Determining Sequential Micellization Steps of Bile Salts with Multi-CMC Modeling

David Rovnyak*, Jiayi He[1], Sophie Kong[2], Kyle W. Eckenroad[3], Gregory A. Manley[4], Raeanne M. Geffert[5], Michael R. Krout**, Timothy G. Strein**

1 Dent Drive, Department of Chemistry, Bucknell University, Lewisburg PA 17837

*communicating author, drovnyak@bucknell.edu
** michael.krout@bucknell.edu, strein@bucknell.edu

1. current address: University of Pennsylvania, Department of Chemistry, 231 S. 34 Street, Philadelphia, PA 19104-6323, hejiayi@sas.upenn.edu

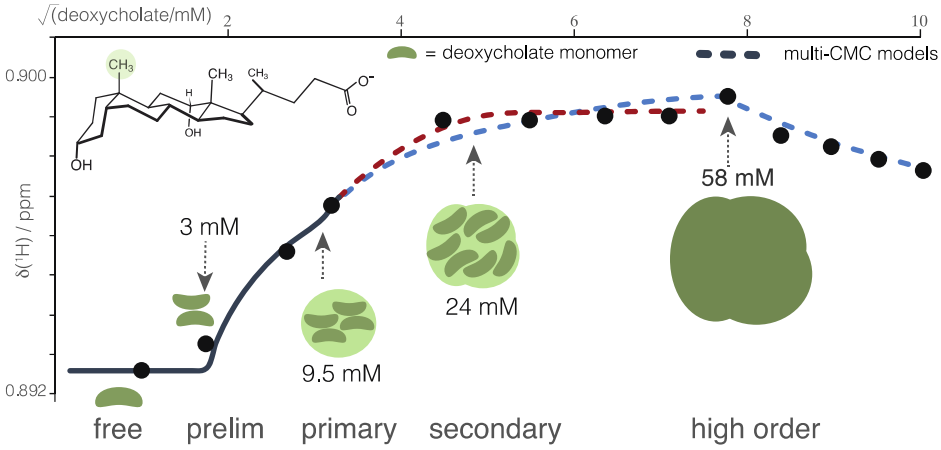
2. current address: Department of Pharmaceutical Chemistry, School of Pharmacy, University of California San Francisco, 1700 4th St, San Francisco CA 94158, Sophie.Kong@ucsf.edu

3. current address: Bristol Myers Squibb, 1 Squibb Drive, 92-218, New Brunswick, NJ, 08901, USA, Kyle.Eckenroad@bms.com

4. current address: AB SCIEX LLC , 500 Old Connecticut Path, Framingham, MA 01701, gmanley@gmail.com

5. current address: The University of North Carolina at Chapel Hill, UNC Eshelman School of Pharmacy, Division of Pharmacotherapy and Experimental Therapeutics
Kerr Hall, Campus Box 7569, Chapel Hill, NC 27599-7569, geffert@unc.edu

TOC FIgure



Key Words: bile salts, critical micelle concentration, nuclear magnetic resonance, NMR, phase separation, mass action

Abstract

Hypothesis

Bile salts exhibit complex concentration-dependent micellization in aqueous solution, rooted in a long-standing hypothesis of increasing size in bile aggregation that has historically focused on the measurement of only one CMC detected by a given method, without resolving successive stepwise aggregates. Whether bile aggregation is continuous or discrete, at what concentration does the first aggregate form, and how many aggregation steps occur, all remain as open questions.

Experiments

Bile salt critical micelle concentrations (CMCs) were investigated with NMR chemical shift titrations and a multi-CMC phase separation modeling approach developed herein. The proposed strategy is to establish a correspondence of the phase separation and mass action models to treat the first CMC; subsequent micellization steps, involving larger micelles, are then treated as phase separation events.

Findings

The NMR data and the proposed multi-CMC model reveal and resolve multiple closely spaced sequential preliminary, primary, and secondary discrete CMCs in dihydroxy and trihydroxy bile salt systems in basic (pH 12) solutions with a single model of one NMR data set. Complex NMR data are closely explained by the model. Four CMCs are established in deoxycholate below 100 mM (298 K, pH 12): 3.8 ± .5 mM, 9.1± .3 mM, 27 ± 2 mM, and 57 ± 4 mM, while three CMCs were observed in multiple bile systems, also under basic conditions. Global fitting leverages sensitivity of different protons to different aggregation stages. In resolving these closely spaced CMCs, the method also obtains chemical shifts of these spectroscopically inaccessible (aka dark) states of the distinct micelles.

Introduction

Amphiphiles generally form persistent soluble aggregates in aqueous solution once they reach a specific concentration known as the critical micelle concentration (CMC). Bile salts are planar (a.k.a. facial) amphiphiles (**Figure 1a**) with non-canonical aggregation behavior [1-4], where a general picture of preliminary, primary, and secondary aggregation has been a dominant lens for understanding bile aggregation for decades. [5, 6] Bile salts have diverse self-assembly and guest solubilization properties with significance spanning liver function and lipid transport [7], topical drug delivery [8, 9], chiral separations [10-15], homeostasis as well as disease pathology [16-20], and solvation of carbon nanotubes [21, 22]. One of the most intriguing features of bile aggregation is the long-recognized dependence of bile aggregate size and structure on the concentration of the bile salt, which remains challenging to characterize, particularly at lower concentrations below about 100 mM. The need to resolve bile aggregation in this regime is addressed here by proposing and applying a multi-CMC model to NMR chemical shift titration data.

Fundamental insights are sought to better understand the complex thermodynamic and chemical factors that determine bile self-aggregation, which can be pursued through the detection and measurement of bile CMCs. [2, 23-28] Representative CMCs measured for cholate and deoxycholate bile salts by diverse methods are given in **Table 1**, where most studies detect only one CMC but a few observe two sequential aggregation steps with distinct CMC values. The variation among the representative cholate (CA) and deoxycholate (DC) CMCs in **Table 1** is partly explained by differences in conditions (pH, salt, temperature, etc.) and the use of different methods [2], but significant variation remains. One intriguing trend is an apparent multimodal character to prior reported CMCs. In diverse cholate studies (**Table 1**), some methods detect a CMC around 6-8 mM while others detect a CMC at 12-16 mM CA. Similarly, DC CMCs are observed circa around 2-3 mM and 6-10 mM DC.

The examples in Table 1 suggest that distinct CMCs permeate prior CMC reports. Fluorescence studies often find a 6-8 mM cholate CMC (**Table 1**), suggesting that pyrene binds to a low-concentration, smaller cholate aggregate. However, probe-free (ITC, solubility) studies observe a cholate CMC ~ 12-16 mM cholate. Also, very large soluble aggregates, often termed secondary and high-order bile micelles, have long been the subject of extensive study [29-32]. Stepwise aggregation through two or even three CMCs has been a longstanding hypothesis in bile chemistry, but is often circumstantially supported. Multiple discrete aggregation steps below 100 mM have not been clearly resolved.

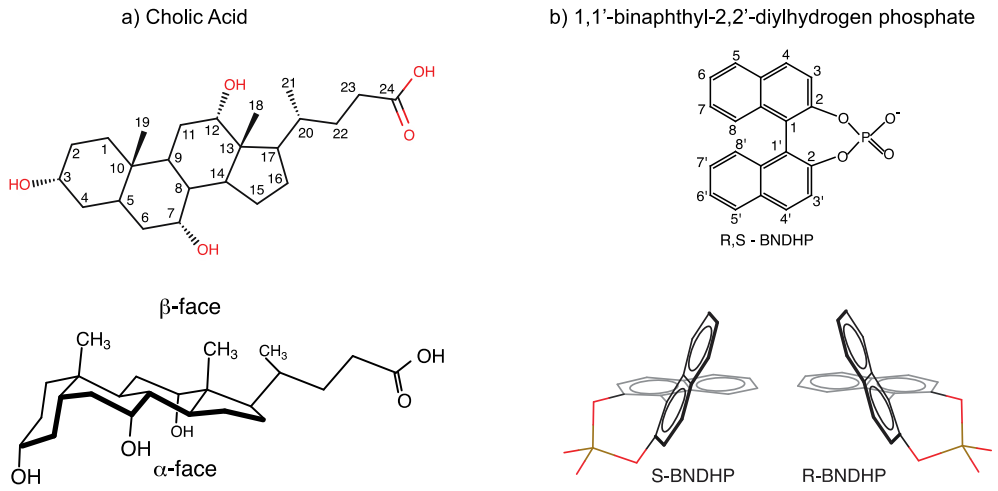


Figure 1. Bile acids such as cholic acid (CA), depicted in (a), are facial amphiphiles with the hydrophobic methyl-lined β face and the hydrophilic hydroxyl lined α face. Bile aggregates are also capable of chirally selective solubilization of guests such as R- and S-BNDHP, depicted in (b).

Bile salt#	Source	Method*	CMC (mM)
Cholate (CA)	Hebling et al. [33]	NMR (298 K, pH 12)	7.0 , 14
	Gouin et al. [34]	FE (pyrene, pH 8) NMR (296 K, pH 8)	13.5 16.0
	Posa et al. [35] Posa et al. [36] Posa et al. [37]	LS (pH 8, 0.1 M NaCl) FE (298 K, pyrene, 0.15 *m* NaCl) NMR (T_1 relaxation, 298 K)	7.6 8.5 12.0
	Mishra et al. [38]	FE (Coumarin 1)	6.0
	Matsuoka et al. [39]	FE (288.2 K, pyrene) FE (298.2 K, pyrene) FE (308.2 K, pyrene)	6.1 , 12.3 6.2 , 12.8 6.3 , 14.1
	Garidel et al. [40]	ITC (288.7 K, pH 7.5, 0.1 M NaCl) ITC (297.3 K, pH 7.5, 0.1 M NaCl) ITC (308.6 K, pH 7.5, 0.1 M NaCl)	12 10 14
	Anderson et al. [41]	ITC (288 K, pH 7.5, 0.1 M NaCl) ITC (298 K, pH 7.5, 0.1 M NaCl) ITC (308 K, pH 7.5, 0.1 M NaCl) ITC (288 K, pH 7.5, 0 NaCl) ITC (298 K, pH 7.5, 0 NaCl) ITC (308 K, pH 7.5, 0 NaCl)	9.2 7.2 9.5 12.6 10.4 16
	Reis et al. [27]	UV (298 K, pH 7) UV (298 K, pH 7, 0.05 M NaCl) UV (298 K, pH 7, 0.1 M NaCl) Pot (298 K, pH 7, 0.1 M NaCl) UV (298 K, pH 7, 0.2 M NaCl) Pot (298 K, pH 7, 0.2 M NaCl)	5.3 5.6 6.0 7.3 5.0 6.85
	Sugioka et al. [42]	FE (298.2 K, pH 7, pyrene) Sol (298.2 K, pH 7)	8.0 12.0
	Subuddhi et al. [43]	FE (288 K, pH 8, DPH) FE (298 K, pH 8, DPH) FE (308 K, pH 8, DPH)	15.0 16.0 17.0
	Jana and Moulik [44]	CON (303 K) ST (303 K) CAL (303 K)	5.2 9.1 7.2
Deoxycholate (DC)	Meier et al. [45]	NMR (298 K, pH 12)	3.0 , 9.0

