Electronic Supplementary Material (ESI) for Chemical Communications.

Pre-nucleation aggregation based on solvent microheterogeneity

Christopher D. Jones,[a] Martin Walker,[b] Yitian Xiao,[a] and Katharina Edkins*[a]

[a] School of Pharmacy, Queen's University Belfast, 97 Lisburn Road, Belfast BT9 7BL, UK

[b] Department of Chemistry, Durham University, South Road, Durham DH1 3LE, UK

Electronic Supporting Information

Contents

Supporting figures for the main manuscript

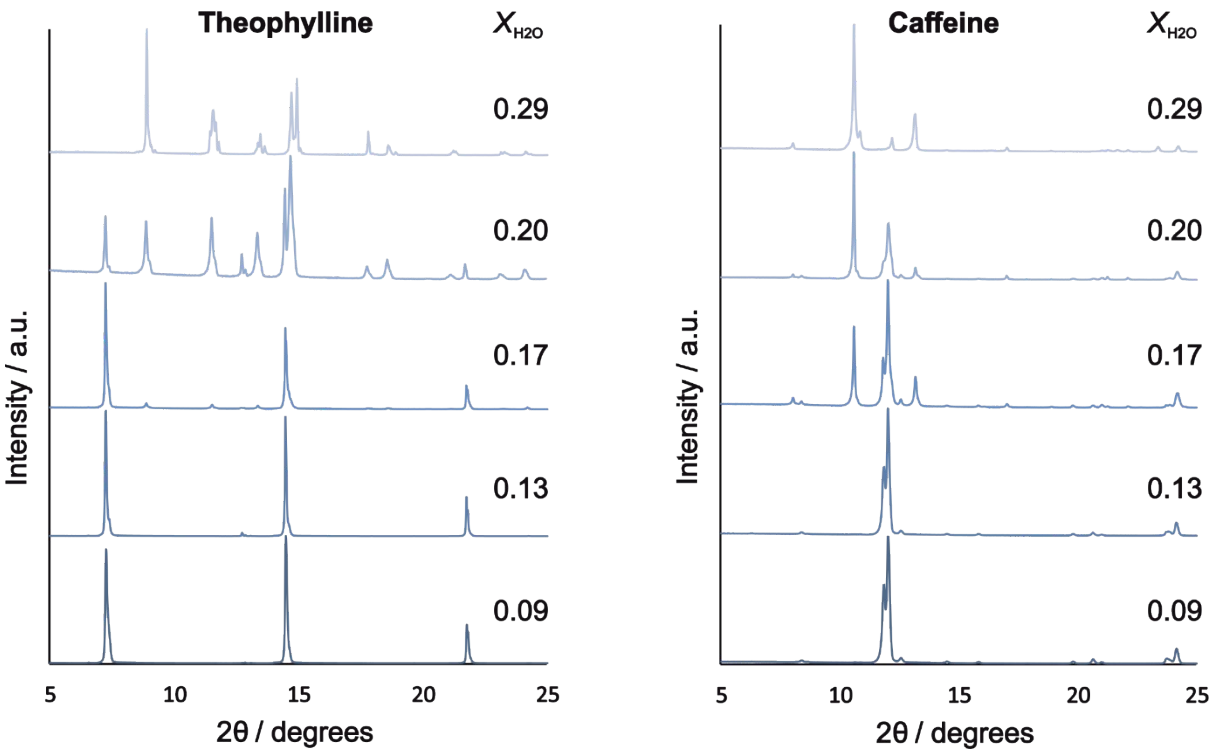

Figure S1 Powder X-ray diffractograms of the crystalline products of slow cooling experiments over one week from aqueous acetonitrile. The top pattern is pure monohydrate of each compound, the bottom pattern is form II for theophylline and form α for caffeine.

