Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Supplementary Information (SI)

Effect of Surface-Active Organic Matter on Carbon Dioxide Nucleation in Atmospheric Wet Aerosols: A Molecular Dynamics Study

Vangelis Daskalakis[1]*, Fevronia Charalambous[1], Fostira Panagiotou[1] and Irene Nearchou[1]

[1]Cyprus University of Technology, Department of Environmental Science and Technology, P.O. Box 50329, 3603 Limassol – Cyprus.

Tel: +357 25002458, Fax. +357 25002820. E-mail: evangelos.daskalakis@cut.ac.cy

*Corresponding Author: Vangelis Daskalakis

Contents

Keywords: carbon dioxide, nucleation, clouds, aerosols, rain acidity.

1. Equilibration – Relaxation Protocols for the samples

Prior to the production runs, all the samples were equilibrated by superheating [1] in two gradual increases in the temperature at constant density ensembles, till evaporation; from 10K to 250K (1.0ns transition) and equilibration at 250K for 1.0ns and from 250K to 600K over 11.0ns, including 0.5ns transition steps and 1.0ns equilibration periods between 250, 260, 270, 275, 298, 350 and 600K, followed by a 2.0ns final equilibration at 600K. Four structures per sample were generated from the 600K-equilibration trajectory that were cooled down and condensed at constant pressure ensembles to 298K, 275 or 250K over a period of 2.0ns and then equilibrated at the respective temperature for another 2.0ns (NpT). The chosen final temperatures correlate with the transition of a cloud drop from the supercooled water environment (250K) found in the cloud tops, towards the lower part of the clouds (275K) or below that and finally the ground (298K). Superheating dynamics resulted in the complete melt down of the ice and effectively mixed the compounds in the gas phase (600K) to destroy any biased CO_2 nucleus formation in the initial setup of the unit cells. This latter rather large temperature is the lowest one to completely mix the aerosol compounds at the short time scale used for equilibration.

2. Correlation between Radial Distribution Functions (RDF) and CO_2 Solubility

The last two time frames of numerous (101) 175ns production trajectories from all samples, where a successful CO_2 nucleation occurred (supersaturation), were used for the correlation. We visually enumerated the dissolved CO_2 molecules in each frame to derive the solubility as %w/w or the molar solubility and we also calculated the CO_2-RDF profile over these last two frames per trajectory chosen. The last two frames contained exactly the same amount of dissolved CO_2 per trajectory. The linear

fitting of the points gives the CO_2 solubility S versus the CO_2-RDF intensity (RDF) at 3.90Å, as S=11.02-0.95RDF. Each point in the correlation is calculated based only on the last two time frames of the 175ns MD trajectories. Despite the dispersive nature of this correlation, we note the negative slope of the linearly fitted points close to -1.0. The latter is an indication that CO_2-RDF intensity at 3.90Å can be used directly to estimate CO_2 solubility for the system setup used.

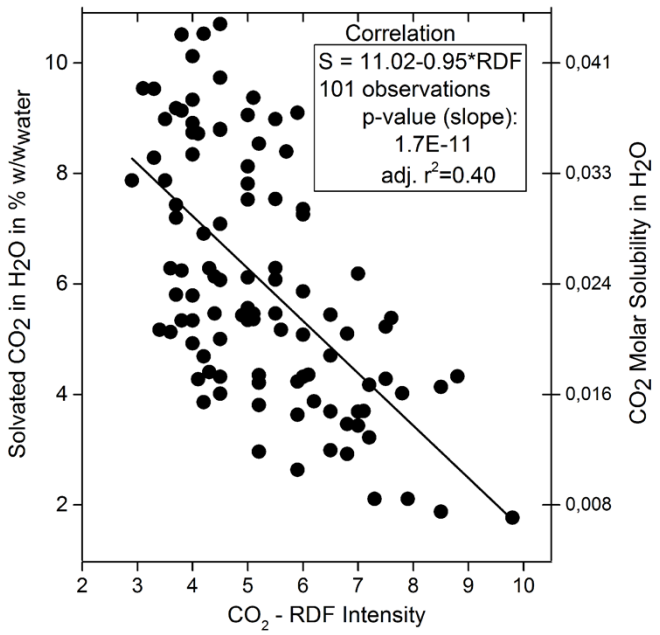

Figure 1S. Correlation diagram between CO_2-RDF intensity of the peak at 3.90Å and % w/w CO_2 solubility (left y axis) or the molar solubility (right y axis) calculated as the number of solvated CO_2 molecules divided by the number of water molecules.

