Electronic Supplementary Material (ESI) for CrystEngComm.

Supplementary Information for: Static and lattice vibrational energy differences between polymorphs

Jonas Nyman and Graeme M. Day

1 Additional results

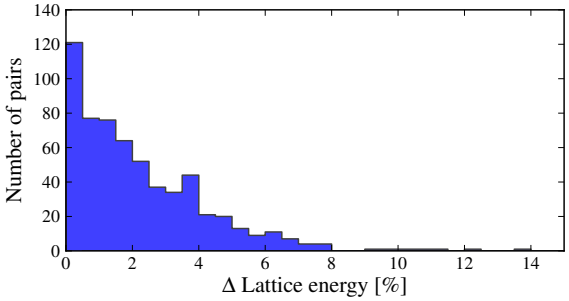

Figure S1: The distribution of differences in relative lattice energy.

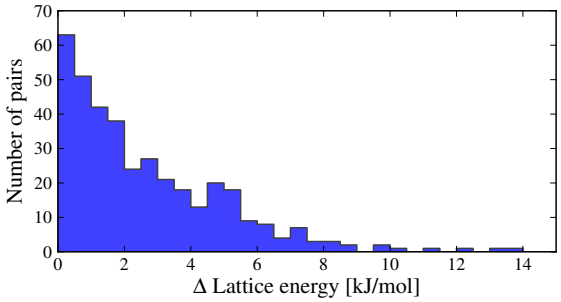

Figure S2: The distribution of lattice energy differences for structures with hydrogen bonding.

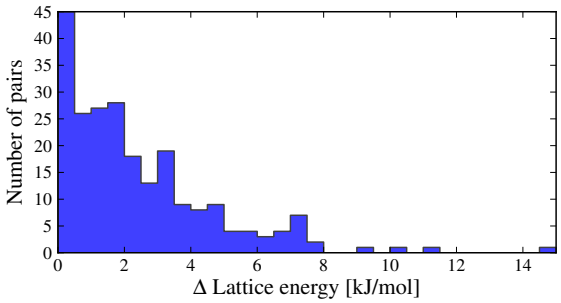

Figure S3: The distribution of lattice energy differences for structures without hydrogen bonding.

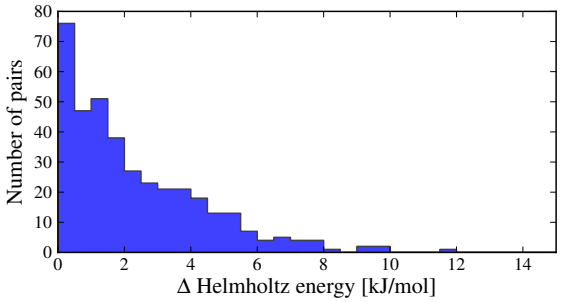

Figure S4: The distribution of free energy differences for structures with hydrogen bonding.

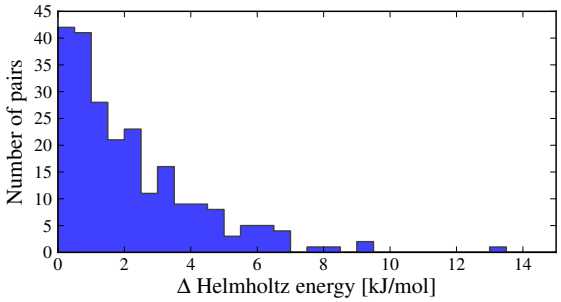

Figure S5: The distribution of free energy differences for structures without hydrogen bonding.

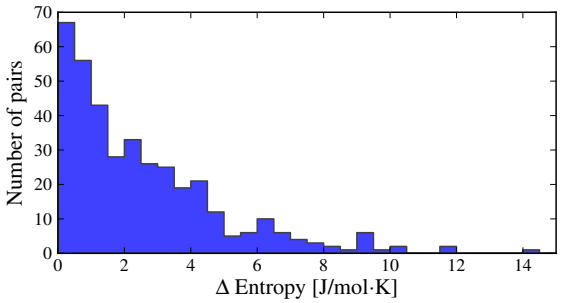

Figure S6: The distribution of entropy differences for structures with hydrogen bonding.