	Posa et al. [36]	FE (298 K, pyrene, 0.15 m NaCl)	3.5
	Reis et al. [27]	UV (298 K, pH 7, 0.1 M NaCl) LS (298 K, pH 7, 0.1 M NaCl)	2.56 2.35
	Matsuoka et al. [39]	FE (288.2 K, pyrene) FE (298.2 K, pyrene) FE (308.2 K, pyrene)	2.3 , 6.1 2.4 , 6.5 3.4 , 6.9
	Kawamura et al. [46]	ESR (298 K, pH 7.8, .06 M NaCl, 5-ds) ESR (298 K, pH 7.8, .06 M NaCl, 12-ds)	2.0 3.0
	Roda et al.	ST (298 K, pH 8), 0 NaCl	10.0
	Garidel et al. [40]	ITC (287.4 K, pH 7.5, 0.1 M NaCl) ITC (297.9 K, pH 7.5, 0.1 M NaCl) ITC (307.6 K, pH 7.5, 0.1 M NaCl)	4.0 4.0 3.8
	Subuddhi et al. [43]	FE (288 K, pH 8, DPH) FE (298 K, pH 8, DPH) FE (308 K, pH 8, DPH)	5.0 6.0 8.0
	Jana and Moulik [44]	CON (303 K) ST (303 K) CAL (303 K)	4.0 2.8 7.9
	Das et al. [47]	ST (298 K, 0 NaCl) FE (298 K, 0 NaCl) ST (298 K, 0.1 NaCl)	2.3 2.0 1.2
	Perinelli et al. [24]	ST (298 K) CON (298 K) DEN (298 K) FE (298 K, pyrene)	2.3 , 4.4 7.1 6.3 8.2
	Kratohvil et al.	LS (298 K, pH 10, 0.15 M NaCl)	2.4

Table 1. A selection of critical micelle concentrations reported for CA and DC bile salt aggregation. In some cases, values have been rounded to one decimal place. FE = fluorescence emission with a probe molecule; NMR = nuclear magnetic resonance; Pot = potentiometry; LS = light scattering; ESR = electron spin resonance with probe 5-ds (5-doxylstearic acid) or 12-ds (12-doxylstearic acid); Sol = solubility ; ST = surface tension; CON = conductance; CAL = isoberibol calorimetry; DEN = densimetry; DPH = 1,6-diphenylhexatriene.
*pH is often not reported; use of sodium salts may imply pH circa 7-8, but bile acids often solubilized in basic pH;
commercial sodium salts of the bile acids often have an undetermined degree of hydrates present that limits the precise determination of bile concentrations; high order CMCs were not included in the table.

Some recent studies have been able to perform more finely grained measurements that clearly delineate sequential CMCs in CA and DC: a preliminary aggregate (2-3 mM DC and 6-8 mM CA) and a primary aggregate (6-10 mM DC and 12-16 mM CA). [33, 39, 45] This work considers NMR parameters for their sensitivity to self-aggregation, while affording structural insights at atomic resolution. [25, 48-52] In our prior work, the data were unambiguous due to the distinct behavior of the adjacent H3 and H4 protons of the binaphthyl BNDHP guest (**Figure 1b**) interacting with cholate and deoxycholate aggregates.[45] Specifically, using single CMC phase separation models, H4-BNDHP reports a CMC at 7 mM cholate (3 mM DC), yet the adjacent H3-BNDHP shows no perturbation until 14 mM cholate (9 mM DC).[45] In the prior work, the single-CMC model frequently deviated from the data at the concentration corresponding to the next sequential CMC. [45]

In order to test the existence of multiple micellization events in NMR data, a multi-CMC model has been developed for this work and applied to prior and new data. First, it is recognized that the initial CMC involves the smallest aggregates and is most likely to

exhibit equilibrium mass action (MA) behavior, whereas subsequent aggregation steps will involve larger aggregates and be more amenable to the phase separation (PS) description. To address the need to balance the MA and PS approaches, a simple correspondence of PS and MA approaches for small aggregation numbers is demonstrated and used to approximate the MA behavior in the first step; then subsequent CMCs are added into the model as phase separation steps. Such an approach closely models seemingly complex chemical shift titration data sets obtained with several bile salt systems, employs relatively small numbers of parameters, and clearly resolves multiple closely-spaced discretized aggregation steps for many bile systems below 100 mM concentrations under basic (pH 12) conditions.

Materials and Methods

Bile acids were obtained from Sigma Aldrich and used without further purification, however glycodeoxycholic acid (GDCA) and glycocholic acid (GCA) were synthesized in house, as described below, due to the cost of obtaining sufficient quantities for the measurements here, and relating also to broader efforts in synthesizing bile acid derivatives.

Ethyl glycodeoxycholate. To a flame-dried 50 mL round-bottom flask with a stir bar under Ar was sequentially added deoxycholic acid (1.702 g, 4.336 mmol, 1.0 equiv), ethyl glycinate hydrochloride (0.7253 g, 5.196 mmol, 1.2 equiv), DMF (17.5 mL, 0.25 M), Et_3N (1.80 mL, 12.91 mmol, 3.0 equiv), and TBTU coupling reagent (1.532 g, 4.771 mmol, 1.1 equiv). Upon DCA consumption by TLC analysis (ca. 1 h; EtOAc + 2% AcOH) the reaction was poured into a separatory funnel and rinsed/diluted with EtOAc (150 mL). The organics were washed with dH_2O (75 mL), then $NaHCO_3$ (50 mL). The combined aq was extracted with EtOAc (20 mL). The combined organics were washed with NaCl (25 mL), dried over $MgSO_4$, filtered, concentrated and placed under high vacuum. The resulting thick pale yellow oil was dissolved in hot EtOAc (ca. 20 mL) and crystallized. The crystals were cooled in an ice/H_2O bath, isolated via vacuum filtration and further dried under high vacuum to provide 1.784 g (3.735 mmol, 86.1% yield) of a white crystalline solid. R_f = 0.26 (19:1 EtOAc/MeOH; *p*-anisaldehyde); ^{1}H NMR (600 MHz, $CDCl_3$): δ 5.93 (br s, 1H), 4.22 (q, *J* = 7.2 Hz, 2H), 4.03 (dd, *J* = 5.1, 2.1 Hz, 2H), 3.98 (br s, 1H), 3.61 (dq, *J* = 11.1, 5.6 Hz, 1H), 2.32 (ddd, *J* = 14.7, 10.2, 4.8 Hz, 1H), 2.16 (ddd, *J* = 14.5, 9.6, 6.4 Hz, 1H), 1.90-1.58 (m, 9H), 1.53-1.51 (m, 3H), 1.46-1.25 (m, 9H), 1.29 (t, *J* = 7.2 Hz, 3H), 1.13 (qd, *J* = 12.3, 4.9 Hz, 1H), 1.06 (td, *J* = 12.0, 6.0 Hz, 1H), 1.01-0.95 (m, 1H), 0.99 (d, *J* = 6.3 Hz, 3H), 0.91 (s, 3H), 0.68 (s, 3H); HRMS (ESI+) *m/z* calcd for $C_{28}H_{48}O_5N$ $[M + H]^+$: 478.3527, found: 478.3535.

Glycodeoxycholic acid (GDCA). To a 500 mL round-bottom flask containing ethyl glycodeoxycholate (2.593 g, 5.427 mmol, 1.0 equiv) and a stir bar under Ar was added MeOH (90 mL), dH_2O (18 mL; 5:1 v/v, 0.05 M), and LiOH (0.6499 g, 27.14 mmol, 5.0 equiv). A reflux condenser was affixed to the flask and the contents were immersed in a 65 °C oil bath until consumption by TLC analysis (ca. 4 h). The reaction was cooled to room temperature and MeOH was removed using a rotovap. To the resulting slurry (pH > 12) was added dH_2O (100 mL) and this aq layer was acidified to pH < 3 with 1 M HCl, resulting in the precipitation of a white solid. After brief sonification, the slurry was cooled in an ice/H_2O bath and isolated via vacuum filtration. The pellet was further dried under high vacuum to provide GDCA (2.2002 g, 4.893 mmol, 90.2% yield) as a white powder. ^{1}H NMR (600 MHz, CD_3OD): δ 3.96 (s, 1H), 3.88 (s, 2H), 3.52 (tt, *J* = 11.1, 4.6 Hz, 1H), 2.31 (ddd, *J* = 13.9, 10.2, 5.1 Hz, 1H), 2.17 (ddd, *J* = 14.0, 9.8, 6.6 Hz, 1H), 1.91-1.76 (m, 7H), 1.64-1.58 (m, 3H), 1.53-1.51 (m, 2H), 1.48-1.25 (m, 9H), 1.16 (qd, *J* = 13.9, 4.5 Hz, 1H), 1.11-1.08 (m, 1H), 1.02 (d, *J* = 6.5 Hz, 3H), 0.98 (td, *J* = 14.1, 3.2 Hz, 1H), 0.93 (s, 3H), 0.71 (s, 3H); ^{13}C{^{1}H} NMR (151 MHz, CD_3OD): δ 177.2, 173.1, 74.1, 72.6, 48.1, 47.6, 43.6, 41.8, 37.5, 37.2, 36.8, 36.4, 35.3, 34.8, 33.8, 33.1, 31.1, 29.9, 28.6, 28.4, 27.5, 24.9, 23.7, 17.7, 13.2; HRMS (DART+) *m/z* calcd for $C_{26}H_{44}O_5N$ $[M + H]^+$: 450.3214, found: 450.3216.

Benzyl glycocholate. The benzyl ester of glycocholic acid was prepared in an analgous reaction using cholic acid (1.700 g, 4.161 mmol, 1.0 equiv) and benzyl glycinate hydrochloride (1.009 g, 5.004 mmol, 1.2 equiv). Upon consumption of CA by TLC analysis (ca. 1 h; 9:1 CH_2Cl_2/MeOH + 5% AcOH) the reaction was worked up exactly as above. To the resulting thick pale yellow oil was added hot EtOAc (ca. 50 mL) and sonicated until a solid formed (ca. 5 min). The thick slurry was cooled in an ice/H_2O bath and the solid was isolated via vacuum filtration. The pellet was further dried under high vacuum to provide 2.065 g (3.716 mmol, 89.3% yield) of a white powdery solid. R_f = 0.31 (9:1 EtOAc/MeOH; *p*-anisaldehyde); ^{1}H NMR (600 MHz, $CDCl_3$): δ 7.36 (m, 5H), 6.04 (t, *J* = 5.2 Hz, 1H), 5.19 (s, 2H), 4.09 (d, *J* = 5.2 Hz, 2H), 3.98 (br s, 1H), 3.85 (br s, 1H), 3.45 (br s, 1H), 2.31 (ddd, *J* = 14.5, 9.8, 4.7 Hz, 1H), 2.26-2.14 (m, 3H), 1.98-1.26 (m, 18H), 1.14 (qd, *J* = 12.2, 6.0 Hz, 1H), 1.01-0.97 (m, 1H), 0.99 (d, *J* = 6.2 Hz, 3H), 0.89 (s, 3H), 0.69 (s, 3H); HRMS (ESI+) *m/z* calcd for $C_{33}H_{50}O_6N$ $[M + H]^+$: 556.3633, found: 555.3634.