In **Figure 3A** (*upper*) of the main manuscript intense CO_2-RDF peaks are indicative of a stable CO_2 nucleus throughout the time window probed (75-175ns), while lower intensities for the 3.90Å peak indicate no nucleation, or less CO_2 molecules in the nucleus, or even instability (solvation of the nucleus over time). These latter factors are not taken into account for the correlation diagram in **Figure 1S** as only two frames from each selected trajectory with a CO_2 nucleus were probed. This should be responsible for its dispersive nature. Due to the large number of frames at each trajectory, as well as the

number of the 175ns production trajectories (around 500) it is difficult to visually enumerate the solvated CO_2 molecules in each frame between 75-175ns and in all trajectories to directly associate CO_2-RDF intensities with the CO_2 solubility. Thus a limited, but adequate, number of frames were chosen (101) for the correlation. The equation of the linear fit can be a useful conversion scheme between long time and sample averaged RDF-Intensities and S. For the conversion, only points above 3.0 units are used from **Figure 3A** (*upper*) in the main manuscript. This latter refers only to points that correspond to samples where a CO_2 nucleus is formed, and thus supersaturation for the dissolved CO_2 is achieved. The solubility values in **Figure 3B** in the main manuscript are directly derived by the above aforementioned methodology and are solely based on the MD simulations in this study, without any additional fitting to experimental or theoretical values in the literature, nor other assumptions.

3. Evaluation of the simulations

MC sample supersaturation water/ CO_2 ratio of 9.3 is just in the threshold for CO_2 nucleation in neat water at 298K and in the presence of surface-active OM at 250K, based on the results obtained with the NE6 [2-3] water potential and the orthorhombic unit cells employed. CO_2 solubility in neat water is high (250K), whereas decrease in CO_2 solubility in the presence of surface-activity is highly exerted at this latter temperature, especially for methylglyoxal. The neat water and methylglyoxal contaminated MC samples can, thus, both be used as a reference to compare between experimental and theoretical results of carbon dioxide nucleation in neat water at 298K or the dynamics at 250K in the presence of surface-active OM with diverse parameters for the water potential employed, the unit cell geometry, or the cut-off distance used.

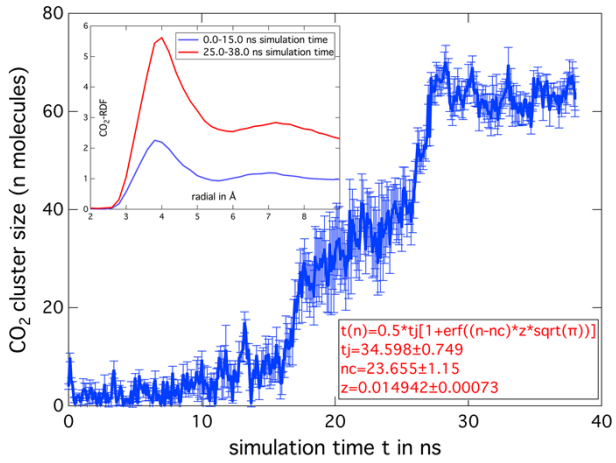

Figure 2S. The kinetics of CO_2 nucleation in neat water at 298K. The fitted parameters, like the nucleation time t_J, the critical cluster size nc and the Zeldovich factor z are also shown. *Inset*: The CO_2-Radial Distribution Function (RDF) averaged over 0.0-15.0ns (red line) and 25.0-38.0ns (blue line).

Nucleation rates and critical cluster size – The most popular and widely used cluster definition model to identify occurring clusters (nuclei) in a MD simulation is the Stillinger criterion [4-5]. Unfortunately, the Stillinger definition tends to strongly overestimate the size of small nuclei, even by a factor of 2 [6] and contains no dynamical information, which has been proposed to be essential in identifying short or long-lived clusters in nucleation studies [7]. In the Stillinger method two molecules are considered to belong to the same cluster once their centres of mass are less than a pre-defined distance (r_s) apart. Passing-by molecules could happen to be within the r_s distance and thus they are counted as belonging to the cluster. In addition, using the Stillinger criterion in highly dense phases leads to serious overcounting of the cluster molecules. An improved implementation of the latter criterion has been developed by ten Wolde and Frenkel [8], using the same r_s definition as Stillinger, but adopting an additional criterion for nucleation quantification; the distribution of number of neighbours per clustering particle for different species (e.g. liquid, gas phase particles). This definition, though, could potentially underestimate the size of a cluster by neglecting the surface particles particularly on small nuclei. Both criteria are therefore not suitable for the purpose of this study, where the solubility of CO_2 is measured in terms of the amount of