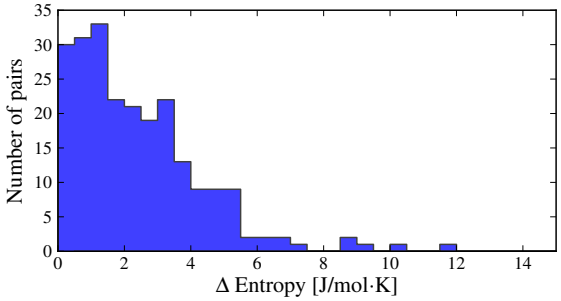

Figure S7: The distribution of entropy differences for structures without hydrogen bonding.

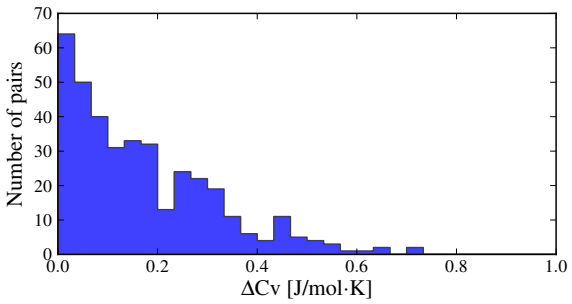

Figure S8: The distribution of heat capacity differences for structures with hydrogen bonding.

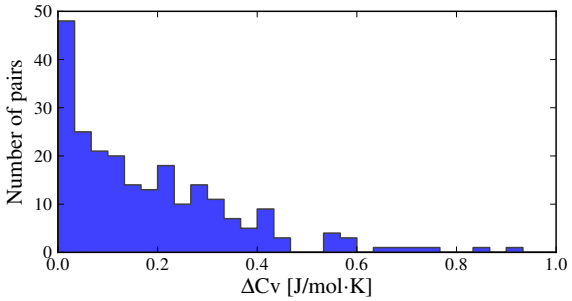

Figure S9: The distribution of heat capacity differences for structures without hydrogen bonding.

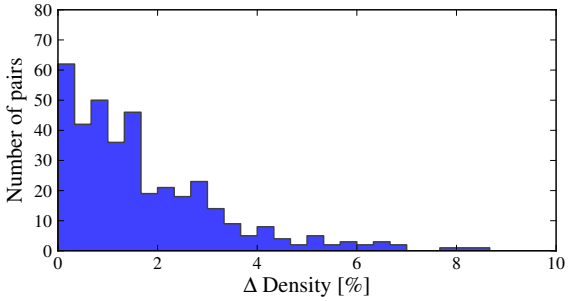

Figure S10: The distribution of density differences for structures with hydrogen bonding.

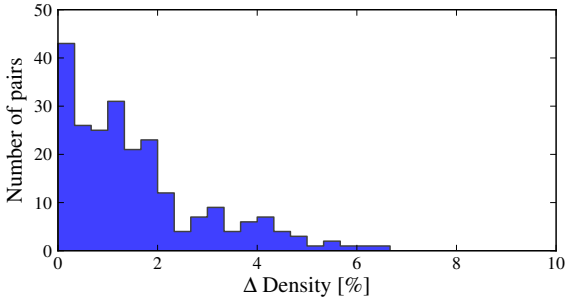

Figure S11: The distribution of density differences for structures without hydrogen bonding.

Since we have not studied conformational polymorphs, we expect the intramolecular energy differences to be small, see Figure S12. A few polymorphs

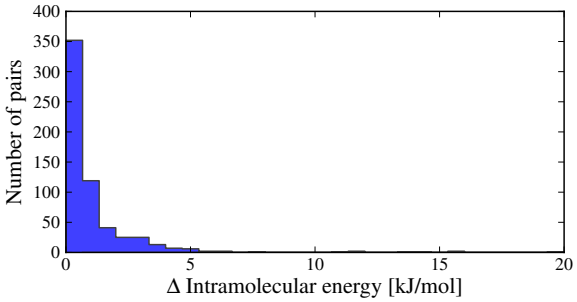

Figure S12: The distribution of intramolecular energies.

do however exhibit different inter- and intramolecular hydrogen bonding motifs, so that intramolecular energy differences can be as large as 19.4 kJ/mol. The CSD structure codes of the outliers, and their calculated intramolecular energy differences, are: HDXMOR/HDXMOR01 (ΔE_{intra} = 14.4 kJ/mol); LEZJAB/LEZJAB01 (ΔE_{intra} = 15.6 kJ/mol); DMANTL01/DMANTL07 (ΔE_{intra} = 15.7 kJ/mol) and IFOVOO/IFOVOO01 (ΔE_{intra} = 19.4 kJ/mol).