Glycocholic acid (GCA). To a 100 mL round bottom flask containing benzyl glycocholate (1.670 g, 2.931 mmol, 1.0 equiv) and a stir bar were added CH_2Cl_2 (23.4 mL) and MeOH (5.9 mL; 4:1 v/v, 0.1 M), followed by Pd/C (5 wt%, 623.2 mg, 0.293 mmol, 0.1 equiv). An H_2 balloon was used to purge the atmosphere and the suspension was stirred under H_2 until consumption by TLC analysis. The resulting suspension was sonicated and then vacuum filtered into a 250 mL round bottom flask rinsing with 6:1 CH_2Cl_2/MeOH,

concentrated and dried under high vacuum. To the crude mixture was added H_2O (50 mL) and 1 M HCl to pH < 3, precipitating a white solid that was isolated and dried as above to give GCA (1.123 g, 2.414 mmol, 82.3% yield), a white powdery solid. ^{1}H NMR (600 MHz, CD_3OD): δ 3.95 (app t, *J* = 3.1 Hz, 1H), 3.88 (s, 2H), 3.79 (app q, *J* = 3.0 Hz, 1H), 3.37 (tt, *J* = 11.2, 4.4 Hz, 1H), 2.35-2.23 (m, 3H), 2.18 (ddd, *J* = 14.1, 9.8, 6.6 Hz, 1H), 2.02-1.72 (m, 7H), 1.67-1.51 (m, 6H), 1.47-1.285 (m, 5H), 1.11 (qd, *J* = 11.9, 5.5 Hz, 1H), 1.03 (d, *J* = 6.5 Hz, 3H), 0.98 (td, *J* = 14.5, 3.7 Hz, 1H), 0.91 (s, 3H), 0.71 (s, 3H); ^{13}C{^{1}H} NMR (151 MHz, CD_3OD): δ 177.2, 173.1, 74.1, 72.9, 69.1, 48.1, 47.5, 43.2, 43.0, 41.7, 41.0, 40.5, 36.8, 36.5, 35.9, 35.9, 33.8, 33.1, 31.2, 29.6, 28.6, 27.9, 24.2, 23.2, 17.7, 13.0; HRMS (DART+) *m/z* calcd for $C_{26}H_{44}O_6N$ $[M + H]^+$: 466.3163, found: 466.3164.

NMR Spectroscopy. Some of the work includes a reanalysis of data obtained previously where noted. [33, 45] All other data were obtained on a 600 MHz spectrometer (vnmrs, Varian Inc., vnmrj 4.2) with an inverse room temperature probe (298 K) with typical $\pi/2(^1H)$ = 7 μs pulse lengths, and using 1D WATERGATE. Processing employed iNMR with digital water suppression, typically 0.5 Hz line broadening, and baseline correction.

Multi-CMC Model

At the critical micelle concentration (CMC) surfactant molecules spontaneously self-aggregate to form larger soluble micelle particles. An observable which is sensitive to the aggregation step must be measured, and a model applied to extract the CMC from the data. A distinction is whether the observable requires a probe molecule (e.g. fluorescence emission) or not (e.g. ITC demicellization). [40, 53-55] This work considers an observable o that is a weighted sum of the fractional occupation of the free and micelle states, leading to a piecewise function for a single CMC, [56, 57]

$$\begin{aligned} o_{obs} &= o_{free} && if \;\; S_{tot} < CMC \;\; , \\ o_{obs} &= f_{free} o_{free} + f_{micelle} o_{micelle} && if \;\; S_{tot} \geq CMC \;\; , \end{aligned} \qquad (1)$$

where S_{tot} is the total concentration of surfactant, f_{free} and f_{micelle} are the fraction of surfactant in the free monomer and micellar aggregate forms, and o_{free} and o_{micelle} are the values of the observable for the monomer and micelle forms. Suitable NMR parameters include the chemical shift and the self-diffusion constant. [56, 57] The mass action (MA) and phase separation (PS) models obtain f_{free} and f_{micelle} in distinct ways and are summarized in the Supplemental Material. Aggregation modeling is an ongoing concern, including recent work by Rusanov [58-60], Shinoda and Hutchinson [61], Oleson and coworkers [54], and references therein.

In order to evaluate Eqn (1) with the PS model, first set $f_{\text{free}} = 1$ if $S_{tot} < CMC$. Next, if $S_{tot} \geq CMC$, the free monomer concentration [S] takes on a constant saturated value

$$S = CMC \quad \text{if} \quad S_{tot} \geq CMC \ . \tag{2}$$

For surfactant concentrations above the CMC, the fraction of free monomer is $f_{\text{free}} = CMC / S_{\text{tot}}$ in the PS model. The PS approach does not treat the aggregation number.

In the mass-action (MA) model, the aggregation number (n) is required, while the counterion occupancy (β) is sometimes included for ionic surfactants. [54] The MA approach defines a critical concentration S_{tot}^{*} at which half the monomers are in the micelle and half are free, and which should not be confused with the CMC. The free monomer concentration S is obtained implicitly with the relation

$$S_{\text{tot}} = S\,(1 + \left(\frac{2S}{S_{tot}^{*}}\right)^{n-1}) \ . \tag{3}$$

If n is large, the CMC ~ $S_{tot}^{*}/2$, while if n is small, a CMC may still be defined in mass action models.[54] The behavior of these models (Eqns 2-3) is illustrated for mock data in **Figure 2**, where the critical concentration S_{tot}* is recognized as the inflection point in the MA model regardless of n. A key property of the models is illustrated, that *they share the same inflection point (**Figure 2a**) and the PS model best approximates MA data when the CMC of the PS model is set to S_{tot}*/2* (**Figure 2b**). As noted, the CMC is still defined in the MA model for very low n, where it occurs at values less than S_{tot}*/2 (**Figure 2a**, filled circles).[54] The PS determined CMC would therefore be slightly larger than that of the mass action approach for small n. For example, for n = 8, the onset of aggregation would be observed closer to ~3.5 mM surfactant.

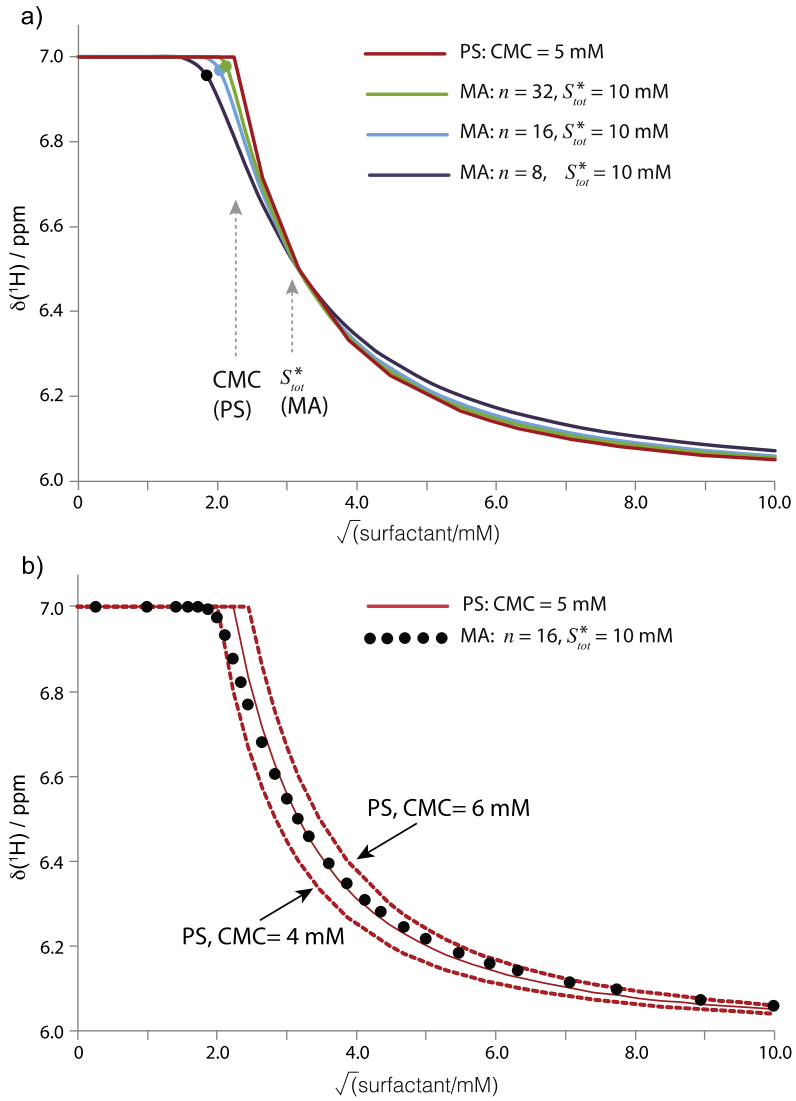


Figure 2. The phase separation (PS) and nonionic mass action (MA) micelle equilibrium models are compared for mock NMR chemical shift parameters (o_{free} = 7 ppm and o_{micelle} = 6 ppm), where the models converge at a common inflection point (S_{tot}*) in (a). A 5 mM CMC in the PS model and a critical concentration S_{tot}*=10 mM in the MA model are used. In (a) filled circles indicate the positions of the CMCs as determined by the MA treatment and are 3.5 mM, 4.1 mM, and 4.4 mM for n = 8, 16, and 32. In (b) for an n = 16 example, the PS model is seen to best approximate MA behavior if the CMC is set to S_{tot}*/2; PS models that deviate by as little as ± 1 mM of the correct value poorly model the MA data.

The PS model requires up to three parameters (o_{free}, o_{micelle}, [S*]), while the MA model adds n (aggregation number) and the counterion occupancy β for the ionic MA model. Importantly, o_{free} is often fixed by the data, reducing the PS model to two-parameters and the non-ionic MA model to three parameters. Multiple CMCs entail an expanded piecewise function. For two sequential CMCs

$$o_{obs} = o_{free} \qquad if\ \ S_{tot} < CMC \tag{4}$$

$$= f_{free} o_{free} + f_1 o_{mic1} \qquad if \quad CMC_1 \leq S_{tot} < CMC_2 \quad , \qquad (4b)$$
$$= f_{free} o_{free} + f_1 o_{mic1} \; + \; f_2 o_{mic2} \quad if \quad S_{tot} \geq CMC_2. \qquad (4c)$$

Here o_{mic1} and o_{mic2} are the values of the observable for species in the first and second micellar forms, respectively. Equation 4 can be generalized to any number of aggregation steps in principle, although this work considers only two or three CMCs.

A generalization of Eqn 2 for two (or more) phase separation CMCs is proposed, where the second micellization described by CMC_2 will also be treated as a phase separation event. Then for $c \geq CMC_2$ the amount of free monomer and the amount of the first micellar phase are considered together to have a constant saturated value: the free monomer is assumed to still be given by CMC_1 and the amount of surfactant in the first micellar phase makes up the difference $(CMC_2 - CMC_1)$. Then all three fractions for Eqn (11c) are:

$$o_{obs} = \frac{CMC_1}{S_{tot}} o_{free} + \frac{CMC_2 - CMC_1}{S_{tot}} o_{mic1} + \frac{c - CMC_2}{S_{tot}} o_{mic2} \quad if \quad S_{tot} \geq CMC_2 \; . \qquad (5)$$

There are as many as five parameters to take in to account in Eqn (4-5), $(o_{free}, o_{mic1}, o_{mic2}, CMC_1, CMC_2)$ where, following the discussion above, o_{free} is expected to be fixed by the data (e.g. **Supplemental Material Section S1**). Further, the values of CMC_1 and CMC_2 are often confined to narrow ranges due to sharp discontinuities in the data as well as global fitting of multiple data sets. Applying Eqns 4-5 to aggregation data may entail determining only o_{mic1} and o_{mic2}. The asymptotic approach to o_{mic1} can be truncated by the onset of the second CMC_2, and so particular attention is often given to adjusting o_{mic1}. For three CMCs, Eqn 5 is extended to:

$$o_{obs} = \frac{CMC_1}{S_{tot}} o_{free} + \frac{CMC_2 - CMC_1}{S_{tot}} o_{mic1} + \frac{CMC_3 - CMC_2}{S_{tot}} o_{mic2} \; + \; \frac{c - CMC_3}{S_{tot}} o_{mic3} \quad if \quad S_{tot} \geq CMC_3 \; . \qquad (6)$$

There may be up to seven adjustable parameters in principle in Eqn (6) $(o_{free}, o_{mic1}, o_{mic2}, o_{mic3}, CMC_1, CMC_2, CMC_3)$, but o_{free} is normally fixed by the data and the CMCs are often narrowly constrained. Additional CMCs could be incorporated if the data justify such a treatment.

