Physical Chemistry Chemical Physics

PCCP

The origin of cooperative solubilisation by hydrotropes

Journal:	*Physical Chemistry Chemical Physics*
Manuscript ID	CP-ART-07-2016-004823.R1
Article Type:	Paper
Date Submitted by the Author:	17-Aug-2016
Complete List of Authors:	Shimizu, Seishi; University of York, Department of Chemistry Matubayasi, Nobuyuki; Osaka University, Graduate School of Engineering Science

The origin of cooperative solubilisation by hydrotropes

Seishi Shimizu[1,*] and Nobuyuki Matubayasi[2,3]

[1]York Structural Biology Laboratory, Department of Chemistry, University of York, Heslington, York YO10 5DD, United Kingdom

[2]Division of Chemical Engineering, Graduate School of Engineering Science, Osaka University, Toyonaka, Osaka 560-8531, Japan

[3]Elements Strategy Initiative for Catalysts and Batteries, Kyoto University, Katsura, Kyoto 615-8520, Japan

KEYWORDS: Hydrotropy; McMillan-Mayer Theory; Water; Cosolvent; Cooperativity; Minimum Hydrotrope Concentration

AUTHOR INFORMATION

Corresponding Author:

Seishi Shimizu

York Structural Biology Laboratory, Department of Chemistry, University of York, Heslington, York YO10 5DD, United Kingdom

Tel: +44 1904 328281, Fax: +44 1904 328281, Email: seishi.shimizu@york.ac.uk

ABSTRACT

The signature of hydrotropic solubilisation is the sigmoidal solubility curve; when plotted against hydrotrope concentration, solubility increases suddenly after the minimum hydrotrope concentration (MHC), and reaches a plateau at higher hydrotrope concentrations. This sigmoidal curve is characteristic of cooperative phenomena, yet the true molecular basis of hydrotropic cooperativity has long remained unclear. Here we develop a theory, derived from the first principles of statistical thermodynamics using partially-open ensembles, to identify the origin of hydrophobic cooperativity. Our theory bears a close resemblance to the cooperative binding model used for protein-ligand binding. The cause of cooperativity is the enhancement of hydrotrope *m*-body interaction induced by the presence of the solute; *m* can be estimated from experimental solubility data.

1. Introduction

Hydrotropes can increase the solubility of hydrophobic solutes up to several orders of magnitude, hence have a number of important industrial applications.[1-6] The signature characteristics of hydrotropes, which sets them apart from other cosolvents, is the sigmoidal solubility curve, or more specifically:[2-6]

1. Solubility hardly increases at low hydrotrope concentrations (<0.5 molar).
2. Above a certain threshold concentration, commonly referred to as the minimum hydrotrope concentration (MHC), solubility increases suddenly.
3. Solubility ceases to increase after a few molars of hydrotropes (saturation of solubilisation).

Such sigmoidal solubilisation curves (Figure 1) are reminiscent of cooperative phenomena.[7]

What is the origin of hydrotropic cooperativity? Because of the apparent similarity between MHC and critical micelle concentration (CMC), [7-9] the self-aggregation of hydrotrope in the bulk phase has often been considered to be the origin of hydrotropic cooperativity.[7-13] However, the fact that MHC is observed even for urea, which forms near-ideal mixture with water, has seriously challenged this hypothesis.[14-17]

Rigorous statistical thermodynamics, on the contrary, has shown that *bulk-phase* self-aggregation is not the cause of MHC.[14-17] Furthermore, such bulk-phase self-aggregation reduces the effective number of hydrotrope molecules, thereby reducing the solubilisation efficiency per hydrotrope molecule.[14-17] Instead, MHC is caused by the *enhanced* hydrotrope self-aggregation in the presence of the solute.[16] This conclusion was reached through the use of the rigorous Kirkwood-Buff (KB) theory,[18-30] which identified, without any approximations, the driving forces of hydrotropic solubilisation,[14-17] which has later been supported by further evidence.[6,31]

Even though our rigorous KB-based approach has revealed the new, universal principles of hydrotropy, there still are two shortcomings: (i) the mathematical form of the theory on the

origin of MHC is complex and is difficult to use;[14-17,32] (ii) the origin of solubilisation plateau remains unexplained.[16,17]

Hence there is a need for a simple theory, which can identify the cause of hydrotropic cooperativity. We start from the first principles of statistical thermodynamics.[30,33,34] Guided by an analogy between ligand binding cooperativity[35-38] and hydrotropic cooperativity, we propose how solubility curve should be analysed, in order to reveal the key aggregation interactions responsible for hydrotropic cooperativity.

2. A statistical thermodynamic foundation for hydrotropic cooperativity

The goal of our theory is to express solubility as a function of hydrotrope concentration. We first note that solubility measurements are usually carried out in the isobaric-isothermal (NPT) ensemble, whereas the concentration polynomial expansion of thermodynamic quantities is carried out in the grand canonical (μVT) ensemble.[30,32-34] Our goal is to obtain such an expansion under an isobaric ensemble. This can be done by the use of the partially-open isobaric ensemble pioneered decades ago by Stockmeyer and Hill.[39-43]

Consider a three component solution consisting of a solute (i=u), water (i=1), and hydrotrope (i=2) molecules. Let μ_i and N_i respectively be the chemical potential and the number of species i, T, V and P respectively be the temperature, volume, and pressure of the system. Let $\rho_i = N_i/V$ be the number density or concentration of species i; when N_i is not kept constant, the ensemble average of N_i is used instead to define ρ_i, such that $\rho_i = \langle N_i \rangle / V$. The convention $\beta = \frac{1}{kT}$ (where

k is the Boltzmann constant) is used throughout. Since T is kept constant throughout this paper, it is often omitted in the subsequent discussions.