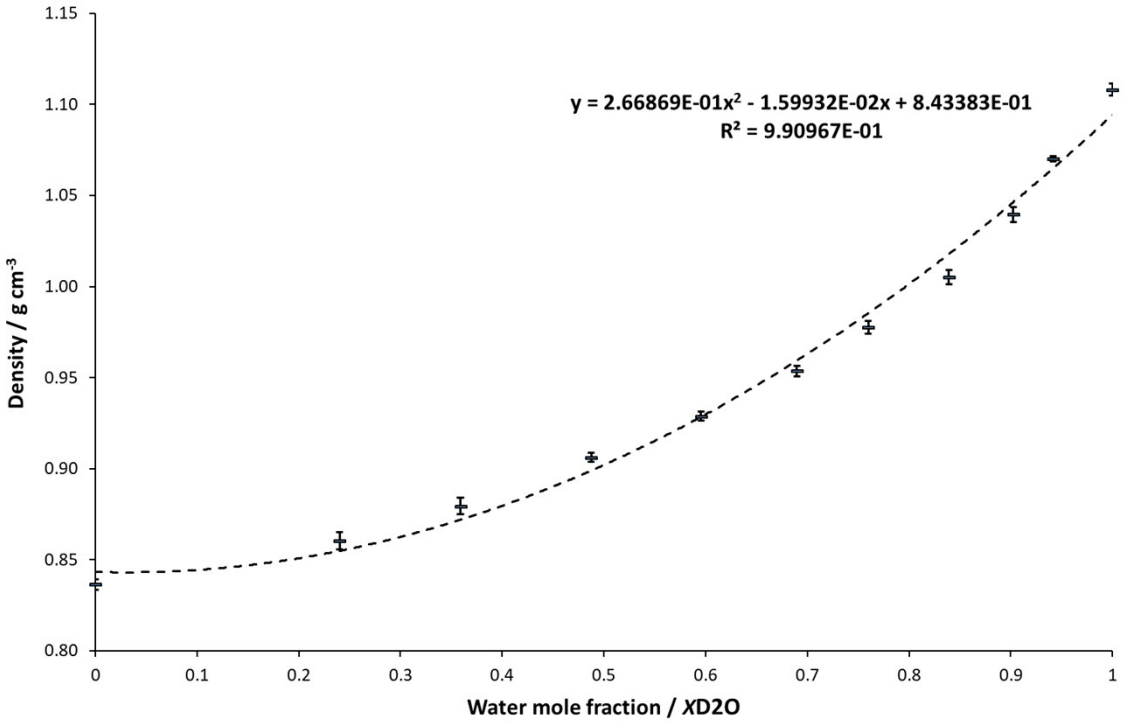

Figure S2 Density measurements of aqueous acetonitrile at varying X_{D2O}.

