

## Why do some molecules form hydrates or solvates?

Simon Boothroyd,<sup>1</sup> Andy Kerridge,<sup>1</sup> Anders Broo,<sup>2</sup> David Buttar<sup>3</sup> and Jamshed Anwar<sup>1,\*</sup>

<sup>1</sup>Chemical Theory and Computation, Department of Chemistry, Lancaster University, Lancaster LA1 4YB U.K.

<sup>2</sup>Pharmaceutical and Analytical R & D, Experimental Formulations, AstraZeneca R&D Mölndal, SE-431 83 Sweden

<sup>3</sup>Pharmaceutical Sciences, AstraZeneca, Silk Road Business Park, Macclesfield, SK10 2NA, U.K.

**KEYWORDS** *Crystal solvate • Crystal hydrate • Crystallisation • Phase diagram • Molecular simulation.*

**ABSTRACT:** The discovery of solvates (crystal structures where the solvent is incorporated into the lattice) dates back to the dawn of chemistry. The phenomenon is ubiquitous with important applications ranging from the development of pharmaceuticals to the potential capture of CO<sub>2</sub> from the atmosphere. Despite this interest, we still do not fully understand why some molecules form solvates. Here, we employ molecular simulation using simple models of solute and solvent molecules, whose interaction parameters can be modulated at will to access a universe of molecules that do and do not form solvates. We investigate the phase behaviour of these model solute-solvent systems as a function of solute-solvent affinity, molecule size ratio, and solute concentration. Our simulations demonstrate that the primary criterion for solvate formation is that the solute-solvent affinity must be sufficient to overwhelm the solute-solute and solvent-solvent affinities. A strong solute-solvent affinity in itself is not a sufficient condition for solvate formation: in the absence of such affinity, a solvate may still form provided that the self-affinities of the solute and the solvent are lower in relative terms. We show that even solvent-phobic molecules can be induced to form solvates by virtue of the pΔV component of the Gibbs potential arising from either a more efficient packing or high pressure overcoming the energy penalty.

### INTRODUCTION

When a solute crystallises from solution, it may do so either as a pure crystal or as a solvate with the solvent molecules being incorporated in the lattice. When the incorporated solvent is water, the solvate crystals are termed hydrates. Solvate formation, and in particular hydrate formation, is a common phenomenon<sup>1,2</sup>. About a third of all organic molecules are able to form hydrates and solvates<sup>3-5</sup>, an example exhibiting extreme promiscuity being the antibacterial sulfathiazole for which over a hundred solvates have been characterised<sup>6</sup>. Solvates can exhibit markedly different physicochemical properties relative to the corresponding anhydrous forms, which include melting point, solubility, crystal habit, and mechanical properties. In the pharmaceutical industry, the choice of whether the form of the active substance is a solvate or anhydrous can affect its bioavailability and the ease (or otherwise) of manufacturing the product, as well as its stability<sup>7</sup>. Hydrate formation is also an issue in the petroleum industry where it can cause blockage of gas pipelines<sup>8</sup>. There are also other hugely-beneficial potential applications ranging from hydrogen and natural gas storage to atmospheric carbon dioxide capture<sup>9-12</sup>.

Despite this extensive interest, the fundamental question of why some molecules form solvates remains an open problem. The thermodynamic perspective is that the solvated forms of these molecules have a lower free energy, but this is not insightful and begs the question why do they have a lower free energy? The thermodynamics approach is exemplified by studies comparing the potential energies (as approximations for free energies) of the various forms with a view to rationalising why a particular molecule forms a hydrate whilst a related one does

not<sup>13-15</sup>. While these methods offer some predictive capability, they inform us only about the system of interest, rather than revealing broader insights. An alternative approach that addresses the posed question somewhat better has attempted to link molecular features to propensity for hydrate formation. A series of surveys of the Cambridge Structural Database (CSD) have revealed a strong correlation with the polar surface area and degree of branching within a molecule and with an increased number of polar functional groups (such as carbonyl (C=O), ether (C-O-C), hydroxyl (O-H) and primary amine (N-H))<sup>16-18</sup>. This suggests that a strong affinity for the solvent may be important and yet there are many examples of substances with high solubility (i.e. those having a strong interaction with the solvent) that do not form solvates. Further, how does one rationalise hydrates of hydrophobic molecules e.g. gas hydrates<sup>9</sup>?