The procedure used in this work is to model the initial CMC first using **Figure 2** as a visual rubric to apply the PS model to data that exhibit mass-action equilibrium behavior (e.g. a gradual change in o_{obs}). The CMC should be estimated slightly beyond the initial roll-off of the data in order to match the inflection point of the PS model to the inflection point of the data for the best fit (**Figure 2b**). Sequential CMCs can then be added into the model, where later CMCs involve larger aggregates and are better approximated by the PS assumption.

The MA model is a strong model for describing bile micellization, but relies on additional parameters in comparison to the PS model. While extending the MA model to multiple CMCs may be feasible, it could involve fitting two-CMC data to as many as 8 independent variables, reducing perhaps to 4-6 parameters under favorable conditions. Chemical shift data, while sensitive to local structure and sequential micellization, do not exhibit sufficient singularities to justify such a large parameter space. In contrast, the PS model will be seen to treat two- and three-CMC data with fewer parameters and obtaining close agreement with the data.

Results

There is an unmet need to holistically treat complex data that are sensitive to two or more aggregation events. For example, a single-CMC model will depart from the data at the concentration corresponding to the next CMC (e.g. Fig 4 of Meier et al [45]). Double-CMC and triple-CMC PS models are introduced through a case study in **Figures 3-6** of the aggregation of DC in basic solution (pH 12), viewed through the guest molecule R-BNDHP and also in probe-free solutions. This portion includes a further analysis of some of the data reported by Meier et al.[45] and Eckenroad et al. [62] The double-CMC and triple-CMC models demonstrate four sequential DC CMCs under these conditions. Additional examples in other bile salts follow in the remaining figures. All results are summarized in **Table 2**.

As described in the prior section, the strategy used here is to match the inflection point of the PS model to that of the data for the first CMC as closely as possible (**Figure 2**) so that the MA behavior of the first aggregation is captured in the model. The initial aggregation of bile salts has small aggregation numbers [28] and is likely to show significant mass action behavior. However, subsequent aggregates are larger and are expected to be more amenable to approximation by the phase separation (PS) model.

An example of the workflow for applying a two-CMC model to chemical shift data is shown in **Figure 3** for H6 of the R-BNDHP (2.5 mM) probe molecule interacting with DC (pH 12). Initially a single CMC phase separation model is applied (**Figure 3a)**. The dashed circle in the low concentration portion of **Figure 3a** shows the gradual mass action 'roll off' behavior: setting the CMC slightly beyond this roll off period, and adjusting δ_{mic1} so that the inflection point of the data and the model agree will fulfill the rubric illustrated in **Figure 2**.

Depending on how aggregation affects their local environments, certain protons are sensitive to different steps of aggregation. In the case of H6 of R-BNDHP, the model deviates from the data at concentrations of about 20-30 mM DC (dashed circle in **Figure 3b**). A single CMC model is insufficient to describe the data. An additional CMC is therefore postulated to occur in the range 20-30 mM. In double-CMC cases, the chemical shift of the guest bound to the second micelle, denoted δ_{mic2} (or of the micelle itself if the experiment is probe-free), is often suggested by the data; here the data at high concentrations appear to approach δ_{mic2} ~7.3 ppm which is a good starting point for δ_{mic2}. Carrying forward CMC_1 and δ_{mic1} from **Figure 3a**, it is now possible to make small adjustments to CMC_2, and δ_{mic2} to obtain a close model to the data as shown in **Figure 3c,** yielding CMC_2 = 26 mM. In practice, CMC_1 and δ_{mic1} may receive minor further

adjustments when the second CMC_2 is incorporated. Notably, o_{mic1} is the chemical shift of H6 in the initial micelle (CMC_1), which cannot be measured directly due to the onset of CMC_2 which interferes with the data. In other words, the initial aggregate at CMC_1 behaves like a dark state and the model provides an indirect means to obtain chemical shifts such as o_{mic1} of this micelle.

The agreement of the double-CMC PS model with the data in **Figure 3c** supports the validity of the model, and the parameters were closely constrained by the raw data. A similar procedure was carried out with the exact same data set, however now examining the chemical shift of H3 of R-BNDHP as a function of DC concentration in **Figure 4**. It is crucial to appreciate that **Figures 3** and **4** are drawn from the same data set. The H3 chemical shift is perturbed in the opposite direction as H6, which means that H3 does not enter the hydrophobic interior of the micelle but instead interacts with the hydrophilic surface of the micelle.[45] The H3 chemical shift shows some weak scatter below about 4 mM, $\delta(^1H3)$ but is essentially insensitive to the 3.7 mM CMC that strongly affected the H6 chemical shift. Instead, H3 of R-BNDHP reports on a CMC ca. 9-10 mM DC, as shown in **Figure 4**. A single-CMC model is first applied initially to the data (**Figure 4a**) at about 9 mM. Yet a single CMC model is unable to explain all of the data in **Figure 4a**, and an additional CMC is postulated: the same 26 mM CMC that was determined in the analysis of $\delta(^1H6)$ in **Figure 3**. With the second CMC fixed at 26 mM, the double CMC model is then applied to $\delta(^1H3)$ in **Figure 4b**. Following small adjustments to CMC_1, $o_{mic1,}$ and $o_{mic2,}$ the model (**Figure 4b)** shows good agreement with the data.

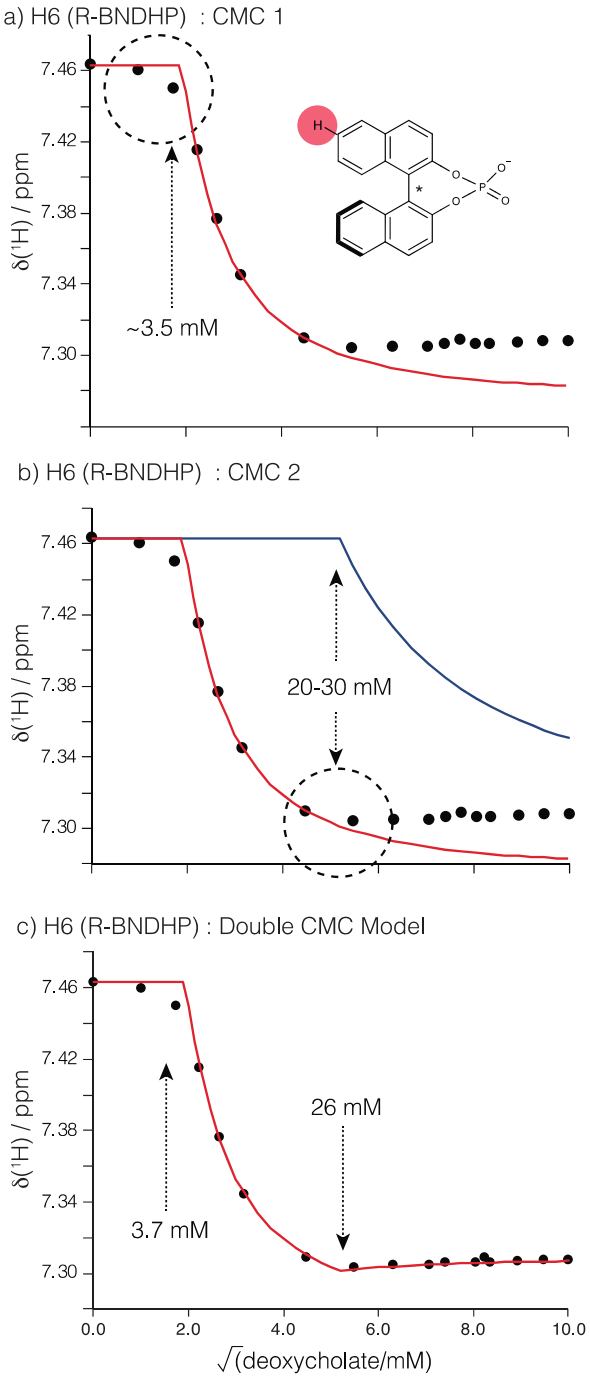


Figure 3. The application of the double-CMC phase separation model to NMR chemical shift perturbations of H6 of the guest molecule R-BNDHP interacting with DC is shown (pH 12). In (a) the single CMC PS model is applied so as to approximate MA behavior as closely as possible resulting in initial determinations of CMC_1 and o_{mic1}; in (b) the second CMC_2 is seen to fall between 20-30 mM and initial guesses of CMC_2 ~ 25 mM and o_{mic2} ~ 7.3 ppm can be made; finally in (c) the parameters CMC_2 and o_{mic2} are adjusted to improve the fit, yielding final estimates of CMC_1 = 3.7 mM, CMC_2 = 26 mM, o_{mic1} = 7.276 ppm, o_{mic2} = 7.309 ppm. The blue curve in (b) shows the single CMC model using only the final parameters determined for the second CMC.

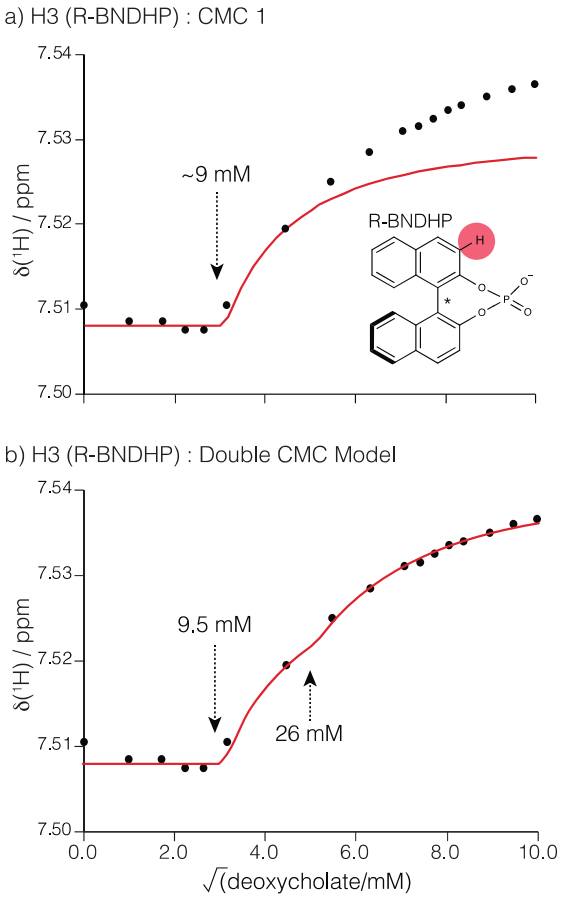


Figure 4. The application of the double-CMC phase separation model to experimental data of NMR chemical shift perturbations of H3 of the guest molecule R-BNDHP interacting with DC is shown. The method outlined in **Figure 3** is followed here with one key exception: CMC_2 is fixed at the 26 mM value previously determined in **Figure 3**. In (a) the single CMC PS model is applied so as to approximate MA behavior; in (b) the second CMC is fixed at 26 mM and an initial guess of o_{mic2} (the chemical shift of H3-R-BNDHP bound to the second micelle) is made; finally in (c) the parameters are adjusted slightly to improve the fit, yielding estimates of CMC_1 = 9.5 mM, CMC_2 = 26 mM, o_{mic1} = 7.530 ppm, o_{mic2} = 7.541 ppm. The blue curve in (b) shows the single CMC model using only the final parameters determined for the second CMC (e.g., 26 mM, 7.541 ppm). All solutions are pH 12.

