CrystalClear: An open, modular protocol for predicting molecular crystal growth from solution – Supporting Information

Peter R. Spackman, Alvin J. Walisinghe, Michael W. Anderson, Julian D. Gale

April 19, 2023

S1 CrystalGrower energy model

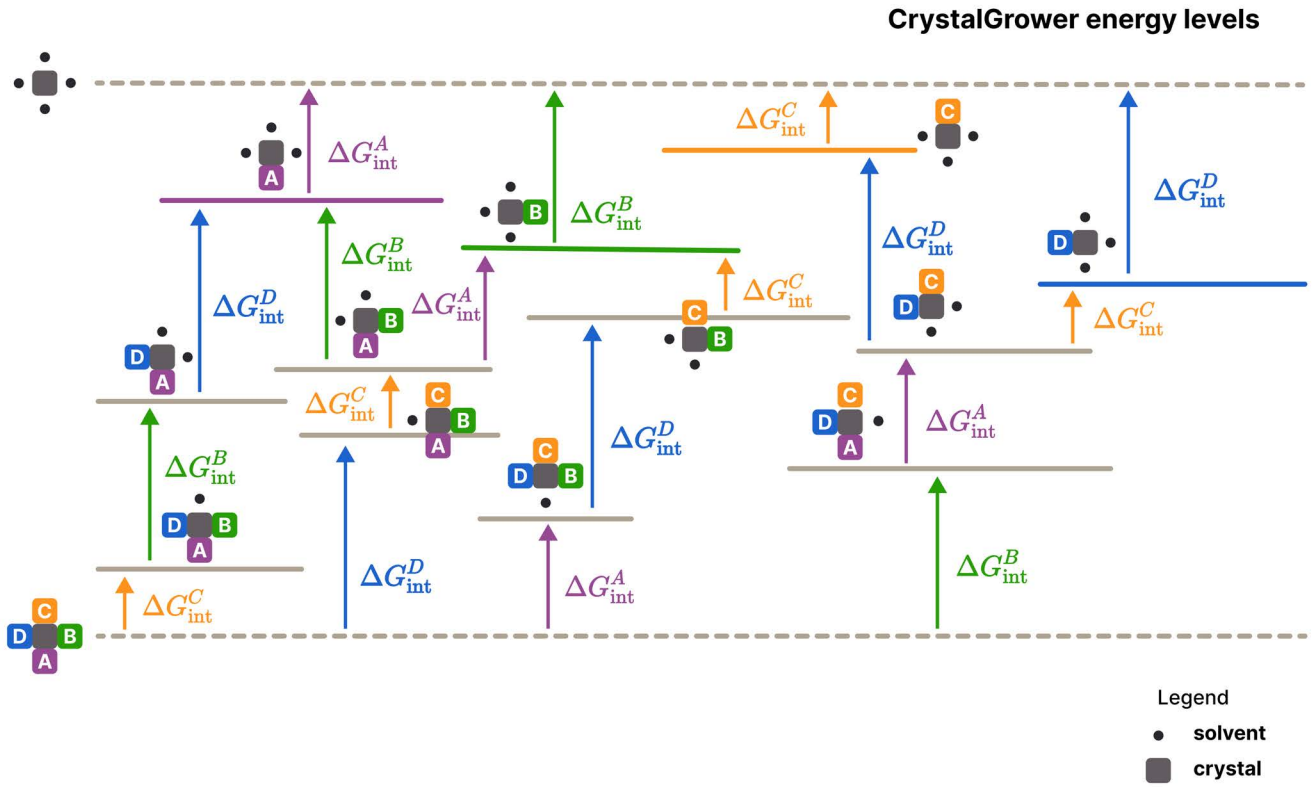


Figure S1: Schematic depiction of the energy levels and their transitions for the CrystalGrower model, for a fictitious 2D system with 4 unique interactions A, B, C and D. Each transition involves replacement of solute with solvent (or *vice versa*) with a subsequent energy change, ΔG_{int}. Each of the energy levels in grey in this figure would correspond to the CrystalGrower energy levels given in green in Figure 1 of the main text.

S2 Case studies

S2.1 Ibuprofen (CSD ref. IBPRAC01)

$$U_{\text{latt}} = -132.981 \text{ kJ mol}^{-1}$$

$$U_{\text{rot}} = -35.368 \text{ kJ mol}^{-1}$$

$$U_{\text{trans}} = -43.573 \text{ kJ mol}^{-1}$$

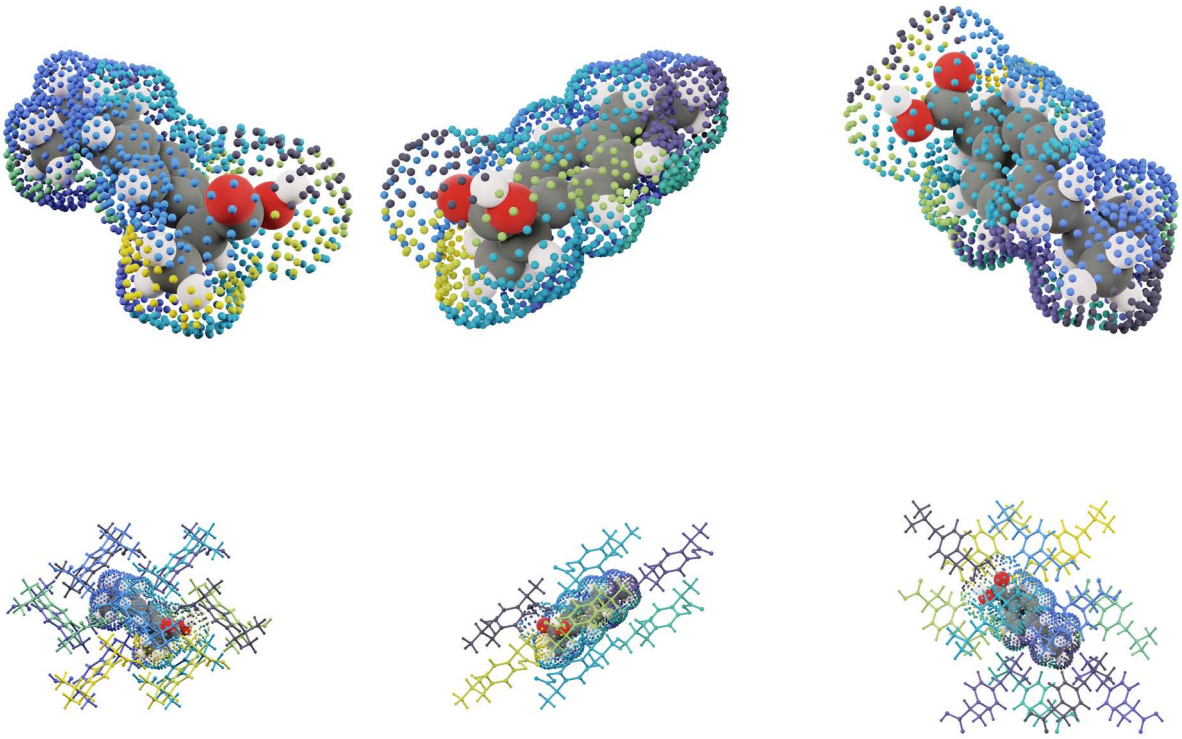


Figure S2: Coulomb solvent accessible surfaces of ibuprofen in acetonitrile (top), coloured by the assigned neighbour (bottom), with the corresponding neighbours displayed as ball-and-stick models. From left to right: viewed down the crystallographic **a**, **b** and **c** axes. Predicted energies for the corresponding dimers and solvent surface partitioning depicted here are given in Table S3, for Tables S1, S2 and S4 the solvation surfaces and partitions will differ slightly, but the dimer interactions in the crystal remain the same.

S2.1.1 Solvent: ethanol

$$\Delta G_{\text{solv}} = -41.844 \text{ kJ mol}^{-1}$$

Table S1: Interaction energies for ibuprofen in ethanol. r_n is the nearest-atom distance between the pair of molecules, r_c is the centroid-centroid distance, U_{crys}^{AB} is the dimer interaction energy in the crystal, $\Delta G_{\text{solv}}^{AB}$ and $\Delta G_{\text{solv}}^{BA}$ are the partitioned solvation free energy from A to B and B to A, respectively, and $\Delta G_{\text{solv}}^{\text{tot}}$ is their antisymmetric sum (see Section 4.5, equation 19), ΔG_{nn} is the dimer energy from non-nearest neighbours that has been assigned to this interaction, and ΔG_{int} is the net interaction energy ($U_{\text{crys}}^{AB} + \Delta G_{\text{solv}}^{\text{tot}} + \Delta G_{nn}$). All energies are provided in kJ mol^{-1}, and distances are in Å. Colours are provided for each row to match the corresponding dimers and surface points assigned in Figure S2.

