Quantification of Monomeric Formaldehyde in Aqueous Solutions by NMR Spectroscopy and Determination of the Chemical Equilibrium Constant of Methylene Glycol Formation

Katharina Koehler,[†] Maximilian Dyga,[‡] Andreas Keller,[‡] Fabian Jirasek,[†] Kerstin Münnemann,[†] and Hans Hasse[*,†]

[†]*Laboratory of Engineering Thermodynamics (LTD), RPTU Kaiserslautern, 67663 Kaiserslautern, Germany*

[‡]*Process Development, BASF SE, Carl-Bosch-Str. 38, 67056 Ludwigshafen am Rhein, Germany*

E-mail: hans.hasse@rptu.de

Abstract

Formaldehyde is an important intermediate in the chemical industry. It is highly reactive and usually handled in aqueous solutions, where it is mainly bound in the oligomeric reaction products methylene glycol and poly(oxymethylene) glycols. Quantitative information on the very low concentration of monomeric formaldehyde in these solutions is difficult to obtain, but crucial for many applications in which this highly reactive species plays a key role. *Rivlin et al.*[1] have shown that NMR spectroscopy can be applied successfully for determining the concentration of monomeric formaldehyde, but they have used either ^{13}C enriched formaldehyde (for ^{13}C NMR) or deuterated water (for ^{1}H NMR) and have studied only dilute formaldehyde solutions at temperatures below 333 K. We show here that the amount of monomeric formaldehyde in aqueous solutions can be determined without the use of deuterated water by ^{1}H NMR spectroscopy and apply the method for studies at temperatures between 293 and 393 K at overall formaldehyde concentrations up to 0.5 g g^{-1}. New data were used to determine the chemical equilibrium constant of methylene glycol formation, for which only strongly scattering data were previously available in the literature.

Introduction

Formaldehyde (FA) is one of the most important intermediates in the chemical industry.[2–4] Due to its high reactivity, formaldehyde is hard to obtain as a pure substance and usually handled in aqueous solutions. In these solutions, it is predominantly bound in the oligomeric reaction products methylene glycol ($HOCH_2OH$, MG_1) and poly(oxymethylene) glycols ($HO(CH_2O)_nH$, $n \geq 2$, MG_n):

$$H_2O + CH_2O \rightleftharpoons HOCH_2OH \tag{I}$$

$$HO(CH_2O)_{n-1}H + CH_2O \rightleftharpoons HO(CH_2O)_nH, \qquad n \geq 2 \tag{II}$$

The oligomerization reactions of formaldehyde with water are reversible.[2] They can be catalyzed by both acids and bases but also take place without the addition of catalysts. Reaction (II) has been studied thoroughly by many authors using nuclear magnetic resonance (NMR) spectroscopy[1,5–16] and reliable information on the distribution of the formaldehyde to the oligomers of different chain lengths is available.[1,9,12–14,17]

In contrast, due to the very low amount of monomeric formaldehyde, it is more difficult to get quantitative information on Reaction (I). Many different methods have been applied in the literature for studying its chemical equilibrium constant: polarography by *Landqvist et al.*[18]

and *Valenta*,[19] UV spectroscopy by *Bieber & Trümpler*,[20] *Iliceto*,[21] *Gruen & McTigue*,[22] *Silling & Akselrod*,[23] *Schecker & Schulz*[24] and *Zavitsas et al.*,[25] thermochemical calculations by *Bryant & Thompson*,[26] absorption and desorption kinetic studies by *Bell & Evans*[27] and *Winkelman et al.*,[28] and radiolysis experiments by *Sutton & Downes.*[29] *Rivlin et al.*[1] applied ^{13}C NMR spectroscopy (using ^{13}C enriched formaldehyde) and ^{1}H NMR spectroscopy (using mostly water) to determine the chemical equilibrium constant of Reaction (I). The chemical equilibrium constant of the Reaction (I) K_{x,MG_1} reported in the literature is based on mole fractions x_i:

$$K_{x,\mathrm{MG}_1} = \frac{x_{\mathrm{MG}_1}}{x_{\mathrm{FA}} \cdot x_{\mathrm{W}}} \tag{1}$$

with the mole fraction of methylene glycol x_{MG_1}, formaldehyde x_{FA} and water x_{W}. The information that is available in the literature on K_{x,MG_1}, is summarized in Figure 1.

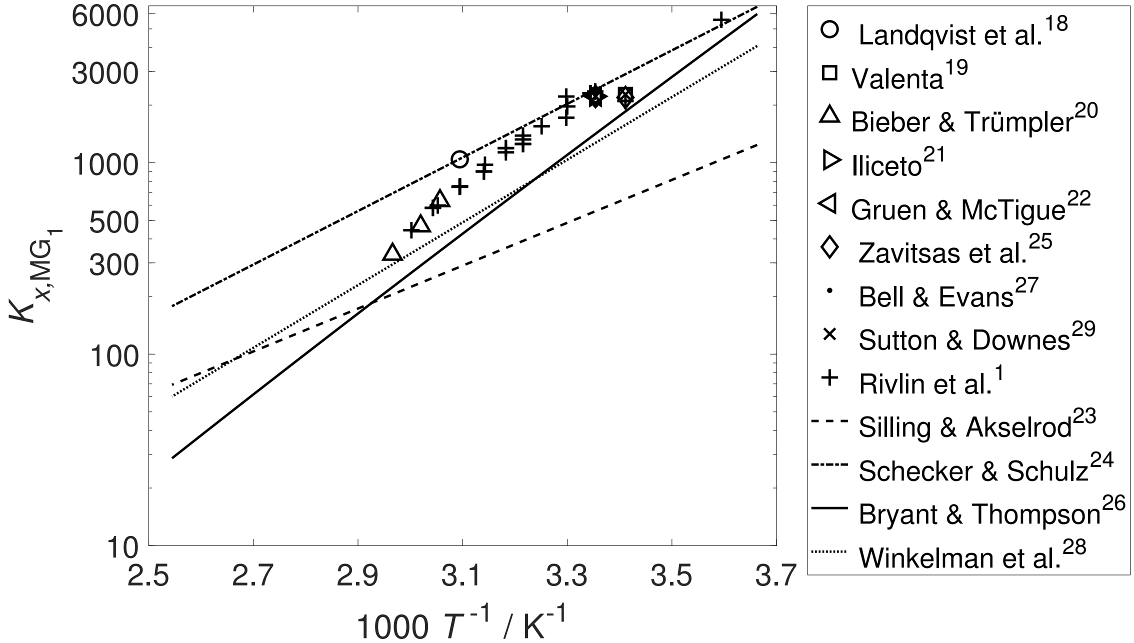


Figure 1: Mole fraction-based chemical equilibrium constant of the formation of methylene glycol K_{x,MG_1} as a function of the inverse temperature. Overview of literature data. Experimental data are shown as symbols.[1,18–22,25,27,29] Some authors only report correlations that were obtained from their experimental data and are shown here as lines.[23,24,26,28]

Due to an incomplete specification of the experimental conditions, the conversion of the data reported in the literature to the form shown in Eq. 1 is not always unambiguous, but the error introduced by the conversion is small compared to the experimental error in all cases. In particular, information on the overall formaldehyde concentration[a] in the studied solutions is often missing, but we assume that most of the studies were carried out in dilute aqueous formaldehyde solutions. Furthermore, some authors have only reported correlations and no experimental data. It can be seen from Figure 1 that experimental data are only available for temperatures between 279 and 333 K. The differences between the data on K_{x,MG_1} obtained by different authors with different methods are important: they sometimes exceed a factor of 5. At ambient temperature, most authors report a value of K_{x,MG_1} of approximately 2200. Accordingly, in dilute aqueous formaldehyde solutions where oligomerization reactions (cf. Eq. (II)) do not occur, only about one in 2200 formaldehyde molecules exists as a monomer, while the remainder are bound as methylene glycol.