molecules within an occurring nucleus in the aerosol bulk aquatic phase. In addition, these criteria are not sensitive to the shape of an occurring CO_2 nucleus. The presence of OM around an occurring CO_2 nucleus can have a significant effect on the calculated number of CO_2 molecules within, as they tend to supress the nucleus size and shape, rendering it denser and smaller. The use of Radial Distribution Functions proves more appropriate for this study, containing also the necessary dynamical information [9], as well as information on the shape and size of an occurring CO_2 nucleus in the liquid water phase [10]. Nevertheless, the adapted ten Wolde and Frenkel definition can be used to accurately calculate the CO_2 nucleation rate J in neat water, as well as the critical cluster size nc to be compared with available experimental observations in the literature. At this critical cluster size nc, a CO_2 nucleus grows spontaneously to form the transition state, past which the embryonic nucleus serves as a seed for the new phase transition and grows uncontrollably towards an equilibrium final size. We have to pinpoint though that the majority of the experimental studies in the literature are based on calculations of nucleation rates for mature CO_2 bubbles occurring in the liquid water phase of carbonated beverages [11-12] and not the considerably smaller bubble components of greater density identified in this study. These smaller bubble components, despite being a crucial key step in mature bubble formation through coalescence [10], they would be more difficult to probe experimentally unless they converge to mature larger bubbles.

Sixteen different configurations were captured from the four 175ns NpT production trajectories of the neat water MC samples at 250K (CO_2 enrichment of 9.3% or supersaturation ratio of water/CO_2 at 9.3). All CO_2 molecules within each of these structures were found fully dissolved in the liquid water phase. Each of these configurations at 250K served as a unique starting point and was subjected to an instantaneous temperature jump to 298K at the NpT ensemble. Molecular Dynamics trajectories were run for 38ns per conformation. For each trajectory we employed the ten Wolde and Frenkel method with r_S=3Å and 3.77 average water neighbours per fully dissolved CO_2 molecule to calculate the size n of

the occurring nuclei versus time n(t). The parameters for r_s and the number of neighbouring water molecules were derived by trial and error and by comparing the occurring nuclei sizes both visually and derived by the ten Wolde and Frankel criterion with different parameters. We have to note that a discrepancy did exist between the actual nuclei sizes measured visually and those calculated by the latter criterion, indeed attributed to the CO_2 molecules on the nucleus surface. In the vicinity of the critical cluster size the t(n) is given by the first-passage time method [13], based on the critical cluster size nc, the so-called Zeldovich factor Z and the nucleation time t_J: $t(n) = 0.5t_J\left[1 + erf((n - nc)Z\sqrt{\pi}\right]$. The time t_J is related to the nucleation rate J and the system average volume V by $J = 1/(t_J V)$. The averaged n(t) curve for the sixteen trajectories, along with standard error bars is depicted in **Figure 2S**. By fitting this curve to the first-passage time method equation, we calculate the parameters t_J, nc and Z to be around 34.6, 23.7 and 0.015 respectively. This gives a CO_2 nucleation rate in neat water of 8.6E-4 CO_2 molecules $ns^{-1}nm^{-3}$ or J=2.6E4 $s^{-1}cm^{-1}$ and $\log_{10}J$=4.4, assuming 2.99E-23 cm^3 per CO_2 molecule at 298K [11]. The MC samples correspond to a water/ CO_2 supersaturation ratio of 9.3 and the rate J falls into the experimental values of J with $\log_{10}J$ between 2-6 [11]. The critical cluster size is calculated at around 24 CO_2 molecules for the small bubble components identified in this study at 298K and a supersaturation ratio of 9.3 and for the homogeneous nucleation case probed, compared to the critical size of 89 CO_2 molecules, reported for a different water/ CO_2 supersaturation ratio and in the heterogeneous nucleation [11]. This discrepancy can be attributed to the failure of the improved Stillinger criterion to account for the surface molecules on the occurring CO_2 nucleus, as well as the nature of the smaller nucleus consisting not a mature CO_2 bubble, but a bubble component, as referred to in this study. There is no information in the literature on experimental values of J or nc in the presence of Volatile Organic Compounds (VOC), whereas no conformation was found where CO_2 molecules are fully dissolved in the presence of several organic species (especially those exerting surface-activity) in any temperature probed in this study between 250-298K. The calculation of nucleation rates in the latter cases exceeds