Colour	r_n	r_c	Symmetry Operation	U_{crys}^{AB}	$\Delta G_{\text{solv}}^{AB}$	$\Delta G_{\text{solv}}^{BA}$	$\Delta G_{\text{solv}}^{\text{tot}}$	ΔG_{nn}	ΔG_{int}
	1.64	11.16	1-x,1-y,1-z	-72.85	-19.30	-19.30	-38.61	-1.50	35.75
	2.26	9.96	-x,-1/2+y,3/2-z	-7.60	0.45	-2.52	-0.58	-0.96	7.98
	2.26	9.96	-x,1/2+y,3/2-z	-7.60	-2.52	0.45	-3.55	-0.95	5.00
	2.32	10.12	-x,-1-y,1-z	-5.93	-1.16	-1.16	-2.32	-0.98	4.59
	2.41	6.28	x,-1/2-y,-1/2+z	-20.93	-4.07	-3.40	-7.80	-1.09	14.21
	2.41	6.28	x,-1/2-y,1/2+z	-20.93	-3.40	-4.07	-7.14	-1.69	15.48
	2.46	7.49	1-x,-y,1-z	-26.28	-10.83	-10.83	-21.66	-2.15	6.77
	2.49	6.83	x,1/2-y,-1/2+z	-22.82	-2.94	-3.76	-6.29	-1.06	17.59
	2.49	6.83	x,1/2-y,1/2+z	-22.82	-3.76	-2.94	-7.11	-1.34	17.05
	2.50	6.96	-x,-y,1-z	-11.85	-1.59	-1.59	-3.17	-1.11	9.79
	2.84	7.82	x,-1+y,z	-8.33	0.33	-0.75	0.12	-1.32	9.77
	2.84	7.82	x,1+y,z	-8.33	-0.75	0.33	-0.96	-1.09	8.46
	2.87	10.78	1-x,1/2+y,1/2-z	-5.83	-0.35	0.13	-0.46	-1.32	6.69
	2.87	10.78	1-x,-1/2+y,1/2-z	-5.83	0.13	-0.35	0.02	-1.48	7.33

S2.1.2 Solvent: ethyl acetate

$$\Delta G_{\text{solv}} = -36.482 \text{ kJ mol}^{-1}$$

Table S2: Interaction energies for ibuprofen in ethyl acetate. r_n is the nearest-atom distance between the pair of molecules, r_c is the centroid-centroid distance, U_{crys}^{AB} is the dimer interaction energy in the crystal, $\Delta G_{\text{solv}}^{AB}$ and $\Delta G_{\text{solv}}^{BA}$ are the partitioned solvation free energy from A to B and B to A, respectively, and $\Delta G_{\text{solv}}^{\text{tot}}$ is their antisymmetric sum (see Section 4.5, equation 19), ΔG_{nn} is the dimer energy from non-nearest neighbours that has been assigned to this interaction, and ΔG_{int} is the net interaction energy ($U_{\text{crys}}^{AB} + \Delta G_{\text{solv}}^{\text{tot}} + \Delta G_{nn}$). All energies are provided in kJ mol^{-1}, and distances are in Å. Colours are provided for each row to match the corresponding dimers and surface points assigned in Figure S2.

Colour	r_n	r_c	Symmetry Operation	U_{crys}^{AB}	$\Delta G_{\text{solv}}^{AB}$	$\Delta G_{\text{solv}}^{BA}$	$\Delta G_{\text{solv}}^{\text{tot}}$	ΔG_{nn}	ΔG_{int}
	1.64	11.16	1-x,1-y,1-z	-72.85	-7.27	-7.27	-14.54	-1.50	59.82
	2.26	9.96	-x,-1/2+y,3/2-z	-7.60	-0.01	-2.36	-1.20	-0.96	7.36
	2.26	9.96	-x,1/2+y,3/2-z	-7.60	-2.36	-0.01	-3.55	-0.95	5.00
	2.32	10.12	-x,-1-y,1-z	-5.93	-1.16	-1.16	-2.33	-0.98	4.58
	2.41	6.28	x,-1/2-y,-1/2+z	-20.93	-5.23	-3.60	-9.64	-1.09	12.38
	2.41	6.28	x,-1/2-y,1/2+z	-20.93	-3.60	-5.23	-8.01	-1.69	14.61
	2.46	7.49	1-x,-y,1-z	-26.28	-7.94	-7.94	-15.88	-2.15	12.56
	2.49	6.83	x,1/2-y,-1/2+z	-22.82	-4.61	-5.73	-9.78	-1.06	14.10
	2.49	6.83	x,1/2-y,1/2+z	-22.82	-5.73	-4.61	-10.90	-1.34	13.26
	2.50	6.96	-x,-y,1-z	-11.85	-1.74	-1.74	-3.49	-1.11	9.48
	2.84	7.82	x,-1+y,z	-8.33	-0.08	-2.86	-1.55	-1.32	8.10
	2.84	7.82	x,1+y,z	-8.33	-2.86	-0.08	-4.32	-1.09	5.09
	2.87	10.78	1-x,1/2+y,1/2-z	-5.83	-1.75	-0.06	-2.66	-1.32	4.49
	2.87	10.78	1-x,-1/2+y,1/2-z	-5.83	-0.06	-1.75	-0.96	-1.48	6.35

S2.1.3 Solvent: acetonitrile

$$\Delta G_{\text{solv}} = -37.005 \text{ kJ mol}^{-1}$$

Table S3: Interaction energies for ibuprofen in acetonitrile. r_n is the nearest-atom distance between the pair of molecules, r_c is the centroid-centroid distance, U_{crys}^{AB} is the dimer interaction energy in the crystal, $\Delta G_{\text{solv}}^{AB}$ and $\Delta G_{\text{solv}}^{BA}$ are the partitioned solvation free energy from A to B and B to A, respectively, and $\Delta G_{\text{solv}}^{\text{tot}}$ is their antisymmetric sum (see Section 4.5, equation 19), ΔG_{nn} is the dimer energy from non-nearest neighbours that has been assigned to this interaction, and ΔG_{int} is the net interaction energy ($U_{\text{crys}}^{AB} + \Delta G_{\text{solv}}^{\text{tot}} + \Delta G_{nn}$). All energies are provided in kJ mol^{-1}, and distances are in Å. Colours are provided for each row to match the corresponding dimers and surface points assigned in Figure S2.

Colour	r_n	r_c	Symmetry Operation	U_{crys}^{AB}	$\Delta G_{\text{solv}}^{AB}$	$\Delta G_{\text{solv}}^{BA}$	$\Delta G_{\text{solv}}^{\text{tot}}$	ΔG_{nn}	ΔG_{int}
	1.64	11.16	1-x,1-y,1-z	-72.85	-10.59	-10.59	-21.18	-1.50	53.18
	2.26	9.96	-x,-1/2+y,3/2-z	-7.60	0.17	-2.33	-0.91	-0.96	7.64
	2.26	9.96	-x,1/2+y,3/2-z	-7.60	-2.33	0.17	-3.41	-0.95	5.14
	2.32	10.12	-x,-1-y,1-z	-5.93	-1.00	-1.00	-2.00	-0.98	4.91
	2.41	6.28	x,-1/2-y,-1/2+z	-20.93	-5.38	-3.22	-9.68	-1.09	12.33
	2.41	6.28	x,-1/2-y,1/2+z	-20.93	-3.22	-5.38	-7.52	-1.69	15.10
	2.46	7.49	1-x,-y,1-z	-26.28	-8.23	-8.23	-16.46	-2.15	11.97
	2.49	6.83	x,1/2-y,-1/2+z	-22.82	-3.92	-5.52	-8.63	-1.06	15.24
	2.49	6.83	x,1/2-y,1/2+z	-22.82	-5.52	-3.92	-10.24	-1.34	13.92
	2.50	6.96	-x,-y,1-z	-11.85	-1.57	-1.57	-3.14	-1.11	9.83
	2.84	7.82	x,-1+y,z	-8.33	0.05	-2.06	-0.95	-1.32	8.70
	2.84	7.82	x,1+y,z	-8.33	-2.06	0.05	-3.06	-1.09	6.36
	2.87	10.78	1-x,1/2+y,1/2-z	-5.83	-1.44	0.12	-2.10	-1.32	5.04

S2.1.4 Solvent: toluene

$$\Delta G_{\text{solv}} = -34.107 \text{ kJ mol}^{-1}$$

Table S4: Interaction energies for ibuprofen in toluene. r_n is the nearest-atom distance between the pair of molecules, r_c is the centroid-centroid distance, U_{crys}^{AB} is the dimer interaction energy in the crystal, $\Delta G_{\text{solv}}^{AB}$ and $\Delta G_{\text{solv}}^{BA}$ are the partitioned solvation free energy from A to B and B to A, respectively, and $\Delta G_{\text{solv}}^{\text{tot}}$ is their antisymmetric sum (see Section 4.5, equation 19), ΔG_{nn} is the dimer energy from non-nearest neighbours that has been assigned to this interaction, and ΔG_{int} is the net interaction energy ($U_{\text{crys}}^{AB} + \Delta G_{\text{solv}}^{\text{tot}} + \Delta G_{nn}$). All energies are provided in kJ/mol, and distances are in Å. Colours are provided for each row to match the corresponding dimers and surface points assigned in Figure S2.