To calculate the solvation free energy of a solute molecule, the pair of systems, with and without the solute, needs to be considered.[16] When the solute is present, it is fixed at the origin and acts as the source for an external field for the water and hydrotrope molecules. In this case, the solution system is inhomogeneous.[16] When the solute is absent, the system consists only of water and hydrotrope; hence this system is homogeneous. The partially-open partition functions for the two-component system with and without the solute molecule can be expressed respectively in the following[16,30,32-34]

$$\Gamma_{\mathrm{u}}(T,P,N_1,\mu_2) = \sum_{N_2\geq 0} \lambda_2^{N_2}\, \mathrm{R}_{\mathrm{u}}(T,P,N_1,N_2) \tag{1}$$

$$\Gamma(T,P,N_1,\mu_2) = \sum_{N_2\geq 0} \lambda_2^{N_2} \mathrm{R}(T,P,N_1,N_2) \tag{2}$$

where λ_i is the fugacity of the species *i,* is defined as[16,30,32-34]

$$\lambda_i = \exp(\beta\mu_i) \tag{3}$$

and the isobaric-isothermal partition functions are defined as

$$\mathrm{R}_{\mathrm{u}}(T,P,N_1,N_2) = \frac{1}{N_1!N_2!}\frac{q_1^{N_1}q_2^{N_2}}{\Lambda_1^{3N_1}\Lambda_2^{3N_2}} \int dV\, e^{-\beta PV} \int d\boldsymbol{x}_u d\boldsymbol{X}^{N_1} d\boldsymbol{X}^{N_2} e^{-\beta U(\boldsymbol{x}_u,\boldsymbol{X}^{N_1},\boldsymbol{X}^{N_2})} \tag{4}$$

$$\mathrm{R}(T,P,N_1,N_2) = \frac{1}{N_1!N_2!}\frac{q_1^{N_1}q_2^{N_2}}{\Lambda_1^{3N_1}\Lambda_2^{3N_2}} \int dV e^{-\beta PV} \int d\boldsymbol{X}^{N_1} d\boldsymbol{X}^{N_2} e^{-\beta U(\boldsymbol{X}^{N_1},\boldsymbol{X}^{N_2})} \tag{5}$$

where Λ_i is the de Broglie wavelength, q_i is the intramolecular partition function, $\boldsymbol{X}^{N_1}$ and $\boldsymbol{X}^{N_2}$ denote collectively the coordinates of the species 1 and 2, respectively, and $\boldsymbol{x}_u$ is the internal coordinates of the solute.

Connection to thermodynamics can be made through the following formulae

$$(N_1\mu_1)_u = -kT \ln \Gamma_u(T, P, N_1, \mu_2) \quad (6)$$

$$(N_1\mu_1) = -kT \ln \Gamma(T, P, N_1, \mu_2) \quad (7)$$

in which T, kept constant throughout the discussion, is omitted. Most importantly, the chemical potential of the fixed solute μ_u^* can be expressed in terms of the partially-open partition functions in the following manner:[16,32]

$$e^{-\beta\mu_u^*} = e^{-\beta[(N_1\mu_1)_u - (N_1\mu_1)]} = \frac{\Gamma_u(T,P,N_1,\mu_2)}{\Gamma(T,P,N_1,\mu_2)} \quad (8)$$

The main concern of this paper is the solubility increase which accompanies the introduction of the hydrotrope molecule. This, in the language of thermodynamics, is due to the change of μ_u^* from its pure water value μ_u^{*0}

$$\Delta\mu_u^* = \mu_u^* - \mu_u^{*0} \quad (9)$$

Solubility increase is expressed, using eqn (9), as follows

$$e^{-\beta\Delta\mu_u^*} = \frac{\Gamma_u(T,P,N_1,\mu_2)}{\Gamma(T,P,N_1,\mu_2)} \frac{\Gamma(T,P,N_1\infty)}{\Gamma_u(T,P,N_1,\infty)} = \frac{\frac{\Gamma_u(T,P,N_1,\mu_2)}{\Gamma_u(T,P,N_1,\infty)}}{\frac{\Gamma(T,P,N_1,\mu_2)}{\Gamma(T,P,N_1\infty)}} \quad (10)$$

where $\Gamma_u(T, P, N_1, \infty)$ and $\Gamma(T, P, N_1\infty)$ refers to the partition functions in pure-water solvent, in which the concentration of the hydrotrope is 0 and its chemical potential μ_2 diverges.

3. The local subsystem open only to hydrotropes

To construct a theory of hydrotropy in an analogous manner to the cooperative binding theory, here we aim to express $e^{-\beta\Delta\mu_u^*}$ in terms of the "local" distribution of hydrotropes around the solute. To this end, let us introduce a "local" subsystem around the solute, which lies within the macroscopic systems introduced in Section 2. Let the boundary of this inhomogeneous subsystem be the range of solute-hydrotrope correlation, and v be the volume of this subsystem,

which is kept constant throughout. Following the classical works of Stockmeyer[37] and Schellman,[38] let this subsystem be partially open, namely open only to the hydrotropes. We shall later show how to specify v from the behaviours of R_u and R.

Let us now see the consequence of introducing the local subsystem. To this end, let us first note that there are $\frac{N_2!}{n_2!(N_2-n_2)!}$ ways of choosing n_2 molecules out of N_2 identical molecules to be placed within the local subsystem. Using the constancy of v in order to move $\int d\boldsymbol{X}^{n_2}$ out of the total volume integral $\int dV$ in the isobaric ensemble, the partially-open partition function can be rewritten as

$$\Gamma_{\mathrm{u}}(T,P,N_1,\mu_2) = \sum_{N_2\geq 0}\lambda_2^{N_2}\sum_{n_2=0}^{N_2}\frac{N_2!}{n_2!(N_2-n_2)!}\mathrm{R}_{\mathrm{u}}(T,P,N_1,N_2)$$

$$= \sum_{n_2\geq 0}\frac{\lambda_2^{n_2}}{n_2!}\frac{q_2^{n_2}}{\Lambda_2^{3n_2}}\int d\boldsymbol{X}^{n_2}\mathrm{R}_{\mathrm{u}}\left(T,P,N_1,N_2^b;\boldsymbol{X}^{n_2}\right)\Gamma_{\mathrm{u}}(T,P,N_1,\infty) \quad (11)$$

where, at the last step, a new variable is introduced as $N_2^b = N_2 - n_2$. Note that the kernel of the integration has been denoted here as $\mathrm{R}_{\mathrm{u}}\left(T,P,N_1,N_2^b;\boldsymbol{X}^{n_2}\right)\Gamma_{\mathrm{u}}(T,P,N_1,\infty)$, because of a clear physical meaning which can be attributed to $\mathrm{R}_{\mathrm{u}}\left(T,P,N_1,N_2^b;\boldsymbol{X}^{n_2}\right)$. This can be appreciated by rewriting, using eqn (11), the numerator of eqn (10) into the following form:

$$\frac{\Gamma_{\mathrm{u}}(T,P,N_1,\mu_2)}{\Gamma_{\mathrm{u}}(T,P,N_1\infty)} = \sum_{n_2\geq 0}\lambda_2^{n_2}\,R_{u,n_2} \quad (12)$$

$$R_{u,n_2} = \frac{q_2^{n_2}}{n_2!\Lambda_2^{3n_2}}\int d\boldsymbol{X}^{n_2}\mathrm{R}_{\mathrm{u}}\left(T,P,N_1,N_2^b;\boldsymbol{X}^{n_2}\right) \quad (13)$$

$\mathrm{R}_{\mathrm{u}}\left(T,P,N_1,N_2^b;\boldsymbol{X}^{n_2}\right)$, according to eqn (13), has a clear physical meaning: the fugacity of inserting n_2 identical hydrotrope molecules fixed at the configuration $\boldsymbol{X}^{n_2}$.