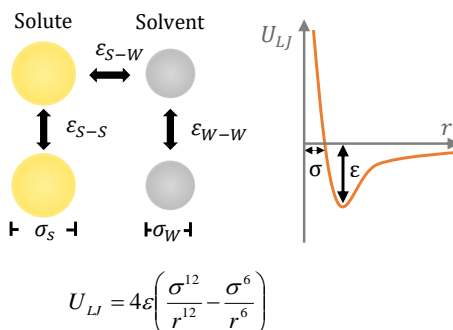
At the heart of the question of why a particular molecule forms a solvate are the molecular interactions: specifically, it is the interplay between the solute-solvent, solute-solute, and solvent-solvent interactions. Coupled to these interactions is the nature of the packing of the molecules in the potential anhydrous and solvate forms. Ideally, we need to explore and understand how the phase diagram of a solute/solvent system varies as a function of the solute-solvent, solute-solute, and solvent-solvent interactions, and molecular packing. How might one achieve this? A cursory review of the problem suggests that this is not feasible. To study the effect of variation in the inter-molecular interactions on phase behaviour requires the consideration of a series of solute and solvent molecules with a variety of molecular structures. The elucidation of the phase diagram for each of these solute-solvent pairs would be a major task in itself, independent of whether it is based on experiment or modelling.

In addition to this, there is the difficulty of de-convoluting the effects of molecular packing from the inter-molecular interactions.

Here we access the phase behaviour of a universe of molecules that do and do not form solvates by means of molecular simulation using simple, coarse-grained models of molecules. These simple models strip away the molecular complexity that otherwise obscures the core issue, while enabling modulation of the inter-molecular interactions by design. Thus, we investigate the crystallisation behaviour of a series of solute-solvent systems as a function of the affinity and molecule size ratio (packing) between the solute and solvent. We show that solvate formation is promoted when the solute-solvent affinity overwhelms the solute and solvent self-affinities, and that a strong solute-solvent affinity is not a sufficient condition in itself. Solvate formation can also occur for solutes with low-solvent affinity by virtue of the *pAV* component of the Gibbs potential arising from either more efficient packing or high applied pressure overcoming the energy penalty.

The phase behaviour of the solute-solvent systems was explored using molecular dynamics (MD) simulations. The solute and solvent molecules were represented by simple, single-particle models based on the Lennard Jones (LJ) interaction. Such models are appropriate as solvate formation is a generic phenomenon, being observed in a wide class of materials. These models have been successfully employed by us earlier to probe crystal nucleation problems including the identification of design rules for nucleation inhibitors<sup>19,20</sup> and for uncovering molecular processes in secondary nucleation<sup>21</sup>. The LJ model is characterized by two parameters:  $\sigma$ , the distance at which the interaction potential is zero, which serves as the effective molecule size; and  $\epsilon$ , the potential energy well-depth that characterizes the affinity between the molecules (Figure 1). Our choice of LJ parameters for the models was not arbitrary but based on the LJ phase diagram, which is known<sup>22</sup>. Thus, the chosen solvent parameters,  $\sigma_{W-W} = 0.47$  nm;  $\epsilon_{W-W} = 3.28$  kJ mol<sup>-1</sup>, define a liquid (the solvent) with a melting point of 273 K (corresponding to water). The solute phase packing parameters ranged from  $\sigma_{S-S} = 0.47$ –1.47 nm, while the affinity was fixed at  $\epsilon_{S-S} = 5.00$  kJ mol<sup>-1</sup>. This chosen solute affinity for  $\sigma_{S-S} = 0.47$  nm defines a solid with a melting point of approximately 421K (an organic solid).

Note that the large affinity-parameter values employed here, up to  $\epsilon = 6.0$  kJ mol<sup>-1</sup>, are well beyond the typical values characterising van der Waals interactions. For comparison, the oxygen-oxygen van der Waals interaction for the TIP3P water model is characterised by  $\epsilon = 0.6364$  kJ mol<sup>-1</sup> [Jorgensen et al. 1983]. The implication is that the LJ model employed in the study serves as a molecular potential, encapsulating both the weak van der Waals and the stronger Coulombic interactions, albeit not strong formal charges. The LJ model as employed here is used in the widely-employed, coarse-grained MARTINI forcefield [Martini] to represent molecular moieties of up to 4 non-hydrogen atoms e.g. -CH2-COOH.