Colour	r_n	r_c	Symmetry Operation	U_{crys}^{AB}	$\Delta G_{\text{solv}}^{AB}$	$\Delta G_{\text{solv}}^{BA}$	$\Delta G_{\text{solv}}^{\text{tot}}$	ΔG_{nn}	ΔG_{int}
	1.64	11.16	1-x,1-y,1-z	-72.85	-5.24	-5.24	-10.49	-1.50	63.87
	2.26	9.96	-x,-1/2+y,3/2-z	-7.60	-0.37	-2.30	-1.71	-0.96	6.85
	2.26	9.96	-x,1/2+y,3/2-z	-7.60	-2.30	-0.37	-3.64	-0.95	4.91
	2.32	10.12	-x,-1-y,1-z	-5.93	-1.30	-1.30	-2.60	-0.98	4.31
	2.41	6.28	x,-1/2-y,-1/2+z	-20.93	-5.79	-3.97	-10.67	-1.09	11.35
	2.41	6.28	x,-1/2-y,1/2+z	-20.93	-3.97	-5.79	-8.85	-1.69	13.77
	2.46	7.49	1-x,-y,1-z	-26.28	-6.58	-6.58	-13.15	-2.15	15.28
	2.49	6.83	x,1/2-y,-1/2+z	-22.82	-4.10	-5.68	-8.98	-1.06	14.90
	2.49	6.83	x,1/2-y,1/2+z	-22.82	-5.68	-4.10	-10.57	-1.34	13.59
	2.50	6.96	-x,-y,1-z	-11.85	-1.95	-1.95	-3.90	-1.11	9.06
	2.84	7.82	x,-1+y,z	-8.33	-0.40	-2.65	-1.93	-1.32	7.72
	2.84	7.82	x,1+y,z	-8.33	-2.65	-0.40	-4.18	-1.09	5.24
	2.87	10.78	1-x,1/2+y,1/2-z	-5.83	-1.43	-0.26	-2.27	-1.32	4.88
	2.87	10.78	1-x,-1/2+y,1/2-z	-5.83	-0.26	-1.43	-1.10	-1.48	6.21

Table S5: Predicted and fitted CrystalGrower energies for ibuprofen, including the selected energy range for fitting. Final 'Fitted' values are the average of those yielding the 5 closest matching shapes. The 'Colour(s)' column indicates the corresponding colour values from Tables S1-S4 and Figure S2, with those interactions which have been grouped having multiple corresponding colours.

	Colour(s)	Interaction Label	r_c (Å)	Interaction Energy (kcal/mol)						
				Ethanol	Ethylacetate	Acetonitrile	Toluene	Search Range		
								Start	End	Step size
Predicted		A	6.28	3.65	3.33	3.38	3.10	2.00	5.00	1.00
		B	6.83	4.24	3.37	3.59	3.50	2.00	6.00	1.00
		C	6.96	2.44	2.37	2.45	2.66	1.00	5.00	1.00
		D	7.49	1.72	3.10	2.96	1.95	0.50	4.50	1.00
		E	7.82	2.28	1.68	1.90	1.65	0.50	4.50	1.00
		F	9.96	1.65	1.58	1.63	1.51	1.80	1.80	N/A
		G	10.12	1.20	1.20	1.27	1.13	1.30	1.30	N/A
		H	10.78	1.78	1.40	1.51	1.31	1.60	1.60	N/A
		I	11.16	8.64	14.40	12.81	15.37	5.00	15.00	2.00
Fitted		A	6.28	4.60	4.40	4.80	4.80			
		B	6.83	4.40	4.00	5.60	5.80			
		C	6.96	2.20	2.40	1.00	1.20			
		D	7.49	2.10	4.10	3.90	3.90			
		E	7.82	3.70	3.70	4.10	4.10			
		F	9.96	1.80	1.80	1.80	1.80			
		G	10.12	1.30	1.30	1.30	1.30			
		H	10.78	1.60	1.60	1.60	1.60			
		I	11.16	5.80	9.40	12.20	13.00			

S2.2 Adipic Acid (CSD ref. ADIPAC)

Table S6: Interaction energies for adipic acid in water. r_n is the nearest-atom distance between the pair of molecules, r_c is the centroid-centroid distance, U_{crys}^{AB} is the dimer interaction energy in the crystal (the first two rows are subsequently modified based on the pK_a), $\Delta G_{\text{solv}}^{AB}$ and $\Delta G_{\text{solv}}^{BA}$ are the partitioned solvation free energy from A to B and B to A, respectively, and $\Delta G_{\text{solv}}^{\text{tot}}$ is their antisymmetric sum (see Section 4.5, equation 19), ΔG_{nn} is the dimer energy from non-nearest neighbours that has been assigned to this interaction, and ΔG_{int} is the net interaction energy ($U_{\text{crys}}^{AB} + \Delta G_{\text{solv}}^{\text{tot}} + \Delta G_{nn}$). All energies are provided in kJ mol^{-1}, and distances are in Å. Colours are provided for each row to match the corresponding dimers and surface points assigned in Figure S3.

Colour	r_n	r_c	Symmetry Operation	U_{crys}^{AB}	$\Delta G_{\text{solv}}^{AB}$	$\Delta G_{\text{solv}}^{BA}$	$\Delta G_{\text{solv}}^{\text{tot}}$	ΔG_{nn}	ΔG_{int}
	1.68	10.01	1+x,y,z	-76.50	-21.75	-21.75	-43.50	-0.67	33.68
	1.68	10.01	-1+x,y,z	-76.50	-21.75	-21.75	-43.50	-0.67	33.68
	2.33	5.15	x,-1+y,z	-8.19	-2.42	-2.42	-4.85	-0.84	4.18
	2.33	5.15	x,1+y,z	-8.19	-2.42	-2.42	-4.85	-0.84	4.18
	2.40	5.65	-x,-1/2+y,3/2-z	-15.62	1.05	-8.04	-2.45	-0.74	13.92
	2.40	5.65	-x,1/2+y,3/2-z	-15.62	-8.04	1.05	-11.54	-0.92	5.01
	2.40	5.65	-x,-1/2+y,5/2-z	-15.62	-8.04	1.05	-11.54	-0.92	5.01
	2.40	5.65	-x,1/2+y,5/2-z	-15.62	1.05	-8.04	-2.45	-0.74	13.92
	2.75	7.67	-1-x,1/2+y,3/2-z	-11.62	-0.73	-0.05	-1.13	-1.37	11.87
	2.75	7.67	-1-x,-1/2+y,3/2-z	-11.62	-0.05	-0.73	-0.45	-1.05	12.23
	2.75	7.67	1-x,-1/2+y,5/2-z	-11.62	-0.73	-0.05	-1.13	-1.37	11.87
	2.75	7.67	1-x,1/2+y,5/2-z	-11.62	-0.05	-0.73	-0.45	-1.05	12.23
	3.52	11.26	1+x,-1+y,z	-2.33	0.25	0.25	0.50	-0.87	3.71
	3.52	11.26	-1+x,1+y,z	-2.33	0.25	0.25	0.50	-0.87	3.71

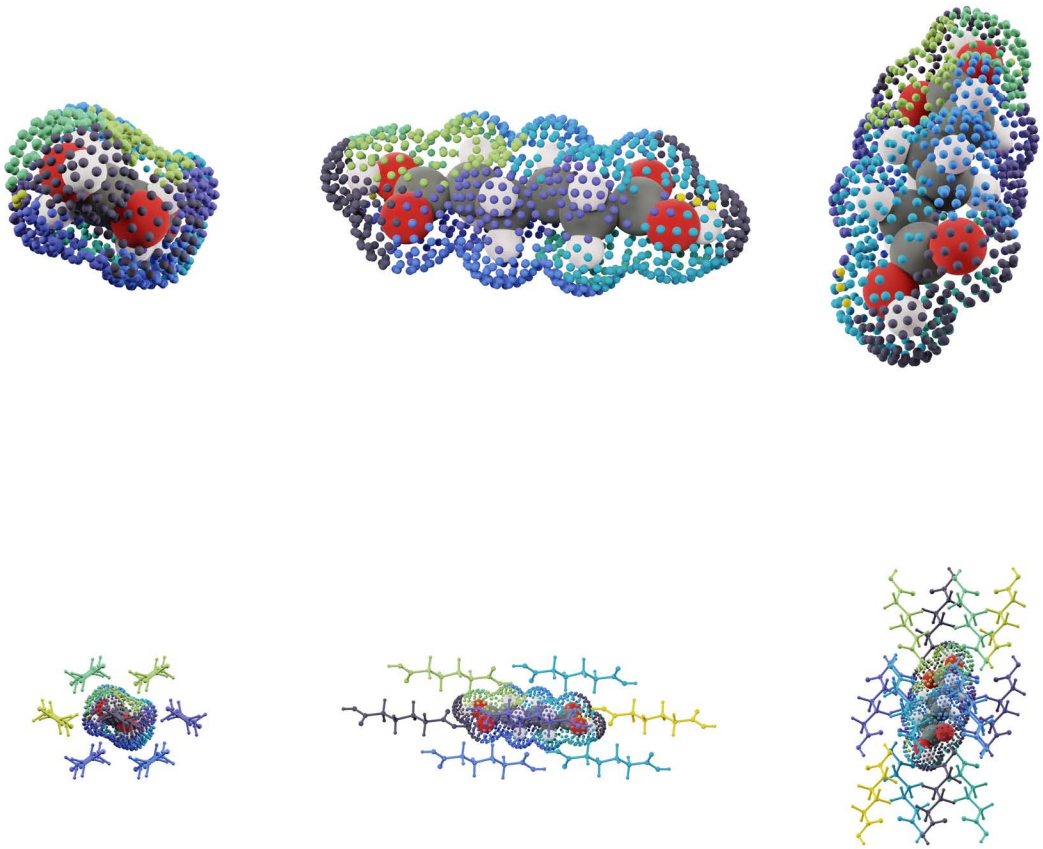


Figure S3: Coulomb solvent accessible surfaces of adipic acid (top), coloured by the assigned neighbour (bottom). From left to right: viewed down the crystallographic **a**, **b** and **c** axes. Predicted energies for the corresponding dimers and solvent surface partitioning depicted here are given in Table S6.