The next highest ΔE_{intra} are: BOHZOO/BOHZOO01 (ΔE_{intra} = 13.3 kJ/mol); FUGJUM/FUGJUM01 (ΔE_{intra} = 11.9 kJ/mol); PABZAM/PABZAM01 (ΔE_{intra} = 11.4 kJ/mol) and FEGWAP/FEGWAP01 (ΔE_{intra} = 11.2 kJ/mol).

2 Selected polymorph pairs

An alphabetically sorted list of CSD reference codes for all polymorphs included in this study can be found in polymorphs.txt

3 Distribution of molecular RMSD in polymorphs

The molecular conformations tends to be very similar in polymorphs. Conformational polymorphism was studied by Cruz-Cabeza and Bernstein[1]. We

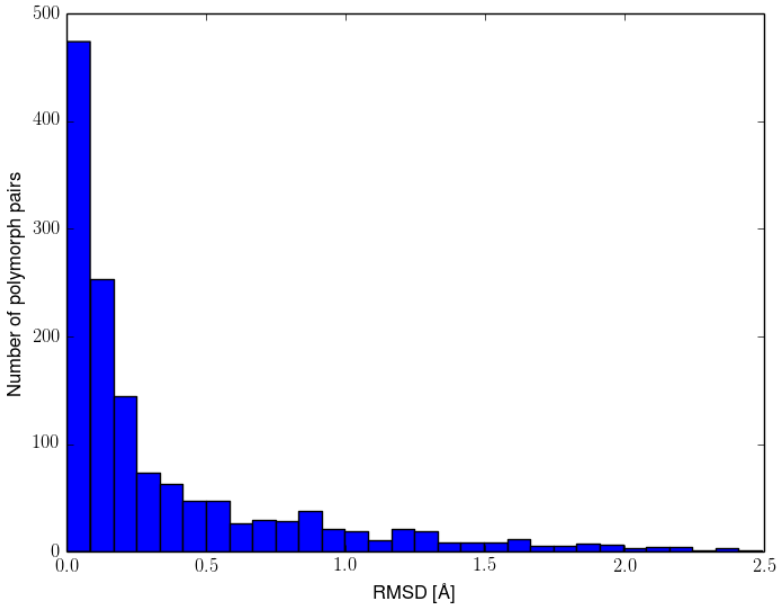

Figure S13: Molecular $RMSD$ between 1397 single-component polymorph pairs.

observe the same distribution of molecular $RMSD$ as did Cruz-Cabeza and Bernstein[1]. We have used a limit of 0.25 Å to remove conformational polymorphs.

4 Co-prime splitting of linear supercells.

The crystal unit cell is expanded into linear supercells (1x1xn) where n is the number necessary to reach the target **k**-point distance. If $n < 5$ the lattice dynamic calculations is performed on the supercell as is. For $n \geqslant 5$ the supercell is split into several smaller supercells (1x1xk, 1x1xℓ, 1x1xm ...) such that $k, \ell, m...$ are all mutually co-prime and $k + \ell + m... \geqslant n$. The long linear supercells are split into 2, 3 or 4 co-prime supercells according to the scheme in Table S1. Note that this is by no means the only possible choice, and we make no claim that this is the best or computationally most efficient splitting. Splitting the supercells means that the sampled **k**-points will no longer be equidistantly placed along the reciprocal axes, but this should have a negligible effect on the results. Unfortunately though, it also means that the convergence will not be monotonic with respect to the target **k**-point sampling distance, making convergence testing difficult.

Table S1: Linear supercells were split into 2, 3 or 4 smaller supercells with mutually co-prime expansion coefficients in this way.

2	→	2
3	→	3
4	→	4
5	→	2, 3
6	→	3, 4
7	→	3, 4
8	→	3, 5
9	→	2, 3, 5
10	→	2, 3, 5
11	→	3, 4, 5
12	→	3, 4, 5
13	→	3, 4, 7
14	→	3, 4, 7
15	→	3, 5, 7
16	→	4, 5, 7
17	→	2, 3, 5, 7
18	→	5, 6, 7
19	→	3, 4, 5, 7
20	→	4, 7, 9
21	→	5, 7, 9
22	→	5, 8, 9
23	→	2, 5, 7, 9
24	→	7, 8, 9
25	→	5, 9, 11
26	→	7, 9, 10

5 Revised Williams99 parameters

All parameters describing interactions between C, N, O and H atoms are described using Williams' original Williams99 forcefield parameters, apart from hydrogen bond H...A interactions, which were reparameterised to work more effectively with the atomic multipole electrostatic model. For H...A interactions, the pre-exponential parameter of the exp-6 model was modified from the Williams99 value. The parameters are given in the following table.