Molecular Dynamics simulations

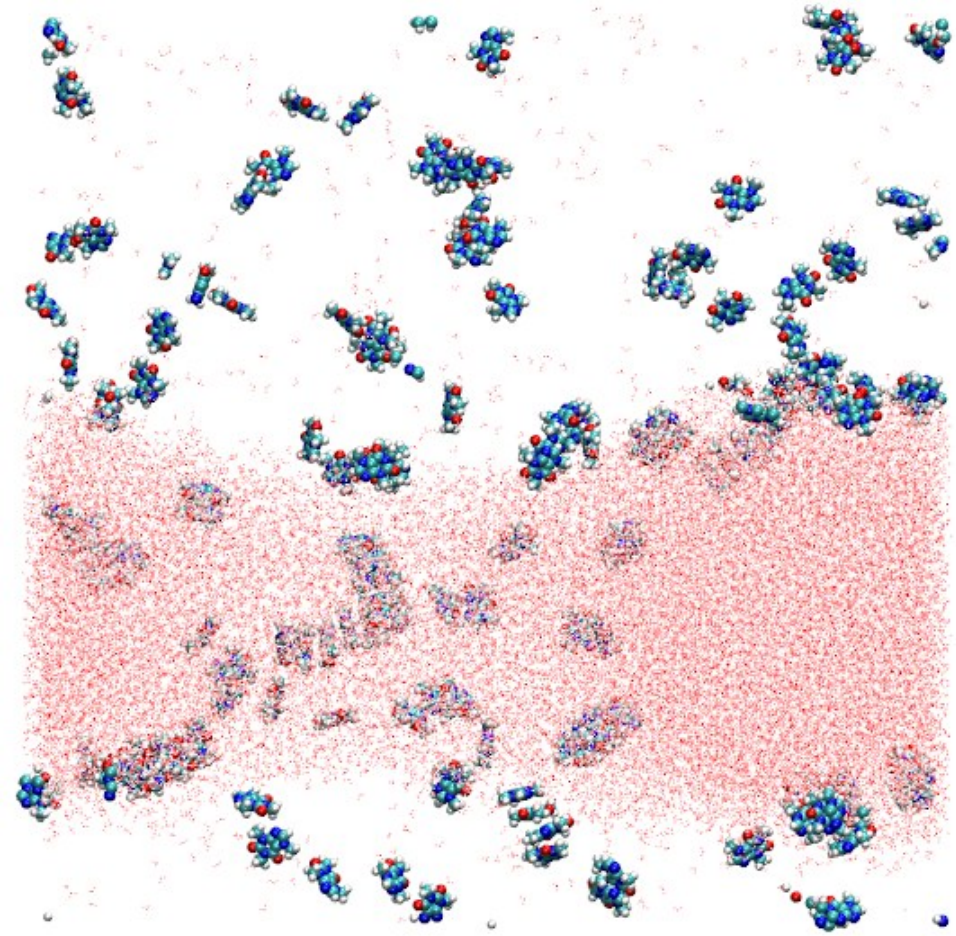

Figure S3 Molecular dynamics model of 0.1 M caffeine at X_{H2O} = 0.5. Water molecules are shown in red wireframe, solute molecules in space-filling representation. Acetonitrile molecules are omitted for clarity.

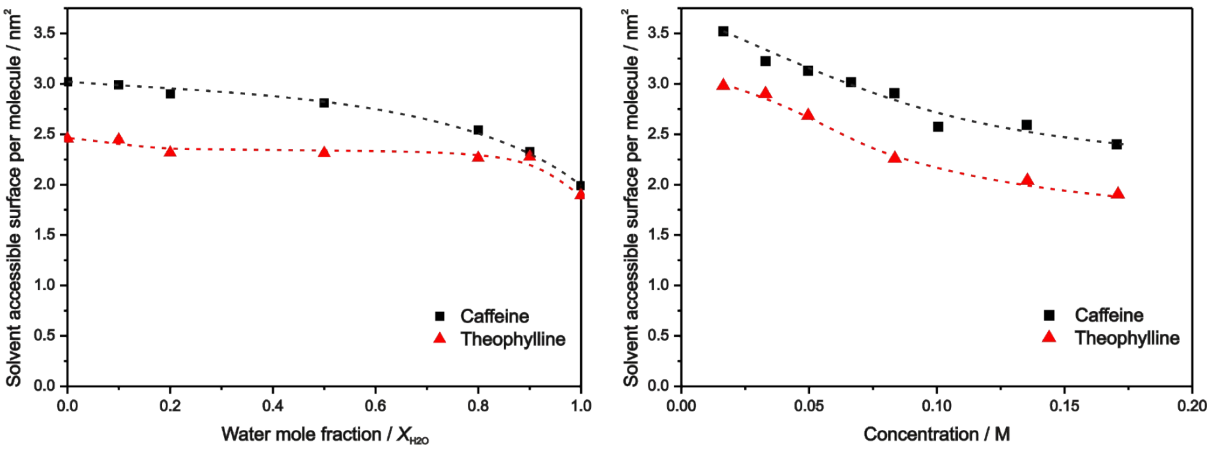

Figure S4 Solvent accessible area per molecule for caffeine (black squares) and theophylline (red triangles) as function of X_{H2O} of a 0.1M solution(left) and as function of solute concentration at X_{H2O} = 0.5 (right). Error bars are smaller than the size of the symbols. Curves are guides to the eye only.