It can be appreciated from **Figure 4** that global modeling of several data sets can be vital in modeling the CMCs, as it would be difficult to constrain the 26 mM CMC from δ(^{1}H3)-BNDHP alone. The analyses in **Figures 3-4** are self-consistent for two of the protons in R-BNDHP, and double-CMC models may also be constructed for H5 and H7 of R-BNDHP as a function of DC with the same CMCs (**Supplemental Material, Section S.2**).

It is clear from the data up to this point that different protons do not all report on the same aggregation events, as they have different structural roles in the aggregates that form, and therefore experience different local environments as aggregates form. A benefit of examining NMR data is the atomic resolution afforded by analyzing the chemical shifts of

individual protons. Whereas the 3.7 mM and 26 mM CMCs are verified by several different protons (H5-H7) of the guest molecule of the same data set, only $\delta(^1H3)$-BNDHP reveals the 9.5 mM CMC. And while the 9.5 mM CMC is unambiguous in the $\delta(^1H3)$-BNDHP data, it is still attractive to seek corroboration with global data model. The H8-R-BNDHP, Me18-DC, and H12-DC chemical shift data were examined with double-CMC models, where these protons (and others) unambiguously verify the 9.5 mM CMC (**Supporting Material, Section S3**). An interesting minor difference is that the initial CMC is modeled at about 3 mM DC for the bile protons, rather than 3.7 mM DC as obtained with the R-BNDHP protons; a stronger influence of mass action behavior on the bile protons could make the use of the PS model more challenging for the first CMC. In all three cases, the data show little curvature following the second CMC_2, from about 9.5-60 mM DC. Such linear behavior cannot be explained in the double-CMC model (**see SI**, **Section S.3**). A similar situation occurred when treating H3-BNDHP in **Figure 4**, where there was very little curvature in the data following the first CMC, which was explained by adding in the known 26 mM CMC to the model.

Since other protons have shown that there is a third CMC at about 26 mM DC, we introduce in **Figure 5** the triple-CMC model to test if the linear trend after 9.5 mM DC could be explained by incorporating the third 26 mM CMC of DC. In **Figure 5**, there are actually sufficient features in the data that the triple-CMC models could be applied without prior knowledge of the CMCs determined in **Figures 3-4**. Still, the prior CMCs were useful starting points in **Figure 5** and were essentially unchanged when the models were applied. Values of o_{mic1} and o_{mic2}, the chemical shifts in the first and second micellar forms, were transferred from the models used previously (Supplementary Material **Section S3**), leaving just o_{mic3}, the chemical shift in the third micellar form, to be determined in **Figure 5**. Final parameters are indicated on the figure. Notably, in an independent series of different samples measured at a different field strength, nearly identical behavior is observed (Supplemental Material **Section S.4)**.

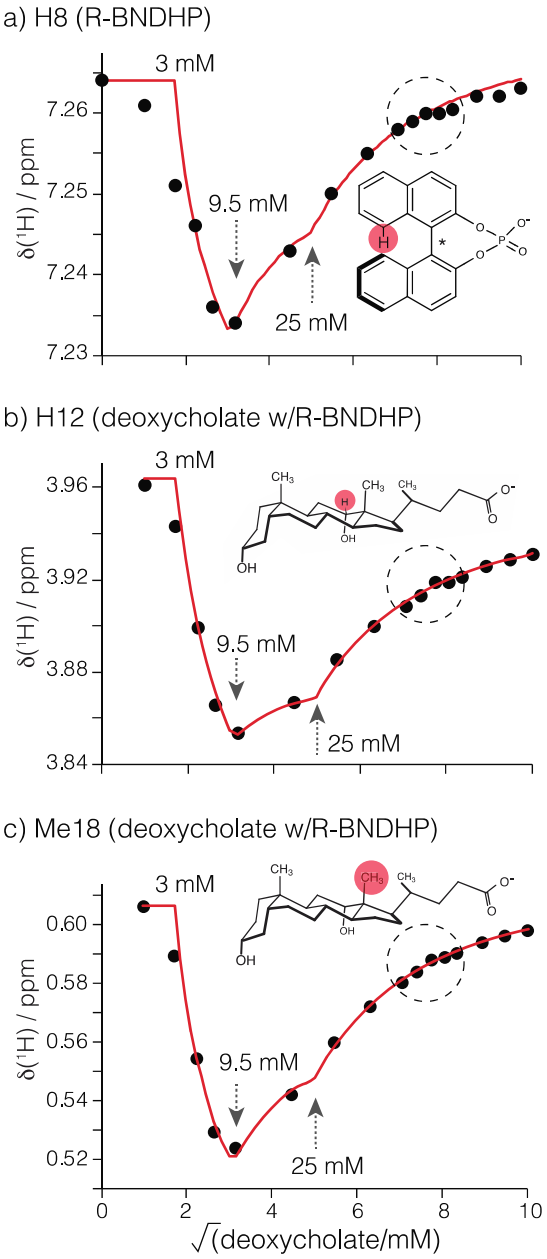


Figure 5. For DC aggregation in the presence of R-BNDHP, the triple CMC PS model (Supplemental Material) is applied to H8 (R-BNDHP), H12(DC), and Me18 (DC), where the three CMCs determined in **Figures 3-4** were used as starting values. Two-CMC models for these protons are found in Supplemental Material, **Section S3**. Only a change in the first CMC was required to model the trends of these bile protons (CMC_1 = 3 mM) versus the guest protons (CMC_1 = 3.7 mM) in the prior figures. See Supplemental Material (**Section S4**) for confirmation in a fully independent trial (900 MHz), which also confirmed the discontinuity near 60 mM DC. For all of the DC data in the presence of R-BNDHP, the weak discontinuity at about 60 mM was present.

The triple-CMC model is seen to be effective in explaining the very complex data in **Figure 5**, using solely the CMCs determined by global fitting simpler trends in other protons from the same data set, although the data in **Figure 5** are sufficient on their own to identify the

three CMCs. In **Figures 3-5** an additional discontinuity can be noticed in the data at about 60 mM DC, indicated by the dashed circles. This feature is weak but ubiquitous for numerous protons in DC data sets with R-BNDHP (**Figure 5 and Supplemental Material Section S4**) and suggests a fourth CMC. The binaphthyl H4-H6 protons were found to be amenable to constructing approximate triple-CMC models of this 60 mM feature (not shown, summarized in **Table 2**).

While there is a preponderance of evidence (**Figures 3-5, and S2-S4**) that many protons of both the probe and the bile salt all indicate a change in their chemical shift trends at about 60 mM DC, we sought to identify stronger reporters which could provide also some physical insight about the micellar structure at ~60 mM DC. Turning to probe-free experiments to test if this feature would still occur, two protons on the β-face of DC, Me19 and H15β, are particularly sensitive to a change in their local environment at about 58 mM DC (**Figure 6**). A four-CMC PS model is certainly plausible in principle to treat such data, but entails more complexity and begs the question if a simpler approach could explain the data. To treat the complex Me19 chemical shift trend, the first three CMCs are verified in **Figure 6a**. Then, in **Figure 6b** the CMC at about 24-25 mM DC is ignored, and the third CMC in the model is instead used to treat the discontinuity at 58 mM, where the model is even seen to capture the unusual datum at about 60 mM, supporting that the 58 mM step is a phase-separation event.

In **Figure 6c**, the H15β chemical shift is also seen to be relatively insensitive to the ~25 mM CMC, and the triple-CMC model is able to capture the 58 mM CMC as well. As with **Figure 6b**, ignoring the ~25 mM CMC means that the model does not reproduce the H15β trend as closely in this regime. The BNDHP probe is known to sample a methyl-lined hydrophobic pocket of bile micelles from NOE analysis [62], and some of the strongest reporters of the ~60 mM CMC are also on the hydrophobic face of the DC molecules. These data strongly support that the deoxycholate micelle undergoes a remodeling at about 60 mM that particularly affects the local environment of the hydrophobic β face.

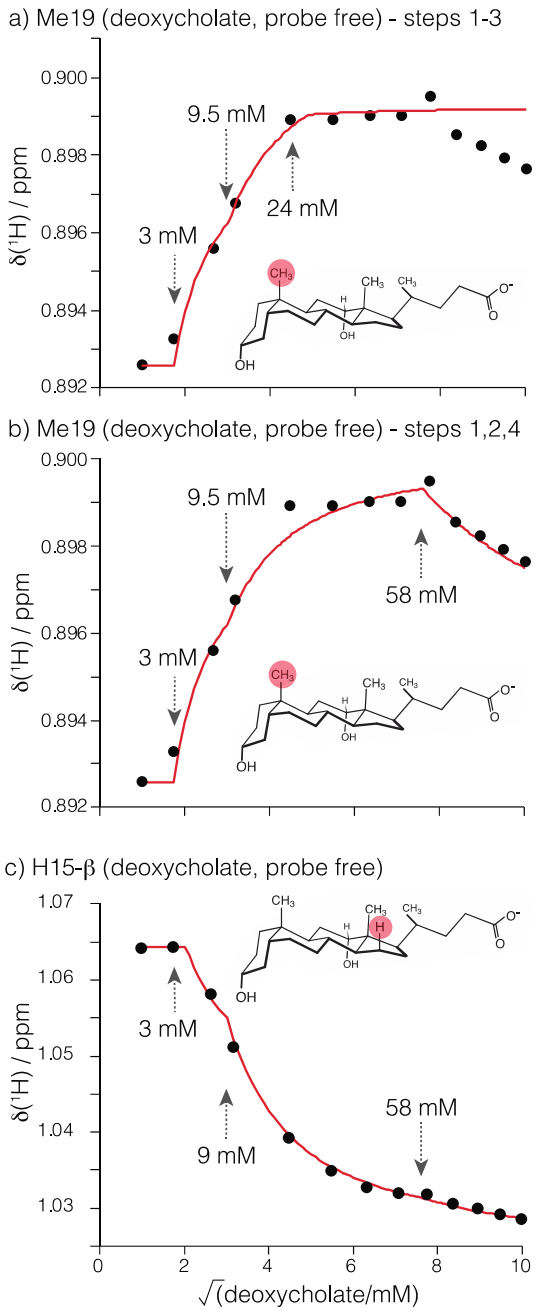


Figure 6. Two protons on the β-face of DC (probe free) are analyzed for a CMC that is modeled at about 58 mM DC here, and that was indicated also by prior data (e.g. **Figure 3-5**). The Me19 chemical shift in particular shows a strong discontinuity at 58 mM DC that is easily noted by visual inspection. To construct a model of the Me19 data, in (a) the triple-CMC model is first verified with the initial CMCs observed in earlier figures. Next, in (b) the third CMC is changed in the triple CMC model to account for the feature at 58 mM; although poorer agreement is noted with the data near 24-25 mM, the behavior after 58 mM, including the unusual datum at 60 mM, is explained by the model. In (c) 15βH is seen to be relatively insensitive to the CMC at about 24-25 mM, and then the triple CMC model can be applied to capture the 58 mM CMC.

