them in the order of descending probability of being the best model. AIC_C is the original AIC equation with a second order correction for small sample size and is given by:

$$AIC_C = n \ln \frac{RSS}{n} + 2k + \frac{2k(k+1)}{n-k-1} \quad (4)$$

Table I. Binding Affinities of Piperidine Conjugates of Glutamyl-chenodeoxycholic Acid to hASBT


Comp.	Group ^a	7-OH	X	R ₁	R ₂	-CONH- ^b	n	Experimental K _i (μM ± SEM)	Predicted K _i (μM)
1	G1	α	N	-C(CH ₃) ₃	-CH ₂ C ₆ H ₅	4	0	1.45 (0.32)	6.41
2	G1	α	N	-CH ₂ C ₆ H ₆	-CH ₂ C ₆ H ₅	4	0	2.26 (0.46)	3.13
3	G2	α	N	-H	-CH ₂ C ₆ H ₅	4	0	1.66 (0.31)	-
4	G1	α	N	-CH ₂ C ₆ H ₅	-H	4	0	3.77 (1.08)	4.77
5	N.I.	β	N	-CH ₂ C ₆ H ₅	-H	4	0	19.5 (3.3)	-
6	G1	α	N	-CH ₂ C ₆ H ₅	-H	4	1	4.32 (1.06)	8.52
7	N.I.	β	N	-CH ₂ C ₆ H ₅	-H	4	1	17.5 (2.7)	-
8	G1	α	N	-CH ₂ C ₆ H ₅	-H	3(β)	0	1.39 (0.27)	4.34
9	G1	α	N	-CH ₂ C ₆ H ₅	-H	3(α)	0	0.953 (0.192)	-
10	G1	α	N	-CH ₂ C ₆ H ₅	-H	3(β)	1	9.92 (1.54)	9.96
11	G1	α	N	-CH ₂ C ₆ H ₅	-H	3(α)	1	15.7 (3.3)	6.33
12	G1	α	N	-CH ₂ C ₆ H ₅	-H	2(β)	1	10.0 (1.4)	10.89
13	G1	α	N	-CH ₂ C ₆ H ₅	-H	2(α)	1	9.72 (1.48)	5.00
14	G2	α	N	-H	-H	4	0	1.76 (0.37)	-
15	N.I.	β	N	-H	-H	4	0	28.1 (4.3)	-
16	G2	α	N	-H	-H	4	1	2.46 (0.50)	-
17	N.I.	β	N	-H	-H	4	1	16.3 (3.1)	-
18	G2	α	N	-H	-H	3(β)	0	4.89 (0.90)	-
19	G2	α	N	-H	-H	3(α)	0	2.19 (0.55)	-
20	G2	α	N	-H	-H	3(β)	1	3.68 (1.23)	-
21	G3	α	N	-H	-H	3(α)	1	31.8 (5.00)	31.01
22	G3	α	N	-H	-H	2(β)	1	18.7 (2.8)	20.66
23	G3	α	N	-H	-H	2(α)	1	17.9 (3.4)	15.91
24	N.I.	α	N	-H	-C(O)CH ₃	4	0	22.6 (5.4)	-
25	N.I.	α	N	-H	-C(O)C ₆ H ₅	4	0	11.9 (1.8)	-
26	N.I.	α	N	-H	-C(O)CH ₂ C ₆ H ₅	4	0	8.30 (1.17)	-
27	N.I.	α	N	-H	-C(O)C(CH ₃) ₃	4	0	6.06 (1.12)	-
28	N.I.	α	N	-H	-C ₆ H ₆	4	0	5.35 (0.93)	-
29	N.I.	α	C	-H	H	-	0	1.74 (0.37)	-

K_i values were predicted using Model 1 (Table III)

^aThis column denotes which group the corresponding compound belongs to (G1, G2, or G3). N.I. indicates that the corresponding compound was 'Not Included' for CSP-SAR model

^bValue indicates position of amide relative to piperidine nitrogen

where n = sample size, k = number of variables including the intercept, RSS = residual sum of squares. For larger sample sizes, the last term of the Eq. 3 vanishes to yield the original AIC equation. The model having the smallest AIC value among all the candidate models is designated as the best model. The AIC weight factor (W_i) was also calculated to measure the probability of the model having the smallest AIC value to be the best in the entire candidate model set (43). W_i was calculated using Eq. 4

$$W_i = \frac{\exp(-\Delta_i/2)}{\sum_{i=1}^N \exp(-\Delta_i/2)} \quad (5)$$

where N = number of candidate models and $\Delta_i = AIC_i - \min(AIC)$.

Candidate models yielding more than 5% probability of being the best model were selected for further analysis. The

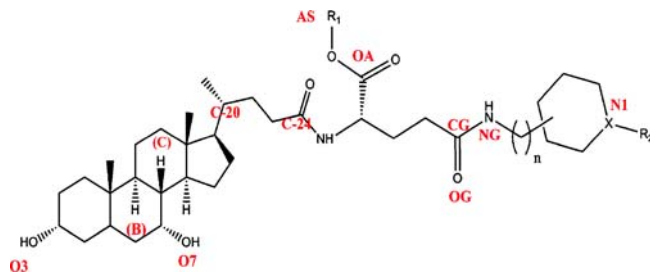


Fig. 1. Structural features in aminopiperidine-glutamyl-CDCA conjugates used for pharmacophore development. Notation is as follows: O3, position of 3-OH; O7, position of 7-OH; BC, centroid of B and C rings of steroidal nucleus; C20, location of C-20; C24, location of C-24 amide carbon; OA, centroid of carboxylic acid oxygens on α -acid; AS, centroid of the substituent (R_1) attached to α -acid; CG, position of the carbon atom of γ -CO; OG, position of the oxygen atom of γ -CO; NG, position of the amide nitrogen attached to γ -CO; NI, position of protonated piperidine nitrogen.

robustness of the predictive power of the selected models was tested by the leave-one-out cross validation analysis. In this method, one of the observations was kept as validation data with the rest of the data used to form the training set. The inhibition constant of the test compound was then predicted using the model based on the training set compounds. This procedure was repeated for all other compounds until each of them served once as a test compound. The predictive power of the model was then assessed by calculating the cross-validated r^2 or Q^2 using Eq. 4.

$$Q^2 = 1 - \frac{\sum (y_{pred} - y_{obs})^2}{\sum (y_{obs} - y_{mean})^2} \quad (6)$$

RESULTS AND DISCUSSION

Presented are results from the synthetic efforts and the subsequent hASBT inhibition experiments, followed by CSP-SAR model results. Synthetic efforts yielded in improved approaches allowing for a relatively large number of conjugates to be studied. hASBT inhibition experiments were first interpreted qualitatively allowing for general associations between compound features and inhibition data to be derived. More rigorous CSP-SAR model results are separated into preliminary model development, quantitative CSP-SAR models, and qualitative CSP-SAR model.

