Supporting Information to:

On the nanoscopic structural heterogeneity of liquid n-alkyl carboxylic acids

Alessandro Mariani,[a,b] Alessandro Innocenti, [a,b] Alberto Varzi, [a,b] and Stefano Passerini* [a,b]

Experimental methods:

All the acids were purchased at Sigma Aldrich and used as received (Methanoic acid >99%; Ethanoic acid >99.98% anhydrous; Propanoic acid >98%; Butanoic acid >99%, Pentanoic acid >99.5%, Hexanoic acid >98%).

The SWAXS patterns were collected using a Xeuss 3.0c (Xenocs – Grenoble, France) equipped with an Eiger2 1M detector (SAXS) and an Eiger2 500k detector (WAXS). The Sample-to-(SAXS) Detector distance was set to 600 mm. A Mo K_α source was exploited with a beam size of 0.7x0.7 mm^2, obtaining a flux of ~10^6 photons/s. The pure acids were injected in 1.5mm o.d. borosilicate capillaries and sealed with solventless hot glue. The sample chamber of the instrument was kept under vacuum (P=50 μbar) during the experiment. Each sample was exposed for 10 minutes to X-rays to ensure a good signal-to-noise ratio. The empty capillary background was also collected and subtracted from the total scattering curve. Data treatment was performed with the Xsact software from Xenocs, PDFgetX3, and some in-house MATLAB code.

Computational methods:

The Molecular dynamics simulations were carried out using Amber 18 software[1] exploiting the GAFF force field[2]. The atomic partial charges were obtained with the RESP algorithm from DFT calculations run with Gaussian09e[3] at the B3LYP/6-311++G** level of theory (see Table S1). The starting random molecular arrangements were obtained by Packmol[4] in boxes with sides of ~ 60 Å. The number of molecules for each system is reported in Table S2. The simulation went through different steps, starting from a geometrical relaxation, followed by gradual heating of the system from 0 to 50K in several NVT sessions. The systems were then equilibrated at 300K for 20ns in NPT ensemble and a further 20ns in NVT. A final productive NVT phase of 10ns was then used for the analysis. For the production phase the timestep used was 2fs thanks to the SHAKE algorithm[5], and the simulation was dumped every 1000 steps, obtaining a final trajectory of 5000 frames spaced by 2ps each.

Periodic boundary conditions (PBC) along all axis were included to avoid spurious border effects. The temperature was kept constant at 300 K using a Langevin thermostat with a coupling constant of $1ps^{-1}$. In NPT steps, the pressure was kept constant at 1 atm with a Berendsen barostat. Cutoffs of 8 Å and 12 Å were set for VdW and Coulombic interactions, respectively. The Ewald summation method was used to account for the PBC. The simulations goodness was checked by comparing the theoretical and experimental densities (Table S3) and structure factors. The analysis (structure function calculation, domains analysis, and radial distribution functions calculation) were performed exploiting TRAVIS with methodologies described elsewhere[6–8]. For each system, three different starting geometries were considered. The results were indistinguishable for each repetition, with errors in the order of <0.5% on density values, S(q) intensities, and RDF intensities. Errors of <0.05% were observed for S(q) and RDF peaks positions.

Table S1: RESP charges used in the simulations. O is the carbonylic Oxygen, O_H is the hydroxylic Oxygen, C_O is the carboxylic Carbon, C_T is the terminal methyl Carbon, $C_{\alpha\text{-}\delta}$ are the Carbon atoms in α-δ position to the carboxylic group. Hydrogen atoms subscripts are the same as the subscripts of the heteroatom to which they are bound. Charges are expressed in electron charges.

Atom	C_1OOH	C_2OOH	C_3OOH	C_4OOH	C_5OOH	C_6OOH
O	-0.464824	-0.523367	-0.564772	-0.524861	-0.513481	-0.507125
O_H	-0.535656	-0.611649	-0.670935	-0.602354	-0.587285	-0.595118
C_O	0.615001	0.802430	0.715480	0.764172	0.714453	0.704504
C_T		-0.499043	-0.296093	-0.193787	-0.246045	-0.137250
C_α	--		0.053531	0.135633	-0.315592	-0.209955
C_β	--	--	--	0.272014	0.105056	0.035634
C_γ	--	--	--	--	0.055699	0.004101
C_δ	--	--	--	--	--	0.203140
H_H	0.396473	0.408896	0.448163	0.411111	0.403760	0.401690
H_T	-0.010994	0.140911	0.082513	0.052523	0.067008	0.030197
H_α	--		0.033543	0.135633	0.100623	0.059477
H_β	--	--	--	-0.043278	-0.019822	0.002852
H_γ	--	--	--	--	0.010405	-0.011154
H_δ	--	--	--	--	--	-0.046283