Some variation between the CMC values used in **Figures 5** and **6** is observed, which likely reflects the probe modifying the micellization. The S-BNDHP enantiomer interacts much more strongly with bile micelles [33, 45] and is associated with slightly lower CMC values here (**SI Section S.5**, and **Table 2**). The S-BNDHP probe appears to significantly stabilize the micellar structure, whereas the weaker binding R-BNDHP enantiomer appears to serve more as a reporter on the bile CMC values, which are similar to values in probe-free conditions (SI **Figure S.5**). These examples include relatively high concentrations of the probe (2.5 mM S- or R-BNDHP), which may help to exaggerate differences between them. All four DC CMCs represent stages of aggregation that cannot be isolated and studied on their own due to the close spacing of these CMCs and their equilibria, and as such can be thought of as 'dark states'. Yet the multi-CMC modeling solves this problem and reveals the specific chemicals shifts (o_{mic1}, o_{mic2}, o_{mic3}, and o_{mic4}) of these distinct micelles, which may lead to a better understanding of the differences in the local chemical environments of aggregate surfaces and interiors.

These dark state chemical shifts offer complex structural information, which can be interpreted qualitatively initially. For example, this work gives some additional structural insight on the preliminary DC micelle at about 2-3 mM. For background, the development of hydrophobic/hydrophilic interactions correlates with upfield/downfield perturbations of proton shifts respectively in the study of water soluble aggregates. [63, 64] As we found previously, large upfield changes for DC-Me18 (Figure 5) and DC-Me19 (not shown) show that the probe strongly interacts with a methyl-lined surface of the early CMC_1 aggregate, confirmed by NOE analysis.[62] However, in the probe free DC data, the early CMC_1 aggregate exhibits a weaker downfield shift of Me19 (**Figure 6**), showing that the early micelle is much less stable in the absence of probe and that the DC methyl groups likely remain solvent exposed in the early DC micelle. Possible models of the preliminary DC CMC_1 include crossed or staggered arrangements of monomers, or edge-associated monomers, or a mixture of these arrangements. Thus the BNDHP guest remodels the DC CMC_1 micelle to orient the DC methyl groups around the binaphthyl ring. Turning to another region of the steroidal backbone in DC, the H15-β signal is perturbed upfield with every stage of aggregation, indicating that the D ring occupies progressively more hydrophobic packing environments at every stage. Finally, notice that Me19 is strongly affected by CMC_3 (24 mM) but H15-β is not, suggesting that CMC_3 is a remodeling of the micelle that leaves the D ring in a conserved local environment, but significantly changes the environment of the A/B rings. As computational (e.f. DFT) methods advance, this work shows that the chemical shifts of distinct aggregates are available to be incorporated as restraints.

Multi-CMC modeling is next applied to other bile acids. Basic solutions of the trihydroxy bile salt sodium cholate (CA) were examined in the absence of probe and revealed CMCs at about 16 mM and 41 mM. Example data and modeling are given in

Figure S.6 of the Supplementary Material for the case of probe-free cholate in basic solution. Prior studies on cholate with the probe molecules R,S-BNDHP clearly show also a preliminary CMC at about 7 mM CA [33], showing that the probe molecules stabilize this early aggregate and make it observable.

An analysis of glycocholic acid (GCA) is summarized in **Figure 7,** which gives the results of fitting the chemical shift of three protons in the absence of probe, and in the presence of each of R- and S-BNDHP. Representative models of the data are given in the Supplementary Material (**Figure S.7**). The R- and S-BNDHP probe molecules were present at just 0.1 mM, and much closer agreement can be seen among all cases with and without probe. Interestingly, the third CMC is shifted to a higher concentration in the presence of S-BNDHP. The S-BNDHP guest is known to bind to bile micelles more strongly, and to interact particularly well with primary micelles of cholate [33], where the primary micelle in glycocholate is found in this work to occur at about 11 mM. The S-BNDHP guest appears to stabilize the primary GCA micelle sufficiently to delay the onset of secondary micellization.

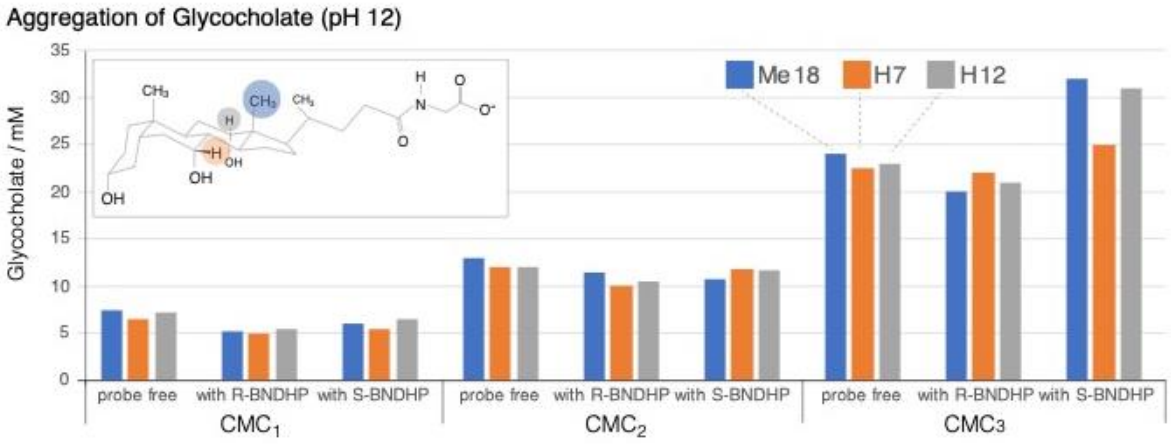
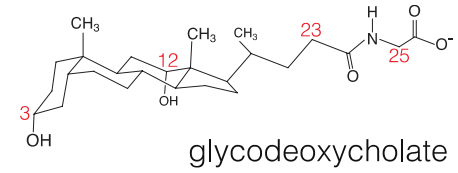
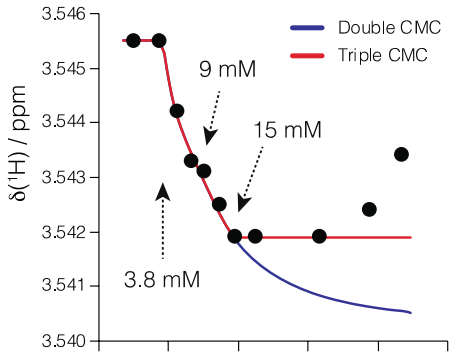


Figure 7: Results of triple-CMC modeling of representative protons on the β-face of a glycocholate concentration series (examples in **Figure S.7**, Supplemental Material).



glycodeoxycholate

a) H25 - steps 1,2,3



b) H25 - steps 1,3,4

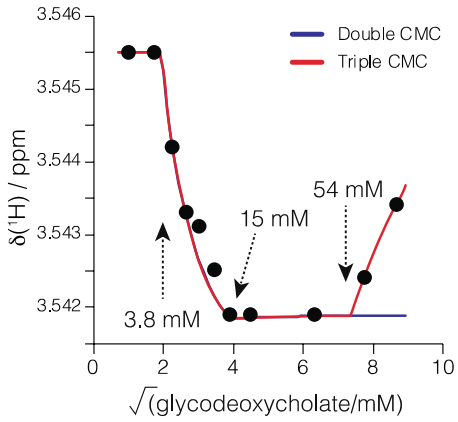


Figure 8: Multi-CMC modeling of several protons is able to resolve 4 CMCs for GDC aggregation (pH 12). Results for $\delta(^1H25)$ are shown here, while results for H3, H12, and H23 are shown in Supplemental Material (**Section S.8**). In (a) and (b), the chemical shift of H25 can be modeled for the first three CMCs in the former, or neglecting the 9 mM CMC in the latter. See **Table 2** for average values determined from these and some additional protons. The data permitted fixing 9 mM and 15 mM from global fits across the four protons considered.

Several chemical shifts of glycodeoxycholate (GDC) are modeled in **Figure 8** and the **Supplemental Material Section S8**. Despite spanning a very narrow chemical shift range, the trend of $\delta(^1H25)$ is remarkable for resolving singularities of all four CMCs (**Figures 8a-b)**. The $\delta(^1H25)$ trend is modeled for the first three CMCs in **Figure 8a**. Since $\delta(^1H25)$ is less sensitive to the 9 mM CMC, a separate model that ignores the 9 mM CMC explains the remaining variation in the data (**Figure 8b**). A structural inference is that the glycine functionalization participates in all aggregates of GDC. Other protons of GDC do not show as dramatic changes as H25, but confirm these CMCs (**Supplemental Material, Section S8**).

Bile acid	Conditions (298 K , pH 12)	Modeled Protons	CMC_1	CMC_2	CMC_3	CMC_4
Cholate[A]	probe free	Me18, Me19, Me21,H15β	-	16	41 ± 4	-
Deoxycholate[B]	probe free	Me18, Me19, Me21, H12β, H23a, H15β	3.8 ± .5	9.1 ± .3	27 ± 2	57 ± 4
	2.5 mM R-BNDHP	-	3.2 ± .4	9.4 ± .2	25.8 ± .7	60 ± 2
	2.5 mM S-BNDHP	-	2.3 ± .4	9.7 ± 1.2	24 ± 2	-
Glycocholate[C]	probe free	Me19, Me18, H3β , H7β, H12β	7.1 ± .4	12.4 ± .5	22.5 ± 1.1	-
	0.1 mM R-BNDHP	Me18, H7β, H12β	5.2 ± .3	10.7 ± .8	21 ±1	-
	0.1 mM S-BNDHP	Me19, Me18, H7β, H12β	5.8 ± .5	11.4 ± .3	29 ± 3	-
Glycodeoxycholate[D]	probe free	H3β, H12β, Me18, Me21, H23α/β, H25	3.8 ± .2	9	15	50 ± 6

Table 2. Summary of CMC values determined in this work with multi-CMC modeling spanning 1-100 mM bile concentrations. Uncertainties are standard deviations for global fits.
A. For cholate, CMC_1 is not supported in these probe-free data at 298 K, but in the presence of a probe (R- or S-BNDHP) CMC_1 = 7 mM has been determined (Hebling et al. [33]); a fixed CMC_2 =16 mM used in global models; CMC_4 not supported in the range 1-100 mM with these data.
B. For deoxycholate with R- or S-BNDHP, over 10 protons spanning the probe and the bile salt, and an independent trial at a different field strength, were used; with R-BNDHP CMC_4 = 60 mM CMC modeled with protons (H4-H6) of R-BNDHP; with S-BNDHP, a high-order CMC_4 ~40-50 mM is suggested but not sufficiently constrained by these data.
C. For glycocholate, CMC_4 is not supported in the range 1-100 mM with these data.
D. For glycodeoxycholate, global fitting fixed the 9 mM and 15 mM CMCs and uncertainties were not determined. The high order CMC_4 = 50 mM is an average of two values (H25 and H23β).

Discussion

A widely held understanding of bile aggregation is that it proceeds through three sequential steps [6, 65]: preliminary micelles, which could be as simple as back-to-back dimers, form first, followed by more stable primary micelles, and then very large secondary micelles. This work strongly supports that bile salt aggregation below 100 mM follows sufficiently discretized sequential CMCs (**Table 2**) in basic solutions (pH 12 , 298 K) to yield the three fundamental CMCs of many bile acids by a multi-CMC modeling strategy proposed here. Further, a fourth high order aggregate is supported here for deoxycholate and glycodeoxycholate. In other words, resolving closely spaced CMCs in this work (**Table 2**) helps clarify the multimodal literature trends noted in **Table 1**.

The ability of bile aggregates to solubilize planar hydrophobic guests at low concentration is well known [31, 33, 62, 66], however this work extends this understanding by further supporting that the first CMC_1 (Table 2) is particularly sensitive to the presence of probe and bile acid. For example, in cholate aggregation, CMC_1= 7 mM is unambiguously observed in the presence of probe [33] but is not resolved in the probe-free data obtained here indicating that CMC_1 can be stabilized by the presence of a probe and is also more favored by dihydroxy bile salts.