The aim of the present study was twofold: first, we wanted to prove that ^{1}H NMR spectroscopy can be used for in situ studies of monomeric formaldehyde in technical aqueous formaldehyde solutions. This has long been considered as unfeasible due to the lack of sensitivity of NMR spectroscopy. Improvements in the NMR technology, in particular cryogenic probe heads with high signal-to-noise ratios, have encouraged us to challenge this prejudice. The information obtained from such measurements is important for many applications, in which formaldehyde (supplied in aqueous solutions) reacts with other educts. It is generally assumed that only the monomeric formaldehyde reacts with the other educts, while the formaldehyde bound in the oligomerization products can only participate after being released from the oligomers by the reverse reactions (II) and (I). Second, we wanted to supply data on the chemical equilibrium constant of methylene glycol formation over a wide temperature range, including high temperatures that are technologically important but have not been studied before. We also wanted to extend the concentration range, in which such studies have been carried out, to higher overall formaldehyde concentrations, which is highly

---

[a] We distinguish here between the overall formaldehyde concentration $\tilde{x}_{\mathrm{FA}}$, which is that of the formaldehyde present, i.e., in monomeric form or bound in the oligomers with water, in any form in and the true concentrations of the formaldehyde-containing species. To make the distinction, we always use the adjective "overall" when applicable, but we omit the adjective "true" in most cases for brevity.

relevant for most applications. The question here was to see if the overall formaldehyde concentration affects the mole fraction-based equilibrium constant K_{x,MG_1} and if the results can be described better by an activity-based equilibrium constant K_{a,MG_1}. Activity coefficient models that can be used for this purpose can be taken from the extensive work on physico-chemical modeling of vapor-liquid equilibria of aqueous formaldehyde solutions.[17,30]

Experiments

Chemicals and Sample Preparation

Paraformaldehyde (≥ 0.95 g g^{-1}) was purchased from Sigma-Aldrich and used without further purification. Water was deionized and purified with a Milli-Q Reference A+ purification system by Merck Millipore. Stock solutions of (formaldehyde + water) were prepared by dissolving paraformaldehyde in water in a thermostatted and pressurized stirred glass vessel at about 403 K and 150 kPa. The clear samples were taken by syringes equipped with a micro-filter to exclude any remaining solid.

Quantitative Analysis

Titration

The overall formaldehyde mass fraction was determined with the Na_2SO_3 titration method[2] using an automated titration station (Metrohm Omnis Sample Robot S). Each sample was analyzed at least twice. The relative deviations between the results were always below 2%. We take this number also as an estimate for the uncertainty of the formaldehyde mass fraction. The overall mass fraction of water was determined from the result for formaldehyde using the summation equation.

^{1}H NMR Spectroscopy

^{1}H NMR spectroscopy was used to determine the true concentrations in aqueous formaldehyde solutions, with a focus on the concentration of monomeric formaldehyde and that of methylene glycol. An NMR spectrometer (Bruker Ascend 400 magnet, Avance III HD 400 console) with a magnetic field strength of 9.4 T was used, which was equipped with a double resonance broadband probe with cryogenically cooled electronics (Bruker Prodigy CryoProbe). The temperature of the spectrometer was controlled by a thermostated nitrogen gas flow, the temperature was calibrated using a platinum resistance thermometer (Pt-100) with an uncertainty of ± 0.05 K, which was calibrated in our laboratory using a certified standard. The uncertainty of the temperature measurement is ± 0.5 K. A ^{13}C decoupled ^{1}H sequence was used to avoid the occurrence of ^{13}C satellites. The acquisition parameters were as follows: 30° flip angle, 6.4 s acquisition time, 30 s relaxation delay, and 128 scans. 12 individual measurements of each sample were acquired from which the mean value and the standard deviation were calculated. The post-processing of the NMR spectra was done with MestReNova (Mestrelab v14.3.0). Manual phase correction and baseline correction (Whittaker smoother) were applied. Also the peak integration was done manually.

Prior to the NMR analysis, the samples were thermostated (4 days at 293.15 K, 2 days at 333.15 K, 1 day at 348.20 K for the measurements at the according temperature and 2 hours at 373.15 K and 393.15 K for the measurements at the according temperatures) to ensure the establishment of the chemical equilibrium before starting the measurements.

An example for a ^{1}H NMR spectrum of an aqueous formaldehyde solution is shown in Figure 2 together with the peak assignment, similar to previous work.[31] All signals stem from CH_2 groups, except for a broad signal that originates from the hydroxyl groups in water and in the formaldehyde reaction products. The latter shifts strongly with temperature and can overlap with the signals from the CH_2 groups hindering quantitative analysis at temperatures ranging from 313.15 to 333.15 K depending on the overall formaldehyde concentration in the system. Experimental quantification of water from the hydroxyl group peak is very imprecise due to the many other species collected in this peak. Therefore, only signals from CH_2 groups have been quantitatively evaluated here. Some of these signals stem from different oligomeric species that cannot be resolved separately. These signals were considered in the evaluation as a sum, which is indicated by plus signs in Figure 2 (e.g., E_{3+} denotes the sum of the signals E_3, E_4, E_5, ...).

