Research

Author for correspondence:
Junhong Chen
e-mail: junhongchen@uchicago.edu

†Present address: X-Ray Science Division, Argonne National Laboratory, 9700 S. Cass Avenue Lemont, IL 60439, USA.

THE ROYAL SOCIETY
PUBLISHING

Molecular interactions in short-chain perfluoroalkyl carboxylic acids and aqueous solutions

Chris J. Benmore[1,3,†], Yuqin Wang[2,4], Seth B. Darling[2,4] and Junhong Chen[2,4]

[1]X-Ray Science Division, Advanced Photon Source, and [2]Chemical Sciences and Engineering Division, Physical Sciences and Engineering Directorate, Argonne National Laboratory, Lemont, IL 60439, USA
[3]Consortium for Advanced Science and Engineering, and [4]Pritzker School of Molecular Engineering, University of Chicago, Chicago, IL 60637, USA

CJB, 0000-0001-7007-7749

The presence of short-chain per- and polyfluoroalkyl substances in water poses a major health and environmental challenge. Here, we have performed high-energy small- and wide-angle X-ray scattering measurements on $CF_3[CF_2]_nCOOH$ (where $n = 1$, 2, 3 represents the chain length) and their aqueous solutions at 10% mole concentrations to characterize their molecular interactions at the atomic and nanometer length scales. The experimental wide-angle structure factors have been modelled using Empirical Potential Structural Refinement. The oxygen–oxygen partial X-ray pair distribution functions show that the coordination number between the hydroxyl oxygen on the acid and surrounding oxygen water molecules increases significantly with acid chain length, rising from 3.2 for $n = 1$ to 4.1 for $n = 3$. The small-angle scattering is dominated by a sharp, high-intensity peak at $Q_1 \sim 0.2\,\text{Å}^{-1}$ and a smaller peak at $Q_2 = 1.2\,\text{Å}^{-1}$ for $n = 3$, both of which decrease with decreasing chain length. The Q_2 peak is attributed to groups of adjacent non-bonded acid molecules, and Q_1 has contributions from both correlations between acid molecules and water–water

interactions. In all cases, the models show nanoscale aggregation occurs in the form of denser channels of winding hydrogen-bonded chains, approximately 20 water molecules in length, surrounding clusters of acid molecules.

This article is part of the theme issue 'Exploring the length scales, timescales and chemistry of challenging materials (Part 2)'.

1. Introduction

Water contamination by per- and polyfluoroalkyl substances (PFAS) has become an increasingly severe environmental and health crisis. Exposure to PFAS chemicals can lead to adverse human health effects, since they are persistent in the environment due to the strong bonding between carbon and fluorine [1–5]. Moreover, since they can accumulate over time, the U.S. Environmental Protection Agency (EPA) has established and recently updated, health advisory levels to provide Americans, including the most vulnerable populations, with a margin of protection from a lifetime of exposure to PFAS [6]. It is estimated that over 200 million people likely receive water with a PFAS concentration at or above $1 \, \mathrm{ng} \, \mathrm{l}^{-1}$ [7]. Structurally, PFAS molecules each possess a hydrophobic 'tail' and a hydrophilic 'head', figure 1. Although long-chain PFAS are most commonly detected in the aquatic environment, their use has been limited by recent regulations, and more attention is being paid to the rise in alternative short-chain compounds. Short-chain PFAS have a high mobility in water bodies, and their removal by adsorption is more challenging [8,9]. The shorter chain length results in less hydrophobic interactions compared with long-chain PFAS, which is the dominant mechanism of capture making the remediation of short-chain PFAS much harder. It is therefore important to quantify both the hydrophobic and hydrophilic interactions of PFAS with water as a function of chain length.

To investigate these interactions associated with fluoroalkyl chains and water at the atomic scale, we have studied the structure of pentafluoropropionic acid, heptafluorobutyric acid, perfluoropentanoic acid and their aqueous solutions at 10% mole concentrations. Rat oral acute toxicity data from the EPA centre for computational toxicology and exposure database indicates that these short-chain fluorinated acids have lethal concentrations at approximately $0.1 \, \mathrm{g} \, \mathrm{l}^{-1}$ [10]. However, diffraction experiments at even these concentrations are untenable due to the weak signal. Nonetheless, the high solubility of short fluoroalkyl acid chains allows us to maximize the X-ray signal between acid and water and investigate the interactions with water. The characterization of the number and strength of hydrogen bonds, and hydrophobic interactions between short-chain PFAS and water molecules at high concentrations provide insights relevant to the environmental challenge of PFAS. It is anticipated that the details of the molecular interactions from such fundamental studies will provide stringent benchmarks for the development of interatomic potential parameters used in computer simulations that can be employed at much lower contaminant concentrations.