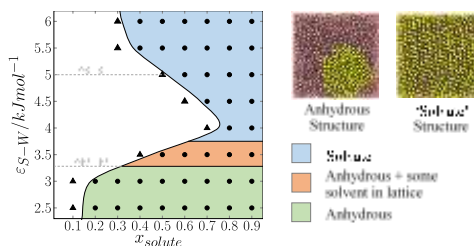


**Figure 1.** The interactions between solute and solvent molecules are characterised by the  $\epsilon$  and  $\sigma$  parameters (left) of the Lennard-Jones potential shown plotted as a function of the separation distance  $r$  (right).

We investigated the crystallization behaviour of the solute for a universe of solute-solvent systems. The solute-solvent affinity was varied to encompass a range of systems,  $\epsilon_{S-W} = 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 5.5$  and  $6.0$  kJ mol<sup>-1</sup>, where the higher values characterize systems with stronger affinity between the solute and the solvent. For each solute-solvent pair, we explored the crystallization behaviour of the solute from a series of solutions of ranging solute concentration,  $x_{solute} = 10, 20, 30, 40, 50, 60, 70, 80, 90$ , and 100 mol %. The system size in all cases was 10,000 particles. The primary question for analysis was: which product crystallised out, the anhydrous form or the solvate?

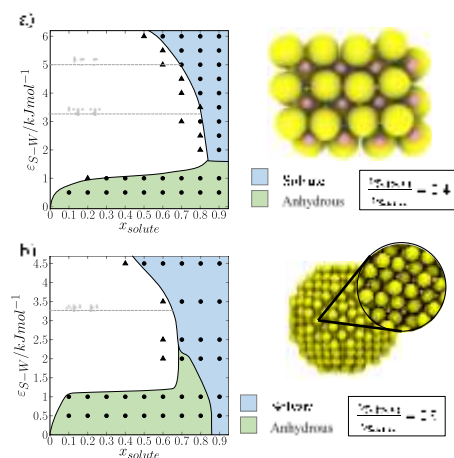
## RESULTS AND DISCUSSION

The first set of simulations explored the crystallization behaviour of solutes for a universe of *equal* particle-size ( $\sigma_{S-S} = \sigma_{W-W} = 0.47$  nm) solute-solvent systems. The dependency of the crystallisation product on the solute-solvent affinity is shown in the phase diagram in Figure 2. A low solute-solvent affinity  $\epsilon_{S-W}$  implies a low solubility. Consequently, at low solute-solvent affinities, the solution becomes supersaturated at low concentrations, limiting the solution region (lower, left region of the plot in Figure 2). At this low solute-solvent affinity the resulting product is the anhydrous structure. As the solute-solvent affinity increases (going up the y-axis in Figure 2), the solubility increases, and the solution region becomes broader.



**Figure 2.** Phase diagram for equal particle-size solute-solvent ( $\sigma_{W-W} = \sigma_{S-S} = 0.47$  nm) systems as a function of solute-solvent affinity  $\epsilon_{S-W}$  and solute concentration  $x_{solute}$  at 283K. The phase diagram exhibits 4 distinct regions: solution (white), solvate (blue), anhydrous (green), and anhydrous with some solvent inclusion (orange). Each data point on the plot represents a simulation result. Circles mark a crystallisation event (structures shown on the right), while a triangle signifies that the system remained a homogeneous solution. We note that the solvate structure is a lattice but is disordered with respect to occupation of the lattice sites. This is expected, since close packing of two distinct but equal-sized particles cannot yield an interpenetrating lattice like that observed for NaCl.

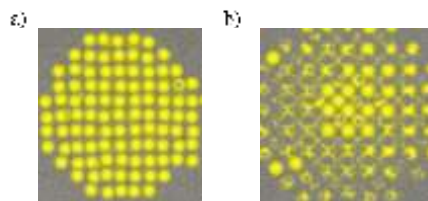
At higher solute-solvent affinities ( $\epsilon_{S-W} > 3.28$  kJ mol<sup>-1</sup>), the solute-solvent affinity surpasses the solvent's affinity for itself and each solute (solvent) particle shows a greater preference to have a solvent (solute) particle as a neighbour. At an affinity of  $\epsilon_{S-W} = 4.0$  kJ mol<sup>-1</sup> and above, the solute and solvent becomes fully integrated to yield a solvate lattice. At still higher solute-solvent affinities, the solute (solvent) particles attract and order the solvent (solute) particles around themselves to such an extent so as to induce crystallisation of the solvate even at low concentrations. Consequently, the solution region in the phase diagram becomes more limited, with the saturation line tending towards lower concentrations (top, left region of plot in Figure 2). These results suggest that the determining factor for solvate formation is the strength of the solute-solvent interaction relative to the solute-solute and the solvent-solvent interactions.