Importantly, R_{u,n_2} is a microscopic quantity, since the range of the integral is over the local system, which is microscopic. In a similar vein, a homogeneous subsystem must be defined in order to complete the link between solubilisation and local hydrotrope distribution. Let us consider the same volume v, which does not contain any solute molecules and sets its origin at the centre-of-mass position of the solute that is to be inserted.

Using the same argument which led to eqn (12) and (13), the following relationships for the bulk solution:

$$\frac{\Gamma(T,P,N_1,\mu_2)}{\Gamma(T,P,N_1\infty)} = \sum_{n_2\geq 0} \lambda_2^{n_2} R_{n_2} \tag{14}$$

$$R_{n_2} = \frac{q_2^{n_2}}{n_2!\Lambda_2^{3n_2}} \int d\boldsymbol{X}^{n_2} \mathrm{R}\left(T, P, N_1, N_2^b; \boldsymbol{X}^{n_2}\right) \tag{15}$$

Solubilisation can now be linked to the local distribution of hydrotropes; this can be achieved by combining eqn (10), (12)-(15) in following form:

$$e^{-\beta\Delta\mu_u^*} = \frac{1+\sum_{n_2\geq 1} R_{u,n_2}\lambda_2^{n_2}}{1+\sum_{n_2\geq 1} R_{n_2}\lambda_2^{n_2}} \tag{16}$$

We emphasise that R_{u,n_2} and R_{n_2} in eqn (16) are microscopic.

Now we derive a rational polynomial expansion of $e^{-\beta\Delta\mu_u^*}$ based upon eqn (16). To do so, we employ the elegant method of the MM theory, namely to consider $N_2^b \to 0$ limit of R_{u,n_2} and R_{n_2}. In our definition of the local subsystem, v represents the range of correlation between solute and hydrotrope-hydrotrope interaction; putting $N_2^b \to 0$ is equivalent to ignoring the contribution from the hydrotrope molecules outside of the correlation range. Hence eqn (16), when

considered under the $N_2^b \to 0$ limit, serves as the basis for the rational polynomial expansion of solubilisation that we sought after.

4. Hydrotropic cooperativity versus binding cooperativity

Our result, eqn (16) (at $N_2^b \to 0$ limit) is mathematically analogous to binding polynomials[35-37] in the theory of cooperative binding. This can be better appreciated by a trivial rewriting of eqn (16),

$$e^{-\beta\Delta\mu_u^*} - 1 = \frac{\sum_{n_2\geq 1} \Delta R_{n_2} \lambda_2^{n_2}}{1+\sum_{n_2\geq 1} R_{n_2} \lambda_2^{n_2}} \quad (17)$$

in which $\Delta R_{n_2} = R_{u,n_2} - R_{n_2}$. This equation is analogous to binding polynomials.

The analogy between hydrotropic solubilisation and cooperative binding theories opens up a new possibility towards revealing the mechanism of hydrotropic cooperativity. (In fact it is considered to be the continuation of the classical attempts to extend the theory of binding to weak-nonspecific interactions characteristic of solvation).[39,44,45] We assume that the summations in the denominator and numerator of eqn (16) are dominated by a few terms for both, which will greatly simplify the analysis. To this end, let us rewrite eqn (17) in the following manner:

$$e^{-\beta\Delta\mu_u^*} - 1 = \sum_{n_2} p_{n_2} \left(\frac{R_{u,n_2} - R_{n_2}}{R_{n_2}}\right) \quad (18)$$

where

$$p_{n_2} = \frac{R_{n_2} \lambda_2^{n_2}}{1+\sum R_{n_2} \lambda_2^{n_2}} \quad (19)$$

signifies the probability of finding n_2 hydrotrope molecules in the local subsystem without the solute. Now, the denominator and numerator of eqn (16) can be shown to be dominated by a few terms if the following conditions are met, regardless of the n_2-dependence of p_{n_2}:

a. $\frac{R_{u,n_2}-R_{n_2}}{R_{n_2}} = 0$ for $n_2 < \overline{n_2} - \Delta n_2$

b. $\frac{R_{u,n_2}-R_{n_2}}{R_{n_2}}$ is large and positive for $\overline{n_2} - \Delta n_2 < n_2 < \overline{n_2} + \Delta n_2$

c. $\frac{R_{u,n_2}-R_{n_2}}{R_{n_2}} = 0$ for $n_2 > \overline{n_2} + \Delta n_2$

d. Δn_2 is at its minimum, or $\Delta n_2 \simeq 0$

In Appendix, we show that these four conditions will lead to a unique determination of v and $\overline{n_2}$, and that v signifies the range within which hydrotrope-hydrotrope interactions are affected by the presence of a solute.

There is also a logically possible yet less likely scenario that p_{n_2} is non-zero only between $n_2 = \overline{n_2} - \Delta n_2$ and $\overline{n_2} + \Delta n_2$, and $\frac{R_{u,n_2}-R_{n_2}}{R_{n_2}}$ is positive in the same range. In either case, if we neglect the peak width such that $\Delta n_2 = 0$, eqn (18) can be rewritten as

$$e^{-\beta\Delta\mu^*_{u,NPT}} - 1 = \frac{\Delta R_{\overline{n_2}}\lambda_2^{\overline{n_2}}}{1+R_{\overline{n_2}}\lambda_2^{\overline{n_2}}} \quad (20)$$

Eqn (20) is analogous to the Hill model of binding cooperativity.[35-37,45] Now we modify eqn. (20) in two different ways. The first is completely general; we rewrite eqn (20) in terms of the hydrotrope activity, a_2. This can be done by using the the standard chemical potential of hydrotrope $\mu_2^o = \mu_2 - RT\ln a_2$ and the corresponding fugacity of pure hydrotrope $\lambda_2^o = \exp(\beta\mu_2^o)$, as

$$e^{-\beta\Delta\mu^*_u} - 1 = \frac{\Delta R'_{\overline{n_2}}a_2^{\overline{n_2}}}{1+R'_{\overline{n_2}}a_2^{\overline{n_2}}} \quad (21)$$