Synthesis. We successfully synthesized a series of 29 piperidine conjugates of CDCA-glu and UDCA-glu in good to excellent yields. In the remainder of the text CDCA-glu denotes either α -benzyl ester glutamic acid CDCA amide or glutamic acid CDCA amide and UDCA-glu denotes either α -benzyl ester glutamic acid UDCA amide or glutamic acid UDCA amide.

N1-boc protecting group offered the possibility for selective removal from intermediates e-n by reaction in a mixture of anhydrous dioxane: 2N HCl in ether (1:4). This method is a modification of a previously reported methodology applied in the selective removal of boc protection in amino acids (21). When the described procedure was directly applied to e-n, products were formed in very low yields and only after very cumbersome chromatographic purification. Based on the observation that e-n were insoluble in ether, N1 was deprotected by first dissolving the intermediate in 1 part of dioxane (usually 5 mL) and then adding 4 parts of a commercial mixture of 2N HCl in ether. The reaction was rapid (30 min) and practically quantitative. The precipitation of the HCl salt not only drove the reaction to completion, but also prevented potential alkylation of hydroxyl groups in the steroidal nucleus due to side products from the cleavage of boc. The reaction was found to be compatible with both α -benzyl and α -acidic intermediates, allowing for cations 4–13 and the zwitterions 14–23 from common intermediates to be obtained (Scheme 2).

Several chemical features of the conjugates were systematically varied and their impact on hASBT binding assessed: 1) position of the piperidine nitrogen (N1) relative to the γ -carbonyl (γ -CO) in the glutamic acid linker (e.g. 4 and 8);

2) distance of N1 to γ -CO and conformational freedom of piperidine probe by inclusion of a methylene bridge between γ -CO and ring (e.g. 8 and 10); 3) charge state of N1 (e.g. 14 and 24); 4) stereochemistry of bridging carbon between γ -CO and probe (e.g. 12 and 13); 5) presence of a bulky group on N1 (e.g. 2 and 4); 6) charge state of α -acid (e.g. 7 and 17); and 7) stereochemistry of 7-hydroxyl (7-OH) in the steroidal nucleus (e.g. 14 and 15).

Table I shows the general structure of compounds and their binding affinities to hASBT. Compounds in Table I are organized based on stereochemistry of 7-OH on steroidal nucleus (α in CDCA, β in UDCA), nature of substituent on α -acid (R_1), nature of substituent in N1 (R_2), and position of bridging carbon between probe and γ -amide relative to N1 ($-\text{CONH}-$). This column also shows the relative stereochemistry of the substituent on the piperidine ring. Columns labeled X and n refer to the presence/absence of nitrogen on the ring and to the presence/absence of methylene bridge between γ -amide and probe, respectively.

hASBT Inhibition SAR. All compounds were found to be potent inhibitors of hASBT, regardless of charge state, with K_i values ranging from 0.953 (9) to 31.8 (21) μM . Analysis of the activity data suggested the following general contributions towards hASBT inhibition. Presence of a benzyl group on N1 did not affect potency (1–4), nor did the presence or absence of the nitrogen in the piperidine ring (14 vs 29). However, amidation of the piperidine nitrogen reduced binding affinity with a small group having a bigger impact (24 vs 27). Benzylation of the α -acid had only a minor effect on activity (~ 2 fold decrease) when compared to the respective free acids. This observation suggests that a negative charge around C-24 is not necessary for binding. A 7-OH in an α orientation was found to be preferable for binding, rather than the β configuration, resembling the difference in activity between unconjugated CDCA and UDCA (7). In fact, this difference mimicked CDCA and UDCA difference when a negative charge was proximal to C-24 (i.e. free α -acid).

A smaller distance between the positive charge and the γ -amide (absence of a methylene bridge) was preferable for binding, especially for 3 α -substituted piperidine (19 vs 21) possibly due to the formation of an internal electrostatic interaction between N1 and the α -acid, as discussed below. Position of the bridging carbon between the γ -amide and the piperidine probe (2, 3, or 4 relative to N1) was found to impact activity, especially for zwitterionic derivatives. The potency ranking was 4>2>3 with stereochemistry modulating activity of the 3-substituted derivatives. Thus 20 (i.e. 3 β) was 10-fold more potent than 21 (3 α).

hASBT Inhibition CSP-SAR Model Development: Preliminary Evaluation. To develop more rigorous models relating structure to affinity, derivatives 1–4, 6, 8–14, 16, and 18–23 were subjected to CSP-SAR model analysis. These derivatives were chosen based on their structural similarity. More specifically, selected compounds contained a basic nitrogen in the piperidine ring. The ten compounds that were excluded either lacked a basic nitrogen or were conjugates of UDCA-glu. Preliminary analysis indicated that the activity of the six compounds lacking a basic nitrogen require a different model. The four UDCA-glu derivatives were excluded since their

activities, relative to the corresponding CDCA-glu conjugates, were simply explained by the stereochemistry of the 7-OH on UDCA *versus* CDCA. A basic nitrogen (*versus* non-basic) and an α -OH configuration at C-7 (*versus* β -OH) promote binding, such that lack of these elements caused lack of model fit.

In the absence of experimental data of the receptor 3D structure as well as that of inhibitor-transporter complexes, ligand-based drug design approaches must be applied (44). The CSP method is a novel ligand-based approach that offers the advantage of including all possible conformations sampled by the training set molecules in model development. The CSP-SAR method using replica-exchange MD simulations is particularly advantageous for compounds with high molecular flexibility such as 1–29 (20) as the replica-exchange methodology greatly facilitates the sampling of conformational space by the ligands under study. This has been shown to be effective for opioid peptides (19) and facilitates application of the CSP method to the bile acid conjugates, whose side chains are relatively large and flexible.