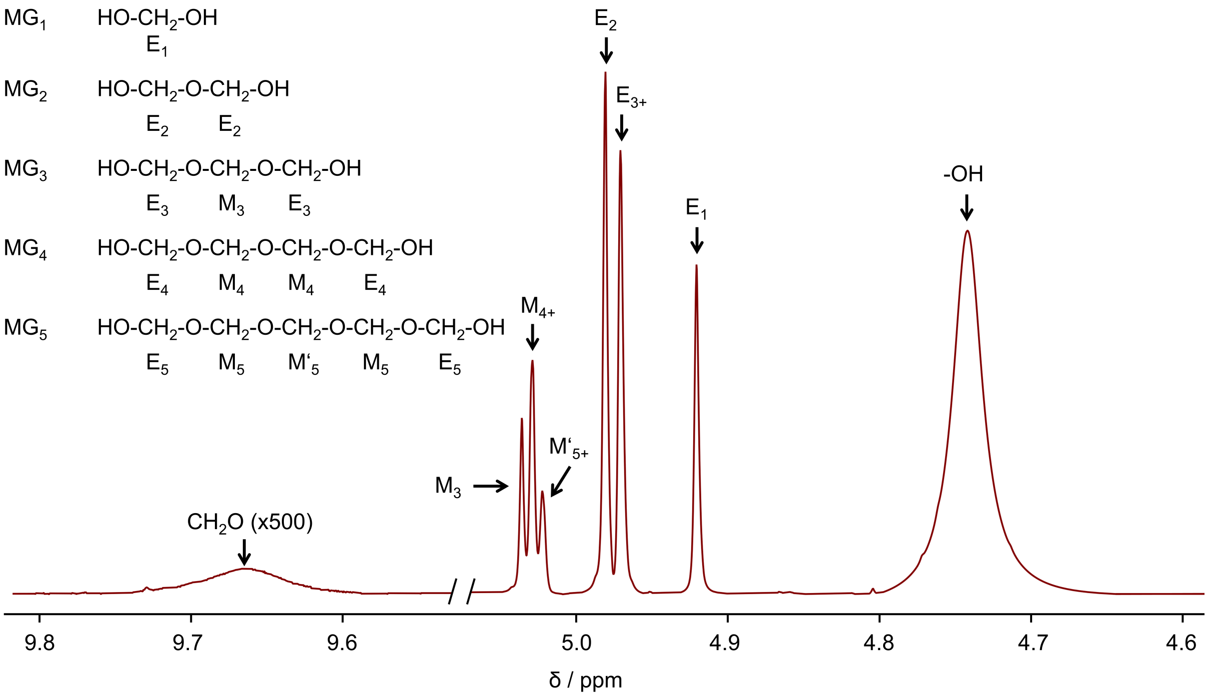


Figure 2: ^{1}H NMR spectrum of an aqueous formaldehyde solution ($\tilde{x}_{FA}= 0.3436$ mol mol^{-1}, $T = 353.15$ K) with peak assignment. The small broad peak on the left stems from monomeric formaldehyde and is magnified by a factor of 500.

The areas of the peaks of monomeric formaldehyde A_{FA} and of methylene glycol A_{MG_1} were treated individually, whereas the areas of all other peaks that stem from CH_2O groups in reaction products other than MG_1 were lumped together with the areas A_{FA} and A_{MG_1} to a sum signal $\tilde{A}_{FA}$. (The signals from the hydroxyl groups that are difficult to quantify were not used). There are several reasons why we did not try to retrieve the quantitative information on the MG_n with $n > 1$ and water from the spectra. First, this has been done already in more comprehensive previous studies, and, second, ^{1}H NMR spectroscopy is not the method of choice for these studies, ^{13}C NMR spectroscopy is better suited[14,32] and good models are available for describing the species distribution of the oligomers that were developed using extensive ^{13}C NMR spectroscopic data, namely the model of *Hahnenstein et al.*[13] We have preferred to rely on this model rather that trying to retrieve the information from the limited corresponding data from the present work. The focus of our work is on monomeric formaldehyde and the methylene glycol formation.

Determination of the Concentration of Monomeric Formaldehyde and the Chemical Equilibrium Constant of the Methylene Glycol Formation

The quantification is based on the assumption that the area under the different peaks is proportional to the mole number of the corresponding CH_2 group, and that the proportionality factor is the same for all CH_2 groups, including that in monomeric formaldehyde.[33,34]
The peak areas A_{FA} and $\tilde{A}_{\mathrm{FA}}$, therefore, yield direct information on how much of the overall formaldehyde is present as monomer:

$$\frac{A_{\mathrm{FA}}}{\tilde{A}_{\mathrm{FA}}} = \frac{n_{\mathrm{FA}}}{\tilde{n}_{\mathrm{FA}}} \tag{2}$$

Furthermore, from A_{FA} and $A_{\mathrm{MG_1}}$, the ratio of the mole numbers of monomeric formaldehyde n_{FA} and methylene glycol $n_{\mathrm{MG_1}}$ can be calculated, which is closely related to the chemical equilibrium constant $K_{x,\mathrm{MG_1}}$ of the methylene glycol formation (cf. Eq. 1).

$$\frac{A_{\mathrm{MG_1}}}{A_{\mathrm{FA}}} = \frac{n_{\mathrm{MG_1}}}{n_{\mathrm{FA}}} \tag{3}$$

However, to obtain $K_{x,\mathrm{MG_1}}$ (cf. Eq. 1) from the information on ($n_{\mathrm{MG_1}}$ / n_{FA}), additional information is needed. There are different ways to incorporate this information and we have decided to use an approach that relies on the analytical value of the overall formaldehyde concentration $\tilde{x}_{\mathrm{FA}}$, which is known with high accuracy. The true mole fraction of formaldehyde was calculated here from (A_{FA} / $\tilde{A}_{\mathrm{FA}}$) as follows:

$$x_{\mathrm{FA}} = \frac{n_{\mathrm{FA}}}{n_{\mathrm{total}}} = \frac{n_{\mathrm{FA}}}{\tilde{n}_{\mathrm{FA}}} \cdot \frac{\tilde{n}_{\mathrm{FA}}}{\tilde{n}_{\mathrm{total}}} \cdot \frac{\tilde{n}_{\mathrm{total}}}{n_{\mathrm{total}}} = \frac{A_{\mathrm{FA}}}{\tilde{A}_{\mathrm{FA}}} \cdot \tilde{x}_{\mathrm{FA}} \cdot \frac{\tilde{n}_{\mathrm{total}}}{n_{\mathrm{total}}} \tag{4}$$

The value of ($\tilde{n}_{\mathrm{total}}$ / n_{total}) was calculated from $\tilde{x}_{\mathrm{FA}}$ and T using the model of *Hahnenstein et al.*[13] The main uncertainty of the numbers for x_{FA} stems from the determination of (A_{FA} / $\tilde{A}_{\mathrm{FA}}$), and here, in particular, from the uncertainty of A_{FA}. The uncertainties introduced by $\tilde{x}_{\mathrm{FA}}$ and ($\tilde{n}_{\mathrm{total}}$ / n_{total}) are minor compared to this. The uncertainty of the number of x_{FA} is highest for high temperatures for which the peak of monomeric formaldehyde is very broad, see Figure 3. The relative uncertainty of the experimental values of x_{FA} is estimated to be about 5%. All three peaks in Figure 3 have the same y-axis scale. Differences in the breadth of the baseline result from the higher noise at high temperatures. The peak broadening at high temperatures is a result of an increasingly faster chemical exchange of the monomeric formaldehyde with the other formaldehyde-containing species.