Where

$$\Delta R'_{\overline{n_2}} = (\lambda_2^o)^{\overline{n_2}} R_{u,\overline{n_2}} - (\lambda_2^o)^{\overline{n_2}} R_{\overline{n_2}} \quad (22)$$

$$R'_{\overline{n_2}} = (\lambda_2^o)^{\overline{n_2}} R_{\overline{n_2}} \quad (23)$$

A simplification of eqn (21)-(23) is possible when the aqueous hydrotrope solution obeys the dilute ideal solution. In this case, eqn (20) can be rewritten in terms of the mole fraction of the hydrotrope, x_2. This can be achieved by using the dilute ideal standard chemical potential $\mu_2^{ox} = \mu_2 - RT \ln x_2$ and the corresponding standard fugacity $\lambda_2^{ox} = \exp(\beta \mu_2^{ox})$, as

$$e^{-\beta \Delta \mu_u^*} - 1 = \frac{\Delta R''_{\overline{n_2}} x_2^{\overline{n_2}}}{1 + R''_{\overline{n_2}} x_2^{\overline{n_2}}} \quad (24)$$

where

$$\Delta R''_{\overline{n_2}} = (\lambda_2^{ox})^{\overline{n_2}} R_{u,\overline{n_2}} - (\lambda_2^{ox})^{\overline{n_2}} R_{\overline{n_2}} \quad (25)$$

$$R''_{\overline{n_2}} = (\lambda_2^{ox})^{\overline{n_2}} R_{\overline{n_2}} \quad (26)$$

Note that $\overline{n_2}$ does not depend on the choice of the standard state of the chemical potential, i.e., $\overline{n_2}$ is the same for eqn (20), (21) and (24).

We emphasize there that we have introduced the conditions a-d to specify when hydrotrope-induced solubilization behaves in a cooperative manner (eqn (20)). In our present theoretical formalism, the validity of these conditions can only be verified through how well eqn (20), (21) or (24) can fit the solubility data, as will be demonstrated in the next section.

5. Linearised plot for analysing solubility data

In the study of ligand binding, key parameters for binding cooperativity can be obtained from experimental binding data through the linear plot, which can visually show how well the experimental data fits the model.[35-37] A formal analogy between hydrotropy and cooperative binding suggests that this powerful method can be extended to hydrotropic cooperativity.

We aim to reproduce the overall shape of the solubility curve by eqn (21) or eqn (24). Here we focus on eqn (24), because aqueous hydrotrope solutions can often be treated as dilute ideal solutions. Eqn (24) contains only three parameters, $\overline{n_2}$, $R''_{\overline{n_2}}$, and $\Delta R''_{\overline{n_2}} = (\lambda_2^{ox})^{\overline{n_2}} R_{u,\overline{n_2}} - (\lambda_2^{ox})^{\overline{n_2}} R_{\overline{n_2}}$. For simplicity, let us first exploit that solubilisation $e^{-\beta\Delta\mu_u^*}$ converges to a plateau, whose value will be referred to as $e^{-\beta\Delta\mu_u^{*,sat}}$ at large x_2, which, according to eqn (24), leads to

$$\frac{R_{u,\overline{n_2}}}{R_{\overline{n_2}}} = e^{-\beta\Delta\mu_u^{*,sat}} \tag{27}$$

Combining eqn (24) and (27), we obtain

$$\ln\left[\frac{1-e^{-\beta\Delta\mu_u^*}}{e^{-\beta\Delta\mu_u^*}-e^{-\beta\Delta\mu_u^{*,sat}}}\right] = \overline{n_2}\,\ln x_2 + \ln R''_{\overline{n_2}} \tag{28}$$

which suggests that by plotting $\ln\left[\frac{1-e^{-\beta\Delta\mu_u^*}}{e^{-\beta\Delta\mu_u^*}-e^{-\beta\Delta\mu_u^{*,sat}}}\right]$ against $\ln x_2$, $\overline{n_2}$ can be determined from the gradient and $\ln R''_{\overline{n_2}}$ from the intercept.

6. Hydrotropic cooperativity from solubility data

Now we apply eqn (28) to analyse experimental data.[15,47-49] As model systems, the solutes butyl acetate (BA) and benzyl benzoate (BB) in water are solubilized by the hydrotropes urea, sodium benzoate (sb) and sodium salicylate (ss). Solubilization reaches plateau at approximately $x_2 \simeq 0.04$ for urea and 0.08 for ss and sb, during which the activity coefficients of water, γ_1,

hardly deviate from 1, meaning that the aqueous hydrotrope solution can be considered as ideal dilute solution. Hence we use the theory for dilute ideal solutions, eqn (24).

Figure 2 presents the linearised plot (eqn (28)), which exhibits near-straight lines below solubility saturation concentration. This shows that eqn (24) is indeed a good approximation. Based upon this linear fit, and the resultant parameters tabulated in Table 1, we have made a direct comparison (Figure 3) between eqn (24) and experimental data. Figure 3 shows that eqn (24) captures the sigmoidal concentration dependence, hence hydrotropic cooperativity, with only three parameters.

Note that the $\overline{n_2}$ values determined from Figure 2 are not integers. This is not surprising, considering that eqn (24) is based upon an approximation $\Delta n_2 = 0$ which means the abrupt onset of the cooperative effect. Hence $\overline{n_2} = 4.37$ for BA in urea, for example, indicates that the cooperative effect is operative around $n_2 = 4$ or 5.

Our theory has clarified that MHC and saturation both arise from the same physical origin: a large positive $(R_{u,n_2} - R_{n_2})$ at around a certain $n_2 = \overline{n_2}$. Considering that R_{n_2} signifies the fugacity of inserting n_2 hydrotropes into the system, there are two possible scenarios:

1. Around $n_2 = \overline{n_2}$, inserting n_2 hydrotropes becomes easier in the presence of the solute compared to the bulk. (Increase of R_{u,n_2} from R_{n_2})
2. Around $n_2 = \overline{n_2}$, inserting n_2 hydrotropes becomes more difficult in the bulk phase. (R_{n_2} as a function of n_2 has a minimum; see below).