In this study, the hASBT inhibitory potency (i.e. K_i) was the activity parameter. 11 structural features (Fig. 1) were used to define the accessible conformational space sampled by each molecule. As discussed in the computational methods section, 9 being the most potent agent was used as the reference molecule to calculate the OC values for all remaining compounds included in the CSP-SAR model development. Table S2 in supplemental material represents the values of physico-chemical descriptors and the overlap coefficients of structural descriptors.

hASBT Inhibition CSP-SAR Model Development: Quantitative Model. Quantitative analysis of the overlap coefficients of various descriptors revealed two general trends. First, the conformational properties of the compounds with protected α -acid and $K_i < 16 \mu\text{M}$ were very similar while being distinct from compounds with free α -acid and $K_i > 16 \mu\text{M}$. Second, the

conformational properties of compounds with free α -acid and $K_i < 16 \mu\text{M}$ were not only different from the corresponding α -benzyl esters, but also highly similar with molecules having $K_i > 16 \mu\text{M}$. This observation led us to segregate compounds into three different groups. G1 compounds were protected α -acids with $K_i < 16 \mu\text{M}$ (1, 2, 4, 6, and 8–13). G2 compounds were free α -acids with $K_i < 16 \mu\text{M}$ (3, 14, 16, and 18–20). G3 compounds were free α -acids with $K_i > 16 \mu\text{M}$ (21–23).

Fig. 2 shows the 1-D probability distributions of the compounds in sets G1, G2, and G3 for the following four structural descriptors: distance between piperidine nitrogen and 7-hydroxyl in the steroidal nucleus (N1-O7; panel A); angle between C-20 in CDCA, piperidine nitrogen, and centroid of α -acid oxygens in glutamic acid linker (C20-N1-OA; panel B); angle between 7-hydroxyl, centroid of α -substituent in glutamic acid linker, and centroid of rings B and C in cholestane skeleton (O7-AS-BC; panel C); and angle O7-AS-C20 (panel D). In general, compounds in G1 were more flexible and sampled a larger conformational space than G2 and G3. Further inspection of Fig. 2 revealed that G2 sampled a conformational space more similar to G3 than that of G1, despite G2 displaying K_i values closer to G1. G2 and G3 showed almost identical conformational distributions for parameters O7-AS-BC and O7-AS-C20 (Fig. 2C and D, respectively). However, G2 and G3 displayed different conformational sampling for parameters N1-O7 and C20-N1-OA (Fig. 2A and B, respectively). This observation demonstrated that G2 and G3 compounds were conformationally very similar, except for the position of the N1 with respect to the rest of the molecule. Interestingly, Fig. 2A revealed a small range in N1-O7 (16 Å–18 Å) that was significantly sampled by G1 and G2, but not by G3. Despite the presence of sampling differences for descriptors in Fig. 2 between G2 and G3, the narrow range of activity data (0.953 μM to 31.8 μM) precluded the development of a CSP-SAR model to distinguish G2 and G3 compounds. Hence, it was decided to exclude G2 from quantitative model development and

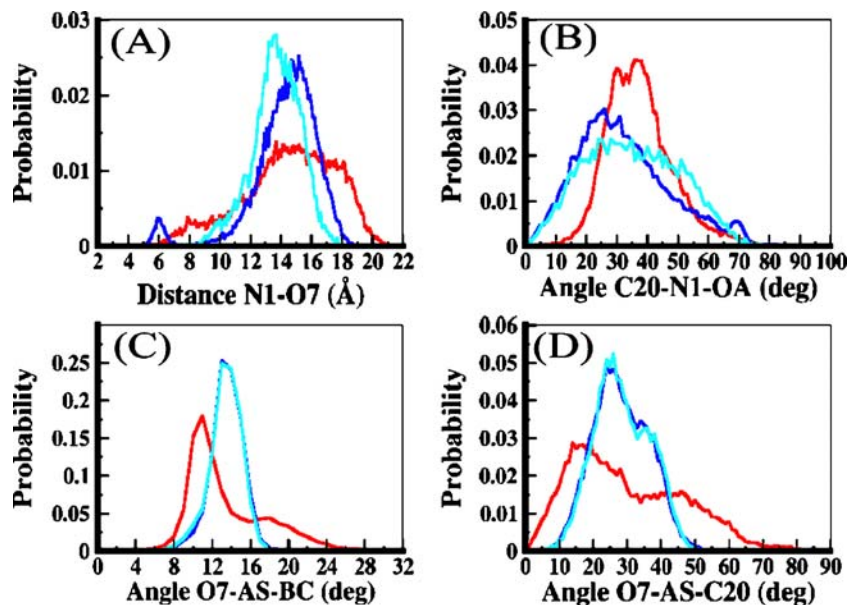


Fig. 2. 1-D probability distribution of **A** distance N1-O7, **B** angle C20-N1-OA, **C** angle O7-AS-BC and **D** angle O7-AS-C20. The red, blue and turquoise lines represent the distribution of compounds in group G1, G2 and G3, respectively.

include only G1 and G3. However, as described below, qualitative CSP-SAR analysis was able to propose a model to explain the high affinity of G2 compounds. Attempts to determine a quantitative model based on G2 and G3 were not successful ($r^2 \leq 0.301$).

Quantitative modeling of the G1 and G3 compounds was initiated using single variable regression analysis of all structural and physico-chemical descriptors with respect to K_i . Analysis yielded 24 structural descriptors with linear $r^2 > 0.01$. Eight distances (N1-OA, N1-O3, N1-O7, N1-CG, O3-AS, O7-AS, AS-BC and AS-C20), 14 virtual angles (O7-AS-BC, BC-O7-AS, O7-BC-AS, O7-AS-C20, C20-O7-AS, O7-C20-AS, O3-AS-BC, BC-O3-AS, O3-BC-AS, OA-N1-CG, C20-N1-OA, CG-OA-N1, OA-CG-N1 and O7-C20-N1), and two virtual dihedral angles (OA-C24-CG-N1 and OG-NG-CG-N1) (Table II). Physico-chemical descriptors SASA-side, SASA, and PSA showed the highest linear correlations. The negative coefficient of SASA-side and SASA suggested that bulky groups on these molecules favor inhibition of hASBT (i.e., decrease K_i). Consistent with this result were the negative coefficients for MW and MR. Notably, the positive

coefficient of PSA suggested that an increase in polar surface area disfavors hASBT inhibition. Similarly, the positive coefficient of PSA-side and negative coefficient of GBener indicated increased polarity to decrease activity. This observation was supported by the negative coefficient of logP (o/w). The flexibility-related descriptors KierFlex and brotN also displayed negative coefficients, indicating that increase flexibility promotes hASBT inhibition. However, their contribution may be related to larger side chains having more rotatable bonds.

