The hydration of formic acid and acetic acid

Stefano Soffientini [a], Leonardo Bernasconi [b], Silvia Imberti [b,*]

[a] Università degli Studi di Milano-Bicocca, Dipartimento di Fisica, Milano, Italy
[b] STFC, Rutherford Appleton Laboratory, Harwell Oxford Campus, OX11 0QX, United Kingdom

ARTICLE INFO

ABSTRACT

Formic acid and acetic acid share the unique feature, amongst carboxylic acids, of crystallising in the form of long chains, containing both O–H··O and C–H··O hydrogen bonds. We have performed a neutron diffraction study of the pure acids and of three mixtures of acid and water (2:1, 1:1 and 1:2). The data from the SANDALS diffractometer at ISIS have been modelled using the Empirical Potential Structure Refinement code, which is able to reproduce a set of configurations compatible with the experimental data. The relative importance of the hydrogen bonds present in the solution is assessed based on geometrical criteria: bond length and directionality as well as number of bonds. At all concentrations, the carbonyl oxygen on the carboxylic group is the most active site for a strong hydrogen bond. The tendency to establish direct interactions between acid molecules in the presence of water is reduced for acetic acid by a larger degree than for formic acid. The overall tendency is for a greater number of hydrogen bonds being formed when the solution is more diluted. The availability of good quality structural data on the liquid states is of great importance for the understanding of spectroscopic experiments and for benchmarking both classic molecular dynamics and ab initio simulations. The results provide a springboard to more realistic models of aerosol formation, which is greatly needed for better understanding of cloud formation processes.

1. Introduction

Anthropogenic aerosol particles such as sulfate and carbonaceous aerosols can influence climate in several ways: they scatter and absorb solar radiation and thermal radiation from the Earth (direct effects) and they also act as cloud condensation nuclei and ice nuclei (indirect effects). Indirect aerosol effects are nowadays one of the greatest sources of uncertainty in assessing the human impact on climate change [1]. While nucleation of sulfate aerosol particles is well established, nucleation from organic precursors has not been proven. One currently held hypothesis states that organic mixtures form initially a different phase (glassy or crystalline), then give rise to heterogeneous nucleation. The alternative one is that the process triggered by non-sulfate organic molecules in the atmosphere is not nucleation but rather condensational growth [2]. In either case, a quantitative knowledge of the fundamental nucleation and crystallisation processes of organic–water mixtures is essential in order to understand and predict the impact of organic pollution on the Earth's atmosphere and climate [3].

Infrared spectra of aqueous formic acid and acetic acid aerosols reveal that their internal structure critically depends on the particle formation conditions (e.g. temperature and humidity variations) and, especially for the solutions, on their composition [4]. But, as stated by Koop et al. [5]: "(…) before these effects are at a point that they can

be described for atmospheric purposes, there is still a large lack of basic physical and chemical understanding in these systems, requiring a considerable piece of basic research before significant progress can be made in this area." Indeed, several basic questions remain unanswered, such as: What are the glass-forming properties of the neat acid and of its mixture with water? What is the stability of these complexes when temperature and humidity level vary? How hygroscopic is an acid–water complex? and closer to the point of this article: What is the structure of an acid–water complex in the liquid and glassy state? How does cooperativity affect this geometry when the number of hydrating water molecules increases?

In the present contribution we have performed a neutron diffraction study of three mixtures of acid and water, at approximately 70%, 50% and 30% acid mole fraction. For each concentration three independent patterns of formic acid/water mixture, and four patterns of acetic acid/water mixtures have been measured, with deuterium substituted for hydrogen to a varying degree. For each concentration, the three (or four) patterns have been used to simultaneously drive a Monte Carlo simulation, through the well tested EPSR method [6]. In this way the structural properties have been calculated as an average over an ensemble of simulation boxes, all compatible with the neutron diffraction experimental data.

Formic acid and acetic acid share the unique feature, amongst carboxylic acids, of crystallising in the form of long chains, containing both O–H··O and C–H··O hydrogen bonds; these chains are then bonded to each other via dispersion forces, forming a secondary and a tertiary

* Corresponding author.
  E-mail address: silvia.imberti@stfc.ac.uk (S. Imberti).

structure [7]. This is at variance with the gaseous state, which is mostly constituted of cyclic dimers [8]. Spectral features of the pure liquids have been measured in the past [9,10], but their interpretation is still somewhat open. This constitutes the focus of many of the ab initio studies on small complexes [11–15]. Several studies have tackled the structure of both neat acids in the liquid state (see for example Chelli et al. [16] and Imberti et al. [17] and references therein), but data on the aqueous system were still mostly missing (with one notable exception in Ref. [18]).

Our main goal is to extract the radial and angular distribution functions of the water–acid system for each concentration and to compare them with those of the pure acids [17]. The relative importance of the various hydrogen bonds present in the solution is assessed based on geometrical criteria: bond length and directionality as well as number of bonds.

2. Materials and methods

The neutron experiments have been performed on the SANDALS diffractometer at the ISIS spallation neutron source, located at the Rutherford Appleton Laboratory (United Kingdom). SANDALS is a total scattering neutron diffractometer optimized for the study of liquids and amorphous samples containing light elements. The physical quantity measured by the diffractometer is the differential scattering cross section $d\Sigma/d\Omega$ as a function of the exchanged wave vector Q (defined as the modulus of the difference between the incident and the scattered neutron wave vectors). Through the basic theory of neutron scattering [19], it is possible to relate this quantity to the static structure factor $F(Q)$, which is the Fourier transform of the atomic pair distribution function $g(r)$. The latter contains the information about the correlation between the positions of two atoms in the system at a given moment.

