Electronic Supplementary Material (ESI) for ChemComm.

Electronic Supplementary Information:

On the Entropy Cost of Making Solvates

Aurora J. Cruz-Cabeza,*[1] Sarah E. Wright[1]
Department of Chemical Engineering and Analytical Science, University of Manchester, Manchester, UK

Alessia Bacchi*[2,3]
Dipartimento di Scienze Chimiche, della Vita e della Sostenibilità Ambientale, Università degli Studi di Parma, Viale delle Scienze, 17A, 43124 Parma, Italy
Biopharmanet-TEC, Università degli studi di Parma, Parco Area delle Scienze 27/A, 43124 Parma, Italy

1. Thermodynamic derivation of the entropic penalty for solvates

The solvation reaction is formulated in Figure 1 where A(s) is the principal component, B(l) is the liquid solvent and AB(s) is the solvate, n and m being the stoichiometries of A and B required to make one mol of A_nB_m, referred to as AB hereafter for simplicity. In order to calculate the free energy of the solvation reaction, we formulate a thermodynamic cycle in which B(l) is first solidified to B(s), and then A(s) and B(s) are sublimed. The free energy of such reaction can be written in terms of such cycle as in equation (1).

$$\Delta G_{formation}^{AB} = (\Delta G_{latt}^{AB} - n\Delta G_{latt}^{A} - m\Delta G_{latt}^{B}) - m\Delta G_{fus}^{B} \qquad (1)$$

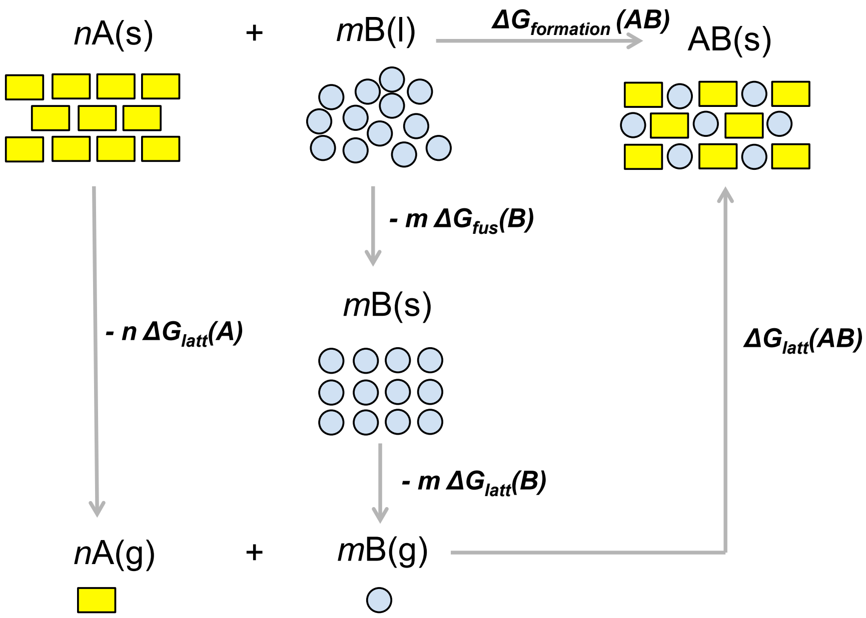

Figure 1. Thermodynamic cycle allowing decomposition of the energy terms involved in the formation of the solvate AB(s) from a solid A(s) and a liquid B(l). n and m is the stoichiometry of A and B respectively.

The term enclosed by parentheses in equation (1) is represented as $\mathrm{DG}^{AB}_{latt\text{-}gain}$ and expresses the free energy gain in forming a mol of the crystalline AB solvate from crystalline solids A and B (with the appropriate stoichiometric ratios n and m). This term can be furthered as $\Delta G^{AB}_{latt-gain} = \Delta H^{AB}_{latt-gain} - T\Delta S^{AB}_{latt-gain}$. It is a common approximation to assume that the entropy difference between crystalline AB and crystalline A and B is small, thus $\Delta S^{AB}_{latt-gain}$~0. The lattice enthalpy difference is also usually approximated to the lattice energy difference of the solids involved since lattice vibrations contributions are small and their calculations can be computationally expensive.[1] Thus $\Delta G^{AB}_{latt-gain} \approx \Delta H^{AB}_{latt-gain} \approx \Delta E^{AB}_{latt-gain}$. The lattice energies of crystalline solids can be computed accurately with DFT-d methods and thus the first term directing solvate formation can be computed readily. In fact, most computational estimations of whether or not a solvate will form,[2],[3] and its stoichiometry,[4,5] have been performed by computing only this lattice energy term.