In the case of monomeric hydrotropes devoid of micelle formation, scenario 1 would be the mechanism of solubilisation, which is consistent with our previous conclusion based upon the KB theory of solution.[14-17] If there is strong anti-cooperativity in hydrotrope self-association, scenario 2 may be possible; however, what is commonly reported in the hydrotrope literature is on the contrary the self-aggregation of hydrotropes.[14-17] In the case of micellar hydrotropes, R_{n_2} takes a maximum value when n_2 is equal to the aggregation number. Scenario 2 is therefore not likely; hence according to Scenario 1, the fugacity of n_2 hydrotropes increases in the presence of a solute due to a strong solute-hydrotrope binding.

7. Discussion

Our theory has identified the cause of hydrotropic cooperativity: the strengthening of $\overline{n_2}$-body hydrotrope aggregation around the solute from their aggregation tendency in the bulk solution. This simple picture clarifies further what has been revealed by our previous work, namely, the increase of the hydrotrope-hydrotrope KB integral induced by the solute molecule.[16,17,32]

The small molecule hydrotrope systems analysed in this paper involved the enhancement of the association of 2 to 5 hydrotrope molecules in the presence of the solutes (Table 1). 3-5 body hydrotrope association is most affected around BA, whereas BB induces the changes of 2-3 body hydrotrope associations. Since the coefficients R_{u,n_2} and R_{n_2} are defined at the $N_2^b \to 0$ limit, this means that the sigmoidal shape of the solubility curve can be understood entirely by how the interaction of 2 to 5 hydrotropes in water is affected when the solute is introduced. This leads to the significant facilitation of simulation-based approaches to hydrotropy. In this context,

understanding aggregation through the combination of statistical thermodynamics with computer simulation is indispensable.[50,51]

We have successfully shown that the sigmoidal shape of the solubility curve can be reproduced with only a few parameters. In contrast to the present theory, which is based firmly upon the first principles of statistical thermodynamics, none of the previous models were able to fit the overall shape of the solubility curve with few parameters, and at the same time to provide a molecular-based insight into the mechanism of solubilisation. Indeed, the association model[12] can fit solubility at lower hydrotrope concentrations, but cannot reproduce the plateau. Alternatively, a polynomial fitting of solubility, incorporating up to the sixth order of hydrotrope concentration was able to fit the data.[52] This approach may be founded firmly upon the MM theory of solutions, yet, as we have previously shown, the higher the order, the increasingly more cumbersome the theoretical expression becomes, hence is difficult to obtain physical insights.[32] In contrast to these previous models, our theory can now describe the entire region of the solubility curve with clear physical insights.

In contrast to our previous statistical thermodynamic approaches,[14-17,32] we employed here the combination of the partially-open ensemble and the local subsystem whose volume remains constant throughout. This is partly reminiscent of the KB theory, which is based upon the combination of the grand canonical ensemble and the local subsystem with constant volume.[16,25] The advantage of our present approach is in its simplicity, because there is no need for a link between an NPT and a grand canonical ensemble through a cumbersome process of changing thermodynamic constraints.[16,32] The disadvantage, however, is the lack of a direct link to the KB

integrals, which have a clear definition in terms of the radial distribution functions.[32] Our R_{u,n_2} and R_{n_2} nevertheless have a clear physical meaning of the fugacity of n_2 hydrotrope molecules around the solute and in the bulk, hence can be accessible through simulation.

Note that, because eqn (17) does not depend on how hydrotrope molecules interact in the bulk phase, it can be used for monomeric and micellar hydrotropes. However, when attempting to identify the dominant $R_{u,\overline{n_2}}$ and $R_{\overline{n_2}}$ terms in view of a further simplification of eqn (17), a careful discussion may be necessary in order to grasp the nature of hydrotrope-hydrotrope interaction.

8. Conclusion

The solubility increase in the presence of hydrotropes exhibits features of cooperative phenomena. The sudden onset of solubilisation at minimum hydrotrope concentration (MHC), as well as solubility saturation at high hydrotrope concentrations makes up the sigmoidal hydrotrope concentration dependence, whose molecular origin has not been captured sufficiently by our previous statistical thermodynamic studies.[21-25,32] A new statistical thermodynamic theory was formulated in this paper to clarify the origin of hydrotropic cooperativity and to establish a close analogy between hydrotropy and cooperative binding phenomena.

In our new theory, hydrotropic cooperativity, which includes MHC and solubility saturation, arise from the same origin: the enhancement of *m*-body hydrotrope aggregation in the presence of the solute. This is the only information needed to reproduce the overall sigmoidal shape of the

solubilisation curve, which can be extracted from experimental solubility data in a method analogous to the linearised plot in cooperative binding theory.

The present theory, because of its generality, can be applied straightforwardly to both monomeric and micellar hydrotropes. The future application of our theory to the micellar system will be useful in clarifying the role of micelle formation in solubilisation. In addition, our theoretical framework may be useful in clarifying the molecular basis of biomolecular solvation in the presence of salts,[53-55] as well as protein denaturation cooperativity in the presence of cosolvents.[56]

Acknowledgements

We thank Steven Abbott for stimulating discussions, and his extensive comments on our manuscript. We also thank Werner Kunz for interesting discussion and warm encouragement. This work is supported by the Grants-in-Aid for Scientific Research (Nos. 15K13550 and 26240045) from the Japan Society for the Promotion of Science, by the Elements Strategy Initiative for Catalysts and Batteries and the Post-K Supercomputing Project from the Ministry of Education, Culture, Sports, Science, and Technology, and by the HPCI System Research Project (Project IDs: hp160013, hp160019, and hp160214).

Appendix

In order to derive eqn (20), which establishes the analogy between hydrotropy and cooperative binding, we have introduced in Section 3 the local subsystem open only to hydrotropes, whose volume is v. Here we show that v can be determined uniquely from the four conditions in Section 4 and that the physical meaning of the boundary of such a subsystem will be made clear.