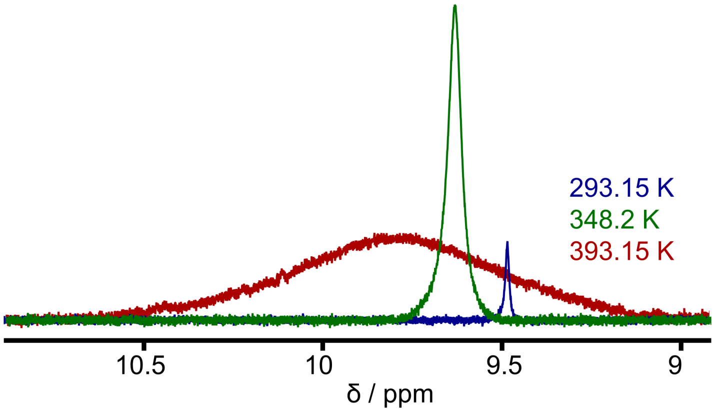


Figure 3: Peak of monomeric formaldehyde in ^{1}H NMR spectra of an aqueous formaldehyde solution ($\tilde{x}_{\mathrm{FA}} = 0.2046$ mol mol^{-1}) at different temperatures. All three peaks have the same y-axis scale.

The value of the mole fraction-based equilibrium constant K_{x,MG_1} of the methylene glycol formation, cf. Eq. 1, was calculated from (A_{MG_1} / A_{FA}) based on Eq. 3.

$$K_{x,\mathrm{MG}_1} \;=\; \frac{n_{\mathrm{MG}_1}}{n_{\mathrm{FA}} \cdot x_{\mathrm{W}}} \;=\; \frac{A_{\mathrm{MG}_1}}{A_{\mathrm{FA}} \cdot x_{\mathrm{W}}} \tag{5}$$

As in the evaluation of Eq. 4, the missing value of the true mole fraction of water x_{W} was calculated from the overall mole fraction of formaldehyde and the temperature using the model of *Hahnenstein et al.*[13] The uncertainty introduced by this is minor: the uncertainty of K_{x,MG_1} is almost entirely determined by the uncertainty of (A_{MG_1} / A_{FA}), and, again, the most important factor thereby is the uncertainty of A_{FA}.

The values reported here for x_{FA} and K_{x,MG_1} were obtained as the average from 12 individual measurements and the uncertainties we report here are the corresponding standard deviations. They are in fair agreement with results obtained from estimates of the uncertainties of the peak areas by error propagation.

Modeling of the Chemical Equilibrium Constant of Methylene Glycol Formation

Mole Fraction-Based Chemical Equilibrium Constant

The mole fraction-based values of the chemical equilibrium constant of methylene glycol formation, K_{x,MG_1}, were determined from the experimental data as described above using Eq. 5. The temperature dependence of the results for K_{x,MG_1} was correlated using the integrated form of the van't Hoff equation:

$$\ln(K_{x,\mathrm{MG}_1}) \;=\; A + \frac{B}{T/\mathrm{K}} \tag{6}$$

The parameters A and B were obtained from a fit to the experimental data using the following objective function:

$$F = \sum_{m=1}^{M} \left(\frac{\ln(K_{x,\mathrm{MG}_1,m}^{\mathrm{calc}}) - \ln(K_{x,\mathrm{MG}_1,m}^{\mathrm{exp}})}{\ln(K_{x,\mathrm{MG}_1,m}^{\mathrm{exp}})} \right)^2 \tag{7}$$

where M is the total number of experimental data points, $K_{x,\mathrm{MG}_1,m}^{\mathrm{calc}}$ is the mole fraction-based chemical equilibrium constant calculated with Eq. 6 and $K_{x,\mathrm{MG}_1,m}^{\mathrm{exp}}$ the experimental mole fraction-based chemical equilibrium constant of methylene glycol formation. The fit was carried out in MATLAB (version R2024a, MathWorks) using the solver fminsolve.

Activity-Based Chemical Equilibrium Constant

The activity-based chemical equilibrium constant of methylene glycol formation K_{a,MG_1} is:

$$K_{a,\mathrm{MG}_1} = \frac{a_{\mathrm{MG}_1}}{a_{\mathrm{FA}} \cdot a_{\mathrm{W}}} \tag{8}$$

where

$$a_i = x_i \cdot \gamma_i \tag{9}$$

is the activity of component i and γ_i is its activity coefficient. It follows from Eqns. 1, 8 and 9 that

$$K_{a,\mathrm{MG}_1} = K_{x,\mathrm{MG}_1} \cdot K_{\gamma,\mathrm{MG}_1} \tag{10}$$

where

$$K_{\gamma,\mathrm{MG}_1} = \frac{\gamma_{\mathrm{MG}_1}}{\gamma_{\mathrm{FA}} \cdot \gamma_{\mathrm{W}}} \tag{11}$$

The numbers for K_{a,MG_1} were calculated from the experimental results for K_{x,MG_1} using numbers for K_{γ,MG_1} calculated with the model of *Kuhnert et al.*[30] using the transformation of the parameters describing the chemical equilibrium in the vapor phase to the liquid phase by *Drunsel et al.*[17] The temperature dependence of the results for K_{a,MG_1} was correlated and fitted using the same approach as for K_{x,MG_1}, see Eqns. 6 and 7.

Results and Discussion

Overview

The numerical results of the experiments of the present work are summarized in Table 1. Additional information is given in Table S1.

Table 1: Overview of the experimental results from the present work: true concentration of monomeric formaldehyde x_{FA} and equilibrium constant of methylene glycol formation K_{x,MG_1} in aqueous formaldehyde solutions with different overall formaldehyde concentrations $\tilde{x}_{\mathrm{FA}}$ at different temperatures T. The experimental uncertainty for x_{FA} and K_{x,MG_1} is reported for each experiment; for $\tilde{x}_{\mathrm{FA}}$ the relative uncertainty is below 2%, for T the absolute uncertainty is below 0.5 K.

T K	$\tilde{x}_{\mathrm{FA}}$ g g^{-1}	$1000 \cdot x_{\mathrm{FA}}$ mol mol^{-1}	K_{x,MG_1}
293.15	0.286	0.042 ± 0.001	2429 ± 40
333.15	0.286	0.239 ± 0.007	442 ± 13
348.20	0.096	0.205 ± 0.007	248 ± 9
	0.290	0.507 ± 0.020	216 ± 7
	0.479	0.977 ± 0.022	161 ± 5
373.15	0.292	1.256 ± 0.065	95 ± 4
	0.469	2.340 ± 0.148	63 ± 4
393.15	0.290	1.853 ± 0.418	47 ± 3
	0.462	6.133 ± 0.785	31 ± 2

Monomeric Formaldehyde

Figure 4 (left) shows the results for the ratio of the true mole number of monomeric formaldehyde and the overall mole number of formaldehyde (i.e., the fraction of formaldehyde that is present in monomeric form) in aqueous formaldehyde solutions of different composition as a function of the temperature. On the right side of Figure 4, the results for the true mole fraction of formaldehyde in the same samples are shown. Figure 4 (left) shows that the fraction of formaldehyde in aqueous solutions present as monomer increases strongly with increasing temperature. This ratio depends also on the overall formaldehyde concentration and is expected to decrease with increasing overall formaldehyde concentrations, as