Table S7: Predicted and fitted CrystalGrower energies for adipic acid, including the selected energy range for fitting. Final 'Fitted' values are the average of those yielding the 5 closest matching shapes. The 'Colour(s)' column indicates the corresponding colour values from Table S6 and Figure S3, with those interactions which have been grouped having multiple corresponding colours

Colour(s)	Interaction Label	r_c (Å)	Interaction Energy (kcal/mol)				
			Predicted	Search Range			Fitted
				Start	End	Step size	
	A	5.15	1.30	1.00	4.00	0.50	3.50
	B	5.65	2.82	1.00	4.00	0.50	3.00
	C	7.67	3.13	1.00	4.00	0.50	1.75
	D	10.01	1.06	0.50	2.00	0.50	0.85
	E	11.26	0.95	0.00	2.00	0.20	0.20

S2.2.1 Experimental Crystallisation

Adipic acid (0.3038 g, Sigma, 99%) was dissolved in 15 ml of warm de-ionised water (MilliQ). The supersaturated solution was filtered (0.2 μm) and left to undergo evaporative crystallisation ($\sigma = 0.1$, temperature = 20 °C). Crystals formed were removed from solution after a period of 72 hours and visualised under an optical microscope. The optical micrograph was further used in the visual recreation of a electronic shape (STL) file.

S2.3 ROY

S2.3.1 R polymorph (CSD ref. QAXMEH26, solvent: benzyl alcohol)

Table S8: Interaction energies for ROY polymorph R in benzyl alcohol. r_n is the nearest-atom distance between the pair of molecules, r_c is the centroid-centroid distance, U_{crys}^{AB} is the dimer interaction energy in the crystal, $\Delta G_{\text{solv}}^{AB}$ and $\Delta G_{\text{solv}}^{BA}$ are the partitioned solvation free energy from A to B and B to A, respectively, and $\Delta G_{\text{solv}}^{\text{tot}}$ is their antisymmetric sum (see Section 4.5, equation 19), ΔG_{nn} is the dimer energy from non-nearest neighbours that has been assigned to this interaction, and ΔG_{int} is the net interaction energy ($U_{\text{crys}}^{AB} + \Delta G_{\text{solv}}^{\text{tot}} + \Delta G_{nn}$). All energies are provided in kJ/mol, and distances are in Å. Colours are provided for each row to match the corresponding dimers and surface points assigned in Figure S4.

Colour	r_n	r_c	Symmetry Operation	U_{crys}^{AB}	$\Delta G_{\text{solv}}^{AB}$	$\Delta G_{\text{solv}}^{BA}$	$\Delta G_{\text{solv}}^{\text{tot}}$	ΔG_{nn}	ΔG_{int}
	2.38	11.91	x,y,-1+z	-5.81	-1.65	-2.93	-3.94	-0.20	2.07
	2.38	11.91	x,y,1+z	-5.81	-2.93	-1.65	-5.21	0.17	0.42
	2.55	10.44	-x,2-y,1-z	-26.90	-5.08	-5.08	-10.17	-0.10	16.83
	2.55	6.54	1-x,1-y,1-z	-32.10	-5.77	-5.77	-11.53	-0.52	21.08
	2.59	10.15	2-x,-y,2-z	-5.90	-1.92	-1.92	-3.83	-0.26	2.33
	2.63	8.06	-1+x,1+y,z	-20.20	-6.57	-6.52	-13.11	-0.01	7.10
	2.63	8.06	1+x,-1+y,z	-20.20	-6.52	-6.57	-13.07	-0.45	7.58
	2.71	6.94	-x,1-y,2-z	-36.25	-3.19	-3.19	-6.39	-0.55	30.41
	2.81	8.98	-x,1-y,1-z	-19.34	-0.87	-0.87	-1.73	-0.53	18.14
	2.85	9.17	1-x,-y,1-z	-3.76	-2.59	-2.59	-5.18	-0.53	-0.89
	2.90	7.79	x,-1+y,z	-6.17	-1.17	-1.43	-2.47	-1.19	4.90
	2.90	7.79	x,1+y,z	-6.17	-1.43	-1.17	-2.73	-0.43	3.88
	3.22	6.54	1-x,-y,2-z	-29.68	-2.65	-2.65	-5.30	-0.87	25.26
	3.34	6.93	1-x,1-y,2-z	-40.22	0.20	0.20	0.40	0.07	40.56

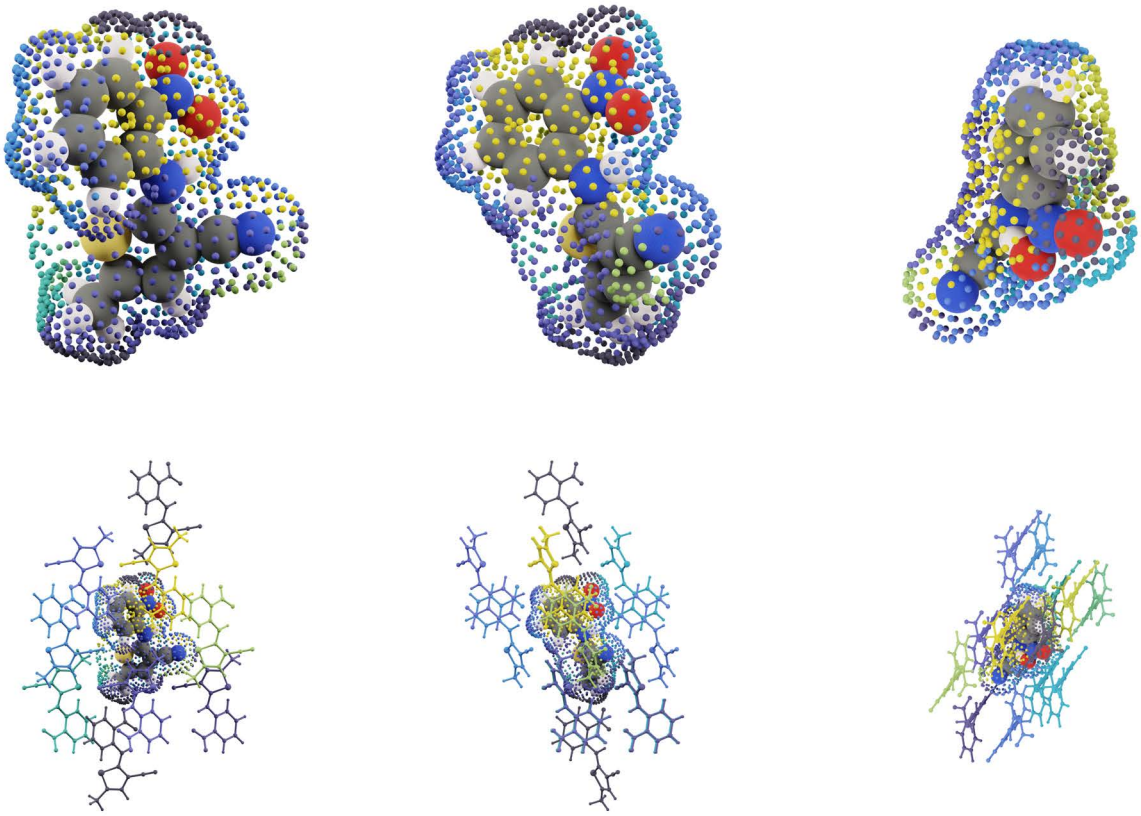


Figure S4: Coulomb solvent accessible surfaces of ROY-R (top), coloured by the assigned neighbour (bottom). From left to right: viewed down the crystallographic **a**, **b** and **c** axes. Predicted energies for the corresponding dimers and solvent surface partitioning depicted here are given in Table S8.

S2.3.2 Y polymorph (CSD ref. QAXMEH25, solvent: ethanol)

Table S9: Interaction energies for ROY polymorph Y in ethanol. r_n is the nearest-atom distance between the pair of molecules, r_c is the centroid-centroid distance, U_{crys}^{AB} is the dimer interaction energy in the crystal, $\Delta G_{\text{solv}}^{AB}$ and $\Delta G_{\text{solv}}^{BA}$ are the partitioned solvation free energy from A to B and B to A, respectively, and $\Delta G_{\text{solv}}^{\text{tot}}$ is their antisymmetric sum (see Section 4.5, equation 19), ΔG_{nn} is the dimer energy from non-nearest neighbours that has been assigned to this interaction, and ΔG_{int} is the net interaction energy ($U_{\text{crys}}^{AB} + \Delta G_{\text{solv}}^{\text{tot}} + \Delta G_{nn}$). All energies are provided in kJ/mol, and distances are in Å. Colours are provided for each row to match the corresponding dimers and surface points assigned in Figure S5.