**Figure 3.** Phase diagram for (a) NaCl-type and (b) channel packing type solute-solvent systems as a function of solute-solvent affinity and solute concentration. The blue region indicates solvate formation and the green the anhydrous form. Each data point on the graph represents a single simulation. Circles mark a crystallisation event (structure shown on the right), while a triangle signifies that the system remained a homogeneous solution.

In the above simulations, the solute and solvent particles were of equal size. We then considered the effects of packing, wherein we increased the solute particle size from  $\sigma_S = 0.47$  nm, firstly to  $\sigma_S = 1.18$  nm, and then to  $\sigma_S = 1.47$  nm, whilst keeping the solvent size fixed at  $\sigma_W = 0.47$  nm. For the first of these systems, the particle sizes ( $\sigma_S = 1.18$  nm,  $\sigma_W = 0.47$  nm; solvent:solute radii ratio  $\sigma_W/\sigma_S = 0.40$ ) were chosen to yield a NaCl-type packing<sup>23</sup> and indeed this is the observed structure. In the second case, the solute molecules are substantially larger than those of the solvent ( $\sigma_S = 1.47$  nm,  $\sigma_W = 0.47$  nm,  $\sigma_W/\sigma_S = 0.32$ ). Both of these systems show a similar behaviour (Figure 3) that in broad terms is not too different from the equal-sized molecules. High solute-solvent affinities (compare top-left of Figures 2 and 3) yield the solvate phase whilst lower solute-solvent affinities yield the anhydrous form. The emergent solvates reveal a face-centred lattice for the solute molecules with the solvent molecules either forming an interpenetrating face-centred lattice (the NaCl structure for  $\sigma_W/\sigma_S = 0.40$ ) or filling the interstitial channels (when  $\sigma_W/\sigma_S = 0.32$ ) (Figure 3). The latter structures are very similar to the class of non-stoichiometric channel solvates,<sup>3,24</sup> [Griesser 2006, Braun et al](#) where the solvent occupies channels formed within the solute lattice and can freely diffuse out depending on the relative vapour pressure of the solvent (relative humidity for a hydrate) in the environment. Indeed, the solvent particles in these simulated channel-solvates exhibit significant diffusion (diffusion coefficient:  $\sim 3.5 - 7.5 \times 10^{-9}$  m<sup>2</sup> s<sup>-1</sup>).

For the system yielding the interstitial channels, we also looked closely at the extreme case of a solute with a very low affinity for solvent,  $\epsilon_{S-W} = 0.3$  kJ mol<sup>-1</sup>, i.e. a solvent-phobic solute (see Figure 4). For this system, the solute-solute affinity was increased to  $\epsilon_{S-S} = 8.0$  kJ mol<sup>-1</sup> and we investigated the system at the low molar solute concentration of 1%. The high solute-solute affinity and low molar concentration favoured the formation of a small solute crystallite in the bulk solvent, making it easier to observe whether the solvent was either included, or excluded from the emergent structure. This system showed phase separation at (ambient) pressure  $p = 0.001$  katm, but yielded a solvate structure at a higher pressure of  $p = 10$  katm. Thus, it is clear that even solvent-phobic solutes can form solvates when driven by the  $p\Delta V$  component of the Gibbs potential  $G$ .