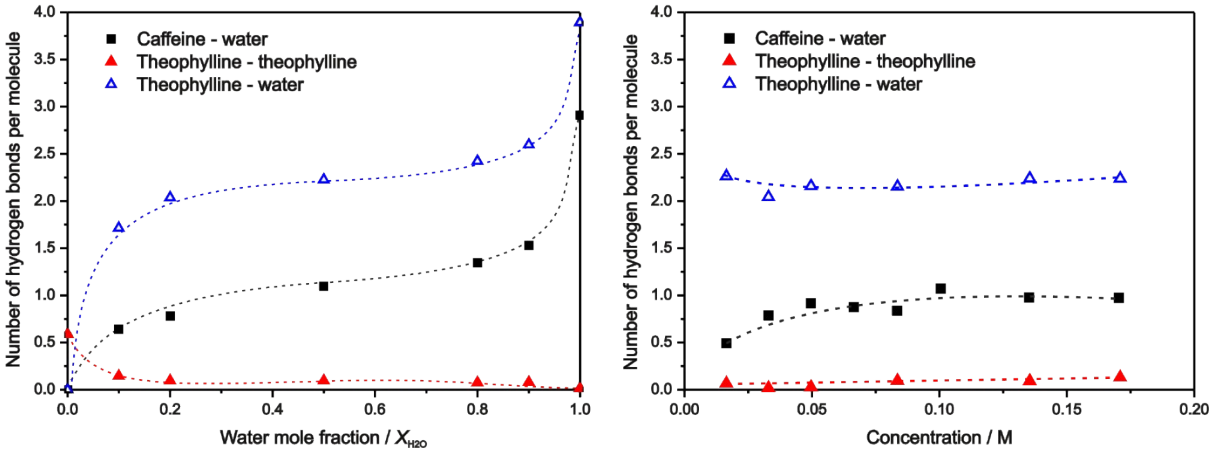

Figure S5 Number of hydrogen bonds per solute molecule between caffeine and water (black squares), theophylline and water (blue empty triangles) and homomeric hydrogen bonds between two theophylline molecules (red filled triangles) as a function of X_{H2O} (left) and concentration at X_{H2O} = 0.5 (right). Error bars are smaller than the size of the symbols. Curves are guides to the eye only.

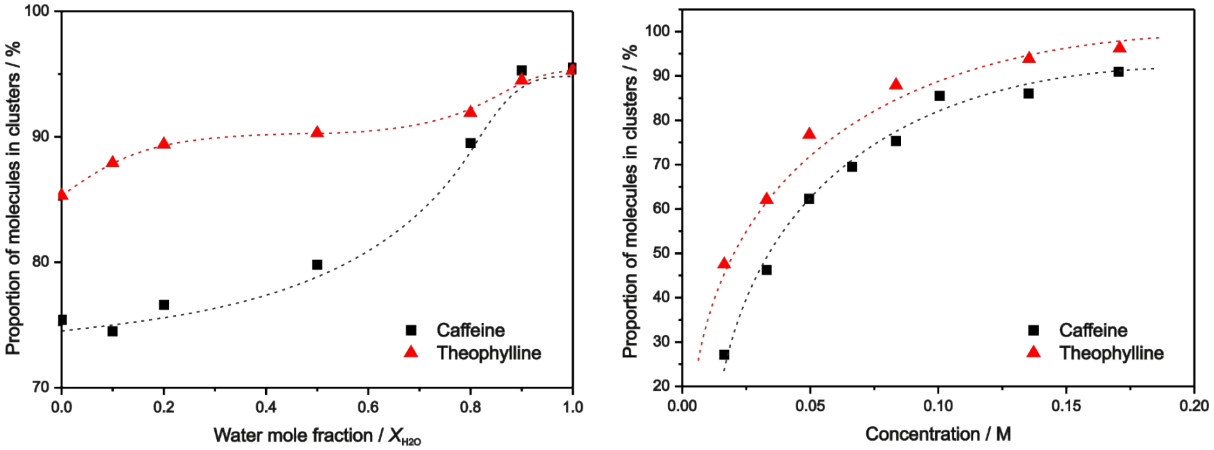

Figure S6 Proportion of solute molecules located in clusters. Red triangles represent caffeine, blue squares represent theophylline. Dotted lines are guides to the eye only.