The last term in equation (1) is the free energy change involved in going from solid B to liquid B times the mols of B involved in the reaction (m). The free energy of this process can be furthered as $\Delta G^B_{fus}[\mathrm{T}]=\Delta H^B_{fus}[T] - T\Delta S^B_{fus}[T]$ at the temperature T at which the solvate formation reaction takes place. Assuming that the heat capacities of the solid and liquid B are constant between the temperature of solvate formation (usually around 300 K) and its fusion temperature, this term can be expressed as in equation (2) where ΔC_p is the difference in heat capacity between liquid and solid B.

$$\Delta G^B_{fus}[T] = \Delta H^B_{fus}[T^B_{fus}] + \Delta C_p(T - T^B_{fus}) - T\Delta S^B_{fus}[T^B_{fus}] - T\Delta C_p ln\left(\frac{T}{T^B_{fus}}\right) \qquad (2)$$

Since at T_{fus}, $\Delta H^B_{fus}[T^B_{fus}] = T^B_{fus}\Delta S^B_{fus}[T^B_{fus}]$, (2) becomes (3).

$$\Delta G^B_{fus}[T] = (T^B_{fus}\Delta S^B_{fus}[T^B_{fus}] - T\Delta S^B_{fus}[T^B_{fus}]\,) + \left(\Delta C_p\left(T - T^B_{fus} - Tln\left(\frac{T}{T^B_{fus}}\right)\right)\right) \qquad (3)$$

The term $\left(T - T^B_{fus} - Tln\left(\frac{T}{T^B_{fus}}\right)\right)$ is usually small but can be significant when T is very different to T^B_{fus}. The differences in heat capacities between solid and liquid B, however, are usually very small thus the second term of equation (3) can be obviated. As an illustration of how small this term is, at 300 K the term

$\Delta C_p\left(T - T_{fus}^B - T\ln\left(\frac{T}{T_{fus}^B}\right)\right)$ is calculated to be 0.00, -0.05, -0.05 and -0.97 kJ/mol for DMSO, ethanol, water and acetone.

Thus, $\Delta G_{fus}^B[T] = -\Delta S_{fus}^B[T_{fus}^B](T - T_{fus}^B)$ and thus the term $-m\Delta G_{fus}^B$ becomes $-m\Delta G_{fus}^B[T] = m\Delta S_{fus}^B[T_{fus}^B](T - T_{fus}^B)$.

Finally, the overall free energy of the reaction can then be formulated as in equation (4) where ΔS_{fus}^B is calculated at T_{fus}^B.

$$\Delta G_{formation}^{AB} \approx \Delta E_{latt-gain}^{AB} + m\Delta S_{fus}^B(T - T_{fus}^B) \tag{4}$$

As discussed above, the first term is the lattice energy gain in going from crystalline A and B into crystalline AB and can be computed readily. The second term, however, is dominated by the entropy loss required for liquid B to go into crystalline B. This term is dependant on liquid B only. Since $\Delta S_{fus}^B = \left(\frac{\Delta H_{fus}^B}{T_{fus}^B}\right)$, equation (4) can be rewritten as equation (5). This second term, which we refer to as the entropy cost or entropy penalty of solvate formation, can be estimated from experimental values of ΔH_{fus}^B and T_{fus}^B. We note that the second term is always positive since we are considering the case of solvate formation where the experimental temperature is always higher than the melting temperature of the solvent.

$$\Delta G_{formation}^{AB} \approx \Delta E_{latt-gain}^{AB} + m\left(\frac{\Delta H_{fus}^B}{T_{fus}^B}\right)(T - T_{fus}^B) \tag{5}$$

2. Entropy Penalty for a Range of common solvents (300 K)

Method

Using the NIST webbook database, we retrieved all enthalpy and temperature of fusion for 78 solvents. If multiple experimental values were available in the NIST database, the most recent value was taken. If the data was not available in the NIST webbook, then the CRC Handbook of Chemistry and Physics was used to obtain the enthalpy and temperature of fusion data.