To do so, let us first appreciate that $\frac{R_{u,\overline{n_2}}}{R_{\overline{n_2}}}$, which can be linked to experimental data via eqn (27), has a clear physical interpretation. Using eqn (13) and (15), we have

$$\frac{R_{u,\overline{n_2}}}{R_{\overline{n_2}}} = \frac{\int d\boldsymbol{X}^{\overline{n_2}} \mathrm{R_u}\left(T,P,N_1,0;\boldsymbol{X}^{\overline{n_2}}\right)}{\int d\boldsymbol{X}^{\overline{n_2}} \mathrm{R}\left(T,P,N_1,0;\boldsymbol{X}^{\overline{n_2}}\right)} \tag{A1}$$

Note that Eq. (A1) is defined at the $N_2^b \to 0$ limit. Eq. (A1) can be rewritten in the following form:

$$\frac{R_{u,\overline{n_2}}}{R_{\overline{n_2}}} = \int d\boldsymbol{X}^{\overline{n_2}} P_{\boldsymbol{X}^{\overline{n_2}}} \frac{\mathrm{R_u}\left(T,P,N_1,0;\boldsymbol{X}^{\overline{n_2}}\right)}{\mathrm{R}\left(T,P,N_1,0;\boldsymbol{X}^{\overline{n_2}}\right)} \tag{A2}$$

where $\frac{\mathrm{R_u}\left(T,P,N_1,0;\boldsymbol{X}^{\overline{n_2}}\right)}{\mathrm{R}\left(T,P,N_1,0;\boldsymbol{X}^{\overline{n_2}}\right)}$ signifies the fugacity of a solute fixed at the origin, when $\overline{n_2}$ hydrotrope molecules in the system take the configuration $\boldsymbol{X}^{\overline{n_2}}$. The probability of observing such a configuration $\boldsymbol{X}^{\overline{n_2}}$ is defined by

$$P_{\boldsymbol{X}^{\overline{n_2}}} = \frac{\mathrm{R}\left(T,P,N_1,0;\boldsymbol{X}^{\overline{n_2}}\right)}{\int d\boldsymbol{X}^{\overline{n_2}} \mathrm{R}\left(T,P,N_1,0;\boldsymbol{X}^{\overline{n_2}}\right)} \tag{A3}$$

This shows that, given the volume of the subsystem v, how hydrotrope molecules are distributed crucially determines $\frac{R_{u,\overline{n_2}}}{R_{\overline{n_2}}}$. We emphasise here that while $\frac{\mathrm{R_u}\left(T,P,N_1,0;\boldsymbol{X}^{\overline{n_2}}\right)}{\mathrm{R}\left(T,P,N_1,0;\boldsymbol{X}^{\overline{n_2}}\right)}$ is not dependent on the

size of the subsystem v, $P_{\boldsymbol{X}^{\overline{n_2}}}$ does depend on v, because of the increase in the number of allowed hydrotrope configurations as the increase of the subsystem size.

We have thus seen that the value of v is crucial to the thermodynamics of cooperative hydrotropy. This makes one wonder if v can be determined uniquely; otherwise there will be multiple combinations of v and $\overline{n_2}$ that satisfies eqn (20) onwards, which makes it impossible to interpret our results at a molecular level. Contrary to such pessimism, we will demonstrate below that it is the condition d that guarantees the uniqueness of v.

To do so, let us take the second local subsystem with volume $v'(> v)$, which contains the first subsystem (volume v) within, as well as n_2 hydrotrope molecules. There are two possible scenarios by which the condition b (which has been introduced in Section 4) is satisfied by this new setup (v', n_2) in addition to the first $(v, \overline{n_2})$:

1. All $n_2 = \overline{n_2}$ and all hydrotrope molecules are located within the local subsystem v.
2. $\overline{n_2}$ hydrotrope molecules are located within v, and $n_2 - \overline{n_2}$ are located in the exterior of v, i.e., $v' - v$

Obviously, the larger v' becomes the more likely scenario 2 becomes. The larger the v' the more different ways by which the scenario 2 is satisfied. Hence the range of n_2 that satisfies the condition b, namely Δn_2, increases.

This is why we have imposed the condition d (which has been introduced in Section 4), by which we choose v in order that Δn_2 is minimum, hence can be approximated safely as 0. Thus

v can be determined by the minimum volume (with minimum Δn_2) that encompasses the range of hydrotrope-hydrotrope interaction.

References

1. C. Neuberg, *Biochem. Z.,* 1916, **76,** 107-176.

2. S. E. Friberg, *Curr. Opin. Coll. Int. Sci.,* 1997, **2,** 490-494.

3. P. Bauduin, A. Renoncourt, A. Kopf, D. Touraud and W. Kunz, *Langmuir,* 2005, **21,** 6769-6775.

4. T. K. Hodgdon and E. W. Kaler, *Curr. Opin. Coll. Int. Sci.,* 2007, **12,** 121-128.

5. C. V. Subbarao, I. P. K. Chackravarthy, A.V.S.L. Sai Bharadwaj and K. M. M. Prasad, *Chem. Eng. Technol.,* 2012, **35,** 225-237.