Diffraction studies on fluorinated liquids include the liquid-state structure of hydrogen fluoride [11], fluorosulfuric acid [12], hexafluoro-iso-propanol (HFIP)–water mixtures [13–16] and halogeno–ethanol–water mixtures [17]. The latter systems also employed X-ray and neutron diffraction with Empirical Potential Structure Refinement (EPSR) modelling that showed microheterogeneities on the length scale of a few nanometers due to clustering. HFIP–water mixtures for example are associated with a change from a tetrahedral-like water network structure to a chain-like water structure at high HFIP molecule concentrations. More broadly, the majority of diffraction, spectroscopic and modelling studies on aqueous solutions have focused on binary alcohol-water mixtures [18–20], usually at high concentrations (greater than 10 mole%) to maximize the signal between the solute and water molecules. Of note, neutron diffraction experiments have indicated that in high-concentration methanol–water mixtures, evidence for separate percolating networks coexists, despite both liquids being fully miscible in all proportions [19]. Tang *et al.* [18] have suggested that nanoscale segregation of water near
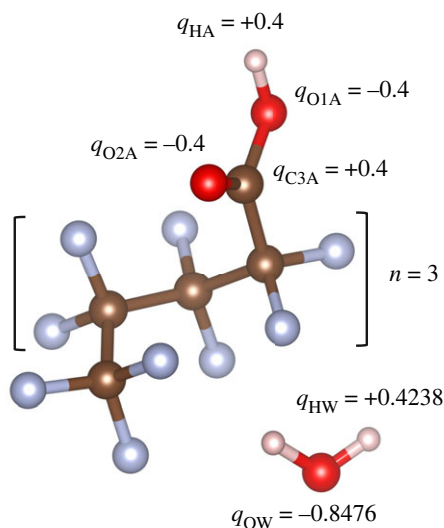
Figure 1. Structure of perfluoroalkyl carboxylic acid and water molecules used in the EPSR simulation including the partial charges and atom labels. Label O1A represents a hydroxyl oxygen and O2A a carbonyl oxygen. Acids with chain lengths of $n = 1$, 2 and 3 were studied.

molecules containing both hydrophobic and hydrophilic groups is commonplace and can lead to anomalous properties in heat capacity, vapour pressure and acidity. In this paper, we apply modern high-energy X-ray small- and wide-angle diffraction techniques combined with EPSR modelling to elucidate the structure of neat fluoroalkyl acids and acids in aqueous solutions to investigate these molecular interactions and their aggregation over a wide range of length scales.

2. High-energy X-ray diffraction

The X-ray pair distribution function (PDF) method is a well-established technique for the characterization of both local and intermediate range ordering of disordered organic materials, providing details of molecular structure at the atomic level [21]. Pure perfluoroalkyl carboxylic acid chains of $CF_3[CF_2]_nCOOH$, where $n = 1$, 2, 3 (figure 1), and their aqueous solutions of 1 acid : 10 water molecules were loaded into 2 mm diameter, thin-walled (0.1 mm) capillaries, and measured at 25°C. The high-energy X-ray small-angle and wide-angle X-ray scattering (SAXS/WAXS) measurements were performed on beamline 6-ID-D at the Advanced Photon Source at Argonne National Laboratory. The set-up and correction procedures have been previously described in detail [22], whereby the SAXS data were normalized to the WAXS data over a substantial overlap region, which for high-energy X-ray diffraction means essentially normalizing to the number of electrons in the system at high-Q. Experiments were carried out using a monochromatic X-ray beam $E = 100$ keV ($\lambda = 0.124$ Å) collimated to a square 0.5 mm cross-section, and the scattered beam measured using a Varex CT4343 area detector. A NIST CeO_2 powder standard was used for the WAXS sample-detector distance calibration, which was set to 360 mm in order to balance resolution and Q-range, where $0.6 < Q(Å^{-1}) < 25.7$. A silver behenate powder standard was used for the SAXS sample-distance calibration, corresponding to 1430 mm, where $0.05 < Q(Å) < 3.3$. The SAXS and WAXS data were analysed as described previously [23] using the software *Fit2D* [24] and *PDFgetX2* [25] and combined to yield S(Q) over the entire Q-range, figure 2. In brief, dark current, geometrical effects, beam polarization, background and attenuation corrections were applied to all datasets to yield the total X-ray structure factors S(Q) and differential PDFs D(r) [22].
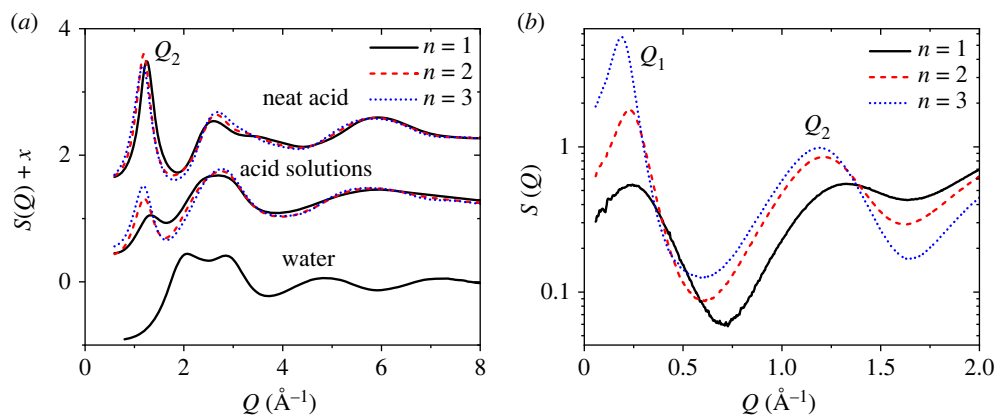
Figure 2. (*a*) The WAXS structure factors of neat acids $CF_3[CF_2]_nCOOH$ where, $n = 1, 2$ and 3, diluted water 10 : 1 acid solutions and pure water. (*b*) The measured SAXS structure factors for the 10 water : 1 acid solutions on a log *y*-scale. Q_1 and Q_2 denote the first two peaks in the SAXS signal.