the scope of this study, as a considerably larger range of temperatures below 250K would be needed to search for CO_2 in a fully dissolved conformation in the presence of surface-active OM. Thus, for the comparison with experiments in the literature only the kinetics of CO_2 nucleation in neat water were probed. The CO_2-Radial Distribution Functions averaged over the 0.0-15.0ns and 25.0-38.0ns simulation time windows and over the sixteen trajectories is also shown as inset of **Figure 2S**. We note that the RDF intensity at 3.90Å is below 3.0 units for the 0.0-15.0ns time window where no nucleus is observed, while it exhibits a strong peak at 3.90Å past the time of 25.0ns, where a mature nucleus has been formed in the liquid phase and in correlation with the ten Wolde and Frankel method [8]. Error bars on the RDF curves prove negligible and they are not designated. We note the considerably greater intensity of the CO_2-RDF at 3.90Å for the 25.0-38.0ns time window, compared to the RDF of the longer time window (100-175ns) for the same curve in **Figure 2A** (main manuscript). At shorter windows the occurring CO_2 nucleus shape and size remain unchanged, indicating stability of the conformation over time, hence the higher intensity peak. For the longer time windows both shape and size fluctuate more.

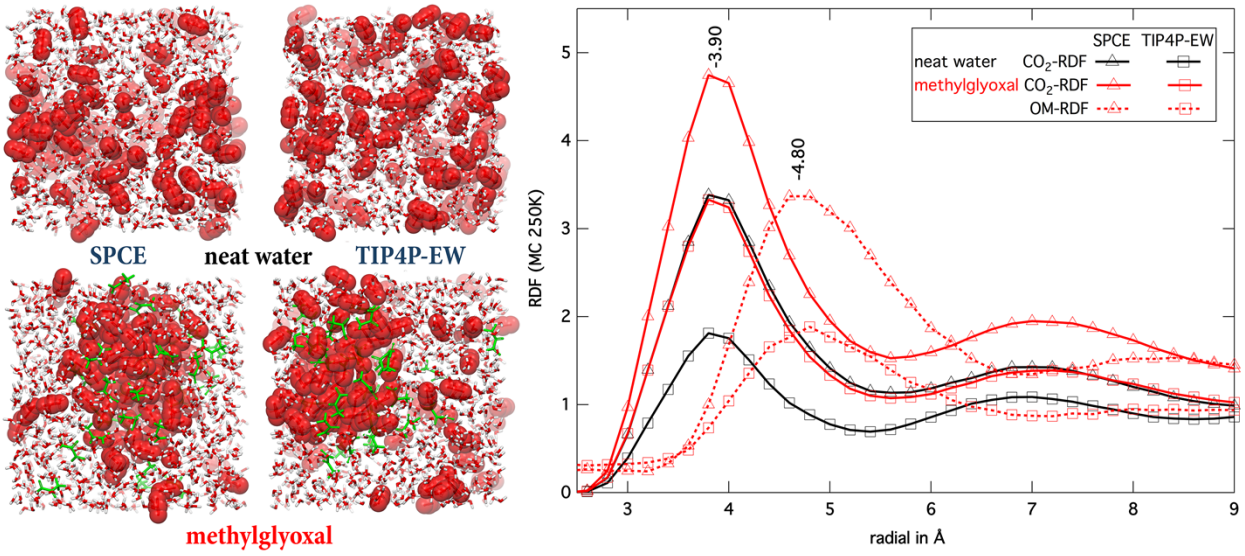

Figure 3S. Snapshots (*left*) at 175ns and CO_2-/ OM-RDF profiles (*right*) between 75-175ns for the CO_2 dynamics in SPCE and TIP4P-EW water in the absence (neat water) and the presence of surface-active OM (methylglyoxal). CO_2 atoms are depicted in red spheres, while organic matter is shown in green. Water is depicted with red oxygens and white hydrogens.