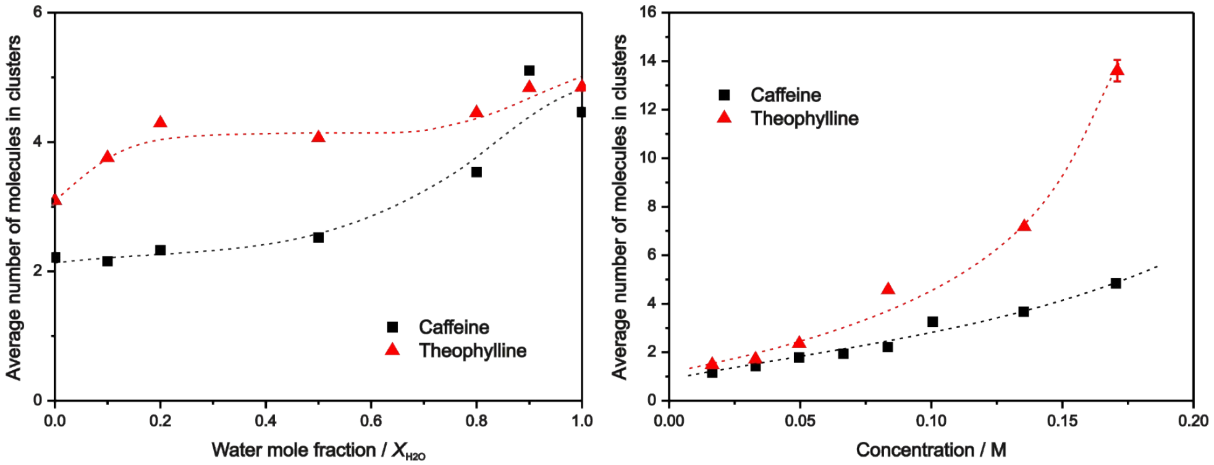

Figure S7 Average number of molecules in molecular clusters for caffeine (black squares) and theophylline (red triangles) as a function of X_{H2O} (left) and concentration at X_{H2O} = 0.5 (right). Error bars are smaller than the size of the symbols. Curves are guides to the eye only.

Experimental section

Materials

All reagents and solvents were obtained from commercial sources and used without further purification. Phase purity of anhydrous caffeine and theophylline was confirmed by powder X-ray diffraction. NMR measurements and crystallisation trials were conducted with freshly opened deuterated solvents, and the acetonitrile water content was assayed by Karl-Fischer potentiometry, using a Metrohm 899 Coulometer with HYDRANAL Coulomat AG anolyte solution. Measured water concentrations were less than 0.2% w/w in all cases.

Powder X-ray diffraction

Supersaturated solutions of caffeine and theophylline, with concentrations of 0.36 and 0.11 M respectively, were prepared by boiling suspensions of the anhydrous solids with weighed 1 cm^3 mixtures of acetonitrile and water in 2 cm^3 sealed vials. The clear, hot solutions were slowly cooled to room temperature and left to undergo crystallisation without addition of seeds to ensure primary nucleation. The samples were stored at room conditions for 1 week before characterization to ensure solvent-mediated phase transition of any intermediate phases. Each crystallisation experiment was repeated three times to ensure reproducibility. Samples for diffraction were loaded onto a silicon sample plate, compressed beneath filter paper to remove excess solvent and analysed immediately. Experiments were performed at 20 ^oC using a PANalytical X'Pert PRO diffractometer fitted with a copper tube operated at 40 kV and 40 mA. A nickel filter was mounted in the primary beam, and 0.04 rad axial Soller slits were used in both beam paths. Scans were conducted over a 2θ range of 5-25^o with a step size of 0.0167^o and a scan speed of 0.0531^o/s.

Nuclear magnetic resonance spectroscopy

NMR spectra were recorded at 20 ^oC using a Bruker Avance 400. Stock solvent mixtures were prepared from weighed quantities of deuterium oxide and acetonitrile-d_3, and used to prepare sample caffeine and theophylline solutions with concentrations in the range 4-100 and 1-40 mM respectively. Solutions were heated in sealed vials to ensure complete dissolution and allowed to cool to room temperature for 2 hours prior to analysis. Binding isotherms were measured by serial dilution of 4-5 solutions of varying concentration.[1] The effect of solvent composition on sample density was accounted for by weighing deuterium oxide/acetonitrile-d_3 mixtures of known volume, and constructing a calibration curve with three replicates per concentration interval. The chemical shifts of methyl protons were fitted simultaneously to an isodesmic polymerisation model[2] *via* a Nelder-Mead algorithm in the online software BindFit.[3]