Colour	r_n	r_c	Symmetry Operation	U_{crys}^{AB}	$\Delta G_{\text{solv}}^{AB}$	$\Delta G_{\text{solv}}^{BA}$	$\Delta G_{\text{solv}}^{\text{tot}}$	ΔG_{nn}	ΔG_{int}
	2.30	7.32	1/2+x,3/2-y,1/2+z	-17.52	-5.96	-4.15	-11.02	-1.64	8.15
	2.30	7.32	-1/2+x,3/2-y,-1/2+z	-17.52	-4.15	-5.96	-9.20	-2.38	10.70
	2.51	8.54	x,y,1+z	-8.63	-4.14	-5.46	-8.95	-1.19	0.88
	2.51	8.54	x,y,-1+z	-8.63	-5.46	-4.14	-10.26	-1.81	0.19
	2.64	7.47	-1/2+x,3/2-y,1/2+z	-31.45	-5.43	-0.15	-8.22	-2.62	25.85
	2.64	7.47	1/2+x,3/2-y,-1/2+z	-31.45	-0.15	-5.43	-2.94	-1.82	30.33
	2.65	8.50	-1+x,y,z	-10.72	-4.07	-1.22	-6.71	-2.38	6.39
	2.65	8.50	1+x,y,z	-10.72	-1.22	-4.07	-3.86	-1.17	8.03
	2.70	9.17	-x,1-y,1-z	-3.88	-2.07	-2.07	-4.14	-2.01	1.75
	2.78	11.67	2-x,1-y,-z	-5.27	-2.05	-2.05	-4.10	-1.49	2.66
	2.78	12.23	-1+x,y,1+z	-8.04	0.16	-2.83	-1.17	-2.10	8.96
	2.78	12.23	1+x,y,-1+z	-8.04	-2.83	0.16	-4.16	-1.69	5.57
	2.81	5.83	1-x,1-y,1-z	-15.50	-5.96	-5.96	-11.93	-0.97	4.54
	3.05	9.98	3/2-x,-1/2+y,1/2-z	-6.24	-0.72	-1.45	-1.81	-1.44	5.86
	3.05	9.98	3/2-x,1/2+y,1/2-z	-6.24	-1.45	-0.72	-2.54	-1.96	5.66
	3.06	6.10	1-x,1-y,-z	-41.18	-0.63	-0.63	-1.26	-2.15	42.07

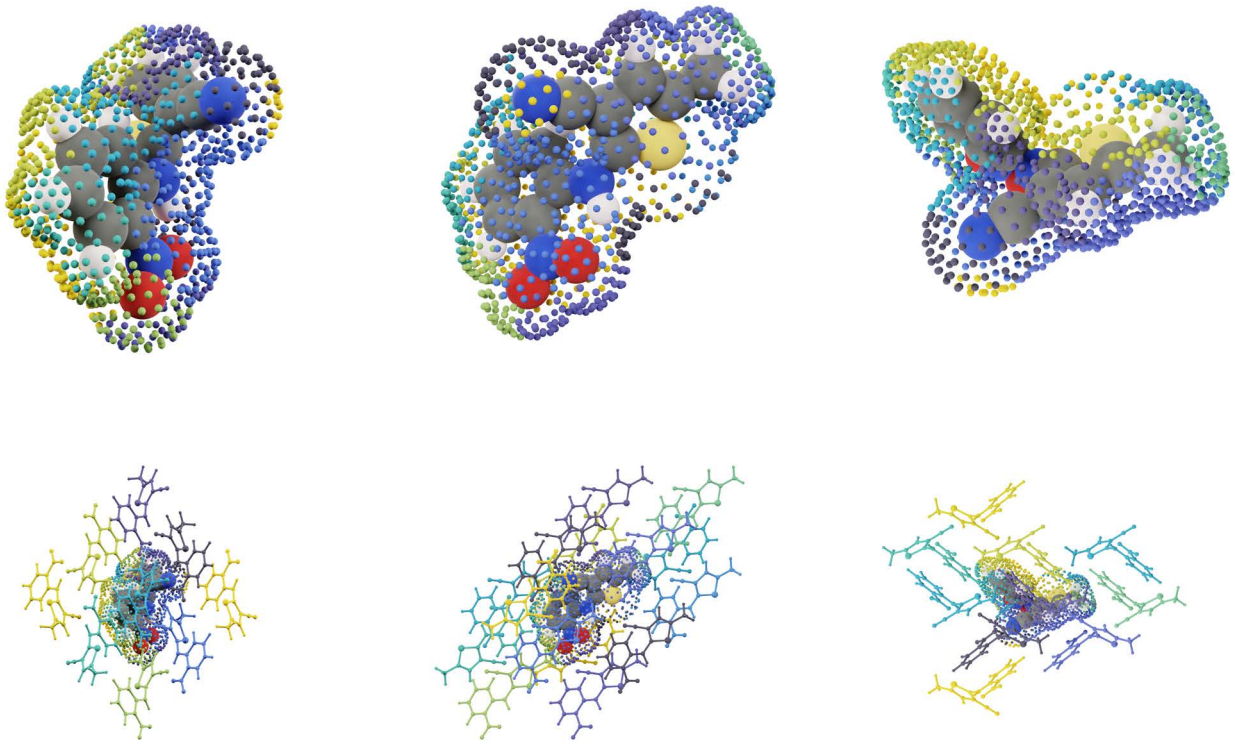


Figure S5: Coulomb solvent accessible surfaces of ROY-Y (top), coloured by the assigned neighbour (bottom). From left to right: viewed down the crystallographic **a**, **b** and **c** axes. Predicted energies for the corresponding dimers and solvent surface partitioning depicted here are given in Table S9.

S2.3.3 YT04 polymorph (CSD ref. QAXMEH12, solvent: ethanol)

Table S10: Interaction energies for ROY polymorph YT04 in ethanol. r_n is the nearest-atom distance between the pair of molecules, r_c is the centroid-centroid distance, U^{AB}_{crys} is the dimer interaction energy in the crystal, $\Delta G^{AB}_{\text{solv}}$ and $\Delta G^{BA}_{\text{solv}}$ are the partitioned solvation free energy from A to B and B to A, respectively, and $\Delta G^{\text{tot}}_{\text{solv}}$ is their antisymmetric sum (see Section 4.5, equation 19), ΔG_{nn} is the dimer energy from non-nearest neighbours that has been assigned to this interaction, and ΔG_{int} is the net interaction energy ($U^{AB}_{\text{crys}} + \Delta G^{\text{tot}}_{\text{solv}} + \Delta G_{nn}$). All energies are provided in kJ/mol, and distances are in Å. Colours are provided for each row to match the corresponding dimers and surface points assigned in Figure S6.

Colour	r_n	r_c	Symmetry Operation	U^{AB}_{crys}	$\Delta G^{AB}_{\text{solv}}$	$\Delta G^{BA}_{\text{solv}}$	$\Delta G^{\text{tot}}_{\text{solv}}$	ΔG_{nn}	ΔG_{int}
	2.41	7.19	3/2-x,1/2+y,1/2-z	-30.64	-6.42	-5.86	-12.56	-0.42	18.50
	2.41	7.19	3/2-x,-1/2+y,1/2-z	-30.64	-5.86	-6.42	-12.00	0.67	17.97
	2.45	6.62	1-x,1-y,-z	-28.16	-11.25	-11.25	-22.49	-0.11	5.78
	2.53	12.31	x,y,-1+z	-7.60	0.55	-2.70	-0.52	-1.07	8.16
	2.53	12.31	x,y,1+z	-7.60	-2.70	0.55	-3.78	0.51	3.32
	2.53	6.99	-1/2+x,3/2-y,-1/2+z	-20.87	-4.86	-1.26	-7.92	-0.45	13.40
	2.53	6.99	1/2+x,3/2-y,1/2+z	-20.87	-1.26	-4.86	-4.32	0.58	15.96
	2.70	7.74	1-x,1-y,1-z	-37.06	-5.46	-5.46	-10.93	0.94	25.19
	2.87	7.27	1/2-x,-1/2+y,1/2-z	-8.26	-3.59	-0.96	-5.87	0.86	1.53
	2.87	7.27	1/2-x,1/2+y,1/2-z	-8.26	-0.96	-3.59	-3.24	-0.08	5.10
	2.93	8.41	1/2+x,3/2-y,-1/2+z	-17.92	-1.62	0.81	-2.02	-1.57	17.47
	2.93	8.41	-1/2+x,3/2-y,1/2+z	-17.92	0.81	-1.62	0.40	0.05	18.28
	3.19	8.23	-1+x,y,z	-5.95	-0.17	-1.65	-1.08	-0.01	4.88
	3.19	8.23	1+x,y,z	-5.95	-1.65	-0.17	-2.56	-0.62	4.01
	3.58	12.28	-x,1-y,1-z	-1.31	-0.10	-0.10	-0.19	0.18	0.94
	3.69	10.61	1-x,2-y,1-z	-1.42	-0.48	-0.48	-0.96	0.03	0.42