NIST Webbook: https://webbook.nist.gov/

William E. Acree, Jr., James S. Chickos, "Phase Transition Enthalpy Measurements of Organic and Organometallic Compounds" in NIST Chemistry WebBook, NIST Standard Reference Database Number 69, Eds. P.J. Linstrom and W.G. Mallard, National Institute of Standards and Technology, Gaithersburg MD, 20899, https://doi.org/10.18434/T4D303, (retrieved November 1, 2019).

CRC Online: http://hbcponline.com/

Fluid Properties. CRC Handbook of Chemistry and Physics [Online], 100th ed; Rumble, J. R., Ed.; CRC Press, 2019-2020. http://hbcponline.com/, (retrieved November 1, 2019)

Results

Table 1 reports the entropic penalty of a vast range of common solvents, showing that this contribution can be very low for solvents such as water or cyclohexane or very important for solvents such as toluene or ethyl acetate.

Table 1. Temperature and enthalpy of fusion (NIST and CRC) and entropy penalty at 300 K for a range of solvents (see Figure 2 in the manuscript)

	Solvent	**T_{fus} (Celsius)**	**T_{fus} (K)**	**ΔH_{fus} (kJ/mol)**	**Entropy Penalty (kJ/mol)**
AcCN	Acetonitrile	-43.7	229.3	8.2	2.5
Amines	Diisopropylamine[†]	2.6	275.6	8.4	0.7
	Aniline	-5.9	267.1	10.5	1.3
Ether O	1,4-dioxane	11.1	284.1	12.8	0.7
	anisole	-36	237	11.7	3.1
	Methoxybenzene	-37	236	12.9	3.5
	1,2-dimethoxyethane[†]	-69	204	12.6	5.9
	t-butylmethylether	-108.4	164.6	7.6	6.3
	1,1-dimethoxymethane[†]	-105.1	167.9	8.3	6.6
	Diethylether	-116.1	156.9	7.2	6.6
	THF	-108.2	164.8	8.5	7.0
	Isopropylether	-85.2	187.8	12.1	7.2
Strong HB Acceptors	Sulfolane	28.7	301.7	1.4	0.0
	DMSO	18.7	291.7	14.4	0.4
	Formamide	2.7	275.7	8.0	0.7
	Nitrobenzene	5.8	278.8	12.1	0.9
	Methyl Acetate	-98.1	174.9	1.5	1.1
	N,N-dimethylacetamide	-18.8	254.2	10.2	1.8
	Nitromethane[†]	-28.7	244.3	9.7	2.2
	Pyridine	-41.5	231.5	8.3	2.5
	3-pentanone	-38.8	234.2	11.6	3.3
	DMF	-60.1	212.9	9.0	3.7
	N-methylpyrrolidone	-24.5	248.5	18.1	3.8
	Furfural	-37.9	235.1	14.4	4.0
	Acetone	-96.4	176.6	5.7	4.0
	Diethylcarbonate	-74.8	198.2	9.2	4.7
	2-butanone	-86.5	186.5	8.4	5.1
	Methylethylketone	-86.5	186.5	8.4	5.1
	Methylacetate	-98.1	174.9	7.5	5.4
	Methylbutylketone	-55.3	217.7	14.9	5.6
	Ethyl Acetate	-83.7	189.3	10.5	6.1

[†] Entropy of fusion data and temperature of fusion for these solvents are obtained from CRC. For the remaining solvents, data was obtained from NIST.