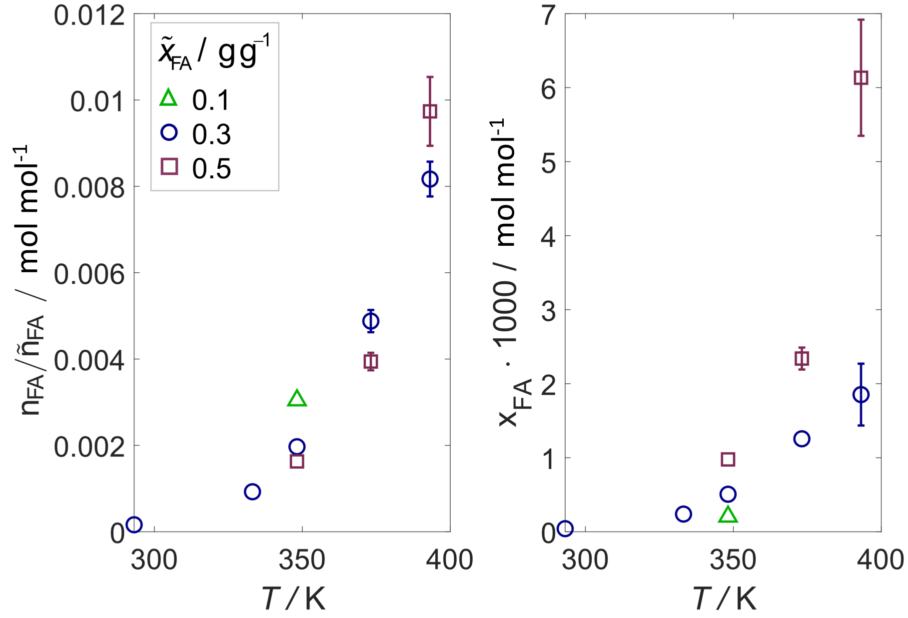


Figure 4: Ratio of the monomeric formaldehyde and the total formaldehyde (left) as well as mole fraction of monomeric formaldehyde (right) as a function of the temperature for different overall formaldehyde mole fractions in aqueous formaldehyde solutions. Error bars are only shown if they exceed the symbol size.

then formaldehyde is increasingly bound in oligomers. This is also confirmed by the results, with the exception of one value for the highest temperature, for which, however, the experimental uncertainty is very high. Figure 4 (right) shows that the true mole fraction of monomeric formaldehyde x_{FA} in aqueous formaldehyde solutions is always very low. It is only $x_{\mathrm{FA}} = 0.00004$ mol mol^{-1} at 293.15 K for an overall formaldehyde concentration of $\tilde{x}_{\mathrm{FA}} \approx 0.3$ g g^{-1} (which is not far below the practical solubility limit of formaldehyde in water at that temperature).[35] As expected, the true mole fraction of formaldehyde x_{FA} increases with increasing overall mole fraction of formaldehyde $\tilde{x}_{\mathrm{FA}}$ and it also increases significantly with increasing temperature. The highest value found in the present study was $x_{\mathrm{FA}} = 0.006$ mol mol^{-1} (for $\tilde{x}_{\mathrm{FA}} \approx 0.5$ g g^{-1} and $T = 393.15$ K).

Equilibrium Constant of Methylene Glycol Formation

Figure 5 shows the results for the mole fraction-based chemical equilibrium constant of methylene glycol formation, K_{x,MG_1}, in aqueous formaldehyde solutions as a function of the temperature. All data from the present work are shown together with the data from *Rivlin et al.*[1] that were also taken with NMR spectroscopy. The equilibrium constant K_{x,MG_1} decreases strongly with increasing temperature, which is in line with the findings shown in Figure 4 that indicate an increase of the concentration on monomeric formaldehyde with increasing temperature. The present results extend the range for which experimental data on K_{x,MG_1} are available substantially, compare Fig. 1. Furthermore, the data from the present work reveal an influence of the overall formaldehyde concentration $\tilde{x}_{\mathrm{FA}}$ on K_{x,MG_1}: the numbers decrease with increasing $\tilde{x}_{\mathrm{FA}}$. This influence has never been described before. Taking the influence of composition into account, the data of *Rivlin et al.*,[1] which were taken at an overall formaldehyde mass fraction of $\tilde{x}_{\mathrm{FA}}$ of about 0.07 g g^{-1}, fit well together with the data from the present work. This underlines that neither using deuterated water instead of water nor using ^{13}C enriched formaldehyde instead of formaldehyde, as *Rivlin et al.*[1] did, has a significant influence on the equilibrium constant. Two types of equilibrium constants were used in the present work: the mole fraction-based K_{x,MG_1} and the activity-based K_{a,MG_1}, cf. Eqns. 1 and 8. The temperature dependence of the experimental data for K_{x,MG_1} and K_{a,MG_1} was correlated with the simple two parametric van't Hoff type ansatz, see Eq. 6. Thermodynamic theory states that K_{x,MG_1} may depend on both the temperature T and the composition (here: on $\tilde{x}_{\mathrm{FA}}$), whereas K_{a,MG_1} depends only on the temperature. Hence, three different correlations were established for $K_{x,\mathrm{MG}_1}(T)$, one for each studied value of $\tilde{x}_{\mathrm{FA}}$. The results are presented in Figure S1 in the Supporting Information. They describe the experimental data very well and lie close to the corresponding lines shown in Figure 5. Also, the results for the parameters A and B of the correlations for $K_{x,\mathrm{MG}_1}(T)$ for each of $\tilde{x}_{\mathrm{FA}}$ are given in Table S2.

In accordance with the theory, only a single correlation was established for $K_{a,\mathrm{MG}_1}(T)$.

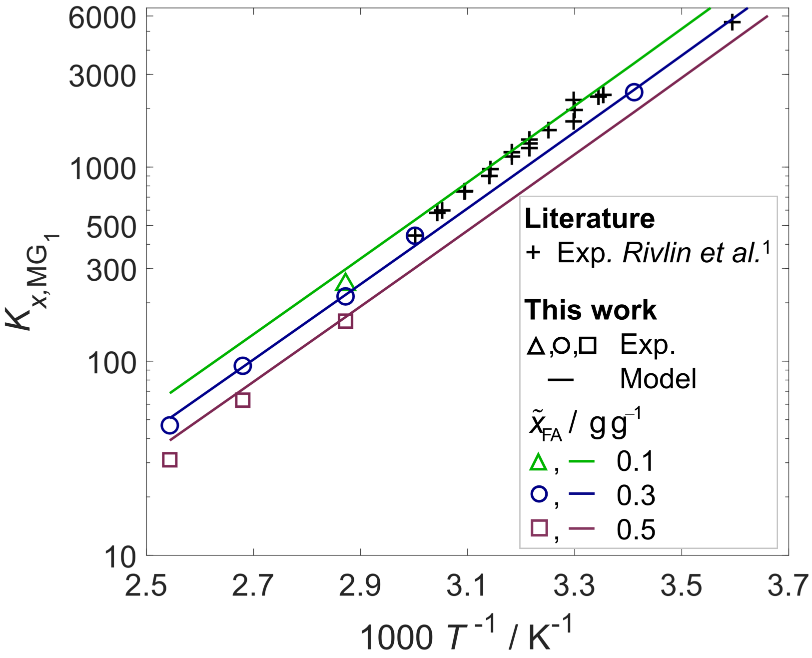


Figure 5: Mole fraction-based equilibrium constants of methylene glycol formation obtained from NMR spectroscopic studies. Symbols: experimental data (colored: this work, black *Rivlin et al.*[1]). Lines denote results of the activity-based model from this work that was fitted only to the data for $\tilde{x}_{\mathrm{FA}} \approx 0.3$ g g^{-1} (blue circles).