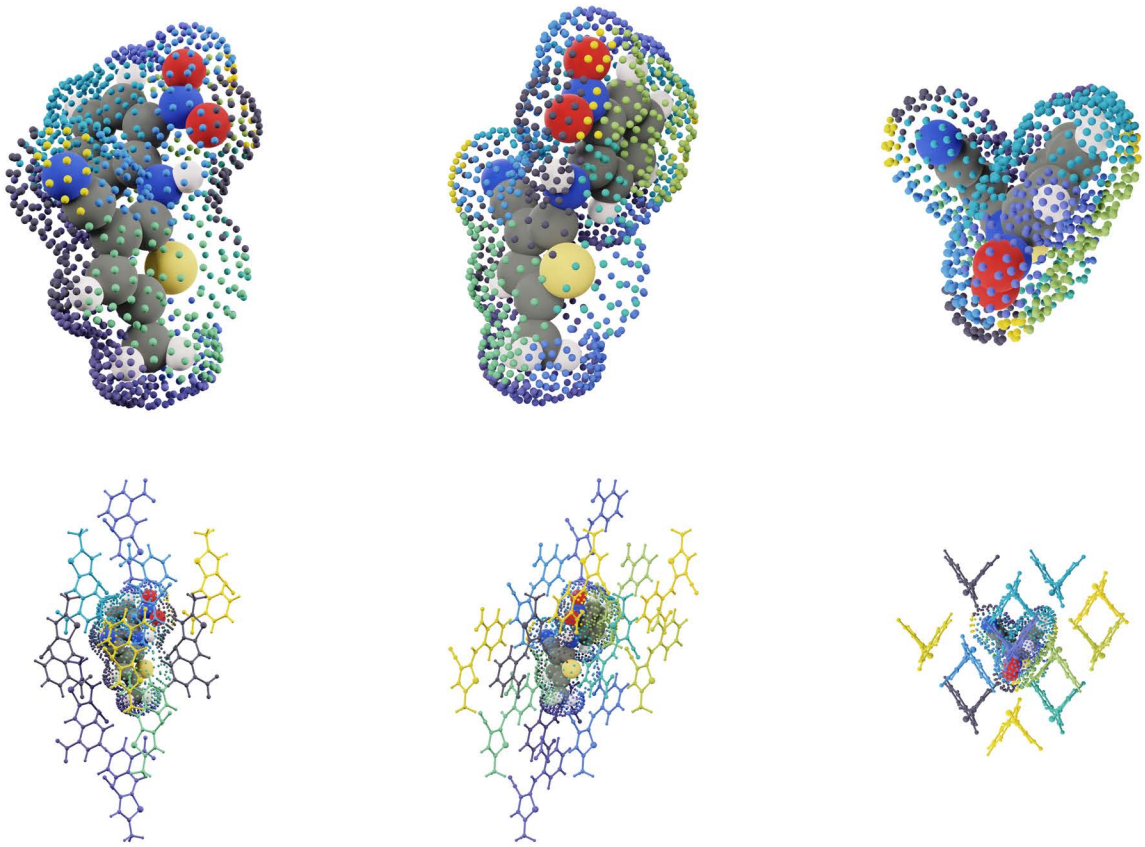


Figure S6: Coulomb solvent accessible surfaces of ROY-YT04 (top), coloured by the assigned neighbour (bottom). From left to right: viewed down the crystallographic **a**, **b** and **c** axes. Predicted energies for the corresponding dimers and solvent surface partitioning depicted here are given in Table S10.

S2.3.4 OP polymorph (CSD ref. QAXMEH27, solvent: methanol)

Table S11: Interaction energies for ROY polymorph OP in methanol. r_n is the nearest-atom distance between the pair of molecules, r_c is the centroid-centroid distance, U^{AB}_{crys} is the dimer interaction energy in the crystal, $\Delta G^{AB}_{\text{solv}}$ and $\Delta G^{BA}_{\text{solv}}$ are the partitioned solvation free energy from A to B and B to A, respectively, and $\Delta G^{\text{tot}}_{\text{solv}}$ is their antisymmetric sum (see Section 4.5, equation 19), ΔG_{nn} is the dimer energy from non-nearest neighbours that has been assigned to this interaction, and ΔG_{int} is the net interaction energy ($U^{AB}_{\text{crys}} + \Delta G^{\text{tot}}_{\text{solv}} + \Delta G_{nn}$). All energies are provided in kJ/mol, and distances are in Å. Colours are provided for each row to match the corresponding dimers and surface points assigned in Figure S7.

Colour	r_n	r_c	Symmetry Operation	U^{AB}_{crys}	$\Delta G^{AB}_{\text{solv}}$	$\Delta G^{BA}_{\text{solv}}$	$\Delta G^{\text{tot}}_{\text{solv}}$	ΔG_{nn}	ΔG_{int}
	2.04	11.68	x,y,-1+z	1.62	-5.88	-3.93	-10.79	-0.29	-12.11
	2.04	11.68	x,y,1+z	1.62	-3.93	-5.88	-8.83	-0.01	-10.44
	2.47	7.36	-x,1-y,1-z	-42.78	-8.69	-8.69	-17.39	-0.82	26.20
	2.54	6.18	-1/2+x,3/2-y,-1/2+z	-37.61	-3.11	-1.22	-5.27	-0.16	32.50
	2.54	6.18	1/2+x,3/2-y,1/2+z	-37.61	-1.22	-3.11	-3.38	-0.03	34.27
	2.58	8.20	-1/2-x,1/2+y,3/2-z	-14.66	-4.33	-6.00	-9.49	-0.46	5.62
	2.58	8.20	-1/2-x,-1/2+y,3/2-z	-14.66	-6.00	-4.33	-11.17	-0.50	3.99
	2.61	8.04	1/2-x,-1/2+y,3/2-z	-10.38	-3.40	-1.82	-6.01	-0.59	4.96
	2.61	8.04	1/2-x,1/2+y,3/2-z	-10.38	-1.82	-3.40	-4.43	-0.55	6.50
	2.71	7.72	-x,2-y,1-z	-8.39	-4.97	-4.97	-9.95	-0.34	-1.23
	2.84	11.20	-1/2-x,-1/2+y,1/2-z	-8.32	-1.86	-0.02	-2.79	-0.80	6.34
	2.84	11.20	-1/2-x,1/2+y,1/2-z	-8.32	-0.02	-1.86	-0.95	-0.44	7.82
	2.84	7.86	-1/2+x,3/2-y,1/2+z	-20.17	1.52	-0.22	2.17	-0.29	22.63
	2.84	7.86	1/2+x,3/2-y,-1/2+z	-20.17	-0.22	1.52	0.42	-0.61	21.21
	3.00	10.43	-x,1-y,2-z	-18.14	-0.50	-0.50	-1.01	-0.30	17.43
	3.79	10.68	-x,2-y,2-z	-1.00	0.01	0.01	0.01	-0.40	1.42

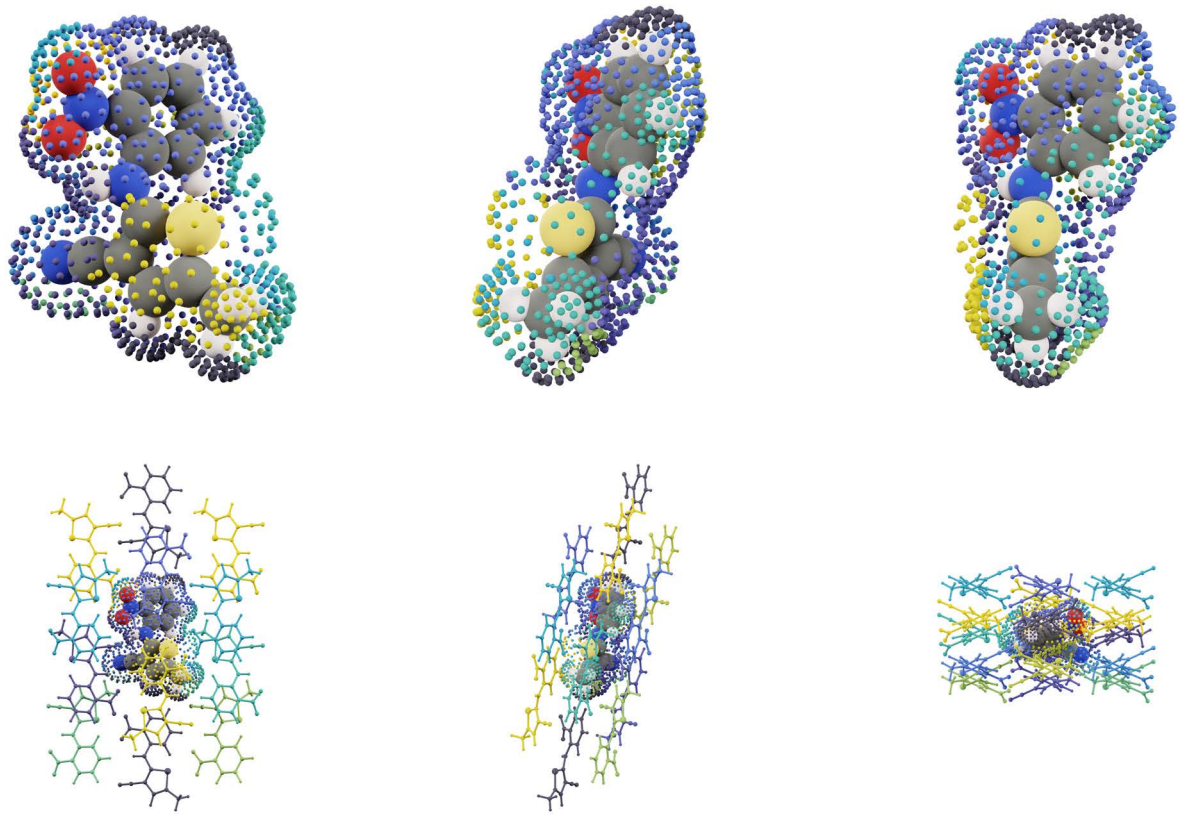


Figure S7: Coulomb solvent accessible surfaces of ROY-OP (top), coloured by the assigned neighbour (bottom). From left to right: viewed down the crystallographic **a**, **b** and **c** axes. Predicted energies for the corresponding dimers and solvent surface partitioning depicted here are given in Table S11.