Surface area terms gave the highest correlations and were subjected to further analysis to investigate the contribution of the surfaces areas of different regions of the molecules to inhibitory activity. Regression analysis of the surface area terms for the entire molecule, as well as the region beyond C-24 side chain, correlated with K_i , where larger surface area of entire molecule or side chain promoted binding. Meanwhile, the total and polar surface area terms for the cholestane moiety alone yielded poor correlation (0.339 and 0.0449, respectively). These observations indicate that the side chain beyond the C-24 region dominated the polar and

Table II. Results from Single Variable Linear Regression of Molecular Descriptors for Piperidine Conjugates of CDCA-glu in Training Set

Molecular descriptor (units)	Linear r^2	Coeff. independent variable ^a	SEM of Coeff. independent variable	Coeff. Intercept ^a	SEM of Coeff. Intercept
SASA-side (\AA^2)	0.673	-0.0745	0.0164	56.1	10.2
SASA (\AA^2)	0.659	-0.0861	0.0196	104	21
PSA (\AA^2)	0.654	0.310	0.133	-49.3	25.7
O7-AS-C20	0.640	-56.0	13.3	52.6	10.1
logP(o/w)	0.637	-6.40	1.52	49.7	9.5
O7-AS-BC	0.636	-40.4	9.7	41.3	7.5
OA-CG-N1	0.631	-29.3	7.1	21.8	3.2
OA-N1-CG	0.623	-26.7	6.6	21.7	3.2
MR (m^3)	0.620	-4.15	1.03	91.8	20.2
MW (g/mol)	0.618	-0.137	0.034	105	24
CG-OA-N1	0.603	-29.3	7.5	21.4	3.3
O7-C20-AS	0.581	-56.8	15.3	53.3	11.6
N1-OA	0.548	-26.9	7.7	20.4	3.4
KierFlex	0.510	-10.4	3.2	131	37
PSA-side (\AA^2)	0.489	0.384	0.124	-25.2	11.7
O7-AS	0.484	-31.0	10.1	31.6	7.1
C20-O7-AS	0.481	-47.4	15.6	46.5	12.0
O3-AS	0.454	-48.2	16.7	44.3	11.9
brotN	0.429	-6.01	2.19	85.2	27.3
AS-BC	0.423	-28.2	10.4	29.4	7.3
AS-C20	0.411	-27.9	10.6	29.7	7.5
BC-O7-AS	0.382	-98.2	39.5	90.2	32.1
OA-C24-CG-N1	0.373	-24.4	10.0	22.0	5.2
O7-BC-AS	0.326	-102	46.4	94.8	38.4
O3-AS-BC	0.286	-97.6	48.7	93.2	41.3
N1-O3	0.268	-35.6	18.6	33.4	12.2
GBener (kcal/mol)	0.258	0.286	0.154	34.9	13.3
C20N1-OA	0.229	-27.1	15.7	28.3	10.6
N1-O7	0.208	-30.2	18.6	28.8	11.5
N1-CG	0.085	-15.1	15.6	13.8	4.25
OG-NG-CG-N1	0.087	-23.8	24.3	22.3	12.3
BC-O3-AS	0.060	-55.2	68.9	55.8	56.5
O7-C20-N1	0.0565	25.7	33.2	-9.32	25.80
O3-BC-AS	0.0544	-53.3	70.2	54.3	57.7

Regression analysis of structural descriptors (e.g. O7-AS-C20) employed the overlap coefficients with respect to 9.

^a Coefficient of independent variable and coefficient of intercept have units of μM divided by units of the molecular descriptor (e.g. $\mu\text{M}/\text{\AA}^2$ for SASA-side).

van der Waals surface interaction for the molecules studied, which impacted K_i ; of course, all modifications of molecules were performed on the side chain, with the cholestane moiety unaltered.

Overlap coefficients for all structural descriptors, except one, showed negative coefficients with respect to inhibitory activity (Table II). This association was expected as higher conformational similarity to the reference compound 9 (i.e. the most potent) promotes inhibitory potency. The only exception was the OC of O7-C20-N1, whose positive coefficient suggested that high similarity with 9 for that angle disfavors binding (see below).

Multivariable regression analysis entailed various combinations of the OCs and/or physico-chemical descriptors. A correlation matrix was derived for all the descriptors (Table S1). Combinations of descriptors with descriptor-descriptor correlation coefficients <0.8 were selected for the multivariable regression analysis. Three candidate quantitative models were obtained. Each included two descriptors (Table III). The p -values of all independent variables and the intercept of each model were less than 0.05. Candidate models were ranked by AIC_C indicators. Three models as assessed by AIC_C weight factor (W_i) had 62.2%, 19.2% and 18.6% probability of being the best model (model 1, 2, and 3, respectively). The quantitative models yielded multivariable r^2 of 0.813, 0.773, and 0.771, respectively.

Fig. 3 depicts the linear regression analysis between experimental and predicted binding affinity values from model 1. The leave-one-out model validation method was applied to the models independently. Q^2 value of model 1 was 0.726, suggesting robustness of the model. Models 2 and 3 yielded $Q^2 < 0.6$ indicating moderate robustness (0.578 and 0.592, respectively). Regression plots for model 2 and 3 are provided in Supplemental Material (Fig. S1).

The best model (model 1) combined the angle O7-AS-BC and the physico-chemical descriptor GBener. Model 2 was a combination of two topological descriptors (O7-AS-C20 and N1-O7). Model 3 is comprised of the physico-chemical parameter $\log P(o/w)$ and the structural descriptor C20-N1-OA. The negative coefficients of physico-chemical descriptors GBener (model 1) and $\log P(o/w)$ (model 3) indicate that hydrophobic character favors hASBT inhibition. Given the presence of a positive charge on N1, these results suggest that higher overall hydrophobicity may compensate for the highly polar nature of this moiety, thereby facilitating binding.

Similar to results from the single variable regression analysis, the OCs of the angles in the models showed negative coefficients. This suggested that conformational properties similar to 9 promote hASBT binding. Angles O7-AS-BC and O7-AS-C20 are related to the orientation of the α -substituent relative to the cholestane skeleton, indicating that this spatial

relationship was important for activity. In model 2, the inclusion of the N1-O7 distance indicates that the spatial relationship of the piperidine ring to the cholestane was also important for activity. Interestingly, the positive coefficient of the distance N1-O7 indicates that compounds with the distribution of this particular distance similar to that of 9 have poorer inhibitory activity, which is unexpected. In 9, the N1-O7 distribution ranges from 10 to 20 Å with the maximum at 18–19 Å (see Fig. S2, Supplemental Material). Notably, 18–19 Å is at the extrema of the distribution for all compounds in G1 (Fig. 2A), suggesting that, indeed, shorter N1-O7 distances favor inhibition. Thus, the model predicts that binding is favored by the close proximity of the piperidine moiety to the cholestane skeleton. In model 3, the structural parameter C20-N1-OA described the side chain conformation involving both the α and γ substituents. The negative coefficient indicated that compounds that sampled regions similar to that of 9 were better inhibitors. For 9 this angle ranged between 20 to 50° (Fig. S2, Supplemental Material), suggesting that preferred side chain conformations are neither fully extended nor more compacted (e.g. optimal C20-N1-OA angle). This observation is consistent with shorter N1-O7 distance (compared to 9) exhibiting favorable potency. Also in model 3 the negative coefficient with $\log P(o/w)$ again indicated that great hydrophobicity improves binding, consistent with the GBener contribution in model 1. Overall, these results from quantitative CSP-SAR models indicated that increased hydrophobicity, α substituent orientation similar to 9, and partially compacted side chain conformation promote inhibitory potency.