3. Empirical Potential Structure Refinement modelling

To investigate the hydrogen bonding and hydrophobic behaviour of the neat short-chain fluorinated acids and aqueous solutions, EPSR modelling was used [26] to obtain atomistic models based on the high-energy X-ray diffraction data. The EPSR simulations were performed on 150–230 acid molecules for the neat liquids (to maintain a cubic box length of approx. 30 Å) using atomic number densities of 0.075 atoms $Å^{-3}$, in order to assess the validity of the method. Larger models containing 300 acid molecules plus 3000 waters (dilute solutions) within a cubic box of length approximately 50 Å were assembled using atomic number densities of 0.09 atoms $Å^{-3}$ for the solutions to capture the observed SAXS features. EPSR is essentially a rigid body simulation but molecular flexibility can also be introduced, so here rotations of bonds along the fluoroalkyl chain and carboxyl group were enabled to provide a better fit. The parameters for the Lennard-Jones parameters and effective charges are illustrated in figure 1 and the electronic supplementary material, information.

Following the initial Monte Carlo equilibration, the empirical potential term was refined to improve agreement with scattering data. Once the goodness-of-fit parameter was minimized between the model and the experimental S(*Q*), structural data were collected over ensembles of at least 2000 configurations, figure 3. Overall, good fits were found for the neat acids; however, finding a balance for fitting both the SAXS and WAXS regions of the solutions simultaneously proved more problematic, particularly for the $n = 1$ solution, although reasonable compromises were eventually found. While the EPSR fit to the data does not necessarily give a unique structural three-dimensional configuration of molecules, it does provide important insight into the types of interactions that are likely in the liquid state. Since X-rays are scattered by electrons, the S(*Q*)s and corresponding PDFs are most sensitive to the heavier carbon, oxygen and fluorine atoms, i.e. the chain backbones and centres of the water molecules.

4. Results

As shown in figure 2, the measured low-*Q* SAXS peak position in the solution spectra decreases systematically with increasing acid chain length, from $Q_1 = 0.245 \pm 0.005 \, Å^{-1}$ (for $n = 1$), to $0.23 \, Å^{-1}$ ($n = 2$) and $0.19 \, Å^{-1}$ ($n = 3$). The intense Q_1 peak for $n = 3$ dominates the X-ray structure factor and is $\times 10.5$ larger than the peak in $n = 1$ and $\times 3.2$ larger than the peak in $n = 2$. The sharpness of the Q_1 peaks indicates the presence of well-defined periodicities in real space that extend beyond the size of our cubic simulation box, i.e. greater than 50 Å length. However, our attempts to scale up to 100 Å box length resulted in unreasonable equilibration times. Similar
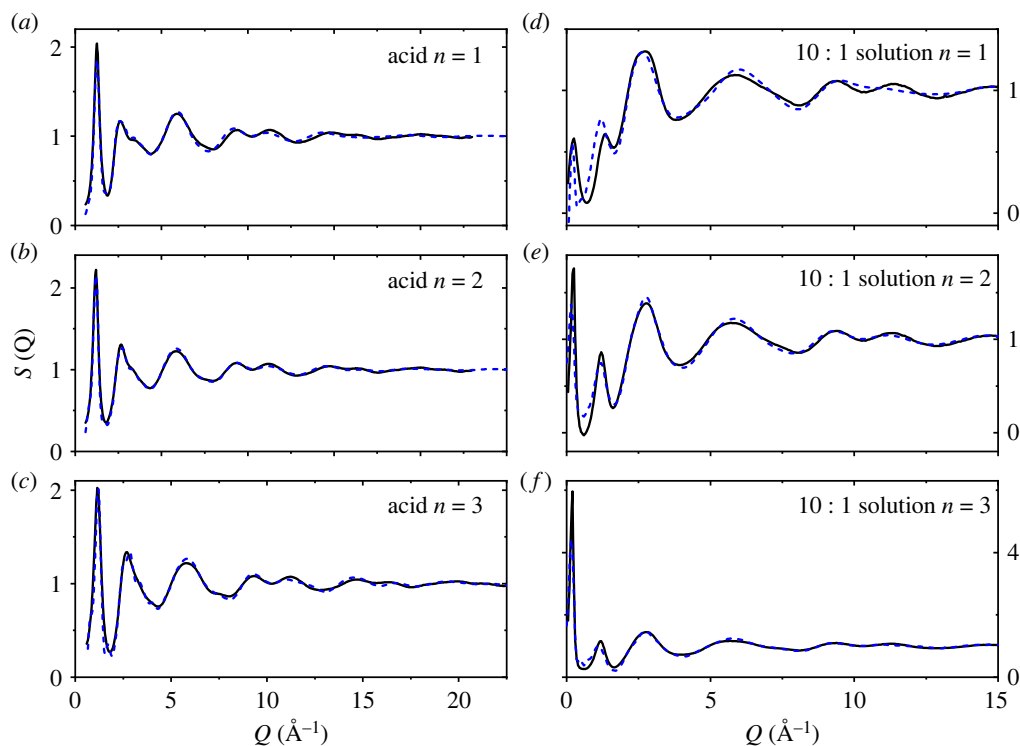
5