6. W. Kunz, K. Holmberg and T. Zemb, *Curr. Opin. Coll. Int. Sci.,* 2016, **22,** 99-107.

7. D. Balasubramanian, V. Srinivas, V. G. Gaikar and M. M. Sharma, *J. Phys. Chem.,* 1989, **93,** 3865-3870.

8. M. G. Neumann, C. C. Schmitt, K. R. Prieto and B. E. Goi, *J. Colloid Interface Sci.,* 2007, **315,** 810–813.

9. J. Y. Kim, S. Kim, M. Papp, K. Park and R. *J. Pharm. Sci.,* 2010, ***99,*** 3953-3965.

10. W. N. Charman, C. S. C. Lai, D. J. Craik, B. C. Finnin and B. L. Reed, *Aust. J. Chem.,* 1993, **46,** 377–385.

11. R. E. Coffman and D. O. Kildsig, *Pharm. Res.,* 1996, **13,** 1460-1463.

12. V. G. Gaikar and P. V. Phatak, Separation Sci. Technol.,1999, **34,** 439-459.

13. Y. Cui, C. Xing, and Y. Ran, *J. Pharm. Sci.,* 2010, **99,** 3048-59.

14. J. J. Booth, S. Abbott, and S. Shimizu, *J. Phys. Chem. B.,* 2012, **116,** 14915-14921.

15. S. Shimizu, J. J. Booth, and S. Abbott, *Phys. Chem. Chem. Phys.,* 2013, **15,** 20625-20632.

16. S. Shimizu and N. Matubayasi, *J. Phys. Chem. B.,* 2014, **118,** 10515–10524.

17. J. J. Booth, M. Omar, S. Abbott and S. Shimizu, *Phys. Chem. Chem. Phys.*, 2015, **17**, 8028–8037.

18. J. G. Kirkwood and B. F. Buff, *J. Chem. Phys.,* 1951, **19,** 774 -777.

19. D. G. Hall, *Trans. Faraday Soc.,* 1971, **67,** 2516-2524.

20. A. Ben-Naim, *J. Chem. Phys.*, 1977, **67**, 4884–4890.

21. S. Shimizu, *Proc. Natl. Acad. Sci., U.S.A.* 2004, **101,** 1195-1199.

22. S. Shimizu and C. L. Boon, *J. Chem. Phys.,* 2004, **121,** 9147-55.

23. S. Shimizu, W. M. McLaren, and N. Matubayasi, *J. Chem. Phys.,* 2006, **124,** 234905/1-234905/5.

24. N. Matubayasi, K. K. Liang, and M. Nakahara, *J. Chem. Phys.,* 2006, **124,** 154908/1-154908/13.

25. S. Shimizu and N. Matubayasi, *J. Phys. Chem. B,* 2014, **118,** 3922-3930.

26. Y. Jiao and P. E. Smith, *J. Chem. Phys.,* **2011,** *135,* 014502/1-014502/13.

27. S. Karunaweera, M. B. Gee, S. Weerasinghe and P. E. Smith, *J. Chem. Theory Comput.,* 2012, **8,** 3493–3503.

28. I. L. Shulgin and E. Ruckenstein, *J. Phys. Chem. B,* 2006, **110,** 12707–12713.

29. E. Ruckenstein and I. L. Shulgin, *Adv. Coll. Interf. Sci.,* 2006, **123-126,** 97–103.

30. A. Ben-Naim, *Molecular Theory of Solutions*. Oxford University Press, New York, 2006.

31. M. L. Klossek, D. Touraud and W. Kunz, *Phys. Chem. Chem. Phys.,* 2013, **15,** 10971-10977.

32. T. W. J. Nicol, N. Matubayasi and S. Shimizu, *Phys. Chem. Chem. Phys.,* 2016, 18, 15205-15217.

33. W. G. McMillan and J. E. Mayer, *J. Chem. Phys.,* 1945, **13,** 276-.

34. T. L. Hill, *Statistical Mechanics*. McGraw-Hill, New York, 1956.

35. T. L. Hill, *Cooperativity Theory in Biochemistry.* Springer, New York, 1985.

36. I. M. Klotz, Ligand-Receptor Energetics: A Guide for the Perplexed. Wiley, New York, 1997.

37. J. Wyman and S. J. Gill, *Binding and Linkage.* Mill Valley, CA: University Science Books, 1990.

38. W. H. Stockmeyer, *J. Chem. Phys.,* 1950, **18,** 58-61.

39. J. A. Schellman, *Biopolymers,* 1987, **26,** 549-559.

40. T. L. Hill, *J. Chem. Phys.,* 1957, **26,** 955-956.

41. T. L. Hill, *J. Am. Chem. Soc.,* 1957, **79,** 4885-4890.

42. T. L. Hill, *J. Chem. Phys.,* 1959, **30,** 93-97.

43. T. L. Hill, *J. Chem. Phys.,* 1961, **34,** 1974-1982.

44. E. F. Casassa, *Adv. Protein Chem.,* 1964, **19,** 287-395.

45. C. Tanford, *J. Mol. Biol.,* 1969, **39,** 539-544.

46. A. V. Hill, *J. Physiol. (Proc. Physiological Soc.),* 1910, **40,** iv-vii.

47. N. N. Gandhi, M. D. Kumar and N. Sathyamurthy, *J. Chem. Eng. Data,* 1998, **43,** 695–699.

48. N. Meyyappan and N.N. Gandhi, *J. Chem. Eng. Data,* 2005, **50,** 796–800.

49. R. H. Stokes, *Aust. J. Chem.,* **1967,** *20,* 2087 – 2100.

50. G. Pellicane, R. L. C. Vink, B. Russo and P. V. Giaquinta, *Phys. Rev. E,* 2013, **88,** 042131/1-042131/8.

51. G. Pellicane and O. D. Pandaram*, J. Chem. Phys.,* 2014, **141,** 044508/1-044508/13.

52. S. T. Kumar, D. G. Prakash and N. Nagendra Gandhi, *Korean J. Chem. Eng.,* 2009, **26,** 1328-1333.

53. M. C. Abramo, C. Caccamo, D. Costa, G. Pellicane and R. Ruberto, *J. Phys. Chem. B,* 2010, **114,** 9109-9118.

54. M. C. Abramo, C. Caccamo, M. Calvo, V. Conti Nibali, D. Costa, R. Giordano, G. Pellicane, R. Ruberto and U. Wanderlingh, *Philos. Mag.,* 2011, **91,** 2066-2076.

55. G. Pellicane and M. Cavero, *J. Chem. Phys.,* 2013, **141,** 115103/1-115103/7.

56. I. M. Klotz, *Proc. Natl. Acad. Sci. U. S. A.,* 1996, **93,** 14411-14415.

Figure 1

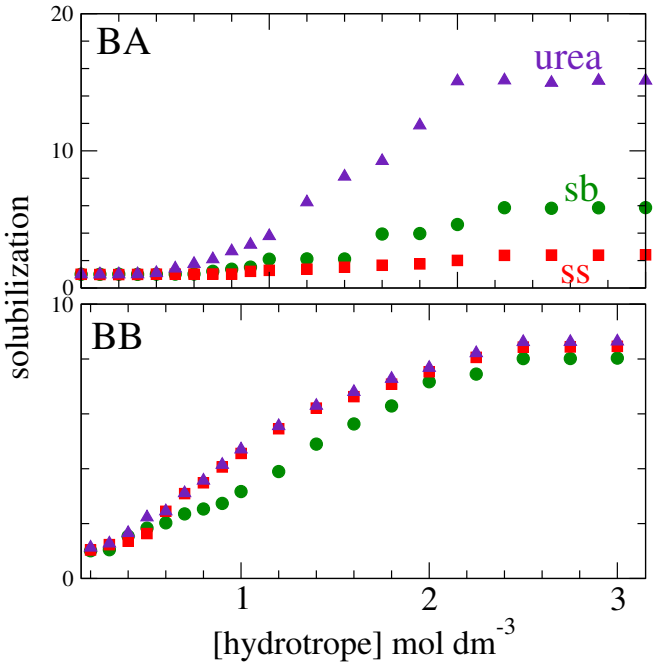

Figure 1. Solubility of butyl acetate (BA) and benzyl benzoate (BB) in the presence of three hydrotropes, sodium benzoate (sb, green circle), sodium salicylate (ss, red square) and urea (blue triangle). Note that the "solubilisation" refers to the molar solubility relative to the value at zero hydrotrope concentration, i.e. $e^{-\beta\Delta\mu_u^*}$ in the main text. Experimental data taken from Refs 47 and 48.