hASBT Inhibition CSP-SAR Model Development: Qualitative Model. Qualitative analysis was performed in order to explain the high activity of G2 compounds. Quantitative models were not achieved apparently due to the high chemical similarity of G2 with poor inhibitors (i.e. G3). Efforts to develop quantitative models included both 1-D and 2-D OC values and physico-chemical descriptors applying the methodology described in the preceding section to all three sets as well as G2 with G3. The later attempt included using compound 3 ($K_i=1.66 \mu\text{M}$) as the reference compound as this is the most potent inhibitor in the G2 set. As stated above no model with $r^2 > 0.314$ was identified (results not shown).

To obtain more detailed conformational information 2-D contour plots involving pairs of various structural descriptors were generated (Fig. 4). In the 2-D plots, red regions illustrate conformations sampled by all compounds in set G1, blue regions are those sampled by G2 and turquoise regions illustrate conformations sampled by G3. Since compounds in G1 and G2 display $K_i < 16 \mu\text{M}$, the regions sampled by the G1 and G2 represent conformations of conjugates that favor hASBT inhibition. However, conformational space

Table III. Modified Akaike Information Criterion (AIC_C) and Weight Factor (W_i) for the Candidate CSP-SAR Models

Model	Descriptor combination	Multivariable r^2	AIC_C	W_i	Q^2	F	K	A	B
1	O7-AS-BC, GBener	0.813	40.9	0.622	0.726	19.6	28.8 (± 7.12)	-80.54 (± 15.6)	-0.506 (± 0.173)
2	O7-AS-C20, N1-O7	0.773	43.2	0.192	0.578	15.3	55.0 (± 8.52)	-93.4 (± 19.4)	42.8 (± 18.7)
3	$\log P(o/w)$, C20-N1-OA	0.771	43.3	0.186	0.592	15.2	60.8 (± 9.28)	-5.94 (± 1.29)	-21.0 (± 9.13)

A and B represent the coefficient of the independent variables (molecular descriptors); K represents the coefficient of the intercept.

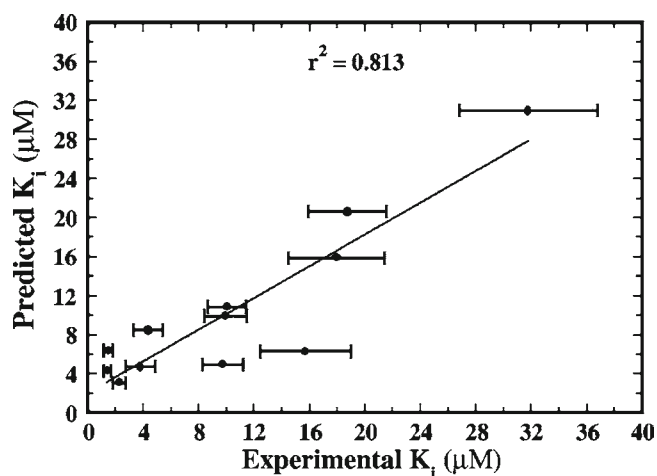


Fig. 3. Experimental vs. predicted values of K_i for the piperidine conjugates of CDCA-glu obtained from model 1.

sampled by G2 and G3 compounds overlapped significantly, consistent with the 1-D results in shown in Fig. 2.

Why do G2 compounds have greater activity than G3 derivatives, given their conformational similarity? We hypothesized that small structural modifications in the γ -substituent impact binding by introducing conformational changes in the side chain beyond C-24. For example, compounds 20 ($K_i=3.7 \mu\text{M}$) and 21 ($K_i=31.8 \mu\text{M}$) only differ in the chirality of the piperidine carbon that is linked to the

side chain (Table I). Such subtle modification is suggested to impact interactions between the γ -substituent and O7, as well as γ -substituent and OA (Fig. 4). For example, Fig. 4A clearly illustrates that conformational space defined by N1-O7 distances within 16–18 Å and C20-N1-OA angles within 25–40° was significantly sampled by G1 and G2 set, unlike G3 compounds. The G3 set samples systematically smaller N1-OA distances, smaller OA-CG-N1 angles, and larger OA-N1-CG angles than G2 (Fig. 4B and C, respectively), while other conformational properties are similar (Fig. 2). The observation that G3 compounds sampled a shorter range of N1-OA distance (<5 Å) than G1 and G2 suggested the existence of an ion-pair or salt-bridge interaction (45) between N1 and α -acid given that the G3 set include the free acid that is not accessible to the G2 compounds. This interaction appears responsible for the restricted movement of γ -substituent with respect to α -substituents thereby leading to decreased affinity. The G1 and G2 compounds also demonstrated the exclusive feature of sampling N1-CG distances within 4.5–5.5 Å, while this range of N1-CG distances was poorly accessed by G3 set. These conformational features, that are exclusively present in G1 and G2 compounds may indicate the conformations of the amino-piperidine conjugates which are favorable for binding to hASBT.

To ascertain the conformational differences between the G1, G2 and G3 sets, compounds 9 (0.953 μM), 20 (3.68 μM), and 21 (31.8 μM) were selected to represent the G1, G2, and