Alcohols & water	Cyclohexanol	25.2	298.2	1.7	0.0
	Water[†]	0	273	6.0	0.6
	Ethyleneglycol	-12.4	260.6	10.0	1.5
	Methanol	-97.7	175.3	3.2	2.3
	1,5-pentanediol	-25	248	15.7	3.3
	2-propanol	-87.8	185.2	5.4	3.4
	1-octanol	-14.6	258.4	25.2	4.1
	Ethanol	-114.2	158.8	4.6	4.1
	2-pentanol	-73	200	8.5	4.2
	2-methyl-1-propanol	-101.8	171.2	6.3	4.8
	1-hexanol	-47.2	225.8	15.5	5.1
	1-propanol	-124.3	148.7	5.4	5.5
	1-pentanol	-77.4	195.6	10.5	5.6
	1-butanol	-89.1	183.9	9.3	5.9
	2-butanol	-89.1	183.9	9.3	5.9
	2-butoxyethanol	-73.5	199.5	11.8	5.9
Acids	Acetic Acid	25.7	298.7	11.7	0.1
	Formic Acid	8.4	281.4	12.7	0.8
	Butanoic Acid	-8.3	264.7	11.1	1.5
	Propanoic Acid	-20.3	252.7	10.7	2.0
	Dichloro acetic acid	-88	185	5.4	3.4
Hydrocarbons	Cyclohexane	6.8	279.8	2.7	0.2
	Cyclopentane	-93.3	179.7	0.6	0.4
	Benzene	5.7	278.7	9.9	0.8
	p-xylene	13.3	286.3	17.1	0.8
	1-methylnaphthalene	-30.3	242.7	5.7	1.3
	o-xylene	-25.2	247.8	13.6	2.9
	Tetralin	-35.6	237.4	12.5	3.3
	m-xylene	-47.7	225.3	11.6	3.8
	Toluene	-95	178	6.6	4.5
	Dodecane	-9.9	263.1	35.7	5.0
	Cumene[†]	-96	177	7.3	5.1
	Methylcyclohexane	-126.4	146.6	6.7	7.0
	Isooctane[†]	-107.4	165.6	9.2	7.5
	Hexane	-95.2	177.8	13.1	9.0
	Heptane	-90.4	182.6	14.0	9.0
	Pentane	-129.5	143.5	8.4	9.2
	2-methylpentane	-108.8	164.2	11.9	9.9
Chlorinated	1,1,1-trichloroethane	-32.9	240.1	1.9	0.5
	Carbon tetrachloride	-24	249	2.7	0.6
	tetrachloroethene	-22.2	250.8	10.9	2.1
	1,2-dichloroethane	-35.8	237.2	8.8	2.3
	trichloroethene	-35.9	237.1	11.4	3.0
	Chlorobenzene	-45.1	227.9	9.6	3.0
	Chloroform	-63.4	209.6	8.8	3.8
	Dichloromethane	-94.8	178.2	6.2	4.2
	1,1-dichloroethene	-96.8	176.2	7.9	5.5

3. Cambridge Structural Database Search and Solvate Occurrences

CSD Structure retrieval

Subsets were generated using ConQuest, searching for structures which contain two chemical units, and then a further search requiring a water molecule or solvent molecule (for the top 15 solvates in the CSD).

Using the CSD python API, it was possible to search for the largest component in the solvate to find other structures containing that component within the rest of the CSD to find structures which have only one chemical unit, the neat form. An additional check to ensure that compounds of opposite chirality are not paired up by generating an InChI key.

We have searched the CSD for all unique crystal structures (redeterminations were removed) containing a solvent molecule in molecular organic crystals. Disorder was removed and only organics were considered. Only 15 subsets of solvates were found to have 200 crystal structures or more. Subsets smaller than 200 were discarded because of the subsets being too small.

Calculations of Solvate Occurrences

Deriving CSD frequencies of occurrence for solvates is straightforward. One could search the CSD for all types of solvates. The frequency of occurence for a particular solvate can be calculated as the number of crystal structures containing such solvent ($N_{solvate}$) over the total number of solvates in the CSD ($N_{total-CSDsolvates}$). Instead of doing it over the total, this can also be done over the total number of solvates for the 15 most common solvents ($N_{total-15solvates}$).

We calculated the frequency of occurrence ($P_{solvate}$) for each of those 15 solvents using the equation:

$$P_{solvate} = N_{solvate}/N_{total-15solvates}$$

The data is presented in table 2 in percentages.

This CSD derived solvate occurrences, however, are biased because certain solvents are used for crystallisation much more frequently than others. It is not surprising thus, how water shows the highest frequency of occurrence. This is simply because water is the most commonly used solvent for crystallisation of polar molecules.

In order to correct for that bias, we would need to calculate the probability of a solvent to be used in crystallisation. Obtaining this data from the CSD is not obvious. Nangia and Desiraju, for example, tried to calculate corrected occurrences for solvates. For this they looked at all crystal structures of solvated published in Acta Cryst for a decade and recorded the crystallisation solvent and outcome.