S2.3.5 ON polymorph (CSD ref. QAXMEH24, solvent: ethanol)

Table S12: Interaction energies for ROY polymorph ON in ethanol. r_n is the nearest-atom distance between the pair of molecules, r_c is the centroid-centroid distance, U_{crys}^{AB} is the dimer interaction energy in the crystal, $\Delta G_{\text{solv}}^{AB}$ and $\Delta G_{\text{solv}}^{BA}$ are the partitioned solvation free energy from A to B and B to A, respectively, and $\Delta G_{\text{solv}}^{\text{tot}}$ is their antisymmetric sum (see Section 4.5, equation 19), ΔG_{nn} is the dimer energy from non-nearest neighbours that has been assigned to this interaction, and ΔG_{int} is the net interaction energy ($U_{\text{crys}}^{AB} + \Delta G_{\text{solv}}^{\text{tot}} + \Delta G_{nn}$). All energies are provided in kJ/mol, and distances are in Å. Colours are provided for each row to match the corresponding dimers and surface points assigned in Figure S8.

Colour	r_n	r_c	Symmetry Operation	U_{crys}^{AB}	$\Delta G_{\text{solv}}^{AB}$	$\Delta G_{\text{solv}}^{BA}$	$\Delta G_{\text{solv}}^{\text{tot}}$	ΔG_{nn}	ΔG_{int}
	2.33	11.27	-x,1-y,1-z	-17.97	-3.48	-3.48	-6.96	-2.10	13.11
	2.39	10.83	1-x,1-y,1-z	-15.14	-2.23	-2.23	-4.46	-2.77	13.44
	2.45	3.95	1+x,y,z	-37.54	-2.10	-3.58	-4.94	-0.66	33.26
	2.45	3.95	-1+x,y,z	-37.54	-3.58	-2.10	-6.43	-0.73	31.84
	2.50	10.39	2-x,1/2+y,3/2-z	-13.68	-5.75	-3.71	-10.48	-1.07	4.27
	2.50	10.39	2-x,-1/2+y,3/2-z	-13.68	-3.71	-5.75	-8.44	-1.16	6.40
	2.56	10.34	1+x,3/2-y,1/2+z	-14.28	-5.41	-5.91	-11.08	-1.69	4.89
	2.56	10.34	-1+x,3/2-y,-1/2+z	-14.28	-5.91	-5.41	-11.57	-1.19	3.90
	2.60	7.50	1-x,1-y,2-z	-8.44	-5.87	-5.87	-11.73	-1.44	-1.84
	2.76	9.78	x,3/2-y,-1/2+z	-14.36	-1.36	-1.06	-2.57	-2.08	13.86
	2.76	9.78	x,3/2-y,1/2+z	-14.36	-1.06	-1.36	-2.27	-2.63	14.72
	3.15	9.53	1-x,-1/2+y,3/2-z	-7.52	-0.70	-0.66	-1.38	-2.11	8.25
	3.15	9.53	1-x,1/2+y,3/2-z	-7.52	-0.66	-0.70	-1.34	-1.78	7.96
	3.31	8.31	2-x,1-y,2-z	-0.40	-0.67	-0.67	-1.34	-0.32	-0.62
	3.46	10.75	-1+x,3/2-y,1/2+z	-7.44	0.01	-0.85	-0.40	-1.80	8.84
	3.46	10.75	1+x,3/2-y,-1/2+z	-7.44	-0.85	0.01	-1.26	-1.42	7.60

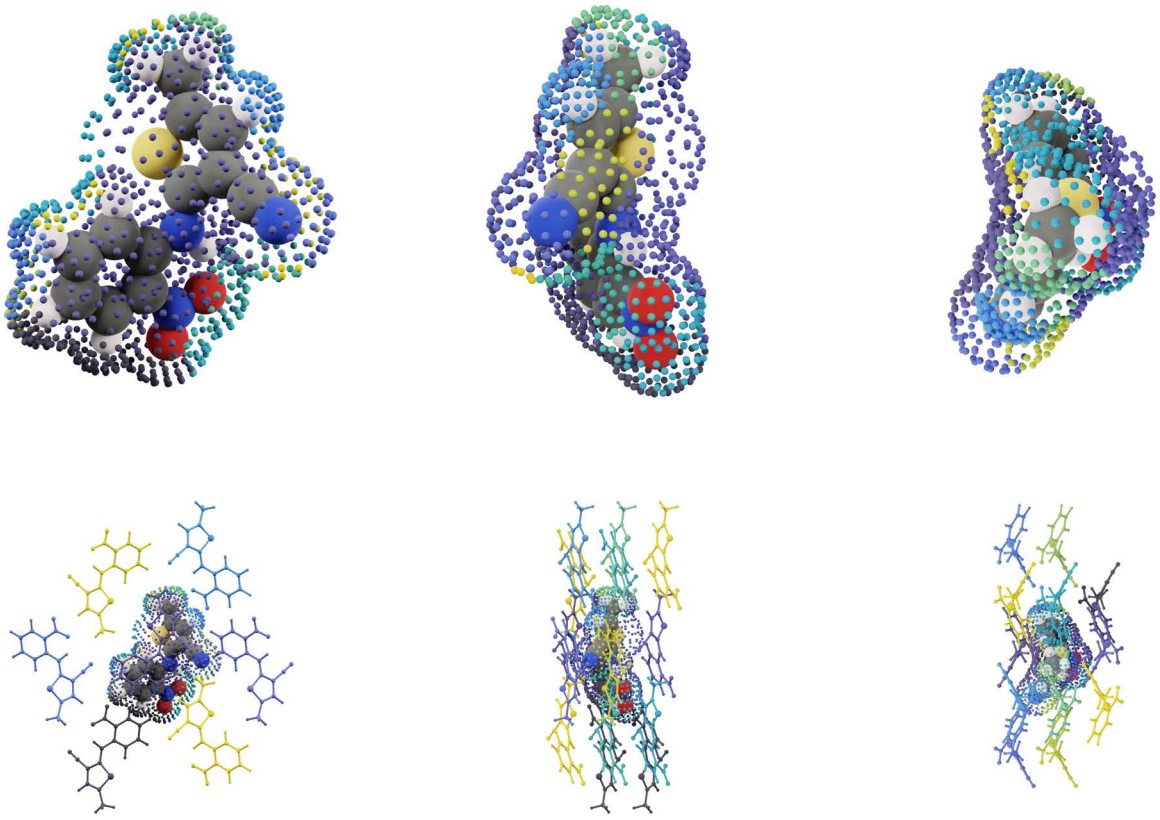


Figure S8: Coulomb solvent accessible surfaces of ROY-ON (top), coloured by the assigned neighbour (bottom). From left to right: viewed down the crystallographic **a**, **b** and **c** axes. Predicted energies for the corresponding dimers and solvent surface partitioning depicted here are given in Table S12.

Table S13: Predicted and fitted ROY interaction energies with the selected energy range for fitting. Final 'Fitted' values are the average of those yielding the 5 closest matching shapes. The 'Colour(s)' column indicates the corresponding colour values from Tables S8-S10 and Figures S4-S6, with those interactions which have been grouped having multiple corresponding colours. Rows marked with '- not fitted -' were neglected due to very low/insignificant interaction energies.