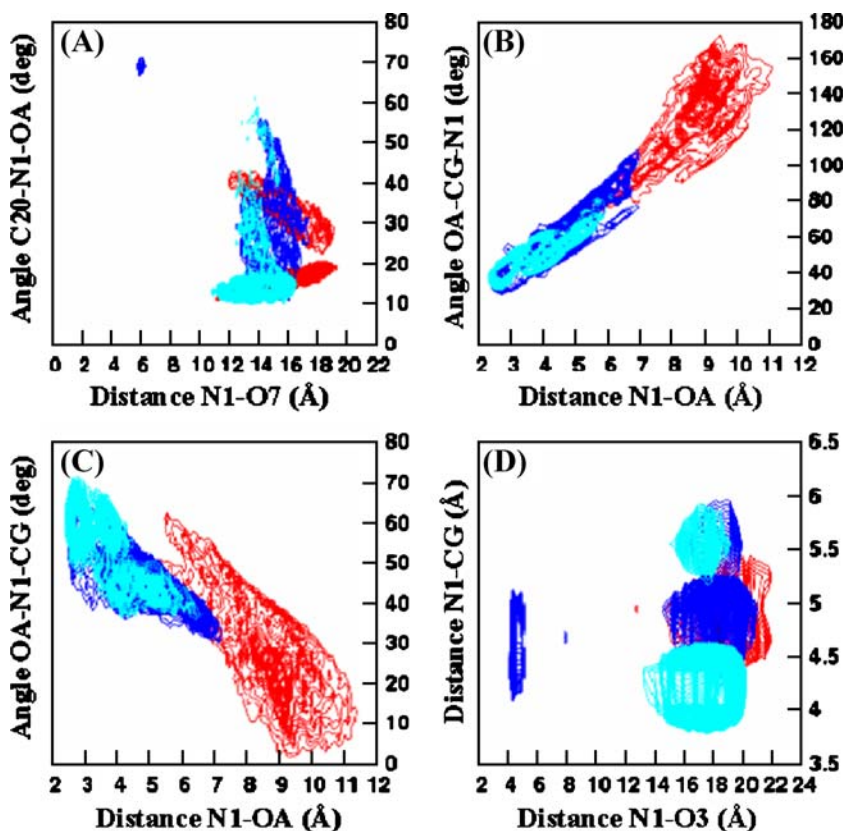


Fig. 4. 2-D contour plots for A N1-O7 & C20-N1-OA, B N1-OA & OA-CG, C N1-OA & OA-N1-CG, and D N1-O3 & N1-CG. The red, blue and turquoise contours represent the conformational space corresponding to that pair of descriptors collectively sampled by the molecules in group G1, G2 and G3, respectively.

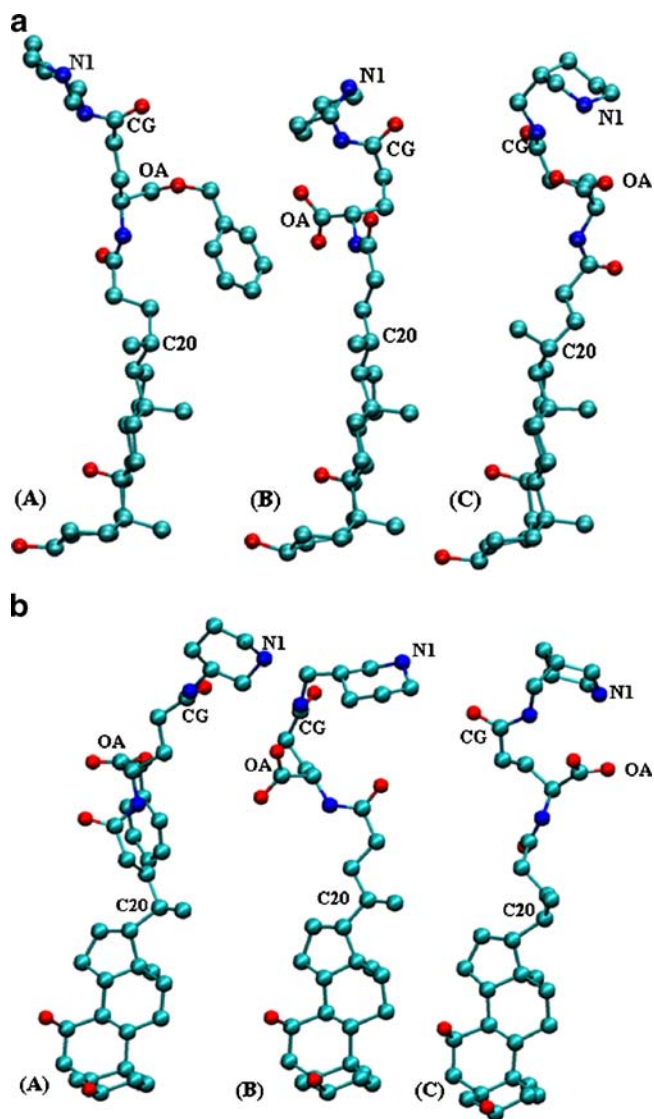


Fig. 5. Representative conformers of compound 9, 20, and 21 highlighting the similarity between 9 and 20, relative to 21 (A, B, and C, respectively). Fig. 5.1 and 5.2 represent two different views of the molecules to facilitate the pictorial comparison. Distance N1-OA >8 Å for 9 and 20, while equal to 3.18 Å for 21. This suggests the presence of salt-bridge between N1 and OA in 21. Similarly, distance N1-CG ranged from 4.5 Å to 5.2 Å for 9 and 20, and equal to 4.35 Å for 21, consistent with Fig. 4D. Angle C20-N1-OA was 34.11°, 37.18°, and 12.78° for 9, 20 and 21, respectively; angle OA-CG-N1 $>110^\circ$ for 9 and 20, while equal to 46.65° for 21.

G3 set, respectively. Being the most potent inhibitor, 9 was selected to represent the G1 set. Compounds 20 and 21 were selected to represent G2 and G3, respectively, because they showed significant difference in activity, despite of their high structural similarity. The only difference between 20 and 21 is the chirality of the linkage between methylene and piperidine ring. Compounds 9, 20, and 21 also have the piperidine nitrogen at 3-position with respect to the amide group. The conformations of the three compounds were selected based on the 2-D distributions of the structural descriptors (Fig. 4). The conformers of 9 and 20 represent the conformational space which is exclusively accessed by G1 and G2 set, and not by G3. On the other hand, the conformer of 21 represents the

most probable conformation of G3 set. By analyzing the structures of 9, 20 and 21 in Fig. 5, the conformational features in the γ -substituent region that impact binding affinity may be elucidated.

Fig. 5 shows two different orientations (Fig. 5.1 and 5.2) of 9 (in G1), 20 (in G2), and 21 (in G3) illustrating the conformational features favoring the affinity of 9 and 20 relative to 21 (A, B and C, respectively). The values for descriptors N1-OA, N1-CG, C20-N1-OA, and OA-CG-N1 for the conformer of 9 depicted in panel (A) were 10.2 Å, 4.8 Å, 34°, and 167°, respectively. The values of the same descriptors for the conformer of 20 in panel (B) were 8.2 Å, 5.2 Å, 37°, and 115°, respectively. Panel (C) represents the conformer of 21 where the same descriptors were 3.2 Å, 4.4 Å, 13°, and 47°, respectively. Analysis of Fig. 5 highlights the conformational similarity between 9 and 20 around the γ -substituent, satisfying the observation derived from the 2-D plots in Fig. 4. N1-OA for 9 and 20 was >8 Å but <5 Å for 21, consistent with the distributions on Fig. 4C. The shorter distance in 21 is associated with an ionic interaction between N1 and OA; a similar interaction was also observed the remaining G3 compounds, 22 and 23.