Here, we propose a new way of calculating how often a solvent is used for crystallisation. In the CSD, the majority of crystal structures are the result of organic synthesis. Those structures are only crystallised once and they may do so with or without the solvent of crystallisation. If we record for each solvent type, the number of solvates which has a corresponding unsolvated form and the number of solvates with no unsolvated forms ($N_{solvate-NU}$) we may then be able to calculate the probability of a crystallisation solvent being used ($P_{solvent-cryst}$) as:

$$P_{solvent-cryst} = N_{solvate-NU}/N_{total-15solvates-NU}$$

Then the CSD solvate occurrence can be corrected by dividing by the probability of the solvent being used for crystallisation:

$$O_{solvate} = P_{solvate\text{-}CSD}/P_{solvent\text{-}cryst}$$

Derived CSD occurrences are all given in table 4. As a validation of our method, we plotted the corrected solvate occurrences derived by Nangia and Desiraju against those derived by us in this study (Figure 3). Whilst the correlation of the corrected solvate occurrences is not perfect, there is certainly a clear trend which is encouraging.

Table 2. Number of structures (N) and uncorrected (P) and corrected (O) frequencies of occurrence for the top 15 solvates in the CSD (see Figure 3 in the manuscript).

Solvent	$N_{Solvate}$	$N_{Solv\text{-}unique}$	$N_{Solv\text{-}WU}$	$P_{solvate}$ (%)	$P_{solvent\text{-}cryst}$ (%)	O_{sovate}
Water	8 273	7 599	674	46.35	46.41	0.999
Methanol	1 793	1 674	119	10.05	10.22	0.983
CCl_2H_2	1 298	1 241	57	7.27	7.58	0.959
CCl_3H	1 208	1 146	62	6.77	7.00	0.967
Acetonitrile	739	663	76	4.14	4.05	1.023
Benzene	669	584	85	3.75	3.57	1.051
Acetone	643	565	78	3.60	3.45	1.044
DMF	643	577	66	3.60	3.52	1.022
DMSO	632	518	114	3.54	3.16	1.119
Ethanol	577	532	45	3.23	3.25	0.995
Ethyl Acetate	349	336	13	1.96	2.05	0.953
THF	289	267	22	1.62	1.63	0.993
Toluene	260	236	24	1.46	1.44	1.011
Diethyl Ether	258	251	7	1.45	1.53	0.943
p-Xylene	218	185	33	1.22	1.13	1.081
N Total	17849	16374	1475	100	100	-

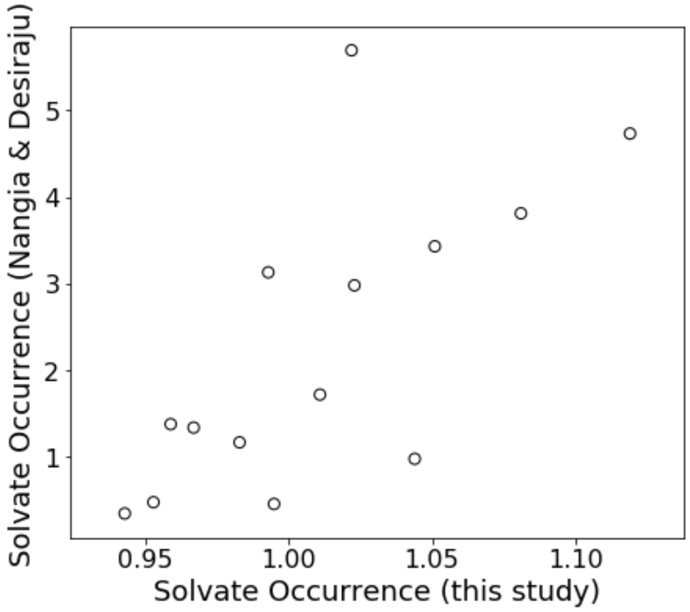

Figure 3. Corrected solvate occurrences as calculated in by Nangia and Desiraju (y-axis) and in this study (x-axis).