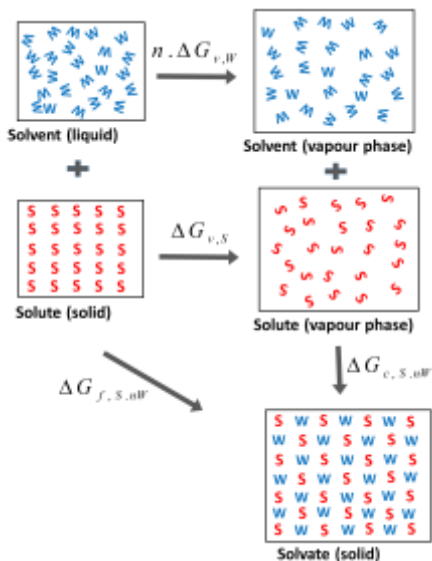
**Figure 4.** A slice taken from the final structure of the solvent-phobic ( $\epsilon_{S-W} = 0.3$  kJ mol<sup>-1</sup>,  $\sigma_W/\sigma_S = 0.32$ ) system. (a) At ambient pressure ( $p = 0.001$  katm) the solvent was observed to be excluded from the solute structure, thus favouring the anhydrous form. (b) Increasing the pressure ( $p = 10$  katm) resulted in the solvent filling

**Commented [AJ1]:** Referee #3 observation. We have solvate at epsilon = 0.5; We also have another point 0.3 which is not a solvate. Include that here and re-draw the line.

the interstitial channels between solute particles, similar to the behaviour observed in channel solvates.

The above results appear to show that a solvate is always formed when the solute-solvent affinity is strong, but can also form when such affinity is lacking. A limited number of additional simulations were carried out for the equal particle-size system but for which the solute-solute affinity was increased to  $\epsilon_{S-S} = 7.0 \text{ kJ mol}^{-1}$  (from  $5.0 \text{ kJ mol}^{-1}$ ) whilst keeping the solute-solvent affinity fixed at  $\epsilon_{S-W} = 4.0 \text{ kJ mol}^{-1}$ . This would be equivalent to the solid having a higher melting point whilst keeping its interaction with the solvent to be the same. One might expect that a such a system, given the strong solute-solvent affinity, would yield a solvate, reproducing the data points for  $\epsilon_{S-W} = 4.0 \text{ kJ mol}^{-1}$  in Figure 2. It did not. Instead, we observed the anhydrous structure. The inference is that a strong solute-solvent affinity in itself is not a sufficient condition for solvate formation. Rather, the solute-solvent affinity must be sufficient to overwhelm the solute and solvent self-affinities.

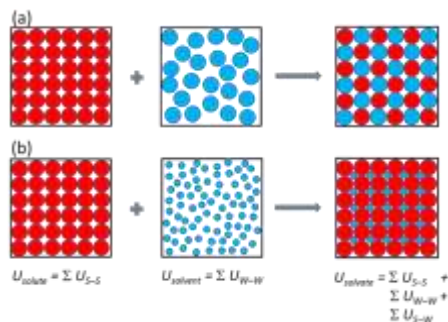
The thermodynamic criteria for solvate formation (see Figure 5) is  $\Delta G_{c,S,nW} < (\Delta G_{v,S} + n \Delta G_{v,W})$ , where  $\Delta G_{v,S}$  and  $\Delta G_{v,W}$  are the molar free energy changes for vaporisation of the solute crystal and the solvent fluid respectively,  $\Delta G_{c,S,nW}$  is the molar free energy change associated with crystallization of the solvate from the vapour phase, and the integer  $n$  is the number of moles of solvent as reflected in the stoichiometry for the reaction of solute plus solvent yielding a solvate:  $S + nW \rightarrow S \cdot nW$ . For a 0K (potential energy) approximation, the solvate formation criteria becomes  $\Delta U_{c,S,nW} < (\Delta U_{v,S} + n \Delta U_{v,W})$  where  $\Delta U_{c,S,nW}$  is the lattice energy of the solvate form  $S \cdot nW$ ,  $\Delta U_{v,S}$  is the lattice energy of the anhydrous form, and  $\Delta U_{v,W}$  is the lattice energy of the solvent crystal (as the solvent would be a solid at 0K).



**Figure 5.** Thermodynamic cycle for the formation of a solvate from the components solute and solvent.  $\Delta G_{f,S,nW}$  is the molar free energy change for solvate formation, and  $\Delta G_{v,S}$  and  $\Delta G_{v,W}$  are the molar free energy changes for vaporisation of the solute crystal and the solvent fluid respectively.  $\Delta G_{c,S,nW}$  is the molar free energy change associated with crystallization of the solvate from the vapour phase, and integer  $n$  reflects the stoichiometry  $S + nW \rightarrow S \cdot nW$ .