Figure 3. Measured X-ray S(Q)'s (black solid lines) and EPSR fits (dashed blue lines) to that of (a–c) neat liquid CF$_3$[CF$_2$]$_n$COOH where, $n = 1$, 2 and 3, and (d–f) diluted water 10 : 1 acid solutions for $n = 1$, 2 and 3.
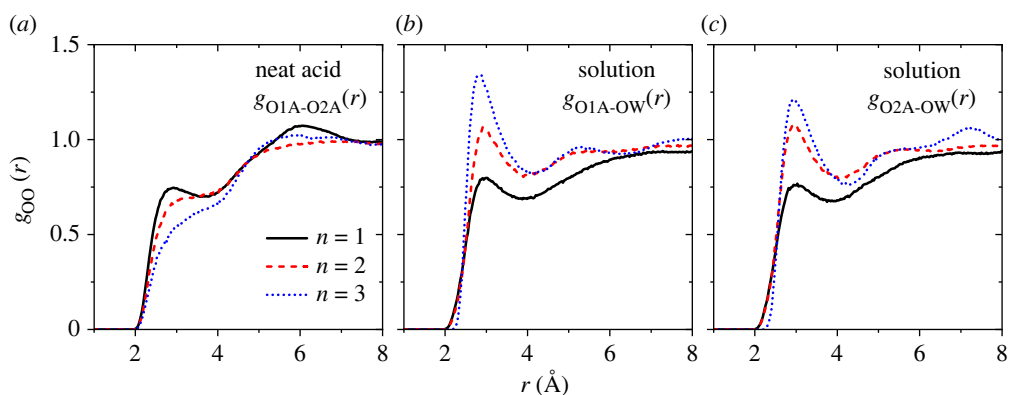

Figure 4. The EPSR oxygen–oxygen partial PDFs for (a) the acid hydroxyl oxygen–acid carbonyl oxygen $g_{O1A-O2A}(r)$ in the neat acids, (b) the acid hydroxyl oxygen–water oxygen $g_{O1A-OW}(r)$ in the solution mixture and (c) the acid carbonyl oxygen–water oxygen $g_{O2A-OW}(r)$ in the solution mixture.

issues have been reported on Monte Carlo fits of other large-scale simulations [27]. The so-called first sharp diffraction peak (FSDP) in the WAXS signal, denoted here as Q_2 in the solution S(Q)'s is known to be associated with intermediate range order in the liquid [21]. The FSDP peak positions are also found to decrease systematically with increasing acid chain length, from $Q_2 = 1.33 \pm 0.01$ Å^{-1} (for $n = 1$), to 1.21 Å^{-1} ($n = 2$) and 1.19 Å^{-1} ($n = 3$).

Details of the molecular interactions between molecules in the pure acids and solutions are contained in the EPSR oxygen–oxygen partial pair correlations in figure 4. The neat acid molecular

models show the liquid structure comprises clusters of winding hydrogen-bonded chains. The hydroxyl oxygen-carbonyl oxygen PDF $g_{O1A-O2A}(r)$ for the neat acid $n = 1$ is centred around $r_{OO} = 2.9$ Å, and the corresponding coordination number is $n_{OO} = 1.0$ integrating out to the first minimum at $r_{int} = 4.0$ Å. As expected from steric hindrance considerations, the value of n_{OO} decreases with increasing acid chain length to 0.7 for $n = 2$ and 0.6 for $n = 3$, resulting in fewer hydrogen-bonded chains (because not every molecule is hydrogen-bonded). In the aqueous acid solutions, the hydrogen bonding between the acid molecules and water molecules increases significantly with increasing acid chain length. For $n = 1$, the carbonyl oxygen acceptor O2A shows the weakest interaction with water oxygen OW at a distance of $r_{OO} = 2.9$ Å and has a coordination number of $n_{OO} \sim 3.2$ (integrating out to $r_{int} = 4$ Å). The value of n_{OO} increases to 3.8 for the $n = 2$ solution and 4.1 for $n = 3$. Concomitantly, the n_{OO} hydroxyl (donor) oxygen O1A–water oxygen coordination number also increases with increasing chain length, from 3.3 for $n = 1$, to 3.8 ($n = 2$) and 3.8 ($n = 3$). This indicates that the carboxylic 'head' on the smallest acid molecules hydrogen bonds less with the surrounding waters, most likely due to geometrical constraints. This is similar to the behaviour found in halogenoethanol-water mixtures, where the neat liquids are hydrogen-bonded to each other but separated due to steric hindrance, and as the water content increases a tetrahedral-like structure of water evolves as the halogenoethanol and water molecules cluster leading to heterogeneous mixing at the molecular level [17].

5. Discussion

Short-chain alcohols such as methanol, ethanol and propanol in aqueous solutions all form heterogeneous clusters at mole fractions greater than 10% [18,28]. Previous X-ray diffraction studies have indicated that hydrogen-bonded alcohol chains and tetrahedral-like water networks coexist at room temperature [28]. Furthermore, Li *et al.* [28,29] have suggested a structural transition occurs between isolated alcohol molecule concentrations that have independent hydration shells, and the inhomogeneous mixing of extensive hydrogen-bonded water networks and the aggregation of solute molecules at higher concentrations. Such nanoscale heterogeneity is highly composition dependent [18] since the maximum number of water molecules in the hydration shell surrounding a single-alcohol molecule increases with the length of the carbon chain of the alcohol [29]. Snapshots of our three-dimensional EPSR simulation boxes for the short-chain perfluoroalkyl carboxylic acid solutions also reveal the nanoscale aggregation of water molecules, figure 5a–c. With increasing acid chain length, the models show a systematic increase in the acid–water O_A–O_W correlations (with similar magnitudes for the acid donor–water acceptor and acid acceptor–water donor hydrogen-bonding interactions). Most of the acid–water interactions involve hydroxyl and/or carboxyl groups hydrogen bonding to water molecules in densely packed channels, although many other interactions in close proximity are non-bonded O_A–O_W interactions.