Table S2: Number of molecules in the simulation boxes

System	# of molecules
C_1OOH	3400
C_2OOH	2240
C_3OOH	1710
C_4OOH	1390
C_5OOH	1170
C_6OOH	1030

Table S3: Comparison between experimental and theoretical densities

	C_1OOH	C_2OOH	C_3OOH	C_4OOH	C_5OOH	C_6OOH
Exp. Density [g/ml]	1.22	1.05	0.99	0.96	0.93	0.93
Sim. Density [g/ml]	1.234	1.068	1.011	0.970	0.941	0.929

Table S4: Fitting parameters for the pre-peaks position

Equation	a	b	c
$y = a + b*x$	1.1580	-0.1110	--
$y = a + b * e^{-c * x}$	-0.0932	1.3790	0.1387

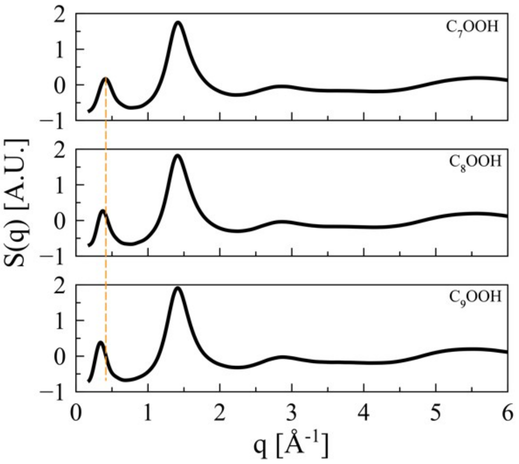

Figure S1: Simulated scattering pattern curve of (top) Heptanoic Acid, (middle) Octanoic Acid, (bottom) Nonanoic Acid. The dashed line is a guide to the eye, highlighting the pre-peak shift.

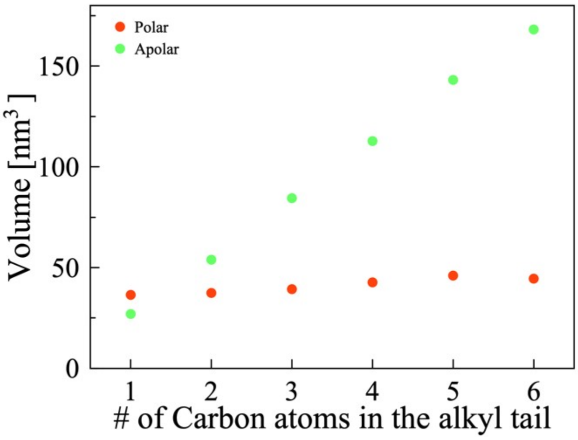

Figure S2: Polar (red) and apolar (green) domains volume as a function of alkyl tail length. The domain volumes were obtained via Voronoi tessellation as implemented in TRAVIS.

References:

1 D. a Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang and R. J. Woods, The Amber biomolecular simulation programs, *J. Comput. Chem.*, 2005, **26**, 1668–1688.

2 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, Development and testing of a general amber force field, *J. Comput. Chem.*, 2004, **25**, 1157–1174.

3 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ort and D. J. Fox, 2009.

4 L. Martínez, R. Andrade, E. G. Birgin and J. M. Martínez, PACKMOL: A package for building initial configurations for molecular dynamics simulations, *J. Comput. Chem.*, 2009, **30**, 2157–2164.

5 J.-P. Ryckaert, G. Ciccotti and H. J. . Berendsen, Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, *J. Comput. Phys.*, 1977, **23**, 327–341.

6 M. Brehm and B. Kirchner, TRAVIS - A Free Analyzer and Visualizer for Monte Carlo and Molecular Dynamics Trajectories, *J. Chem. Inf. Model.*, 2011, **51**, 2007–2023.

7 O. Hollóczki, M. Macchiagodena, H. Weber, M. Thomas, M. Brehm, A. Stark, O. Russina, A.

Triolo and B. Kirchner, Triphilic Ionic-Liquid Mixtures: Fluorinated and Non-fluorinated Aprotic Ionic-Liquid Mixtures, *ChemPhysChem*, 2015, **16**, 3325–3333.

8 M. Brehm, M. Thomas, S. Gehrke and B. Kirchner, TRAVIS—A free analyzer for trajectories from molecular simulation, *J. Chem. Phys.*, 2020, **152**, 164105.