Within the spectrum of molecular interactions and packing ratios characterising solvate formation, one can identify two limiting cases (Figure 6): (a) when there is strong solute-solvent affinity, and (b) when the packing of the solute molecules is essentially independent of the solvent. Expressing the 0K stability criterion,  $U_{solvate} < (U_{solute} + U_{solvent})$ , in terms of *component atom-atom interactions* yields  $(\Sigma U_{S-S}(solvate) + \Sigma U_{W-W}(solvate) + \Sigma U_{S-W}(solvate)) < (\Sigma U_{S-S}(solute) + \Sigma U_{W-W}(solvent))$ . For the equal molecule-size system with strong solute-solvent affinity, case (a), the dominating interactions within the solvate are those between the solute and solvent, as each solute (solvent) molecule is surrounded by solvent (solute) particles. The solute-solute and solvent-solvent interactions in the solvate are marginal. Consequently, for this case (to a first approximation), the stability criterion reduces to  $\Sigma U_{S-W}(solvate) < (\Sigma U_{S-S}(solute) + \Sigma U_{W-W}(solvent))$ . For such a system, we can map the Lennard-Jones affinities onto the stability criterion by considering interactions between particles as pseudo-bonds. As a first order approximation, we restrict the interactions to the first coordination sphere. For the solute in a fcc lattice, there are 12 ‘bonds’ and we approximate the strength of each by  $\epsilon_{S-S}$ . Likewise, there are about 12 ‘bonds’ for the liquid, for each of which we assume the strength  $\epsilon_{W-W}$  (alt-

though the actual interaction is a little weaker, given that the particle separation distance is slightly greater in the liquid state). To form a solvate, the 12 solute-solute and the 12 solvent-solvent 'bonds' must be broken and be replaced with 12 new solute-solvent 'bonds', each with an approximate strength of  $\epsilon_{S-W}$ . The approximate stability criterion for the Lennard-Jones system then becomes  $24\epsilon_{S-W} > 12\epsilon_{S-S} + 12\epsilon_{W-W}$ , that is,  $2\epsilon_{S-W} > \epsilon_{S-S} + \epsilon_{W-W}$ . (Note the switch in the inequality operator from less-than to greater-than, since  $\epsilon$  is not the interaction energy but the energy well-depth parameter). Substituting the self-affinity parameters utilised for the solute and solvent,  $\epsilon_{W-W} = 3.28 \text{ kJ mol}^{-1}$  and  $\epsilon_{S-S} = 5.00 \text{ kJ mol}^{-1}$ , the criterion indicates solvate stability above the solute-solvent affinity  $\epsilon_{S-W} = 4.1 \text{ kJ mol}^{-1}$ . This is entirely consistent with the switch-over point for solvate formation, namely about  $4 \text{ kJ mol}^{-1}$ , observed in Figure 2.



**Figure 6.** Two limiting cases of solvate formation, represented schematically: (a) Equal molecule-size system with strong solute-solvent affinity; (b) solvate formation where the solute packing is essentially the same in the anhydrous and solvate forms and independent of the solvent.

For the limiting case (b) where the solute structure of the anhydrous form is essentially identical to that in the solvate (as in a non-stoichiometric channel solvate, e.g. the system shown in Figure 3b),  $\Sigma U_{S-S}(\text{solute}) \approx \Sigma U_{S-S}(\text{solvate})$  and the solvent-solvent interaction in the solvate is marginal i.e.  $\Sigma U_{W-W}(\text{solvate}) \rightarrow 0$ . In this case the stability criterion reduces to  $\Sigma U_{S-W}(\text{solvate}) < \Sigma U_{W-W}(\text{solvent})$ , that is, the solute-solvent interaction must be stronger than the solvent-solvent interaction. This is intuitive being akin to the interplay between the cohesive forces of a fluid and the adhesive forces that determine whether, for example, water will wet a nanopore (hydrophilic surface) or bridge it (hydrophobic surface - exploited in high-tech wetwear that is waterproof and yet breathable). This issue is manifested by the low solute-solvent affinity system where the solute is essentially solvent-phobic (Figure 4). At low pressures, the system phase-separates into the anhydrous form and solvent. At the higher pressure of  $p = 10 \text{ katm}$ , the  $p\Delta V$  component of the Gibbs potential overwhelms the solvent-solvent affinity forcing the solvent into the lattice to form a solvate.