(a) Nanoscale aggregation

Inspection of the EPSR Q-space models reveals that the Q_2 peak arises mainly from $S_{C-C}(Q)$ (carbon–carbon) interactions, i.e. acid–acid correlations originating from the molecular backbones. It can be seen in figure 5d that adjacent (non-bonded) acid molecules are relatively evenly distributed throughout the simulation box giving rise to the periodicity associated with Q_2. The Q_1 peak in the SAXS signal is more complex and mostly originates from both the $S_{C-C}(Q)$ and $S_{OW-OW}(Q)$ (water–water) partial structure factors with a periodicity of $2\pi/(Q_1) \sim 30$ Å. This corresponds to density fluctuations from both acid–acid and water–water correlations, respectively, due to clustering, that become more strongly correlated as the acid molecule size increases. The peak in $S_{C-C}(Q)$ is attributed to groups of acid molecules and the peak in $S_{OW-OW}(Q)$ from the aggregation of water molecules surrounding these clusters.

Next, we investigate the changing network topology with acid chain length. Figure 6a shows that the water network, denoted by the O_W-O_W partial PDF is significantly broadened compared
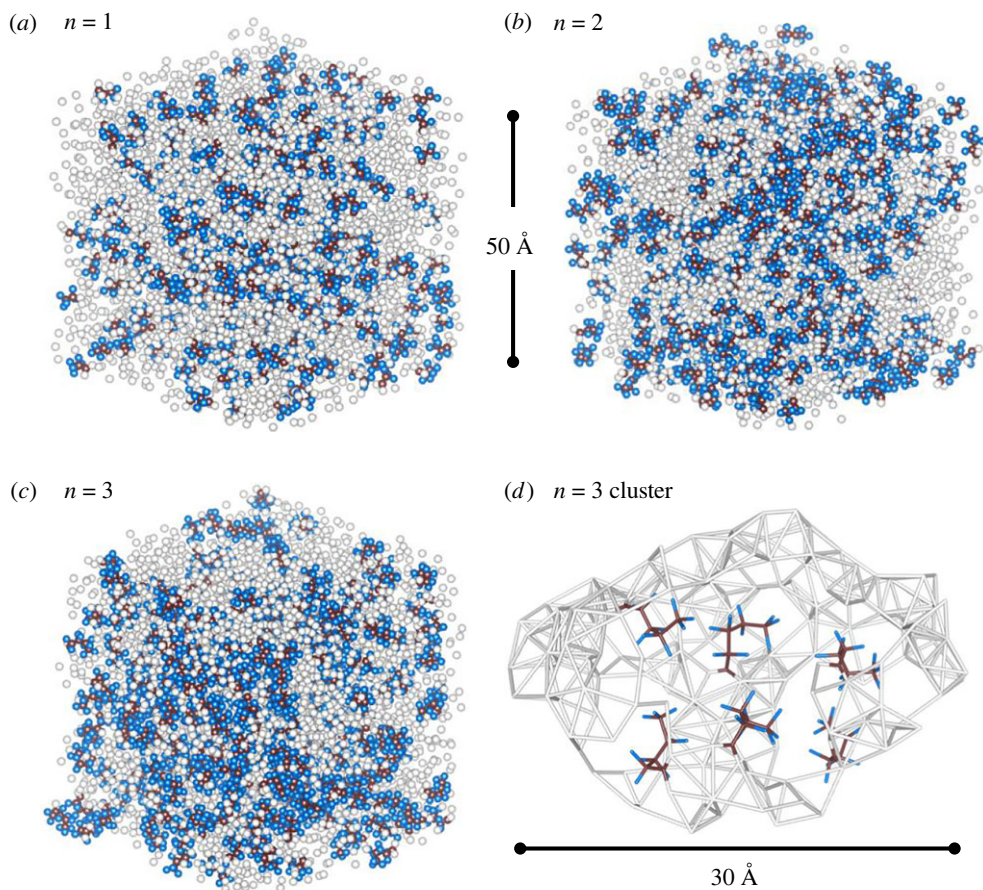
(*a*)   *n* = 1   (*b*)   *n* = 2

50 Å

(*c*)   *n* = 3   (*d*)   *n* = 3 cluster

30 Å

Figure 5. (*a–c*) The full three-dimensional EPSR simulation boxes for $n = 1, 2$ and 3 water 10 : 1 acid solutions, where water oxygens are plotted as white spheres and the acid molecules are brown (carbon) and blue (fluorine) for clarity. Hydrogen atoms have been omitted for clarity; (*d*) stick representation of a water cluster surrounding six $n = 3$ acid molecules, showing the regularity of the acid spacing and an adjacent dense water channel at the top.

with pure liquid water, indicating a highly disordered network containing several non-bonded correlations. The O_W-O_W coordination number in an open tetrahedral-like pure water is 4.3 (integrating out to $r_{int} = 3.3$ Å [30]). This compares with lower n_{OWOW} values of 3.7 (for $n = 1$), 3.6 ($n = 2$) and 3.8 ($n = 3$) in the acid solutions, with typical errors of ±0.3 over the same range. Integrating out to the minima in $g_{OWOW}(r)$ at $r_{int} = 4$ Å the coordination number n_{OWOW} increases to approximately 6.4 water molecules, with 4.4 ± 0.5 bonds/molecule if the constraint of an intermediate hydrogen bond is imposed, i.e. on average there are approximately two non-bonded water molecules within the first coordination shell. In addition, the intensity of the first (approx. 2.7 Å) and second (at approx. 5.2 Å) shells in $g_{OWOW}(r)$ for the acid solutions increases with longer chain lengths, which can be associated with an increase in the local ordering of water molecules. Overall this indicates the presence of densely packed water channels containing a mixture of bent hydrogen bonds and non-bonded molecular interactions.