Table S2: Revised H...A parameters in the exp-6 intermolecular model used. C = 0 for all interactions.

hydrogen	acceptor	A (eV)
H2	N1	149
H2	N2	166
H2	N3	163
H2	O1	129
H2	O2	105
H3	N1	70
H3	N2	118
H3	O1	127
H3	O2	133
H4	N1	141
H4	N2	77
H4	N3	56
H4	N4	112
H4	O1	34
H4	O2	198

6 Model potential parameters for halogens

Halogen atoms tend to have an anisotropic van der Waals radius[2]. To account for this, intermolecular potentials with an anisotropic repulsion term has been developed[3,4]. A local unit vector $\mathbf{e}_z$ is defined at each anisotropic site, parallel to the covalent bond joining the halogen to its bonded atom, pointing away from the bond. A second unit vector, $\mathbf{e}_{ik}$, is the vector between the interacting atoms. DMACRYS describes repulsion anisotropy using a modified exp-6 potential of the form:

$$V = G\exp\left(-B^{\iota\kappa}(r_{ij} - \rho^{\iota\kappa}(\Omega_{ik}))\right) - C^{\iota\kappa}/r^6, \tag{1}$$

where $\rho^{\iota\kappa}(\Omega_{ik})$ describes the anisotropy of repulsion, and is defined as:

$$\rho^{\iota\kappa}(\Omega_{ik}) = \rho_0^{\iota\kappa} + \rho_1^{\iota}(\mathbf{e}_z^i \cdot \mathbf{e}_{ik}) + \rho_1^{\kappa}(-\mathbf{e}_z^k \cdot \mathbf{e}_{ik}) + \rho_2^{\iota}(3[\mathbf{e}_z^i \cdot \mathbf{e}_{ik}]^2 - 1)/2 + \rho_2^{\kappa}(3[\mathbf{e}_z^k \cdot \mathbf{e}_{ik}]^2 - 1)/2 \tag{2}$$

ρ_0 describes the isotropic repulsion, ρ_1 parameters describe a shift of the centre of repulsion and ρ_2 parameters describe a quadrupolar distortion of the atom. Parameters for Cl and F were taken from Day's specifically developed potential for molecule XIII in the 4th blind test of crystal structure prediction[5]. Halogen parameters were empirically fitted to reproduce the crystal structures of a set of halogenated aromatic molecules. Details are available in the ESI to the 4th blind test paper. The parameters, in input format for DMACRYS are provided below:

```
BUCK   F_01  F_01
   3761.006673    0.240385    7.144500   0.0  70.0
ANIS   F_01  F_01
 0 0 2 0 2 -0.035000
 0 0 0 2 2 -0.035000
ENDS
BUCK   Cl01  Cl01
   5903.747391    0.299155   86.716330   0.0  70.0
ANIS   Cl01  Cl01
 0 0 2 0 2 -0.093860
 0 0 0 2 2 -0.093860
ENDS
```

7 Additional convergence plots

Convergence of free energy, heat capacity, zero point energy and entropy with respect to **k**-point sampling and the number of **k**-points was studied. Since we in this article are only interested in the difference between crystal structures, we do not show results for the convergence of the "absolute" quantities, but such convergence tests have been made previously[6]. Here we show two typical results from our convergence testing.

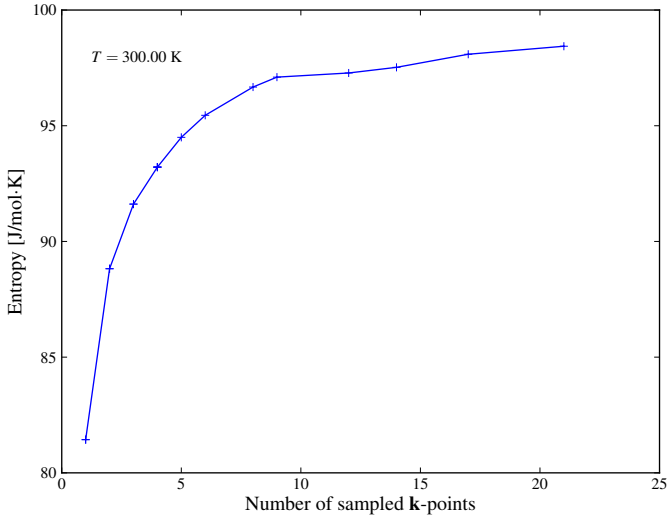

Figure S14: Convergence of the entropy for a crystal structure of theophylline with respect to the number of sampled **k**-points. Plot reproduced from Nyman[6]