In seminal work by Small and coworkers, the methyl chemical shifts of bile salts were demonstrated to be reporters of bile aggregation, helping to establish, for example, that hydrophobic β-face packing (i.e. back-to-back) is important in bile micellization. [65] With higher resolution, it can be seen now that the chemical shifts of the methyls, steroidal ring protons, and chain protons of the bile acids as well as protons of guest molecules are rich sources of information for characterizing CMCs with diverse bile salts. The sensitivity of different regions of the bile acids to different CMCs conveys additional structural information, which is outside the scope of this work.

Recognizing that the first CMC, involving the smallest aggregates, is likely to display a significant degree of MA equilibrium behavior, this work focused on establishing a criterion for applying the PS model to the first CMC such that it closely modeled MA behavior (**Figure 2**). Subsequent CMCs are then incorporated, which are better approximated by the PS model. The multi-CMC strategy employed here was able to explain complex aggregation trends in chemical shift data so closely that it suggested the presence of a CMC where the model deviates from the data (e.g. **Figure 3** and others).

Although the three-CMC model could have in principle up to six adjustable parameters (o_{mic1}, o_{mic2}, o_{mic3}, CMC_1, CMC_2, CMC_3), the values of the CMCs are often tightly constrained by sharp changes in the slope of the data, and by global fitting of multiple protons, leading to small adjustments of a few parameters such as $o_{mic1}, o_{mic2}, \mathrm{and}\ o_{mic3}$. Chemical shift markers such as DSS and TSP are not used in order to avoid the potential to interact with micelles. Further, a broader recommendation of this work is that the two- and three-CMC phase separation models should be applied only where the CMCs are relatively well constrained by the data. Global modeling further strengthens confidence in the models, where NMR data are particularly compelling since multiple chemical shifts can be treated in the same data set.

While it is evident from the data presented here that global modeling of several protons will yield consensus CMC values, a small number of protons showed very complex trends

that could not be satisfactorily modeled in this work. Such cases may be due to local regions of structural heterogeneity or to local dynamics and merit further study.

One extension of the work would be to implement a hybrid model in which the first aggregation is treated explicitly with the mass action approach and subsequent steps with the PS model. Such a hybrid approach could better deal with the MA behavior of the first CMC on the one hand, but would also require incorporating its aggregation number n and counterion occupancy β.

As noted, the multi-CMC model is applied by a rater as described in the results section. Computational fitting is certainly possible, but would be highly constrained by user input in order to account for global fitting, to help identify a good match of the inflection points of the PS and MA behavior for the first CMC, and to ensure that subsequent CMCs also remain constrained. Exploring computational strategies, which could include machine learning in order to deal with these challenges, could certainly be fruitful.

Conclusion

Accurately describing bile aggregation has persisted as a difficult problem. Aggregation may not adhere perfectly to either the mass action or phase separation models [67], spans very low to very high concentrations [4], and some measurements appear resistant to a traditional CMC [1, 68]. Aggregation and counterion numbers are method dependent and exhibit variation [28], internal and surface structures vary from dihydroxy to trihydroxy bile salts [45, 46, 69], and the wealth of methods and conditions employed add enormous complexity.[23] In an effort to unravel such complex trends, this work shows that many bile salts have closely spaced discrete CMCs below 100 mM under the conditions studied here (298 K, pH 12). Moreover, the model produces the chemical shifts of the given proton in each of the micellar states, even though each micelle type can never be individually isolated and is spectroscopically dark.

This work shows that all three CMCs (preliminary/primary/secondary) for dihydroxy and trihydroxy bile salts can be resolved for bile salts, sometimes requiring the aid of probe molecules to stabilize the preliminary micelle of at least one trihydroxy bile salt (e.g. cholate [33]). In DC and GDC dihydroxy bile salts, a discrete fourth ‘high order’ CMC below 100 mM was clearly obtained. To address the challenges of mixed equilibrium and phase separation behavior, a hybrid mass action and phase separation approach successfully models the low-concentration aggregation step prior to switching to phase-separation treatments of subsequent steps. Complex aggregation-dependent NMR data

are closely explained with this straightforward model, relying on small numbers of parameters that are tightly constrained by global treatments of multiple protons.

Investigating additional bile salts under broader conditions is justified by these results. And although strong agreement is found between this newly developed model and a wide variety of experimental data, some data remained too complex to treat, and a drawback of the present method is that it does not reveal aggregation numbers or counterion occupancies, and may not capture continuous aggregation mechanisms. Future work is merited to refine the hybrid MA/PS approach, while the structural information of the 'dark state' chemical shifts can be investigated. The generality of the methods developed here may be explored for other experimental measurements beside NMR titration data as well as other self-aggregating systems.

Acknowledgements

The authors acknowledge the support of the National Science Foundation (NSF RUI #1800401). The acquisition of the 600 MHz NMR instrument was supported by National Science Foundation (MRI #0521108) and Bucknell University. The authors acknowledge the support of NSF (MRI # 2018547) for acquisition of the Q-TOF instrument for the HRMS data. The authors thanks Brian Breczinski for supervision of the Bucknell NMR facility and Dr. Morgan Olsen for supervision of the Q-TOF instrument.

References

[1] A. Jover, F. Fraga, F. Meijide, J. Vázquez Tato, J. Cautela, A. Del Giudice, M.C. di Gregorio, Revealing the complex self-assembly behaviour of sodium deoxycholate in aqueous solution, J Colloid Interface Sci. 604 (2021) 415-428, 10.1016/j.jcis.2021.06.140.
[2] M.C. Carey, in, L. Barbara, R.H. Dowling, A.F. Hofmann, E. Roda (Eds.), Springer Netherlands, Dordrecht, 1983, p. 19.
[3] M.C. Carey, D.M. Small, Micelle formation by bile salts. Physical-chemical and thermodynamic considerations, Arch Intern Med. 130 (1972) 506-527.
[4] D. Madenci, S.U. Egelhaaf, Self-assembly in aqueous bile salt solutions, Current Opinion in Colloid & Interface Science. 15 (2010) 109-115, //doi.org/10.1016/j.cocis.2009.11.010.
[5] N.A. Mazer, M.C. Carey, R.F. Kwasnick, G.B. Benedek, Quasielastic light scattering studies of aqueous biliary lipid systems. Size, shape, and thermodynamics of bile salt micelles, Biochemistry (N Y). 18 (1979) 3064-3075, 10.1021/bi00581a024.
[6] D.M. Small, Size and structure of bile salt micelles Influence of structure, concentration, counterion concentration, pH, and temperature, Adv. Chem. Ser. 84 (1968) 31-52.
[7] M.E. Smith, D.G. Morton, in, M.E. Smith, D.G. Morton (Eds.), Churchill Livingstone, 2010, p. 85.
[8] M.J. Alvarez-Figueroa, C. Muggli-Galaz, P.M. González, Effect of the aggregation state of bile salts on their transdermal absorption enhancing properties, Journal of Drug Delivery Science and Technology. 54 (2019) 101333, 10.1016/j.jddst.2019.101333.
[9] N.F. Goldshleger, A.S. Lobach, V.E. Baulin, A.Y. Tsivadze, Supramolecular gels based on bile salts, Russian Chem Rev. 86 (2017) 269-297.

[10] X. Guo, Q. Liu, S. Hu, W. Guo, Z. Yang, Y. Zhang, Thermodynamic models to elucidate the enantioseparation of drugs with two stereogenic centers by micellar electrokinetic chromatography, Journal of Chromatography A. 1512 (2017) 133-142, 10.1016/j.chroma.2017.07.025.
[11] C.D. Tran, I. Mejac, Chiral ionic liquids for enantioseparation of pharmaceutical products by capillary electrophoresis, J Chromatogr A. 1204 (2008) 204-209, 10.1016/j.chroma.2008.04.013 [doi].
[12] H.X. Chen, M.H. Huang, X.X. Zhang, Micellar electrokinetic chromatography analysis of tetrahydrogestrinone and related anabolic androgenic compounds, Anal Methods. 5 (2013) 5019-5023.
[13] H.X. Chen, Q.P. Deng, L.W. Zhang, X.X. Zhang, Quantification of testosterone and epitestosterone in biological samples by capillary electrophoresis with immunoaffinity extraction, Talanta. 78 (2009) 464-470, 10.1016/j.talanta.2008.11.048 [doi].
[14] S. Hu, X. Guo, H. Shi, R. Luo, Separation mechanisms for palonosetron stereoisomers at different chiral selector concentrations in MEKC, Electrophoresis. 36 (2015) 825-829, 10.1002/elps.201400341.
[15] E. Sanchez-Lopez, A. Salgado, A.L. Crego, M.L. Marina, Investigation on the enantioseparation of duloxetine by capillary electrophoresis, NMR, and mass spectrometry, Electrophoresis. 35 (2014) 2842-2847, 10.1002/elps.201300656 [doi].
[16] A. Perino, H. Demagny, Laura Velazquez-Villegas, K. Schoonjans, Molecular physiology of bile acid signaling in health, disease, and aging, Physiol Rev. 101 (2021) 683-731, 10.1152/physrev.00049.2019.
[17] D.S. Wishart, Metabolomics for Investigating Physiological and Pathophysiological Processes, Physiol Rev. 99 (2019) 1819-1875, 10.1152/physrev.00035.2018.
[18] R.M. Emma, R.B. Stephen, Bile acids and the metabolic syndrome, Ann Clin Biochem. 56 (2019) 326-337, 10.1177/0004563218817798.
[19] J.Y.L. Chiang, J.M. Ferrell, Bile Acids as Metabolic Regulators and Nutrient Sensors, Annu Rev Nutr. 39 (2019) 175-200, 10.1146/annurev-nutr-082018-124344.
[20] Y. Shulpekova, E. Shirokova, M. Zharkova, P. Tkachenko, I. Tikhonov, A. Stepanov, A. Sinitsyna, A. Izotov, T. Butkova, N. Shulpekova, V. Nechaev, I. Damulin, A. Okhlobystin, V. Ivashkin, A Recent Ten-Year Perspective: Bile Acid Metabolism and Signaling, Molecules. 27 (2022) , 10.3390/molecules27061983.
[21] R.M. Jain, M. Ben-Naim, M.P. Landry, M.S. Strano, Competitive binding in mixed surfactant systems for single-walled carbon nanotube separation, J Phys Chem C. 119 (2015) 22737-22745.
[22] A.A. Green, M.C. Duch, M.C. Hersam, Isolation of single-walled carbon nanotube enantiomers by density differention, Nano Research. 2 (2009) 69-77.
[23] B. Natalini, R. Sardella, A. Gioiello, F. Ianni, A. Di Michele, M. Marinozzi, Determination of bile salt critical micellization concentration on the road to drug discovery, J Pharm Biomed Anal. 87 (2014) 62-81, 10.1016/j.jpba.2013.06.029 [doi].
[24] D.R. Perinelli, M. Cespi, N. Lorusso, G.F. Palmieri, G. Bonacucina, P. Blasi, Surfactant Self-Assembling and Critical Micelle Concentration: One Approach Fits All?, Langmuir. 36 (2020) 5745-5753, 10.1021/acs.langmuir.0c00420.
[25] X. Cui, S. Mao, M. Liu, H. Yuan, Y. Du, Mechanism of Surfactant Micelle Formation, Langmuir. 24 (2008) 10771-10775, 10.1021/la801705y.
[26] M.M. Mabrouk, N.A. Hamed, F.R. Mansour, Physicochemical and electrochemical methods for determination of critical micelle concentrations of surfactants: a comprehensive review, Monatshefte für Chemie - Chemical Monthly. 153 (2022) 125-138, 10.1007/s00706-022-02891-2.
[27] S. Reis, C.G. Moutinho, C. Matos, B. de Castro, P. Gameiro, J.L.F.C. Lima, Noninvasive methods to determine the critical micelle concentration of some bile acid salts, Analyt. Biochem. 334 (2004) 117-126.
[28] A. Coello, F. Meijide, E.R. Nunez, J.V. Tato, Aggregation behavior of bile salts in aqueous solution, J Pharm Sci. 85 (1996) 9-15, 10.1021/js950326j [doi].
[29] C.J. O'Connor, R.G. Wallace, Physico-Chemical Behavior of Bile Salts, Adv. Colloid Interface Sci. 22 (1985) 1-111.
[30] K. Mukherjee, A. Barman, R. Biswas, Impact of the aggregation behaviour of sodium cholate and sodium deoxycholate on aqueous solution structure and dynamics: A combined time resolved fluorescence and dielectric relaxation spectroscopic study, Journal of Molecular Liquids. 222 (2016) 495-502, 10.1016/j.molliq.2016.07.053.