(b) Solvation shells

By contrast, the hydrophobic-water partial PDF $g_{F-OW}(r)$ in the solutions shows the first peak around 3.0 Å and the corresponding n_{FO-W} coordination numbers decrease slightly from 1.8 ($n = 1$), 1.8 ($n = 2$) to 1.5 ($n = 3$, integrating out to $r_{int} = 3.5$ Å). That is to say, each F atom on the chain is surrounded by approximately two water molecules on average, figure 6*b*. The first
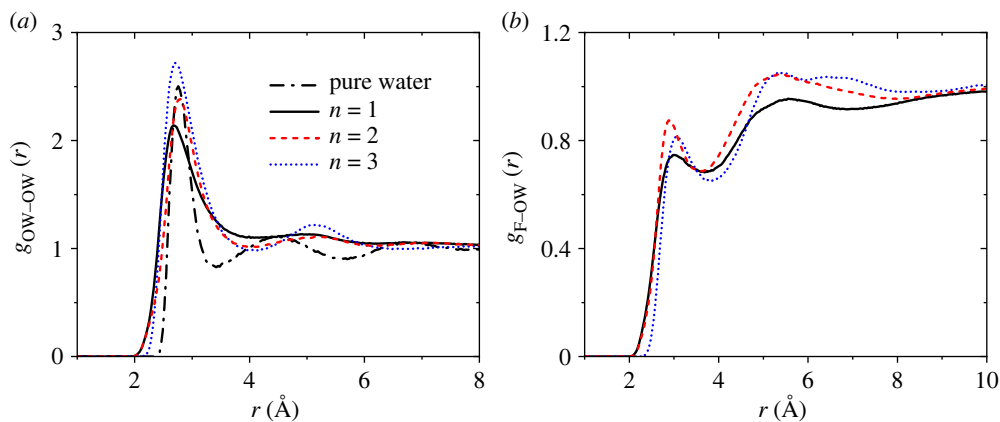
Figure 6. The EPSR partial PDFs for (*a*) water oxygen–water oxygen O_W–O_W and (*b*) the acid chain fluorine–oxygen water F–O_W.
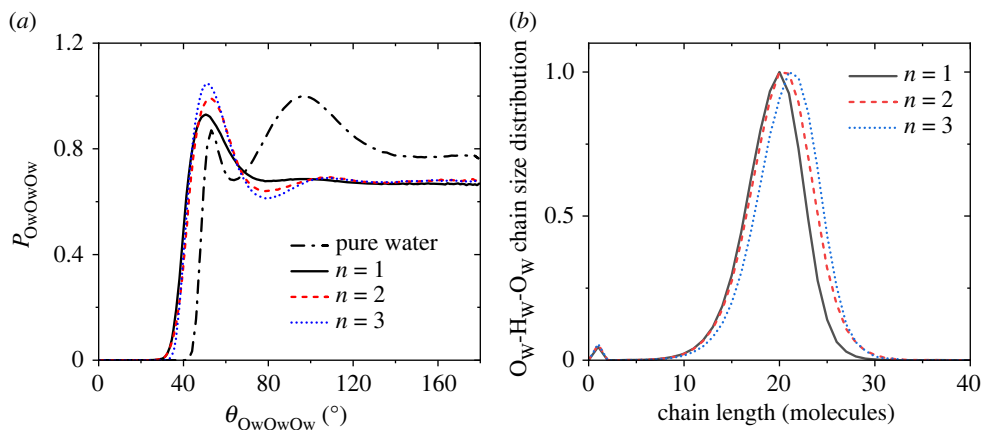

Figure 7. (*a*) The normalized -O_W-O_W-O_W- bond-angle distribution for water molecules in the $n = 1$, 2 and 3 acid solutions compared with pure water. (*b*) The -O_W-H_W-O_W- chain size distributions from the EPSR models in the water 10 : 1 acid solutions for $n = 1$, 2 and 3. Obtained using cut-offs of $r_{O-H} = 2.4$ Å and $r_{O-O} = 4.0$ Å.

and second (5.4 Å) nearest neighbour F-O_W peaks in $g_{F-OW}(r)$ increase in intensity from $n = 1$ to $n = 2$. For the longest chain length ($n = 3$ solution), there is a slightly stronger repulsion as the first peak distance increases by approximately 0.1 Å, and there is a loss of intensity in the 4.3 Å region between the first and second nearest neighbour shells.

The nanoscale heterogeneity leads to the question of how much the structure of hydrogen-bonded water is affected by the presence of short-chain perfluoroalkyl carboxylic acids? The -O_W-O_W-O_W- bond-angle distributions for water molecules in the $n = 1$, 2 and 3 acid solutions in figure 7a show the water molecule arrangement to be notably different from the tetrahedral-like distribution found in pure water. For pure water at room temperature, the -O_W-O_W-O_W- bond-angle distribution has a main tetrahedral peak at 96° (close to the perfectly tetrahedral peak at 109.5°) and a smaller non-bonded interstitial peak at 53° [31]. By contrast, the -O_W-O_W-O_W- bond-angle distributions for acid solutions have the main peak around 52° and only a very small broad peak persists at approximately 108°. These bond-angle distributions sharpen with increasing chain length. The presence of short-chain perfluoroalkyl carboxylic acids therefore

causes a major perturbation to the water structure as solvation shells are formed by chains of water molecules with bent hydrogen bonds due to the restricted geometrical arrangements.