To test its ability to predict the influence of the composition of the mixture on the values of K_{x,MG_1}, the correlation of $K_{a,\mathrm{MG}_1}(T)$ was fitted only to the experimental data for $\tilde{x}_{\mathrm{FA}} \approx 0.3$ g g^{-1}. As expected, these data are described very well by the correlation, but the activity-based model also predicts the results for the other studied values of $\tilde{x}_{\mathrm{FA}}$ reasonably well, see Figure 5. Obviously, the influence of the composition of the mixture is captured well by the K_γ term (cf. Eq. 11). The resulting parameters of the correlation of the activity-based model for $K_{a,\mathrm{MG}_1}(T)$ are given in Table 2.

Table 2: Activity-based equilibrium constant of the formation of methylene glycol K_{a,MG_1} (cf. Eq. 6) determined from the NMR spectroscopic data of the present work.

$\ln(K_{a,\mathrm{MG}_1})$	$= A + \frac{B}{T/\mathrm{K}}$
A	-7.9568
B	4.3183

From the temperature dependence of the activity-based equilibrium constant $K_{a,\mathrm{MG}_1}(T)$, we find, using the Gibbs-Helmholtz relation, a reaction enthalpy of methylene glycol formation of -37.5 kJ mol^{-1}; i.e., the reaction is exothermal.

Our results can be compared to those from a widely used physico-chemical model of vapor-liquid equilibria in formaldehyde-containing systems,[4,14,36–42] which was developed by *Maurer* in 1985[43] and has been updated and extended since.[4,14,36–42] A central element of that model is an activity-based correlation of the chemical equilibrium of methylene glycol formation. In Figure S1, we show that the numbers of K_{x,MG_1} that are obtained from the current version of this model are distinctly too low. They lie about a factor of 3 below those from the present work. This means that the model strongly overestimates the concentration of monomeric formaldehyde, which is probably compensated in the vapor-liquid equilibrium model (which describes the vapor-liquid equilibria very well) by values of the vapor pressure of methylene glycol that are too low. (Methylene glycol is unstable as a pure substance, so that its vapor pressure curve had to be estimated by *Maurer*[43]). It would therefore be highly desirable to reparameterize that model based on the new results obtained in the present work. This would, however, require a very substantial effort as the monomeric formaldehyde plays a central role in the model. Hence, basically all model parameters would have to be updated, including all extensions to systems containing other components besides formaldehyde and water, such as methanol, trioxane, and many others.[4,14,17,30,32,35–51] Such an update was out of the scope of the present work.

Different groups have worked on developing physico-chemical models of the properties of reactive systems containing formaldehyde and water, e.g. Refs.[48,49,51] The new results presented here call for these models to be revisited in cases where the concentration of monomeric formaldehyde is important. Clearly, works that use results from such models are also affected.

Conclusion

Formaldehyde is one of the most important intermediates in the chemical industry and almost always used in aqueous solutions. In these solutions, formaldehyde reacts with water to methylene glycol and, subsequently, to poly(oxymethylene) glycols. The formation of these products leads to very low concentrations of the monomeric formaldehyde in aqueous solutions, which are, nonetheless, important, as in the processes, in which formaldehyde is used as an educt, it is the monomeric formaldehyde that reacts with the other educts. The concentration of monomeric formaldehyde is also a central feature in all physico-chemical models of formaldehyde solutions. Despite the importance of methylene glycol formation and the concentration of monomeric formaldehyde in aqueous formaldehyde solutions, and despite extensive studies on the topic that are available in the (mostly older) literature, the quantitative information on the topic was not satisfactory, as data from different sources scatter widely. This is why we have carried out a ^{1}H NMR spectroscopic study of the issue with a cryogenically cooled probe head that enables obtaining good quantitative signals for the small peak of monomeric formaldehyde without using deuterated water. This also paves the way to quantitative studies of the concentration of monomeric formaldehyde in other technical solutions, in which formaldehyde is a reactive species, including reaction kinetic studies, in which in the past, generally, direct quantitative information on the true concentration of formaldehyde was lacking.

Our results confirm those from a study of *Rivlin et al.*,[1] who have also used NMR spectroscopy, but with deuterated water (^{1}H NMR) or ^{13}C enriched formaldehyde (^{13}C NMR). However, our results substantially extend the temperature range, in which data on the concentration of monomeric formaldehyde and the equilibrium constant of methylene glycol formation are available. We also demonstrate for the first time that the mole fraction-based equilibrium constant of methylene glycol formation is concentration-dependent and show that, by using an activity-based model of the chemical equilibrium, concentration dependence can be well predicted. The new data are important for developing physico-chemical models of aqueous formaldehyde solutions, which should give reliable predictions for the small but highly important concentrations of the monomeric formaldehyde.

Acknowledgement

We thank Bernhard Brunner and Stephan Maurer from BASF, Ludwigshafen am Rhein, for fruitful discussions and BASF SE for financial support. This work was carried out in the core facility LASE-MR of RPTU, which is supported by Deutsche Forschungsgemeinschaft (DFG), grant INST 248/370-1.

Supplementary Information Available

Details on the primary experimental data and the correlations of mole fraction-based equilibrium constants of the methylene glycol formation are given in the Supplementary Information.

References

(1) Rivlin, M.; Eliav, U.; Navon, G. NMR Studies of the Equilibria and Reaction Rates in Aqueous Solutions of Formaldehyde. *The Journal of Physical Chemistry B* **2015**, *119*, 4479–4487.

(2) Walker, J. *Formaldehyde*, 3rd ed.; Reinhold Publ. Corp: New York, 1964.

(3) Pekala, R. Organic Aerogels from the Polycondensation of Resorcinol with Formaldehyde. *Journal of Materials Science* **1989**, *24*, 3221 – 3227.

(4) Dyga, M.; Keller, A.; Hasse, H. Vapor–Liquid Equilibria and Chemical Equilibria in the System (Formaldehyde + Water + Isoprenol). *Industrial & Engineering Chemistry Research* **2021**, *60*, 4471 – 4483.

(5) Skell, P.; Suhr, H. Untersuchungen an wässerigen Formaldehydlösungen. *Chemische Berichte* **1961**, *94*, 3317–3327.