	Colour(s)	Interaction Label	r_c (Å)	Interaction Energy (kcal/mol)				
				Predicted	Search Range			Fitted
					Start	End	Step size	
YT04 polymorph		A	6.62	1.38	0.50	2.50	1.00	1.50
		B	6.99	3.51	1.50	5.50	1.00	4.50
		C	7.19	4.36	2.50	6.50	1.00	4.50
		D	7.27	0.79	0.80	0.80	N/A	0.80
		E	7.74	6.02	4.00	8.00	1.00	6.00
		F	8.23	1.06	1.00	1.00	N/A	1.00
		G	8.41	4.27	2.50	6.50	1.00	3.50
		H	10.61	0.10	- not fitted -			
		I	12.28	0.23	- not fitted -			
		J	12.31	1.37	0.50	2.50	1.00	1.50
R polymorph		A	6.54	6.04	4.00	8.00	1.00	4.40
		B	6.54	5.04	3.00	7.00	1.00	6.20
		C	6.93	9.69	7.00	11.00	1.00	9.00
		D	6.94	7.27	5.00	9.00	1.00	7.80
		E	7.79	1.05	1.00	1.00	N/A	1.00
		F	8.06	1.75	1.50	1.50	N/A	1.50
		G	8.98	4.34	2.00	6.00	1.00	6.00
		H	9.17	-0.21	0.10	0.10	N/A	0.10
		I	10.15	0.56	0.50	0.50	N/A	0.50
		J	10.44	4.02	2.00	6.00	1.00	4.60
		K	11.91	0.30	0.20	0.20	N/A	0.20
Y polymorph		A	5.83	1.08	0.10	3.10	1.00	2.9
		B	6.10	10.06	8.00	12.00	1.00	9.60
		C	7.32	2.25	0.50	4.50	1.00	1.70
		D	7.47	6.71	4.50	8.50	1.00	7.90
		E	8.50	1.72	0.50	2.50	1.00	0.90
		F	8.54	0.13	0.10	0.10	N/A	0.10
		G	9.17	0.42	0.50	0.50	N/A	0.50
		H	9.98	1.38	0.50	3.50	1.00	0.70
		I	11.67	0.64	0.50	0.50	N/A	0.50
		J	12.23	1.74	0.50	3.50	1.00	0.70

S2.4 Solubility predictions

Figure S9 shows the correlation of predicted vs. experimental solubilities, and Table **??** shows the time taken for each calculation.

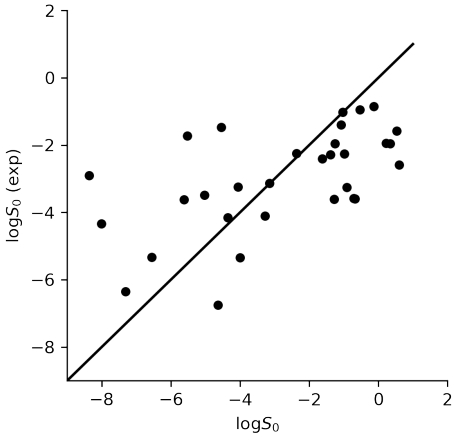


Figure S9: Scatter plot of experimental vs. predicted solubilities ($\log S_0$) for the set of 32 drug and drug-like molecules calculated in this work. Individual values are given in Table 1 of the main text. The solid line (black) represents a perfect $1:1$ correlation.

S3 Experimental morphologies used in fitting procedure

The following equilibrium morphologies (shown in Figure S10) for each system, reconstructed from experimental micrographs and indexed images using intersecting planes in VESTA, are provided accompanying this work. The files available are as follows:

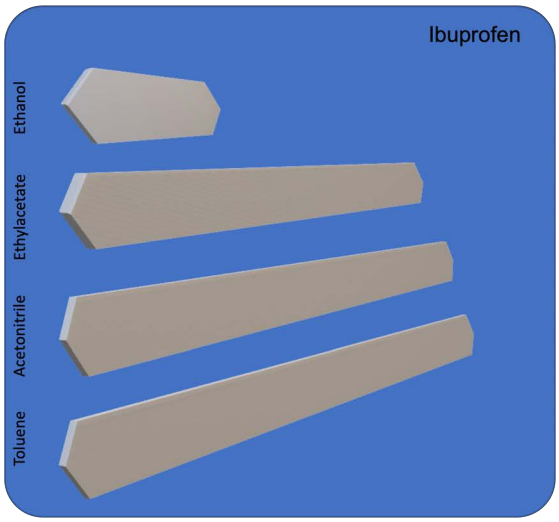
- `adipic_acid.stl` - adipic acid crystal morphology.
- `ibuprofen_ethanol.stl` - ibuprofen crystal morphology, grown from ethanol.
- `ibuprofen_ethylacetate.stl` - ibuprofen crystal morphology, grown from ethyl acetate.
- `ibuprofen_acetonitrile.stl` - ibuprofen crystal morphology, grown from acetonitrile.
- `ibuprofen_toluene.stl` - ibuprofen crystal morphology, grown from toluene.
- `roy_yt04.stl` - ROY YT04 polymorph crystal morphology.
- `roy_r.stl` - ROY R polymorph crystal morphology.
- `roy_y.stl` - ROY Y polymorph crystla morphology.

S4 Video of crystal growth

The video file `GrowthVideo.mp4` provided, shows the simulated growth of the crystals in this work (fitted energies) using CrystalGrower, under different starting supersaturations ($\Delta\mu$). The same profile was used for all simulations starting at a high supersaturation ($\Delta\mu = x$) then rapidly decreasing to a low supersaturation/equilibrium (automatically found). The temperature was set to 25° C and each simulation was carried out for 10 million iterations in CrystalGrower. The video also illustrates the use of a high starting supersaturation as this allows the system to overcome all 2D nucleation barriers. For example, at the current fitted energies, growth was not observed for ibuprofen at $\Delta\mu = 10$ kcal/mol and 20 kcal/mol, and ROY at $\Delta\mu = 10$ kcal/mol.

Table S14: Calculation timings for molecular crystals used as part of the solubility dataset. All calculations were performed on a 2.4 GHz AMD EPYC Processor (with IBPB). Calculations were parallelized across 12 cores, but the time here has been normalized to reflect the time taken on one core. Timings here are provided simply to give a rough figure for reference, as timings will undoubtedly vary significantly based on a number of factors.

Molecule	CSD ref. code	Time taken	No. atoms	No. unique dimers
1-naphthol	NAPHOL01	43m	19	109
4-aminosalicylic acid	AMSALA01	38m	18	121
acetaminophen	HXACAN01	29m	20	113
acetanilide	ACANIL01	14m	19	54
adenosine	ADENOS10	27m	32	42
alclofenac	FICJAC	58m	26	102
allopurinol	ALOPUR	14m	14	63
alprazolam	MENMIB	4h 35m	35/35	159
benzamide	BZAMID02	16m	16	63
benzocaine	QQQAXG02	20m	23	61
benzoic acid	BENZAC02	14m	15	59
chlorprothixene	CMAPTX	5h 33m	39	67
chlorzoxazone	NEWKOP	27m	15	130
clozapine	NDNHCL01	2h 26m	42	59
cocaine	COCAIN10	3h 21m	43	60
diclofenac	SIKLIH01	4h 55m	30	84
flufenamic acid	FPAMCA	6h 6m	30	140
ibuprofen	IBPRAC01	40 mins	33	96
mefenamic acid	XYANAC	2h 15m	33	109
metclopramide	METPRA	1h 3m	36	85
naphthalene	NAPHTA04	22m	18	68
naproxen	COYRUD11	34m	31	75
nicotinic acid	NICOAC02	26m	14	134
nitrofurantoin	LABJON	15m	23	23
papaverine	MVERIQ01	4h 43m	46	64
phenobarbital	PHBARB	2h 1m	29	85
phthalic acid	PHTHAC01	26m	18	42
propranolol	IMITON	5h 10m	40	73
pyrazinamide	PYRZIN	26m	14	133
pyrimethamine	MUFMAB	1h 30m	30/30	166
salicylic acid	SALIAC03	15m	16	61
trimethoprim	AMXBPM10	2h 34m	39	95

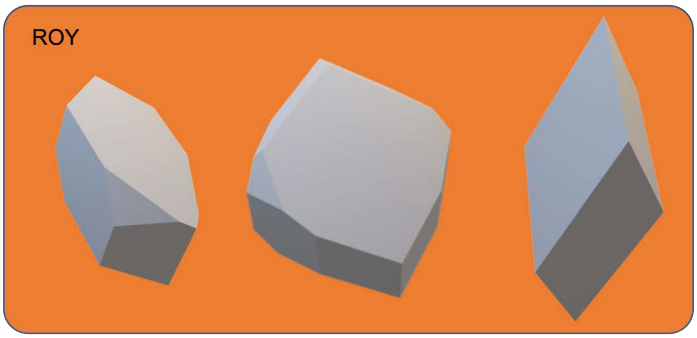


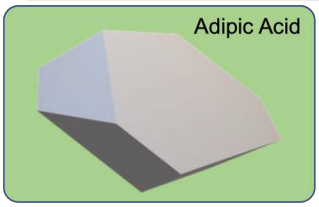
	Ethanol	Ethylacetate	Acetonitrile	Toluene
100	0.0889	0.0235	0.0078	0.0084
002	0.4444	0.2222	0.2222	0.2222
011	1.1111	2.1333	3.2000	3.0000
111	1.0000	1.0000	1.0000	1.0000

	YT04
101	0.4348
020	0.8696
11$\bar{1}$	1.0000
1101	0.8261

	Y
011	1.0000
111	0.6164
101	0.2740
110	0.8493
120	0.9589

	R
010	1.0000
111	0.6164
10$\bar{1}$	0.2740





	Water
100	0.4156
10$\bar{2}$	0.8684
011	1.0000

Figure S10: Reconstructions of crystal morphologies based on experimental micrographs (see Section 4.9 for references). The relative distances from the origin for individual crystallographic surface planes (symmetry related planes included) are given in the respective tables.