To investigate the distorted hydrogen-bonded water arrangement further, we calculated the -O_W-H_W-O_W- chain size distributions from the EPSR models in the water 10 : 1 acid solutions for $n = 1, 2$ and 3 (using cut-offs of $r_{O-H} = 2.4$ Å and $r_{O-O} = 4.0$ Å). The chain lengths were calculated using the shortest path criterion, i.e. the path with the least number of linkages between two molecules. All acid solutions show a primary chain size distribution centred on approximately 20 molecules that shifts slightly to larger chain lengths for longer acid chain molecules, figure 7b. Only few (less than 1%) water molecules were found to be completely isolated (chain length = 1).

6. Conclusion

Small-and-wide angle X-ray scattering measurements on short-chain perfluoroalkyl carboxylic acids and solutions have been interpreted with EPSR modelling, which revealed distinct long-range interactions between both non-bonded chains of acid molecules and distorted hydrogen-bonded solvation shells formed by water molecules. This water structure is very different than the usually preferred tetrahedral-like network topology. The results reveal a rich variety of structures on varying length scales. Most notably, in tandem with the regularity between non-bonded acid molecules in the liquid, nanoscale heterogeneity was observed to occur in all three model solutions studied, and the nanoscale aggregation of molecules increased with increasing acid chain length. Within the solvation shells, carboxyl and hydroxyl groups on the short-chain perfluoroalkyl carboxylic acids preferentially hydrogen-bonded to adjacent dense water-rich regions. The presence of long-range structural ordering and aggregation effects in these molecular liquid mixtures is most likely highly composition- and size-dependent, so further study is necessary to better understand these relations in detail. It is anticipated that these preliminary results will aid in the benchmarking and development of realistic interatomic potentials, so that accurate computer simulations can be performed at more dilute concentrations relevant to current environmental challenges.

References

1. Melzer D, Rice N, Depledge MH, Henley WE, Galloway TS. 2010 Association between serum perfluorooctanoic acid (PFOA) and thyroid disease in the U.S. national health and nutrition examination survey. *Environ. Health Perspect.* **118**, 686–692. (doi:10.1289/ehp.0901584)
2. Shankar A, Xiao J, Ducatman A. 2011 Perfluoroalkyl chemicals and chronic kidney disease in US adults. *Am. J. Epidemiol.* **174**, 893–900. (doi:10.1093/aje/kwr171)
3. Bassler J, Ducatman A, Elliott M, Wen S, Wahlang B, Barnett J, Cave MC. 2019 Environmental perfluoroalkyl acid exposures are associated with liver disease characterized by apoptosis and altered serum adipocytokines. *Environ. Pollut.* **247**, 1055–1063. (doi:10.1016/j.envpol.2019.01.064)

4. Graber JM *et al.* 2019 Per and polyfluoroalkyl substances (PFAS) blood levels after contamination of a community water supply and comparison with 2013–2014 NHANES. *J. Expo. Sci. Environ. Epidemiol.* **29**, 172–182. (doi:10.1038/s41370-018-0096-z)

5. Barry V, Winquist A, Steenland K. 2013 Perfluorooctanoic acid (PFOA) exposures and incident cancers among adults living near a chemical plant. *Environ. Health Perspect.* **121**, 1313–1318. (doi:10.1289/ehp.1306615)

6. U.S Environmental Protection Agency. Drinking Water Health Advisories for PFOA and PFAS. https://www.epa.gov/ground-water-and-drinking-water/drinking-water-health-advisories-pfoa-and-pfos.

7. Bell EM *et al.* 2021 Exposure, health effects, sensing, and remediation of the emerging PFAS contaminants – Scientific challenges and potential research directions. *Sci. Total Environ.* **780**, 146399. (doi:10.1016/j.scitotenv.2021.146399)

8. Gagliano E, Sgroi M, Falciglia PP, Vagliasindi FGA, Roccaro P. 2020 Removal of poly- and perfluoroalkyl substances (PFAS) from water by adsorption: role of PFAS chain length, effect of organic matter and challenges in adsorbent regeneration. *Water Res.* **171**, 115381. (doi:10.1016/j.watres.2019.115381)

9. Wang Y, Darling SB, Chen J. 2021 Selectivity of per- and polyfluoroalkyl substance sensors and sorbents in water. *ACS Appl. Mater. Interfaces* **13**, 60 789–60 814. (doi:10.1021/acsami.1c16517)

10. Williams AJ *et al.* 2017 The CompTox Chemistry Dashboard: a community data resource for environmental chemistry. *J. Cheminformatics* **9**, 61. (doi:10.1186/s13321-017-0247-6)

11. McLain SE, Benmore CJ, Siewenie JE, Urquidi J, Turner JFC. 2004 On the structure of liquid hydrogen fluoride. *Angewandte Chemie* **43**, 1952–1955. (doi:10.1002/anie.200353289)

12. McLain SE, Benmore CJ, Turner JFC. 2002 The structure of liquid fluorosulfuric acid investigated by neutron diffraction. *J. Chem. Phys.* **117**, 3816–3821. (doi:10.1063/1.1495395)

13. Yoshida K, Yamaguchi T, Adachi T, Otomo T, Matsuo D, Takamuku T, Nishi N. 2003 Structure and dynamics of hexafluoroisopropanol-water mixtures by x-ray diffraction, small-angle neutron scattering, NMR spectroscopy, and mass spectrometry. *J. Chem. Phys.* **119**, 6132–6142. (doi:10.1063/1.1602070)