In conclusion, we have shown that the primary criterion for solvate formation is that the solute-solvent affinity must be sufficient to overwhelm the solute-solute and solvent-solvent affinities. A strong solute-solvent affinity in itself is not a sufficient condition. Solute molecules even with a low affinity for a solvent can form solvates, provided that the self-affinities of the solute and the solvent are lower in relative terms. Indeed, as demonstrated, essentially solvent-phobic molecules can form solvates when driven by the  $p\Delta V$  term i.e. under pressure. In going forward, it would be insightful to carry out atomistic lattice or free energy calculations on hydrate systems (using e.g. Cambridge Crystallographic Data Centre data and tools), partitioning the energy into molecule-molecule (solute-solute, solute-solvent, and solvent-solvent) interactions to see how the insights ascertained here play out in realistic systems. Finally, we note that whilst the focus of the paper is solvate formation, the inferences are also applicable to co-crystal formation for binary systems<sup>25</sup>, where the second molecule in the lattice is not the solvent but another solute (solid phase) molecule.

## METHODOLOGY

Molecular dynamics simulations were carried out using the DL-POLY 4.06 software package<sup>26</sup> in the NPT ensemble using a Nosé-Hoover thermostat and barostat. All simulations were run at 283 K and a pressure of 0.001 katm unless otherwise indicated. The interactions (van der Waals) were truncated at  $2.5 \times \sigma_5$ . All simulations were run for a minimum of 5 million steps using a 30 fs time step. The mass of all particles was set to  $72 \text{ g mol}^{-1}$ . System size was 10,000 particles. Initial configurations comprised randomised coordinates.

## AUTHOR INFORMATION

### Corresponding Author

\* Jamshed Anwar, j.anwar@lancaster.ac.uk

### Funding Sources

EPSRC Doctoral Training Account Case studentship with AstraZeneca as industry partner.

## ACKNOWLEDGMENT

SB would like thank the EPSRC, Lancaster University, and AstraZeneca, for the award of his PhD Case studentship.

## REFERENCES

- [1] Nangia, A. & Desiraju G.R. Pseudopolymorphism: occurrences of hydrogen bonding organic solvents in molecular crystals. *Chem. Commun.*, **7**, 605–606 (1999).
- [2] Gorbitz, C. H., Hersleth, H-P., On the inclusion of solvent molecules in the crystal structures of organic compounds, *Acta Crystallographica Section B Structural Science*, *Acta Cryst.* B56, 526-534 (2000)
- [3] Brittain H. G. *Polymorphism in Pharmaceutical Solids* (Informa Healthcare, New York, 2009).
- [4] Threlfall, T. L. Analysis of organic polymorphs. A review. *Analyst*, **120**, 2435–2460 (1995).
- [5] Stahly, G. P. Diversity in Single- and Multiple-Component Crystals. The Search for and Prevalence of Polymorphs and Cocrystals. *Cryst. Growth Des.*, **7**, 1007–1026 (2007).