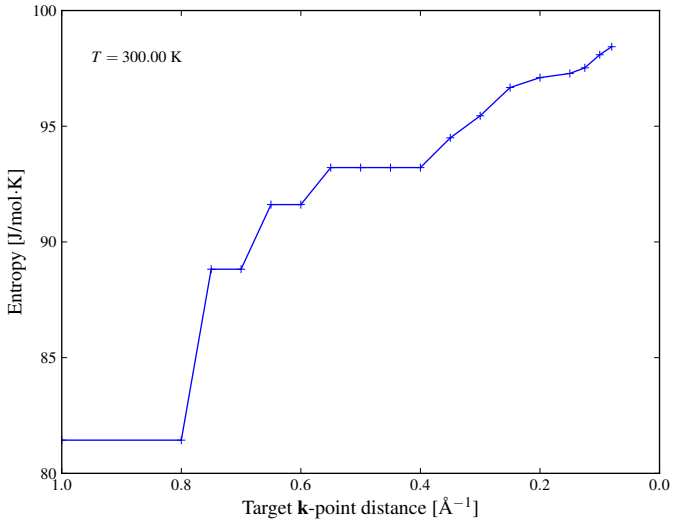

Figure S15: The same data as in the diagram above, but now over the **k**-point distance. Plot reproduced from Nyman[6]

To make the "absolute" thermodynamic quantities converge is difficult and require a careful **k**-point sampling. With our co-prime split linear supercell

sampling, the convergence is no longer monotonic. However, the difference in entropy, vibrational energy *etc* between polymorphs converges slightly faster.

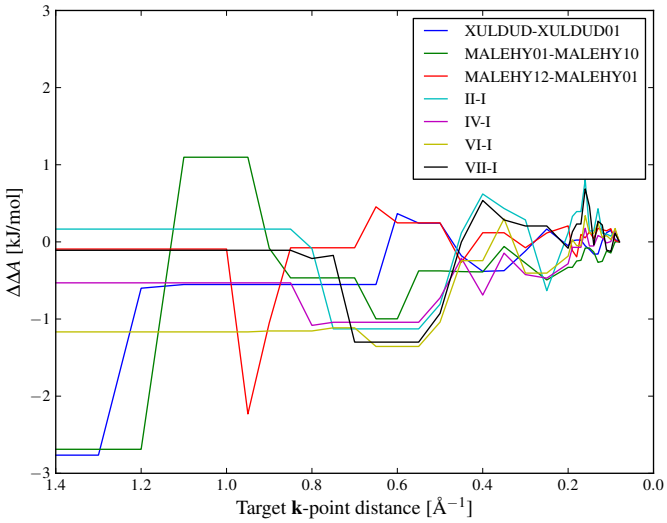

Figure S16: Convergence of ΔA for a set of theophylline (Roman numerals), maleic hydrazide, and 3,4-cyclobutylfuran polymorphs with respect to **k**-point distance.

We have estimated the error in ΔS at 300 K due to the **k**-point sampling to ± 1 J/mol·K at a **k**-point distance of 0.12 Å. The error in Helmholtz energy caused by the incomplete **k**-point sampling at the same temperature is about 1 kJ/mol, probably smaller than the error in lattice energy. The zero point energy and heat capacity converges much more rapidly. The errors for these quantities are negligible.

Errors arise both because of the finite sampling, and the selective way of sampling along only three directions. The error caused by neglecting the "diagonal" **k**-vectors is however small and likely systematic, not affecting the delta-quantities significantly.

The thermodynamic properties are functions of the phonon density of states, and an alternative way to judge the convergence is to compare phonon DOS calculated at different **k**-point samplings. In the article, we show the phonon density of states for two theophylline and two maleic hydrazide polymorphs calculated at a **k**-point sampling of 0.12 Å^{-1}. Below, we show the density of states of the same structures, but at a denser target **k**-point distance of 0.08 Å^{-1}. The density of states calculated at 0.12 and 0.08 Å^{-1} are very similar and well converged at these samplings.