Figure 2

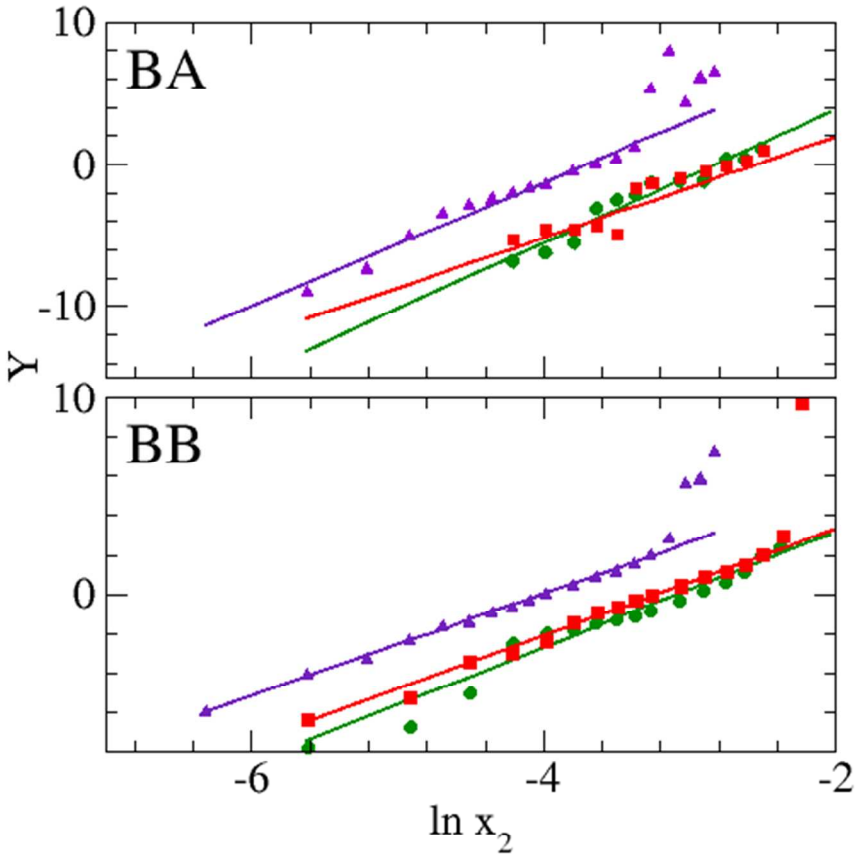

Figure 2. Fitting experimental solubility data to eqn (28), in order to determine the necessary parameters tabulated in Table 1. Note that the vertical axis refers to $Y \equiv \ln\left[\frac{1-e^{-\beta\Delta\mu_u^*}}{e^{-\beta\Delta\mu_u^*}-e^{-\beta\Delta\mu_u^{*sat}}}\right]$, i.e., l.h.s. of eqn (28). The hydrotropes used to solubilize BA and BB are sodium benzoate (sb, green circle), sodium salicylate (ss, red square) and urea (blue triangle).

Figure 3

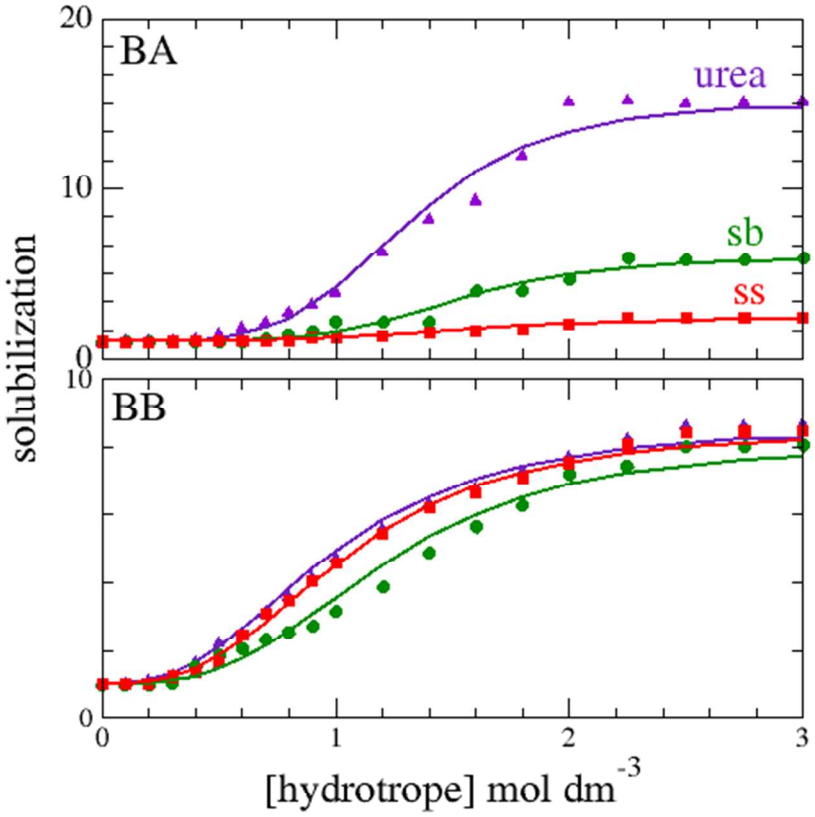

Figure 3. Comparison of our theory (eqn (24)) against experimental data, with the parameters summarised in Table 1. The hydrotropes used to solubilize BA and BB are sodium benzoate (sb, green circle), sodium salicylate (ss, red square) and urea (blue triangle).

Table 1. Fitting parameters for the linearized plot (Figure 2). The definition of each parameter can be found in eqn (28).

Solute	Hydrotrope	$\overline{n_2}$	$\ln R''_{\overline{n_2}}$	$e^{-\beta\Delta\mu_u^{*,sat}} = \frac{R''_{u,\overline{n_2}}}{R_{\overline{n_2}}}$
Butyl acetate	Sodium benzoate	4.68	13.1	5.86
Butyl acetate	Sodium salicylate	3.55	9.82	2.41
Butyl acetate	Urea	4.37	16.1	15.2
Benzyl benzoate	Sodium benzoate	2.93	8.98	8.03
Benzyl benzoate	Sodium salicylate	2.72	8.73	8.46
Benzyl benzoate	Urea	2.61	10.4	8.65

Table of contents graphics