[31] C. Ju, C. Bohne, Probing bile salt aggregates by fluorescence quenching, Photochem Photobiol. 63 (1996) 60-67, 10.1111/j.1751-1097.1996.tb02992.x.
[32] F. Lopez, J. Samseth, K. Mortensen, E. Rosenqvist, J. Rouch, Micro- and Macrostructural Studies of Sodium Deoxycholate Micellar Complexes in Aqueous Solutions, Langmuir. 12 (1996) 6188-6196, 10.1021/la960006v.
[33] C.M. Hebling, L.E. Thompson, K.W. Eckenroad, G.A. Manley, R.A. Fry, K.T. Mueller, T.G. Strein, D. Rovnyak, Sodium cholate aggregation and chiral recognition of the probe molecule (R,S)-1,1'-binaphthyl-2,2'-diylhydrogenphosphate (BNDHP) observed by 1H and 31P NMR spectroscopy., Langmuir : the ACS journal of surfaces and colloids. 24 (2008) 13866-74, 10.1021/la802000x.
[34] S. Gouin, X.X. Zhu, Fluorescence and NMR Studies of the Effect of a Bile Acid Dimer on the Micellization of Bile Salts, Langmuir. 14 (1998) 4025-4029, 10.1021/la971155w [doi].
[35] M. Poša, S. Kevrešan, M. Mikov, V. Ćirin-Novta, C. Sârbu, K. Kuhajda, Determination of critical micellar concentrations of cholic acid and its keto derivatives, Colloids and Surfaces B: Biointerfaces. 59 (2007) 179-183, 10.1016/j.colsurfb.2007.05.008.
[36] M. Poša, V. Tepavčević, L. Grbović, M. Mikulić, K. Pavlović, Hydrophobicity and self-association (micellization) of bile salts with a lactone or lactam group in a steroid skeleton, J Phys Org Chem. 34 (2021) e4133, 10.1002/poc.4133.
[37] B. Jojart, B. Viskolcz, M. Posa, S.N. Fejer, Global optimization of cholic acid aggregates, J Chem Phys. 140 (2014) 144302, 10.1063/1.4869832 [doi].
[38] S.S. Mishra, S. Mohanty, J. Mishra, U. Subuddhi, Photophysical Properties of Coumarin 1 in Bile Salt Aggregates: An Insight into the Role of Bile Salt Structure on the Aggregation Behavior, Langmuir. 35 (2019) 16555-16567, 10.1021/acs.langmuir.9b02664.
[39] K. Matsuoka, Y. Moroi, Micelle formation of sodium deoxycholate and sodium ursodeoxycholate (part 1), Biochim Biophys Acta. 1580 (2002) 189-199, S1388198101002037 [pii].
[40] P. Garidel, A. Hildebrand, R. Neubert, A. Blume, Thermodynamic characterization of bile salt aggregation as a function of temperature and ionic strength using isothermal titration calorimetry, Langmuir. 16 (2000) 5267-5275, 10.1021/la9912390.
[41] S.L. Anderson, D. Rovnyak, T.G. Strein, Direct Measurement of the Thermodynamics of Chiral Recognition in Bile Salt Micelles, Chirality. 28(4) (2016) 290-298.
[42] H. Sugioka, K. Matsuoka, Y. Moroi, Temperature effect on formation of sodium cholate micelles, J Colloid Interface Sci. 259 (2003) 156-162, 10.1016/S0021-9797(02)00191-1.
[43] U. Subuddhi, A.K. Mishra, Micellization of bile salts in aqueous medium: A fluorescence study, Colloids and Surfaces B: Biointerfaces. 57 (2007) 102-107, 10.1016/j.colsurfb.2007.01.009.
[44] P.K. Jana, S.P. Moulik, Interaction of bile salts with hexadecyltrimethylammonium bromide and sodium dodecyl sulfate, J Phys Chem. 95 (1991) 9525-9532, 10.1021/j100176a089.
[45] A.R. Meier, J.B. Yehl, K.W. Eckenroad, G.A. Manley, T.G. Strein, D. Rovnyak, Stepwise Aggregation of Cholate and Deoxycholate Dictates the Formation and Loss of Surface-Available Chirally Selective Binding Sites, Langmuir. 34 (2018) 6489-6501, 10.1021/acs.langmuir.8b00467.
[46] H. Kawamura, Y. Murata, T. Yamaguchi, H. Igimi, M. Tanaka, G. Sugihara, J.P. Kratohvil, Spin-label studies of bile salt micelles, J Phys Chem. 93 (1989) 3321-3326.
[47] S. Das, J. Dey, T. Mukhim, K. Ismail, Effect of sodium salicylate, sodium oxalate, and sodium chloride on the micellization and adsorption of sodium deoxycholate in aqueous solutions, J Colloid Interface Sci. 357 (2011) 434-439, 10.1016/j.jcis.2011.02.020.
[48] K.F. Morris, A.L. Froberg ., B.A. Becker, V.K. Almeida, J. Tarus, C.K. Larive, Using NMR to Develop Insights into Electrokinetic Chromatography, Anal Chem. 77 (2005) 254A-263A, 10.1021/ac0534071.
[49] K.F. Morris, E.J. Billiot, F.H. Billiot, K.B. Lipkowitz, W.M. Southerland, Y. Fang, Investigation of Chiral Molecular Micelles by NMR Spectroscopy and Molecular Dynamics Simulation , Open Journal of Physical Chemistry. 2 (2012) 240-251, 10.4236/ojpc.2012.24032.
[50] J.F. Yan, M.B. Palmer, A Nuclear Magnetic Resonance Method for Determination of Critical Micelle Concentration, J Colloid Interface Sci. 30 (1969) 177-182, 10.1016/S0021-9797(69)80004-4.
[51] O. Söderman, P. Stilbs, NMR studies of complex surfactant systems, Prog Nucl Magn Reson Spectrosc. 26 (1994) 445-482, 10.1016/0079-6565(94)80013-8.
[52] U.R.M. Kjellin, J. Reimer, P. Hansson, An investigation of dynamic surface tension, critical micelle concentration, and aggregation number of three nonionic surfactants using NMR, time-resolved

fluorescence quenching, and maximum bubble pressure tensiometry, J Colloid Interface Sci. 262 (2003) 506-515, 10.1016/S0021-9797(03)00168-1.
[53] A. Hildebrand, P. Garidel, R. Neubert, A. Blume, Thermodynamics of demicellization of mixed micelles composed of sodium oleate and bile salts, Langmuir. 20 (2004) 320-328, 10.1021/la035526m.
[54] N.E. Olesen, P. Westh, R. Holm, Determination of thermodynamic potentials and the aggregation number for micelles with the mass-action model by isothermal titration calorimetry: A case study on bile salts, J Colloid Interface Sci. 453 (2015) 79-89, https://doi.org/10.1016/j.jcis.2015.03.069.
[55] A. Kroflic, B. Sarac, M. Bester-Rogac, Thermodynamic characterization of 3-[(3-cholamidopropyl)-dimethylammonium]-1-propanesulfonate (CHAPS) micellization using isothermal titration calorimetry: temperature, salt, and pH dependence, Langmuir. 28 (2012) 10363-10371, 10.1021/la302133q [doi].
[56] A. Pluckthun, E.A. Dennis, 31P nuclear magnetic resonance study on the incorporation of monomeric phospholipids into nonionic detergent micelles, J Phys Chem. 86 (1981) 678-683.
[57] O. Söderman, P. Stilbs, W.S. Price, NMR studies of surfactants, Concepts in Magnetic Resonance Part A. 23A (2004) 121-135, 10.1002/cmr.a.20022.
[58] A.I. Rusanov, The Mass-Action-Law Theory of Micellization Revisited, Langmuir. 30 (2014) 14443-14451, 10.1021/la503770a.
[59] A.I. Rusanov, The mass action law theory of micellar solutions, Adv Colloid Interface Sci. 45 (1993) 1-78, https://doi.org/10.1016/0001-8686(93)80027-9.
[60] A.I. Rusanov, The wonderful world of micelles, Colloid Journal. 76 (2014) 121-126, 10.1134/S1061933X14020070.
[61] K.ō Shinoda, E. Hutchinson, PSEUDO-PHASE SEPARATION MODEL FOR THERMODYNAMIC CALCULATIONS ON MICELLAR SOLUTIONS1, J Phys Chem. 66 (1962) 577-582, 10.1021/j100810a001.
[62] K.W. Eckenroad, G.A. Manley, J.B. Yehl, R.T. Pirnie, T.G. Strein, D. Rovnyak, An Edge Selection Mechanism for Chirally Selective Solubilization of Binaphthyl Atropisomeric Guests by Cholate and Deoxycholate Micelles, Chirality. 28 (2016) 525-533, 10.1002/chir.22609.
[63] T.J. Wenzel, Discrimination of Chiral Compounds using NMR Spectroscopy, Wiley-Interscience, 2007.
[64] K.W. Eckenroad, L.E. Thompson, T.G. Strein, D. Rovnyak, Proton NMR assignments for R,S-1,1'-binaphthol (BN) and R,S-1,1'-binaphthyl-2,2'-diyl hydrogen phosphate (BNDHP) interacting with bile salt micelles, Magn Reson Chem. 45 (2007) 72-75, 10.1002/mrc.1916.
[65] D.M. Small, S.A. Penkett, D. Chapman, Studies on simple and mixed bile salt micelles by nuclear magnetic resonance spectroscopy, Biochimica Et Biophysica Acta. 176 (1969) 178-189.
[66] O.L. Waissbluth, M.C. Morales, C. Bohne, Influence of planarity and size on guest binding with sodium cholate aggregates, Photochem Photobiol. 82 (2006) 1030-1038, 10.1562/2006-02-14-RA-803.
[67] C.J. O'Connor, B.T. Ch'ng, R.G. Wallace, Studies in bile salt solutions: 1. Surface tension evidence for a stepwise aggregation model, Journal of Colloid and Interface Science. 95 (1983) 410-419, 10.1016/0021-9797(83)90200-X.
[68] A. Seret, M. Bahri, The CMC-like behaviour of bile salts as probed by photoexcited Rose Bengal, Colloids Surf Physicochem Eng Aspects. 339 (2009) 153-158, 10.1016/j.colsurfa.2009.02.015.
[69] H. Kawamura, M. Manabe, T. Narikiyo, H. Igimi, Y. Murata, G. Sugihara, M. Tanaka, Spin and fluorescence label studies on bile salt micelles, Journal of Solution Chemistry. 16 (1987) 433-441, 10.1007/BF00648594.