The 3D QSAR model described by Baringhaus *et al.* involved inhibition of rabbit ASBT, and the model was developed using either negatively charged or neutral compounds (8). According to their model, 5-member ring D, methyl group at 21 position, 7 or 12-OH in α -orientation, and a negative charge on the side chain were important for binding. Meanwhile, 3-OH and the *cis*-configuration of A and B rings of the steroidal nucleus were not relevant for binding. This model did not contain any information about the bile acid conjugates with positively charged side chain. In contrast to the model of Baringhaus *et al.*, the model presented here explained the importance of the 3D orientation of the positively charged side chain for favorable activity. The presence of salt bridge between the piperidine nitrogen and α -acid diminished activity. CSP-SAR model also identified some structural descriptors involving 3-OH (i.e. distance O3-AS, angle O3-AS-BC, and distance N1-O3) to describe biological activity (Table II), even though 3-OH was not relevant in Baringhaus *et al.*

In conclusion, a series of 29 aminopiperidine conjugates of CDCA-glu and UDCA-glu were synthesized in good yield and found to be highly potent inhibitors of hASBT. Inhibitory potency ranged from about 1–40 μ M. A CSP-SAR model was developed that successfully identified structural and physicochemical features that favor binding based on a quantitative model that included G1 and G3 compounds. In addition a qualitative CSP-SAR model distinguished compounds in G1 and G2 from those in G3. The CSP-SAR model demonstrated the presence of a salt-bridge interaction between N1 and α -acid for G3 compounds. This interaction is not accessible to the G2 compounds due to subtle, but important changes in the connectivity of the piperidine ring to the γ position. For example, 20 in the G2 set and 21 in G3 only differ by the chirality of the linkage of the piperidine to the methylene, leading to an order of magnitude difference in activity. The putative ionic interaction in the G3 set leads to a conformational restriction on the N1-OA distance, resulting in poorer binding affinity. Given that the binding affinity of the present set of compounds ranged over a relatively small range, the

fact that the CSP-SAR method was able to yield useful models relating structure and physio-chemical properties to activity indicates the potential of this approach for studies of hASBT binding and, ultimately, transport.

ACKNOWLEDGMENTS

This work was supported in part by National Institutes of Health grants DK67530 (to JEP) and the Computer-Aided Drug Design Center, University of Maryland, Baltimore for computational resources.