14. Yoshida K, Yamaguchi T, Bowron DT, Finney JL. 2021 The structure of aqueous solutions of hexafluoro-iso-propanol studied by neutron diffraction with hydrogen/deuterium isotope substitution and empirical potential structure refinement modeling. *Phys. Chem. Chem. Phys.* **23**, 13 561–13 573. (doi:10.1039/D1CP00950H)

15. Yamaguchi T, Imura S, Kai T, Yoshida K. 2013 Structure of hexafluoroisopropanol–water mixtures by molecular dynamics simulations. *Z. Naturforsch.* **68**, 145–151. (doi:10.5560/zna.2012-0100)

16. Yoshida K, Yamaguchi T, Bowron DT, Finney JL. 2004 Cluster structure in helix-promoting hexafluoro-iso-propanol–water mixtures. *J. Neutron Res.* **12**, 305–309. (doi:10.1080/10238160412331299500)

17. Takamuku T, Kumai T, Yoshida K, Otomo T, Yamaguchi T. 2005 Structure and dynamics of halogenoethanol−water mixtures studied by large-angle X-ray scattering, small-angle neutron scattering, and NMR relaxation. *J. Phys. Chem. A* **109**, 7667–7676. (doi:10.1021/jp051470o)

18. Tang H, Cai J, Zhu CY, Chen GJ, Wang XH, Sun CY. 2022 Review on the clustering behavior in aqueous solutions. *J. Mol. Liq.* **367**, 120382. (doi:10.1016/j.molliq.2022.120382)

19. Dougan L, Bates SP, Hargreaves R, Fox JP, Crain J, Finney JL, Réat V, Soper AK. 2004 Methanol-water solutions: a bi-percolating liquid mixture. *J. Chem. Phys.* **121**, 6456–6462. (doi:10.1063/1.1789951)

20. Radola B, Picaud S, Vardanega D, Jedlovszky P. 2015 Molecular dynamics simulations of the interaction between water molecules and aggregates of acetic or propionic acid molecules. *J. Phys. Chem. B* **119**, 15 662–15 674. (doi:10.1021/acs.jpcb.5b08110)

21. Benmore C, Gallington L, Soignard E. 2019 Intermediate range order in supercooled water. *Mol. Phys.* **117**, 2470–2476. (doi:10.1080/00268976.2019.1567853)

22. Skinner LB, Benmore CJ, Parise JB. 2012 Area detector corrections for high quality synchrotron X-ray structure factor measurements. *Nucl. Instrum. Methods Phys. Res. Sect. A* **662**, 61–70. (doi:10.1016/j.nima.2011.09.031)

23. Benmore CJ *et al.* 2020 Extended range X-ray pair distribution functions. *Nuclear Instruments & Methods in Physics Research Section A-accelerators Spectrometers Detectors and Associated Equipment* **955**, 163318. (doi:10.1016/j.nima.2019.163318)

24. Hammersley AP, Svensson SO, Hanfland M, Fitch AN, Hausermann D. 1996 Two-dimensional detector software: from real detector to idealised image or two-theta scan. *High Pressure Res.* **14**, 235–248. (doi:10.1080/08957959608201408)

25. Qiu X, Thompson JW, Billinge SJL. 2004 PDFgetX2: a GUI-driven program to obtain the pair distribution function from X-ray powder diffraction data. *J. Appl. Crystallogr.* **37**, 678. (doi:10.1107/S0021889804011744)

26. Soper AK. 1996 Empirical potential Monte Carlo simulation of fluid structure. *Chem. Phys.* **202**, 295–306. (doi:10.1016/0301-0104(95)00357-6)

27. Misawa M, Dairoku I, Honma A, Yamada Y, Sato T, Maruyama K, Mori K, Suzuki S, Otomo T. 2004 Mass-fractal clustering and power-law decay of cluster size in 1-propanol aqueous solution. *J. Chem. Phys.* **121**, 4716–4723. (doi:10.1063/1.1780931)

28. Takamuku T, Saisho K, Nozawa S, Yamaguchi T. 2005 X-ray diffraction studies on methanol–water, ethanol–water, and 2-propanol–water mixtures at low temperatures. *J. Mol. Liq.* **119**, 133–146. (doi:10.1016/j.molliq.2004.10.020)

29. Li F, Duan J, Tian S, Ji H, Zhu Y, Wei Z, Zhao D. 2020 Short-chain per- and polyfluoroalkyl substances in aquatic systems: occurrence, impacts and treatment. *Chem. Eng. J.* **380**, 122506. (doi:10.1016/j.cej.2019.122506)

30. Skinner LB, Huang C, Schlesinger D, Pettersson LGM, Nilsson A, Benmore CJ. 2013 Benchmark oxygen-oxygen pair-distribution function of ambient water from x-ray diffraction measurements with a wide Q-range. *J. Chem. Phys.* **138**, 074506. (doi:10.1063/1.4790861)

31. Shi C, Alderman OLG, Benmore CJ. 2019 Network topology of deeply supercooled water. *Mol. Phys.* **117**, 3217–3226. (doi:10.1080/00268976.2019.1649492)

32. Benmore CJ, Wang Y, Darling SB, Chen J. 2023 Molecular interactions in short-chain perfluoroalkyl carboxylic acids and aqueous solutions. Figshare. (doi:10.6084/m9.figshare.c.6742537)