The ISIS pulsed neutron source produces a time-structured beam with a frequency of 50 Hz. The diffractometer faces a liquid methane moderator operating at 110 K, which provides neutrons in the wavelength range between 0.05 and 4.95 Å. The sample is placed in an evacuated chamber after a collimated flight path, 11.02 m from the moderator face. SANDALS works in transmission geometry, the samples are contained in flat TiZr alloy cells (internal thickness and wall thickness 1 mm) and these are mounted on an automated sample changer. Each detector on SANDALS provides an independent measure of the structure factor $F(Q)$, within the Q range allowed by its fixed angular position. The wavelength of the detected neutron is determined by its time of flight from the moderator. The instrument has 633 ZnS scintillator detectors covering an angular range between 3.5 and 40°. The forward scattering geometry minimizes the inelasticity corrections that are applied to the data from light element containing samples. The combination of the wide wavelength range and angular coverage provides a very wide Q range, running from 0.2 to 50 Å^{-1}.

The H/D isotopic substitution method [20] has been applied, in order to provide several independent determinations of the structure factor of the acetic acid (a.a.) and formic acid (f.a.) solutions. The samples have been purchased from Sigma-Aldrich. In the case of a multicomponent system, it can be shown [21] that the structure factor extracted from a neutron diffraction experiment is a linear combination of the partial structure factors $S_{\alpha\beta}(Q)$ in the following way:

$$F(Q) = \sum_{\alpha\beta \geq \alpha} \left(2 - \delta_{\alpha\beta}\right) c_\alpha c_\beta b_\alpha b_\beta S_{\alpha\beta}(Q) \tag{1}$$

where α and β are two atom types present in the system, $\delta_{\alpha\beta}$ is the Kronecker delta function, c is their concentration and b their scattering length [22]. Since the scattering length b varies from one isotope to another, the shape of the measured function $F(Q)$ can vary quite dramatically with changing the isotopes involved. In particular, by substituting [1]H ($b = -3.740$ fm) with [2]H ($b = 6.671$ fm), it is possible to highlight H–X correlations, where X is a non-substituted atom (e.g. oxygen).

The standard corrections and normalizations have been applied to the data through the set of programs gathered under the graphical interface *GudrunN*. The theoretical background to the operations performed by the program are described in reference [21]. Both *GudrunN* and its X-ray equivalent *GudrunX* [23] have been written by Alan K. Soper (ISIS, STFC) and are described in detail in the manual *GudrunN and GudrunX. Programs for correcting raw neutron and X-ray diffraction data to differential scattering cross section* [21].

2.1. Data modelling

Models of the experimental data have been constructed using the Empirical Potential Structure Refinement (EPSR) program. The method has been described in detail elsewhere (see Soper [24] and references therein) and therefore only a brief summary will be given here. The algorithm is based on a classical Monte Carlo simulation of the molecular system under study at fixed concentration and density, and employs an iterative algorithm that aims at building an atomistic three dimensional model consistent with the scattering data. The simulation proceeds through a number of stages. In the first stage the simulation follows the standard Monte Carlo algorithm for simulations of molecular structures [19], based on a pairwise reference potential (Lennard-Jones plus Coulomb). In the following stage a perturbation potential is determined from the difference between the calculated and experimental structure factors. The simulation box is equilibrated with the reference plus empirical potential and a new empirical potential is calculated and added to the previous one. This iterative process drives the model into agreement with the data. Once agreement is reached between the experimental and model structure factors, the final stage of the procedure is undertaken and structural information is collected in the form of ensemble averages. Each simulation box contained 2000 molecules. All the simulation details are summarised in Table 1.

In principle it is possible to calculate the number of molecules which dissociate in water based on tabulated pKa values, according to the Henderson–Hasselbalch equation. In reality, at the high concentrations studied here, there is a considerable disagreement between calculated and measured pH values. This fact points towards a rich dissociation scenario, which may include the formation of protonated acid molecules and other complexes. For these reasons, the simulations presented here, but in fact mostly only the 0.30 formic acid in water one, can only be a somewhat simplified model of the real system.

Each molecule is obtained as a semi-rigid [24] atomic configuration, defined by geometry, such as bond lengths and angles. The intramolecular structure of these molecules is known and a number of sources in literature carry this information. In this work the intramolecular structure of the acid molecules (detailed in Table 2) has been optimized in vacuum using the free computational chemistry software Ghemical 2.0 [25]. It is known that the intramolecular structure in vacuum may differ from the one in the bulk liquid, but in the case of the simulations shown in this study, the difference was not appreciable e.g. there was no significant misfit in the high-Q region of the diffraction patterns. For the water molecules, Single Point Charge Extended [26] has been used for both geometry and potentials. A picture of the acid molecules with atom's labelling is included for convenience (Fig. 1).

The seeding (*reference*) potentials have been derived from different sources in the literature. For acetic acid Jorgensen's OPLS parametrization [27] has been used; this was originally aimed at the reproduction of thermodynamic properties of liquids, such as density and latent heat of vaporisation, but has proven to work well also for structural properties. For formic acid the potential has been obtained from the work by Jedlovszky et al. [28,29]. This potential has been derived specifically for reproducing structural patterns reliably. Both these potentials have been previously tested in an EPSR simulation with success [17]. The actual values used for the potentials are summarised in Table 3. All of the EPSR simulations have been run under the same conditions, allowing for a maximum amplitude of the nonparametrized empirical