(6) Ihashi, Y.; Sawa, K.; Morita, S. Studies of the Dissolved Formaldehyde in Deuterium Oxide by the High Resolution Proton Magnetic Resonance. *The Journal of the Society of Chemical Industry, Japan* **1965**, *68*, 1427–1430.

(7) Koberstein, E.; Müller, K.-P.; Nonnenmacher, G. Molekulargewichtsverteilung von Formaldehyd in wäßrigen Lösungen. *Berichte der Bunsengesellschaft für physikalische Chemie* **1971**, *75*, 549–553.

(8) Fiala, Z.; Navrátil, M. NMR Study of Water-Methanol Solutions of Formaldehyde. *Collection of Czechoslovak Chemical Communications* **1974**, *39*, 2200 – 2205.

(9) Slomin, I. S. G.; Aleksejeva, S.; Akselrod, B. Y.; Urman, Y. G. Spectroscopic Determination of Equilibrium Constants of the Hydration and Protonation of Formaldehyde. *Vysokomolekulyarnye Soedineniya, Seriya B* **1975**, *17*, 919 – 922.

(10) Dankelman, W.; Daemen, J. M. H. Gas Chromatographic and Nuclear Magnetic Resonance Determination of Linear Formaldehyde Oligomers in Formalin. *Analytical Chemistry* **1976**, *48*, 401–404.

(11) Le Botlan, D. J.; Mechin, B. G.; Martin, G. J. Proton and Carbon-13 Nuclear Magnetic Resonance Spectrometry of Formaldehyde in Water. *Analytical Chemistry* **1983**, *55*, 587–591.

(12) Hahnenstein, I.; Hasse, H.; Kreiter, C. G.; Maurer, G. ^{1}H- and ^{13}C- NMR-Spectroscopic Study of Chemical Equilibria in Solutions of Formaldehyde in Water, Deuterium Oxide, and Methanol. *Industrial & Engineering Chemistry Research* **1994**, *33*, 1022–1029.

(13) Hahnenstein, I.; Albert, M.; Hasse, H.; Kreiter, C. G.; Maurer, G. NMR Spectroscopic and Densimetric Study of Reaction Kinetics of Formaldehyde Polymer Formation in

Water, Deuterium Oxide, and Methanol. *Industrial & Engineering Chemistry Research* **1995**, *34*, 440–450.

(14) Albert, M.; García, B. C.; Kreiter, C.; Maurer, G. Vapor-Liquid and Chemical Equilibria of Formaldehyde-Water Mixtures. *AIChE Journal* **1999**, *45*, 2024–2033.

(15) Ott, M.; Fischer, H. H.; Maiwald, M.; Albert, K.; Hasse, H. Kinetics of Oligomerization Reactions in Formaldehyde Solutions: NMR Experiments up to 373 K and Thermodynamically Consistent Model. *Chemical Engineering and Processing: Process Intensification* **2005**, *44*, 653–660.

(16) Maiwald, M.; Fischer, H. H.; Kim, Y.-K.; Albert, K.; Hasse, H. Quantitative High-Resolution On-Line NMR Spectroscopy in Reaction and Process Monitoring. *Journal of Magnetic Resonance* **2004**, *166*, 135–146.

(17) Drunsel, J.-O.; Renner, M.; Hasse, H. Experimental Study and Model of Reaction Kinetics of Heterogeneously Catalyzed Methylal Synthesis. *Chemical Engineering Research and Design* **2012**, *90*, 696–703.

(18) Landqvist, N.; Nygaard, A. P.; Bonnichsen, R.; Virtanen, A. I. On the Polarography of Formaldehyde. *Acta Chemica Scandinavica* **1955**, *9*, 867–892.

(19) Valenta, P. Oszillographische Strom-Spannungs-Kurven III. Untersuchung des Formaldehyds in gepuffertem Milieu. *Collection of Czechoslovak Chemical Communications* **1960**, *25*, 853–861.

(20) Bieber, R.; Trümpler, G. Angenäherte, spektrographische Bestimmung der Hydratationsgleichgewichtskonstanten wässriger Formaldehydlösungen. *Helvetica Chimica Acta* **1947**, *30*, 1860–1865.

(21) Iliceto, A. Sul sistema acque-formaldeide. Nota VI. Equilibri della fasi liquida e gassosa. *Gazzeta Chimica Italia* **1954**, *84*, 536 – 552.

(22) Gruen, L. C.; McTigue, P. T. Hydration Equilibria of Aliphatic Aldehydes in H_2O and D_2O. *Journal of the Chemical Society* **1963**, 5217–5223.

(23) Silling, M. I.; Akselrod, B. Y. Spectrophotometric Determination of Equilibrium Constants of the Hydration and Protonation of Formaldehyde. *Russian Journal of Physical Chemistry* **1968**, *42*, 1479–1482.

(24) Schecker, H.-G.; Schulz, G. Untersuchungen zur Hydratationskinetik von Formaldehyd in wäßriger Lösung. *Zeitschrift für Physikalische Chemie* **1969**, *65*, 221–224.

(25) Zavitsas, A. A.; Coffiner, M.; Wiseman, T.; Zavitsas, L. R. Reversible Hydration of Formaldehyde. Thermodynamic Parameters. *The Journal of Physical Chemistry* **1970**, *74*, 2746–2750.

(26) Bryant, W. M. D.; Thompson, J. B. Chemical Thermodynamics of Polymerization of Formaldehyde in an Aqueous Environment. *Journal of Polymer Science Part A-1: Polymer Chemistry* **1971**, *9*, 2523–2540.

(27) Bell, R. B.; Evans, P. G. Kinetics of the Dehydration of Methylene Glycol in Aqueous Solution. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **1966**, *291*, 297–323.

(28) Winkelman, J.; Voorwinde, O.; Ottens, M.; Beenackers, A.; Janssen, L. Kinetics and Chemical Equilibrium of the Hydration of Formaldehyde. *Chemical Engineering Science* **2002**, *57*, 4067–4076.

(29) Sutton, H. C.; Downes, T. M. Rate of Hydration of Formaldehyde in Aqueous Solution. *Journal of the Chemical Society, Chemical Communications* **1972**, *1*, 1 – 2.

(30) Kuhnert, C.; Albert, M.; Breyer, S.; Hahnenstein, I.; Hasse, H.; Maurer, G. Phase Equilibrium in Formaldehyde Containing Multicomponent Mixtures: Experimental Results for Fluid Phase Equilibria of (Formaldehyde + (Water or Methanol) + Methylal) and

(Formaldehyde + Water + Methanol + Methylal) and Comparison with Predictions. *Industrial & Engineering Chemistry Research* **2006**, *45*, 5155–5164.

(31) Maiwald, M.; Fischer, H. H.; Ott, M.; Peschla, R.; Kuhnert, C.; Kreiter, C. G.; Maurer, G.; Hasse, H. Quantitative NMR Spectroscopy of Complex Liquid Mixtures: Methods and Results for Chemical Equilibria in Formaldehyde-Water-Methanol at Temperatures up to 383 K. *Industrial & Engineering Chemistry Research* **2003**, *42*, 259–266.