4. Calculation of Lattice Energy Gains upon Solvate Formation

Calculation of Lattice energies

Lattice energies are calculated by the following equation, where E_{cryst} is the energy of a crystal, Z is the number of molecules in the unit cell and E_{mol} is the energy of an isolated molecule.

$$E_{latt} = \frac{E_{cryst}}{Z} - E_{mol} \qquad (5)$$

Lattice energy calculations are completed using the Vienna Ab initio Simulation Package (VASP 5.4.4) [6–9]. The approach implemented in VASP used the PBE functional[10] with PAW pseudopotentials[11,12] and the Grimme's van der Waals corrections (d2)[13]. The Monkhorst-Pack approximation[14] was used with a variety of k-point grids to sample the Brillouin zone. Structural relaxations were stopped once the calculated force on every atom was less than 0.003 eV/Å.

Three calculation steps are made to obtain the lattice energy, optimisation of the unit cell allowing for the cell-parameters to relax, reoptimisation of the unit cell with the unit cell parameters fixed (Ecryst) and the optimisation of a single molecule in a fix supercell of 30Å x 30Å x 30Å (Emol).

Input files for the calculation and the supercell for the final step have been generated using the pymatgen[15] and the CSD python modules.

Lattice energy gains for various Solvate:Unsolvated Pairs

We have calculated the lattice energy gains upon solvation on a number of solvate:unsolvated pairs for a rannge of solvents: Water (Hydrates), Toluene, Ethyl Acetate and DMSO.

Table 3. Calculated lattice energies, lattice energy gains and entropy penalty (EP) at 300 K for a number of solvates. All values are in kJ/mol.

Solvent B	**Solvate Refcode**	**A:B Stoi. n:m**	E_{latt}^{AB}	**A Refcode**	nE_{latt}^{A}	**B* Refcode**	mE_{latt}^{B}	ΔE_{latt}^{AB}	**EP**
DMSO	SEDQEZ	1:1	-263.6	SEDQAV	-199.4	DMETSO05	-70.3	6.1	0.4
DMSO	IMIGAO	1:1	-191.7	CELJIN02	-120.8	DMETSO05	-70.3	-0.6	0.4
DMSO	PAKZUZ	1:1	-227.6	PALBAI	-152.9	DMETSO05	-70.3	-4.4	0.4
DMSO	LINLOJ	1:1	-245.6	LINLUP	-165.7	DMETSO05	-70.3	-9.6	0.4
DMSO	ZEPSIZ	1:1	-240.9	ZEPSEV	-159.5	DMETSO05	-70.3	-11.1	0.4
Water	XOHFEH	1:1	-229.9	XOHFAD	-164.3	Ih	-73.8	8.2	0.6
Water	IHUQEJ	1:1	-395.2	IHURUA	-328.3	Ih	-73.8	6.9	0.6
Water	MOKRUA	1:1	-268.1	XEBYUA01	-201.1	Ih	-73.8	6.8	0.6
Water	IKEXOM	1:1	-180.4	VOLZEC	-111.9	Ih	-73.8	5.3	0.6
Water	ENECAD	1:1	-317.7	UGIVAI	-249.1	Ih	-73.8	5.2	0.6
Water	REVMIP	1:1	-207.6	GAYJAT	-138.9	Ih	-73.8	5.1	0.6
Water	YUXGUV	1:1	-198.4	ZZZEEU05	-129.2	Ih	-73.8	4.6	0.6
Water	SAJRUQ	1:1	-240.3	SAJQUP	-171.2	Ih	-73.8	4.7	0.6
Water	XOHDAB	1:1	-188.4	XOHCUU	-118.5	Ih	-73.8	3.9	0.6
Water	ERUHUV	1:1	-207.2	VIFSAF	-136.2	Ih	-73.8	2.8	0.6
Water	NIGZEJ01	1:1	-216.8	POVFIP	-144.7	Ih	-73.8	1.7	0.6
Water	JIXCOI	1:1	-216.8	VIDMAX	-144.6	Ih	-73.8	1.6	0.6
Water	IKUROV	1:1	-205.8	IKURUB	-133.5	Ih	-73.8	1.5	0.6
Water	PAGYUS	1:1	-221.0	ALXANM01	-147.4	Ih	-73.8	0.2	0.6
Water	GLUCMH12	1:1	-276.2	LOWFUA	-199.6	Ih	-73.8	-2.8	0.6
Water	CITARC01	1:1	-262.8	CITRAC11	-183.5	Ih	-73.8	-5.5	0.6