REFERENCES

- St-Pierre MV, Kullak-Ublick GA, Hagenbuch B, Meier PJ. Transport of bile acids in hepatic and non-hepatic tissues. *J Exp Biol* 2001;204:1673–86.
- Gilardi F, Mitro N, Godio C, Scotti E, Caruso D, Crestani M, De Fabiani E. The pharmacological exploitation of cholesterol 7 α -hydroxylase, the key enzyme in bile acid synthesis: from binding resins to chromatin remodelling to reduce plasma cholesterol. *Pharmacol Ther* 2007;116:449–72. doi:10.1016/j.pharmthera.2007.08.003.
- Pauli-Magnus C, Stieger B, Meier Y, Kullak-Ublick GA, Meier PJ. Enterohepatic transport of bile salts and genetics of cholestasis. *J Hepatol* 2005;43:342–57. doi:10.1016/j.jhep.2005.03.017.
- Sievanen E. Exploitation of bile acid transport systems in prodrug design. *Molecules* 2007;12:1859–89. doi:10.3390/12081859.
- Wong MH, Oelkers P, Craddock AL, Dawson PA. Expression cloning and characterization of the hamster ileal sodium-dependent bile acid transporter. *J Biol Chem* 1994;269:1340–7.
- Zhang EY, Phelps MA, Banerjee A, Khantwal CM, Chang C, Helsper F, Swaan PW. Topology scanning and putative three-dimensional structure of the extracellular binding domains of the apical sodium-dependent bile acid transporter (SLC10A2). *Biochemistry* 2004;43:11380–92. doi:10.1021/bi049270a.
- Balakrishnan A, Wring SA, Polli JE. Interaction of native bile acids with human apical sodium-dependent bile acid transporter (hASBT): influence of steroidal hydroxylation pattern and C-24 conjugation. *Pharm Res* 2006;23:1451–9. doi:10.1007/s11095-006-0219-4.
- Baringhaus KH, Matter H, Stengelin S, Kramer W. Substrate specificity of the ileal and the hepatic Na⁺/bile acid cotransporters of the rabbit. II. A reliable 3D QSAR pharmacophore model for the ileal Na⁺/bile acid cotransporter. *J Lipid Res* 1999;40:2158–68.
- Craddock AL, Love MW, Daniel RW, Kirby LC, Walters HC, Wong MH, Dawson PA. Expression and transport properties of the human ileal and renal sodium-dependent bile acid transporter. *Am J Physiol* 1998;274:G157–69.
- Lack L, Weiner IM. Intestinal bile salt transport: structure–activity relationships and other properties. *Am J Physiol* 1966;210:1142–52.
- Braun A, Yesilaltay A, Acton S, Broschat KO, Krul ES, Napawan N, Stagliano N, Krieger M. Inhibition of intestinal absorption of cholesterol by ezetimibe or bile acids by SC-435 alters lipoprotein metabolism and extends the lifespan of SR-BI/apoE double knockout mice. *Atherosclerosis* 2008;198:77–84. doi:10.1016/j.atherosclerosis.2007.10.012.
- Tremont SJ, Lee LF, Huang HC, Keller BT, Banerjee SC, Both SR, Carpenter AJ, Wang CC, Garland DJ, Huang W, Jones C, Koeller KJ, Kolodziej SA, Li J, Manning RE, Mahoney MW, Miller RE, Mischke DA, Rath NP, Fletcher T, Reinhard EJ, Tollefson MB, Vernier WF, Wagner GM, Rapp SR, Beaudry J, Glenn K, Regina K, Schuh JR, Smith ME, Trivedi JS, Reitz DB. Discovery of potent, nonsystemic apical sodium-codependent bile acid transporter inhibitors (Part 1). *J Med Chem* 2005;48:5837–52. doi:10.1021/jm040215+.
- Vicens M, Macias RI, Briz O, Rodriguez A, El-Mir MY, Medarde M, Marin JJ. Inhibition of the intestinal absorption of bile acids using cationic derivatives: mechanism and repercussions. *Biochem Pharmacol* 2007;73:394–404. doi:10.1016/j.bcp.2006.10.014.
- Vicens M, Medarde M, Macias RI, Larena MG, Villafaina A, Serrano MA, Marin JJ. Novel cationic and neutral glycocholic acid and polyamine conjugates able to inhibit transporters involved in hepatic and intestinal bile acid uptake. *Bioorg Med Chem* 2007;15:2359–67. doi:10.1016/j.bmc.2007.01.027.
- 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. doi:10.1016/j.atherosclerosis.2003.08.019.
- Balakrishnan Aa, Polli JE. Apical sodium dependent bile acid transporter (ASBT, SLC10A2): a potential prodrug target. *Mol Pharm* 2006;3:223–30. doi:10.1021/mp060022d.
- Balakrishnan A, Wring SA, Coop A, Polli JE. Influence of charge and steric bulk in the C-24 region on the interaction of bile acids with human apical sodium-dependent bile acid transporter. *Mol Pharm* 2006;3:282–92. doi:10.1021/mp0600135.
- Bernard D, Coop A, MacKerell AD Jr. 2D conformationally sampled pharmacophore: a ligand-based pharmacophore to differentiate delta opioid agonists from antagonists. *J Am Chem Soc* 2003;125:3101–7. doi:10.1021/ja027644 m.
- Bernard D, Coop A, MacKerell AD Jr. Quantitative conformationally sampled pharmacophore for delta opioid ligands: reevaluation of hydrophobic moieties essential for biological activity. *J Med Chem* 2007;50:1799–809. doi:10.1021/jm0612463.
- Bernard D, Coop A, MacKerell AD Jr. Conformationally sampled pharmacophore for peptidic delta opioid ligands. *J Med Chem* 2005;48:7773–80. doi:10.1021/jm050785p.
- Han G, Tamaki M, Hruba VJ. Fast, efficient and selective deprotection of the tert-butoxycarbonyl (Boc) group using HCl/dioxane (4 m). *J Pept Res* 2001;58:338–41. doi:10.1034/j.1399-3011.2001.00935.x.
- Chandrasekaran S, Kluge AF, Edwards JA. Studies in beta-lactams 6. Synthesis of substituted beta-lactams by addition of nitromethane to 6-oxopenicillanates and 7-oxocephalosporanates. *J Org Chem* 1977;42:3972–74. doi:10.1021/jo00444a046.
- White E. Organic syntheses. Collective volume. New York: Wiley; 1973.
- Bailey S, Gautier ECL, et al. Nicotinamide derivatives useful as PDF4 inhibitors. In U.S.P. Office (ed.), *US Patent Application Publication*, Pfizer Inc, USA; 2005. p. 56.
- Bruhaspathy M, Bhattacharyya S, Williamson J. Chemoselective reductive alkylation of ammonia with carbonyl compounds: synthesis of primary and symmetrical secondary amines. *Tetrahedron* 2005;60:1463–71.
- 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. doi:10.1007/s11095-005-5274-8.
- Gonzalez P, Polli JE. Impact of impurity on kinetic estimates from transport and inhibition studies. *J Pharmacol Exp Ther* 2008;326:296–305. doi:10.1124/jpet.107.135863.
- Balakrishnan A, Hussainzada N, Gonzalez P, Bermejo M, Swaan PW, Polli JE. Bias in estimation of transporter kinetic parameters from overexpression systems: Interplay of transporter expression level and substrate affinity. *J Pharmacol Exp Ther* 2007;320:133–44. doi:10.1124/jpet.106.107433.
- Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J Comput Chem* 1983;4:187–217. doi:10.1002/jcc.540040211.
- Feig M, Karanicolas J, Brooks CL 3rd. MMTSB Tool Set: enhanced sampling and multiscale modeling methods for applications in structural biology. *J Mol Graph Model* 2004;22:377–95. doi:10.1016/j.jmgm.2003.12.005.
- García AE, Sanbonmatsu KY. Exploring the energy landscape of a beta hairpin in explicit solvent. *Proteins Struct Funct Bioinform*

- 2001;42(3):345–54. doi:10.1002/1097-0134(20010215)42:3<345::AID-PROT50>3.0.CO;2-H.
32. Sugita Y, Okamoto Y. Replica-exchange molecular dynamics method for protein folding. *Chem Phys Lett* 1999;314:141–51. doi:10.1016/S0009-2614(99)01123-9.
 33. Allen MP, Tildesley DJ. *Computer simulation of liquids*. Oxford: Clarendon; 1987.
 34. Lee MS, Feig M, Salsbury FR, Brooks CL. New analytic approximation to the standard molecular volume definition and its application to generalized born calculations. *J Comput Chem* 2003;24:1348–56. doi:10.1002/jcc.10272.
 35. Lee MS, Salsbury FR, Brooks CL. Novel generalized Born methods. *J Chem Phys* 2002;116:10606–14. doi:10.1063/1.1480013.
 36. Ryckaert JP, Ciccotti G, Berendsen JC. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys* 1977;23:327–41. doi:10.1016/0021-9991(77)90098-5.
 37. Lee B, Richards FM. The interpretation of protein structures: estimation of static accessibility. *J Mol Biol* 1971;55:379–400. doi:10.1016/0022-2836(71)90324-X.
 38. Chothia C. Structural invariants in protein folding. *Nature* 1975;254:304–8. doi:10.1038/254304a0.
 39. Chemical Computing Group. *Molecular Operating Environment*. Montreal, Quebec, Canada, <http://www.chemcomp.com> (2009).
 40. Reiser B, Faraggi D. Confidence intervals for the overlapping coefficient: the normal equal variance case. *Statistician* 1999;48:413–8. doi:10.1111/1467-9884.00199.
 41. Akaike H. Likelihood of a model and information criteria. *J Econom* 1981;16:3–14. doi:10.1016/0304-4076(81)90071-3.
 42. McQuarrie ADR, Tsai CL. *Regression and time series model selection*. Singapore: World Scientific; 1998.
 43. Buckland ST, Burnham KP, Augustin NH. Model selection: An integral part of inference. *Biometrics* 1997;53:603–18. doi:10.2307/2533961.
 44. Ekins S, Mirny L, Schuetz EG. A ligand-based approach to understanding selectivity of nuclear hormone receptors PXR, CAR, FXR, LXR α , and LXR β . *Pharm Res* 2002;19:1788–800. doi:10.1023/A:1021429105173.
 45. Barlow DJ, Thornton JM. Ion-pairs in proteins. *J Mol Biol* 1983;168:867–85. doi:10.1016/S0022-2836(83)80079-5.