- [6] Bingham, A. L., Hughes, D. S., Hursthouse, M. B., Lancaster, R. W., Tavener, S., Threlfall, T. L., Over one hundred solvates of sulfathiazole. *Chem. Commun.*, 603–604 (2001)
- [7] (a) Khankari, R. K. & Grant, D. J. W. Pharmaceutical hydrates. *Thermochimica Acta*, **248**, 61–79 (1995); (b) Datta, S., Grant, D. J. W., Crystal structures of drugs: advances in determination, prediction and engineering. *Nature Review Drug Discovery* **3**, 42–57 (2004).
- [8] E. D. Sloan and C. A. Koh, *Clathrate hydrates of natural gases*, CRC Press, Boca Raton, FL, 2008.
- [9] Herslund, P. J., Thomsen, K., Abildskov, J., von Solms, N. Modelling of tetrahydrofuran promoted gas hydrate systems for carbon dioxide capture processes. *Fluid Phase Equilibria*, **375**, 45–65 (2014).
- [10] Yang, M., Song, Y., Jiang, L., Zhao, Y., Ruan, X., Zhang, Y., Wang, S. Hydrate-based technology for CO<sub>2</sub> capture from fossil fuel power plants. *Applied Energy*, **116**, 26–40 (2014).
- [11] Sugahara, T., Haag, J. C., Pinnelli, S. P., Warntjes, A. A., Sloan, E. D., Sum, A. K. Increasing hydrogen storage capacity using tetrahydrofuran. *J. Am. Chem. Soc.*, **131**, 14616–14617 (2009).
- [12] Sugahara, T., Haag, J. C., Warntjes, A. A., Sloan, E. D., Koh, C. A. Large-Cage Occupancies of Hydrogen in Binary Clathrate Hydrates Dependent on Pressures and Guest Concentrations. *J. Phys. Chem. C*, **114**, 15218–15222 (2010).
- [13] Braun, D. E., Karamertzanis, P. G., Price, S. L. Which, if any, hydrates will crystallise? Predicting hydrate formation of two dihydroxybenzoic acids. *Chem. Comm.*, **47**, 5443–5445 (2011).
- [14] Hulme A. T. & Price, S. L. Toward the Prediction of Organic Hydrate Crystal Structures. *J. Chem. Theor. Comput.*, **3**, 1597–1608 (2007).
- [15] Cruz-Cabeza, A. J., Day G. M., Jones, W. Towards prediction of stoichiometry in crystalline multicomponent complexes. *Chem.–Eur. J.*, **14**, 8830–8836 (2008).
- [16] Takiyeddin, K., Khimiyak, Y. Z., Fábrián, L. Prediction of Hydrate and Solvate Formation Using Statistical Models. *Cryst. Growth Des.*, **16**, 70–81 (2016).
- [17] Infantes, L., Chrisholm, J., Motherwell, S. Extended motifs from water and chemical functional groups in organic molecular crystals. *CrystEngComm*, **5**, 480–486 (2003).
- [18] Infantes, L., Fabian, L., Motherwell, W. D. S. Organic crystal hydrates: what are the important factors for formation. *CrystEngComm*, **9**, 65–71 (2007).
- [19] Anwar, J., Boateng, P. K. Computer Simulation of Crystallization from Solution. *J. Am. Chem. Soc.*, **120**, 9600–9604 (1998).
- [20] Anwar, J., Boateng, P. K., Tamaki, R. Mode of action and design rules for additives that modulate crystal nucleation. *Angew. Chem. Int. Ed. Engl.*, **48**, 1596–1600 (2009).
- [21] Anwar, J., Khan, S., Lindfors, L. Secondary Crystal Nucleation: Nuclei Breeding Factory Uncovered. *Angew. Chem. Int. Ed.*, **54**, 14681–14684 (2015).
- [22] Agrawal, R., Kofke, D. A. Thermodynamic and structural properties of model systems at solid-fluid coexistence. *Molecular Physics*, **85**, 43–59 (1995). Thermodynamic and structural properties of model systems at solid-fluid coexistence: I. Fcc and bcc soft spheres R Agrawal, DA Kofke, *Molecular physics* **85** (1), 23–42
- [23] Filion, L., Dijkstra, M. Prediction of binary hard-sphere crystal structures. *Phys. Rev. E*, **79**, 046714-046719 (2009).
- [24] Ahlqvist, M. U. A., Taylor, L. S. Water dynamics in channel hydrates investigated using H/D exchange. *International Journal of Pharmaceutics* **241**, 253–261 (2002)
- [25] Schultheiss, N., Newman, A. Pharmaceutical Cocrystals and Their Physicochemical Properties. *Cryst. Growth Des.*, **9**, 2950–2967 (2009).
- [26] Todorov, I. T., Smith, W., Trachenko, K., Dove, M. T. DL\_POLY\_3: new dimensions in molecular dynamics simulations via massive parallelism. *J. Mat. Chem.*, **16**, 1911–1918 (2006).
- Braun, D. E., Koztecki, L. H., McMahon, J. A., Price, S. L., Reutzel-Edens, S. M. Navigating the Waters of Unconventional Crystalline Hydrates. *Mol. Pharmaceutics* **12**, 3069–3088 (2015).
- Griesser, U. The importance of solvates, in polymorphism in the pharmaceutical industry, ed. R. Hilfiker, Wiley-VCH, Weinheim, 2006, pp. 211–234.
- Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., Klein, M. L. Comparison of simple potential functions for simulating liquid water. *Journal of Chemical Physics*, **79**, 926–935 (1983).
- Marrink, S. J., Risselada, H. J., Yefimov, S., Tieleman, D. P., de Vries, A. H., The MARTINI force field: coarse grained model for biomolecular simulations. *The Journal of Physical Chemistry B*. **111**, 7812–7824 (2007).



**To form a solvate or not? That is the question**