Solvent B	Solvate Refcode	A:B Stoi. n:m	E_{latt}^{AB} (kJ/mol)	A Refcode	nE_{latt}^{A} (kJ/mol)	B* Refcode	mE_{latt}^{B} (kJ/mol)	ΔE_{latt}^{AB} (kJ/mol)	EP
DMSO	KOYPIX	1:2	-315.1	SIFLOI02	-176.0	DMETSO05	-140.6	1.5	0.8
DMSO	MIFXUU	1:2	-297.3	PMELIM	-151.7	DMETSO05	-140.6	-5.0	0.8
DMSO	YINROE	1:2	-386.0	YINSAR01	-231.1	DMETSO05	-140.6	-14.3	0.8
Toluene	NIHMUN	1:1	-284.5	SUPBEK03	-224.7	TOLUEN03	-60.6	0.8	4.5
Toluene	GUNKOP	1:1	-318.1	JEYDEW01	-255.1	TOLUEN03	-60.6	-2.4	4.5
Toluene	PEGJUH	1:1	-253.4	AYUHIM	-186.0	TOLUEN03	-60.6	-6.8	4.5
Toluene	XENMUA	1:1	-216.7	XIWXOR	-149.2	TOLUEN03	-60.6	-6.9	4.5
Toluene	EQUREO	1:1	-225.7	EQURIS	-157.4	TOLUEN03	-60.6	-7.7	4.5
Toluene	XITYEE	1:1	-235.4	ACIJIF	-167.1	TOLUEN03	-60.6	-7.7	4.5
Toluene	OFEDOU	1:1	-268.3	OFEDUA	-199.2	TOLUEN03	-60.6	-8.5	4.5
Toluene	NIHMUN	1:1	-284.1	SUPBEK	-212.1	TOLUEN03	-60.6	-11.4	4.5
Toluene	XEBDOA	1:1	-303.2	XEBDEQ	-226.5	TOLUEN03	-60.6	-16.1	4.5
EthylAcetate	TOHXAQ	1:1	-219.36	TOHVIW	-151.99	AXUCUS	-60.9	-6.5	6.1
EthylAcetate	TUJHIR	1:1	-314.66	TUJHEN	-245.29	AXUCUS	-60.9	-8.5	6.1
EthylAcetate	YINRAQ	1:1	-305.73	YINSAR01	-230.88	AXUCUS	-60.9	-14.0	6.1
EthylAcetate	MODTAB	1:1	-272.84	PUPLES	-193.54	AXUCUS	-60.9	-18.4	6.1

* Since both DMSO, toluene and ice are polymorphic, the lattice energy of the most stable form was used for DMSO and toluene whilst only ice Ih was computed.

5. Statistics on Disorder in Solvates in the CSD

Methods

Subsets were generated using the CSD Python API, searching for structures which contain two chemical units containing a water molecule or solvent molecule (for the top 15 solvates in the CSD). Only organics (only containing H, D, C, N, O, S, P and the halides) with 3D coordinates and must not contain errors, ions and must not be polymeric. These searches were completed twice, once allowing disorder and again not allowing disorder.

Results

The amount of disorder found in solvates varies from 17% for Acetic Acid, to 50% for toluene. Whilst disorder may be present in either the main component or the solvent molecule, common patterns were observed in the solvent disorder. For example, disorder in toluene solvates may be observed as the toluene being flipped, the methyl group rotating, or may be found in two different positions in the crystal structure. In hexane the disorder is due to conformational changes of the hexane, or the hexane molecule may be found in two different positions in the crystal structure. Acetic acid may have disorder around its methyl group, may be flipped or rotated slightly.

Table 4. Solvents which form more than 200 solvates and how much disorder is found in these subsets.

Solvent	Total	No Disorder	% Disordered
Acetic Acid	237	196	17%
Water	11612	9480	18%
Methanol	2567	1999	22%
DMF	935	713	24%
Pyridine	260	198	24%
Benzene	977	726	26%
Acetone	960	711	26%
1,4-dioxane	436	313	28%
Diethylether	328	233	29%
Acetonitrile	1165	815	30%
Ethyl Acetate	572	388	32%
Ethanol	996	669	33%
Isopropanol	201	132	34%
DMSO	1223	718	41%
DCM	2498	1449	42%
Chloroform	2528	1412	44%
THF	521	290	44%
Hexane	349	179	49%
Toluene	634	318	50%

References

(1) Nyman, J.; Day, G. M. Static and Lattice Vibrational Energy Differences between Polymorphs. *CrystEngComm* **2015**, *17* (28), 5154–5165. https://doi.org/10.1039/C5CE00045A.