Molecular dynamics simulations

MD simulations were performed in Gromacs 4.9.1[4] using the General Amber Force Field (GAFF).[5] The Antechamber package[6] was used to assign bonded and non-bonded interaction

parameters, and calculate atomic charges via the semi-empirical AM1 method with bond charge correction (AM1-BCC).[7] Caffeine and theophylline solutions of 100 mM concentration and varying water concentrations were simulated by solvating 60 solute molecules in a 10 nm cubic periodic box, with a random configuration of acetonitrile and TIP5P water[8] molecules matching the experimental density of the solvent mixture. Solutions of varying concentration were simulated in a cubic box of three times the volume, with initial edge lengths of approximately 14.4 nm, in order to produce accurate clustering statistics at the lowest solute concentrations. The initial structures were energy-minimised over 10 000 steps *via* a steepest-descent procedure with no bond constraints. Equilibration was performed over 100 ps at a constant temperature of 293 K and reference pressure of 1 bar, using a velocity-rescale thermostat[9] and isotropic Berendsen barostat[10] with time constants of 0.1 and 1.0 ps, respectively. Production runs were performed over 10 ns under the same NPT conditions using a Nosé-Hoover thermostat[11] and isotropic Parrinello-Rahman barostat,[12] both with time constants of 2.0 ps. All simulation steps were performed with fixed-bond constraints, a time step of 2 fs, initial velocities randomly assigned according to a Maxwell distribution at 293 K and a compressibility set to the experimental value.[13] Electrostatic interactions were modelled with the Particle-Mesh Ewald method,[14] using a cut-off distance of 1.2 nm, grid spacing of 0.1 nm and cubic interpolation. Van der Waals forces were modelled with a cut-off of 1.2 nm, and long-range dispersion corrections were applied for energy and pressure. Interactions between solute molecules were averaged over the final 5 ns of simulation time *via* the *distance*, *hbond* and *sasa* commands in Gromacs. Solvent-accessible areas were measured using a probe radius of 0.14 nm, while hydrogen bond calculations were performed with length and angle thresholds of 0.35 nm and 30°, respectively. Simulations were visualised in VMD.[15]

References

1. Thordarson, P., Determining association constants from titration experiments in supramolecular chemistry. *Chem. Soc. Rev.* **2011,** *40* (3), 1305-1323.
2. Martin, R. B., Comparisons of Indefinite Self-Association Models. *Chem. Rev.* **1996,** *96* (8), 3043-3064.
3. Chen, S. H.; Mallamace, F.; Liu, L.; Liu, D. Z.; Chu, X. Q.; Zhang, Y.; Kim, C.; Faraone, A.; Mou, C. Y.; Fratini, E.; Baglioni, P.; Kolesnikov, A. I.; Garcia-Sakai, V., Dynamic crossover phenomenon in confined supercooled water and its relation to the existence of a liquid-liquid critical point in water. In *Complex Systems*, Tokuyama, M.; Oppenheim, I.; Nishiyama, H., Eds. Amer Inst Physics: Melville, 2008; Vol. 982, pp 39-52.
4. Van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C., GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **2005,** *26* (16), 1701-1718.
5. Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J., The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005,** *26* (16), 1668-1688.
6. Wang, J. M.; Wang, W.; Kollman, P. A.; Case, D. A., Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graphics Modell.* **2006,** *25* (2), 247-260.
7. Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I., Fast, efficient generation of high-quality atomic Charges. AM1-BCC model: I. Method. *J. Comput. Chem.* **2000,** *21* (2), 132-146.

8. Mahoney, M. W.; Jorgensen, W. L., A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *J. Chem. Phys.* **2000,** *112* (20), 8910-8922.
9. Bussi, G.; Donadio, D.; Parrinello, M., Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007,** *126* (1), 014101.
10. Berendsen, H. J. C.; Postma, J. P. M.; Vangunsteren, W. F.; DiNola, A.; Haak, J. R., Molecular-dynamics with coupling to an external bath. *J. Chem. Phys.* **1984,** *81* (8), 3684-3690.
11. Evans, D. J.; Holian, B. L., The Nose–Hoover thermostat. *J. Chem. Phys.* **1985,** *83* (8), 4069-4074.
12. Parrinello, M.; Rahman, A., Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **1981,** *52* (12), 7182-7190.
13. Grant-Taylor, D. F.; Macdonald, D. D., Thermal pressure and energy–volume coefficients for the acetonitrile + water system. *Can. J. Chem.* **1976,** *54* (17), 2813-2819.
14. Darden, T.; York, D.; Pedersen, L., Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993,** *98* (12), 10089-10092.
15. Humphrey, W.; Dalke, A.; Schulten, K., VMD - Visual Molecular Dynamics. *J. Mol. Graphics* **1996,** *14*, 33-38.