The Choice of Force Field – The same protocols, as for the simulations with NE6, were employed for equilibration and relaxation, as well as for the production runs with the SPCE [14] and the TIP4P-EW [15] water potentials. In the latter two cases, cubic unit cells of the same volume, compared with the orthorhombic scheme adopted with NE6 [2-3], and larger cut-off distances were chosen, truncating electrostatic interactions between 1.20 and 1.08nm. Cubic cell width varied between 31.8-32.8Å at 250K. Selected snapshots at the end of these 175ns production trajectories are shown in **Figure 3S**, along with the associated CO_2-/ OM-RDF profiles, averaged over 75-175ns and over the four production trajectories per sample. The shapes and sizes of the occurring nuclei in the metylglyoxal contaminated samples in **Figure 3S** compare with those with NE6 at orthorhombic unit cell geometry (see **Figure 1** in the main manuscript). Despite this similarity, no direct comparison can be made between the RDF-/ OM-CO_2 intensities for the orthorhombic (NE6) and the cubic (SPCE or TIP4P-EW) runs, as the unit cell shape should also have an effect on the absolute RDF profiles, especially for OM-RDF. In a cubic unit cell geometry, and an occurring smaller nucleus, the surface-active molecules are distributed on the whole surface of the nucleus, but in a laterally narrower, orthorhombic unit cell geometry, or a considerably larger CO_2 nucleus separating the liquid water phase in two, the surface-active OM is mainly found concentrated on the two accessible interfaces between CO_2 and liquid water, giving intense OM-RDF peaks at 4.80Å. Nevertheless, it is evident by both the visual monitoring of each trajectory and the CO_2-RDF profiles in **Figure 3S** that, even at the cubic unit cell runs and the larger cut-off distances employed, CO_2 nucleation is more favourable in the presence of the surface-active methylglyoxal. Nucleation for CO_2 does occur in SPCE or TIP4P-EW water in the presence of methylglyoxal, whereas in the neat water although some small nuclei appear throughout the trajectories, they prove unstable dissolving constantly in the liquid water phase. This results in the lower intensity CO_2-RDF profiles in neat SPCE or TIP4P-EW water in the time window between 75-175ns. We have to pinpoint, though, that a slight decrease in CO_2 solubility is observed especially in SPCE neat water, compared to NE6, as small but

highly unstable CO_2 nuclei occur within the cubic unit cell at random time intervals. However, surface-activity proves again important, even with SPCE and TIP4P-EW, for stabilizing such small nuclei and efficiently reducing the CO_2 solubility. In correlation with the simulations employing NE6, nucleation occurs also for the surface-active methylglyoxal. It is found on the surfaces of the occurring CO_2 nuclei (see the lower two snapshots in **Figure 3S**). An intense peak is exhibited in the OM-RDF profile at 4.80Å for the SPCE dynamics and the considerably larger CO_2 nucleus almost separating the liquid water phase in two (**Figure 3S**). The dispersed nature of methylglyoxal molecules on the surface of the comparably smaller occurring nucleus in TIP4P-EW water gives a lower intensity OM-RDF profile, although OM is found on this surface throughout all trajectories. We conclude that the laterally narrow orthorhombic unit cell adopted for the production simulations (NE6) and possible finite-size effects or the rather smaller cut-off distance play negligible role on the studied dynamics and the main conclusions of this study. The choice of a laterally narrow orthorhombic unit cell for the production runs (NE6) proves satisfactory, in line with a previous study [10] and the correct dynamical picture represented by the RDF profiles for CO_2 and OM for the orthorhombic geometry.

References

1. K. Mochizuki, M. Matsumoto and I. Ohmine, *Nature,* 2013, 498, 350-354.
2. H. Nada and P. J. M. Jan van der Eerden, *J. Chem. Phys.,* 2003, 118, 7401-7413.
3. H. Nada and Y. Furukawa, *J. Cryst. Growth,* 2005, 283, 242-256.
4. F. H. Stillinger, *J. Chem. Phys.* 1963, 38, 1486-1494.
5. Y. C. Chiew, G. Stell, and E. D. Glandt, *J. Chem. Phys.* 1985, 83, 761-767.
6. J. Wedekid and D. Reguera, *J. Chem. Phys.*, 2007, 127, 154516.
7. L. A. Pugnaloni and F. Vericat, *J. Chem. Phys.* 2002, 116, 1097-1108.
8. R. ten Wolde and D. Frenkel, *J. Chem. Phys*. 1998, 109, 9901-9918.
9. B. Senger, P. Schaaf, D. S. Corti, R. Bowles, D. Pointu, J. C. Voegel and H. Reiss, *J. Chem. Phys.,* 1999, 110, 6438-6450.
10. M. Uddin, D. Coombe, *J. Phys. Chem. A*, 2014, 118, 1971-1988.
11. P. M. Wilt, *J. Colloid and Interface Sci*. 1986, 530-538.
12. O. R. Enriquez, C. Sun, D. Lohse, A. Prosperetti and D. van der Meer, *J. Fluid Mech.* 2014, 741, R1.
13. J. Wedekind, R. Strey, and D. Reguera, *J. Chem. Phys.*, 2007, 126, 134103.
14. H. J. C. Berendsen, J. R. Griger, and T. P. Straatsma, *J. Phys. Chem.,* 1987, 91, 6269-6271.
15. W. Horn, W. C. Swope, J. W. Pitera, J. D. Madura, T. J. Dick, G. L. Hura and T. Head-Gordon, *J. Chem. Phys.*, 2004, 120, 9665-9678.