(2) Prediction and Observation of Isostructurality Induced by Solvent Incorporation in Multicomponent Crystals | Journal of the American Chemical Society https://pubs.acs.org/doi/abs/10.1021/ja065845a (accessed Nov 3, 2019).

(3) Braun, D. E.; Karamertzanis, P. G.; Price, S. L. Which, If Any, Hydrates Will Crystallise? Predicting Hydrate Formation of Two Dihydroxybenzoic Acids. *Chem. Commun.* **2011**, *47* (19), 5443–5445. https://doi.org/10.1039/C1CC10762C.

(4) Cruz-Cabeza, A. J.; Day, G. M.; Jones, W. Towards Prediction of Stoichiometry in Crystalline Multicomponent Complexes. *Chem. – Eur. J.* **2008**, *14* (29), 8830–8836. https://doi.org/10.1002/chem.200800668.

(5) Cruz-Cabeza, A. J.; Karki, S.; Fábián, L.; Friščić, T.; Day, G. M.; Jones, W. Predicting Stoichiometry and Structure of Solvates. *Chem. Commun.* **2010**, *46* (13), 2224–2226. https://doi.org/10.1039/B922955H.

(6) Kresse, G.; Hafner, J. Ab Initio Molecular Dynamics for Liquid Metals. *Phys. Rev. B* **1993**, *47* (1), 558–561. https://doi.org/10.1103/PhysRevB.47.558.

(7) Kresse, G.; Hafner, J. Ab Initio Molecular-Dynamics Simulation of the Liquid-Metal–Amorphous-Semiconductor Transition in Germanium. *Phys. Rev. B* **1994**, *49* (20), 14251–14269. https://doi.org/10.1103/PhysRevB.49.14251.

(8) Kresse, G.; Furthmüller, J. Efficiency of Ab-Initio Total Energy Calculations for Metals and Semiconductors Using a Plane-Wave Basis Set. *Comput. Mater. Sci.* **1996**, *6* (1), 15–50. https://doi.org/10.1016/0927-0256(96)00008-0.

(9) Kresse, G.; Furthmüller, J. Efficient Iterative Schemes for Ab Initio Total-Energy Calculations Using a Plane-Wave Basis Set. *Phys. Rev. B* **1996**, *54* (16), 11169–11186. https://doi.org/10.1103/PhysRevB.54.11169.

(10) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865–3868. https://doi.org/10.1103/PhysRevLett.77.3865.

(11) Blöchl, P. E. Projector Augmented-Wave Method. *Phys. Rev. B* **1994**, *50* (24), 17953–17979. https://doi.org/10.1103/PhysRevB.50.17953.

(12) Kresse, G.; Joubert, D. From Ultrasoft Pseudopotentials to the Projector Augmented-Wave Method. *Phys. Rev. B* **1999**, *59* (3), 1758–1775. https://doi.org/10.1103/PhysRevB.59.1758.

(13) Grimme, S. Semiempirical GGA-Type Density Functional Constructed with a Long-Range Dispersion Correction. *J. Comput. Chem.* **2006**, *27* (15), 1787–1799. https://doi.org/10.1002/jcc.20495.

(14) Monkhorst, H. J.; Pack, J. D. Special Points for Brillouin-Zone Integrations. *Phys. Rev. B* **1976**, *13* (12), 5188–5192. https://doi.org/10.1103/PhysRevB.13.5188.

(15) Ong, S. P.; Richards, W. D.; Jain, A.; Hautier, G.; Kocher, M.; Cholia, S.; Gunter, D.; Chevrier, V. L.; Persson, K. A.; Ceder, G. Python Materials Genomics (Pymatgen): A Robust, Open-Source Python Library for Materials Analysis. *Comput. Mater. Sci.* **2013**, *68*, 314–319. https://doi.org/10.1016/j.commatsci.2012.10.028.