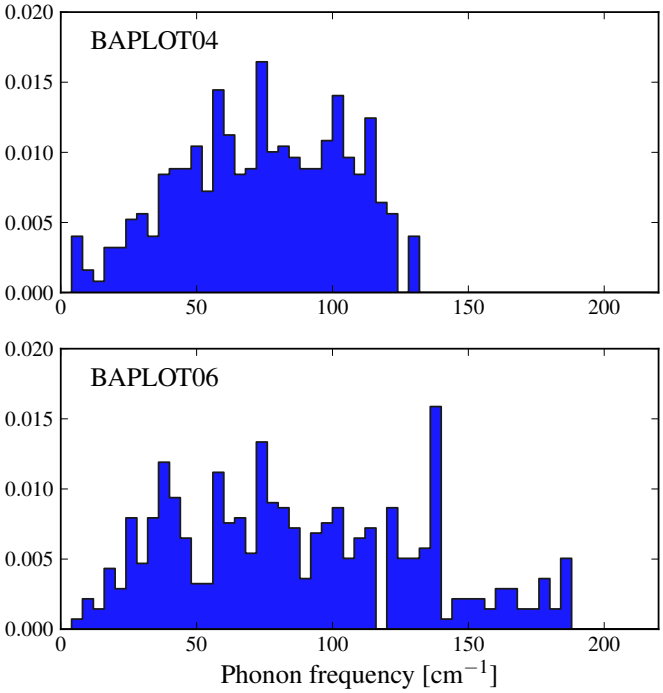

Figure S17: Phonon density of states for theophyline polymorphs I and II calculated with a target **k**-point distance of 0.08 Å^{-1}.

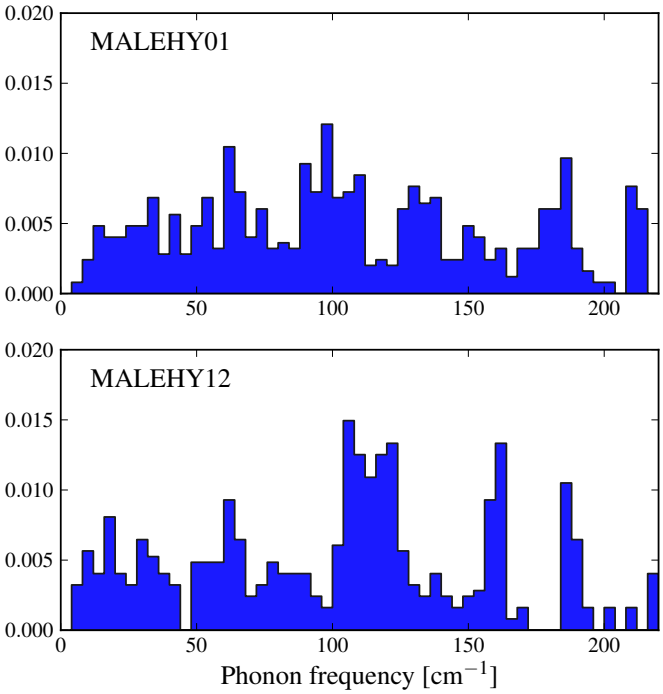

Figure S18: Phonon density of states for two maleic hydrazide polymorphs calculated with a target **k**-point distance of 0.08 Å^{-1}.

References

[1] A. J. Cruz-Cabeza and J. Bernstein, *Chemical reviews*, 2013, **114**, 2170–2191.

[2] S. C. Nyburg and C. H. Faerman, *Acta Crystallographica Section B*, 1985, **41**, 274–279.

[3] G. M. Day and S. L. Price, *Journal of the American Chemical Society*, 2003, **125**, 16434–16443.

[4] S. L. Price, M. Leslie, G. W. A. Welch, M. Habgood, L. S. Price, P. G. Karamertzanis and G. M. Day, *Phys. Chem. Chem. Phys.*, 2010, **12**, 8478–8490.

[5] G. M. Day, T. G. Cooper, A. J. Cruz-Cabeza, K. E. Hejczyk, H. L. Ammon, S. X. M. Boerrigter, J. S. Tan, R. G. Della Valle, E. Venuti, J. Jose, S. R. Gadre, G. R. Desiraju, T. S. Thakur, B. P. van Eijck, J. C. Facelli, V. E. Bazterra, M. B. Ferraro, D. W. M. Hofmann, M. A. Neumann, F. J. J. Leusen, J. Kendrick, S. L. Price, A. J. Misquitta, P. G. Karamertzanis, G. W. A. Welch, H. A. Scheraga, Y. A. Arnautova, M. U. Schmidt, J. van de Streek, A. K. Wolf and B. Schweizer, *Acta Crystallographica Section B*, 2009, **65**, 107–125.

[6] J. Nyman, *In Silico Predictions of Porous Molecular Crystals and Clathrates*, 2014, M.Phil thesis, University of Southampton.