(32) Dyga, M.; Keller, A.; Hasse, H. ^{13}C-NMR Spectroscopic Study of the Kinetics of Formaldehyde Oligomerization Reactions in the System (Formaldehyde + Water + Isoprenol). *Industrial & Engineering Chemistry Research* **2022**, *61*, 224–235.

(33) Claridge, T. D. W. *High-Resolution NMR Techniques in Organic Chemistry*, 3rd ed.; Elsevier Science Technology: Chantilly, 2016.

(34) Friebolin, H. *Ein- Und Zweidimensionale NMR-Spektroskopie: Eine Einführung*, 5th ed.; Wiley-VCH, 2013.

(35) Breitkreuz, C. F.; Burger, J.; Hasse, H. Solid–Liquid Equilibria and Kinetics of the Solid Formation in Binary and Ternary Mixtures Containing (Formaldehyde + Water + Methanol). *Industrial & Engineering Chemistry Research* **2022**, *61*, 1871–1884.

(36) Hasse, H.; Hahnenstein, I.; Maurer, G. Revised Vapor-Liquid Equilibrium Model for Multicomponent Formaldehyde Mixtures. *AIChE Journal* **1990**, *36*, 1807–1814.

(37) Hasse, H.; Maurer, G. Vapor–Liquid Equilibrium of Formaldehyde-Containing Mixtures at Temperatures below 320 K. *Fluid Phase Equilibria* **1991**, *64*, 185–199.

(38) Albert, M.; Hahnenstein, I.; Hasse, H.; Maurer, G. Vapor–Liquid Equilibrium of Formaldehyde Mixtures: New Data and Model Revision. *AIChE Journal* **1996**, *42*, 1741–1752.

(39) Albert, M.; Coto García, B.; Kuhnert, C.; Peschla, R.; Maurer, G. Vapor–Liquid Equilibrium of Aqueous Solutions of Formaldehyde and Methanol. *AIChE Journal* **2000**, *46*, 1676–1687.

(40) Albert, M.; Hasse, H.; Kuhnert, C.; Maurer, G. New Experimental Results for the Vapor–Liquid Equilibrium of the Binary System (Trioxane + Water) and the Ternary System (Formaldehyde + Trioxane + Water). *Journal of Chemical & Engineering Data* **2005**, *50*, 1218–1223.

(41) Albert, M.; Hahnenstein, I.; Hasse, H.; Maurer, G. Vapor–Liquid and Liquid–Liquid Equilibria in Binary and Ternary Mixtures of Water, Methanol, and Methylal. *Journal of Chemical & Engineering Data* **2001**, *46*, 897–903.

(42) Schmitz, N.; Breitkreuz, C. F.; Ströfer, E.; Burger, J.; Hasse, H. Vapor–Liquid Equilibrium and Distillation of Mixtures Containing Formaldehdye and Poly(oxymethylene) dimethyl ethers. *Chemical Engineering and Processing - Process Intensification* **2018**, *131*, 116–124.

(43) Maurer, G. Vapor-Liquid Equilibrium of Formaldehyde- and Water-Containing Multicomponent Mixtures. *AIChE Journal* **1986**, *32*, 932–948.

(44) Burger, J.; Ströfer, E.; Hasse, H. Production Process for Diesel Fuel Components Poly(oxymethylene) dimethyl ethers from Methane-Based Products by Hierarchical Optimization with Varying Model Depth. *Chemical Engineering Research and Design* **2013**, *91*, 2648–2662.

(45) Schmitz, N.; Friebel, A.; von Harbou, E.; Burger, J.; Hasse, H. Liquid-Liquid Equilibrium in Binary and Ternary Mixtures Containing Formaldehyde, Water, Methanol, Methylal, and Poly(oxymethylene) dimethyl ethers. *Fluid Phase Equilibria* **2016**, *425*, 127–135.

(46) Schmitz, N.; Ströfer, E.; Burger, J.; Hasse, H. Conceptual Design of a Novel Process for the Production of Poly(oxymethylene) Dimethyl Ethers from Formaldehyde and Methanol. *Industrial & Engineering Chemistry Research* **2017**, *56*, 11519–11530.

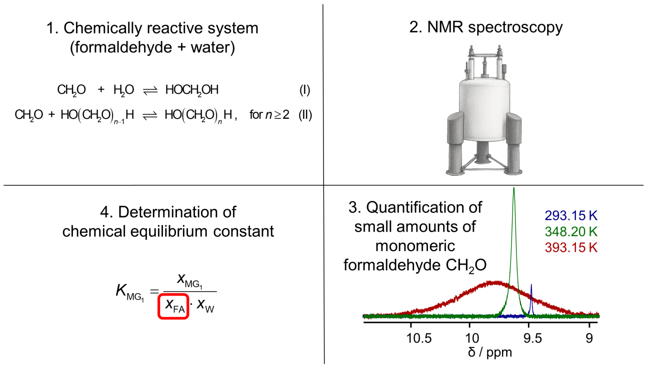
(47) Weidert, J.-O.; Burger, J.; Renner, M.; Blagov, S.; Hasse, H. Development of an Integrated Reaction–Distillation Process for the Production of Methylal. *Industrial & Engineering Chemistry Research* **2017**, *56*, 575–582.

(48) Liu, H.; Bai, Z.; Liu, Y.; Guo, X.; Fu, Y. Prediction of the Vapor–Liquid Equilibrium of Formaldehyde–Water–Trioxane Ternary System by the Conductor-Like Screening Model for Real Solvents. *International Journal of Chemical Engineering and Applications* **2017**, *8*, 82–86.

(49) Ferre, A.; Burger, J. Coadsorption Equilibria on Molecular Sieves 3A and Densities of Liquid Mixtures Containing Formaldehyde, Methanol, and Water at 295.15 and 313.15 K. *Industrial & Engineering Chemistry Research* **2021**, *60*, 15256–15263.

(50) Dyga, M.; Keller, A.; Hasse, H. Distillation Behavior of the System (Formaldehyde + Water + Isoprenol). *Chemical Engineering Science* **2022**, *263*, 118082.

(51) Winklbauer, L.; Burger, J. Chemical equilibrium and liquid–liquid equilibrium in ternary mixtures of formaldehyde, water, and isobutanol/ n-butanol between 295 K and 348 K. *Fluid Phase Equilibria* **2025**, *594*, 114364.


1. Chemically reactive system (formaldehyde + water)
$CH_2O + H_2O \rightleftharpoons HOCH_2OH$ (I)
$CH_2O + HO(CH_2O)_{n-1}H \rightleftharpoons HO(CH_2O)_nH$, for $n \geq 2$ (II)
2. NMR spectroscopy
3. Quantification of small amounts of monomeric formaldehyde CH_2O
293.15 K
348.20 K
393.15 K
10.5
10
9.5
9
δ / ppm
4. Determination of chemical equilibrium constant
$K_{MG_1} = \frac{x_{MG_1}}{\boxed{x_{FA}} \cdot x_W}$