The observed and energetically feasible crystal structures of 5-substituted uracils†

**Sarah A. Barnett,‡ Ashley T. Hulme, Nizar Issa, Thomas C. Lewis,§
Louise S. Price, Derek A. Tocher and Sarah L. Price***

A search of the Cambridge Structural Database for crystal structures of 5-substituted uracils shows that, although there is a recurrent motif with symmetric hydrogen bonding and interdigitation of the 5-substituent R, a range of other hydrogen bonded ribbons, sheets and three-dimensional motifs are possible. In order to try and rationalize this, we have performed a combination of experimental studies and computational searches for low energy structures for the 12 simple 5-substituted uracils with R = H, CH_3, CH_2CH_3, $CHCH_2$, CN, OH, NH_2, NO_2, F, Cl, Br and I. Crystallization experiments on these compounds yielded the first single crystal X-ray determinations of 5-ethyluracil and 5-cyanouracil, as well as low temperature redeterminations of the disordered structures of 5-chlorouracil and 5-bromouracil. The lattice energies were calculated for the known crystal structures and compared with the computed lattice energy landscape for each molecule (except R = Br and I). Although the symmetric ribbon motif often dominates the computed crystal energy landscape, all of the molecules show a variety of different hydrogen bonding structures within a small energy range (5 kJ mol⁻¹) of the global minimum and exhibit quite a diverse range of energetically competitive motifs. Thus, the range of crystallization outcomes, from polymorphism and other multiple forms, to the difficulty in growing single crystals (R = $CHCH_2$ and NH_2) probably reflects the sensitivity of the various hydrogen bonding motifs to the substituent and limited range of crystallization conditions that can be applied.

1. Introduction

The design of organic crystal structures by crystal engineering[1–6] has traditionally used commonly occurring motifs, or synthons, often containing multiple conventional hydrogen bonds. These motifs have normally been identified and ranked[7,8] through searches of the Cambridge Structural Database (CSD).[9] It is generally assumed that preferred motifs represent stronger intermolecular interactions,[10] although the motifs that appear in any given crystal structure represent the balance of all the intermolecular forces between all parts of the molecule[11] and, possibly, kinetic effects. Hence the calculation of the thermodynamically feasible structures, the crystal energy landscape,[12] by crystal structure prediction techniques has the potential to unravel some of the factors involved. For example, the predominance of the $R_2^2(8)$ doubly hydrogen bonded carboxylic dimer motif, is not a property of the carboxylic acid group alone.[1] The influence of other interactions is shown[13] by benzoic acid having only this motif, but formic, acetic and tetrolic acid also have catemeric hydrogen bonded motifs on their crystal energy landscapes. Similarly, exceptions to Etter's rule[14] that all good hydrogen bond acceptors will be used in hydrogen bonding, can occur when use of all carbonyl acceptors is not consistent with steric requirements,[15] with different acceptors being unused in nearly equi-energetic hypothetical structures. In this paper, we examine the favourability of different hydrogen bonding motifs for the 5-substituted uracils, a system of biological relevance, where some crystal structures have uracil C=O acceptors that do not participate in conventional hydrogen bonds.

The diversity of hydrogen bonding motifs in 5-substituted uracils was first brought to our attention by our studies of the anti-cancer agent 5-fluorouracil, where a new polymorph,[16] four solvates[17–20] and a thymine solid solution were found,[21] covering a range of hydrogen bonding motifs. In particular, the trifluoroethanol solvate adopts a hydrogen bonded ribbon structure that was prevalent on the crystal energy landscape of pure 5-fluorouracil, but not observed in the polymorphs. This is consistent with the observation made for 5-fluorocytosine[22] and hydrochlorothiazide[23] that the hydrogen bonding motifs seen on the crystal energy landscape have the potential to

Department of Chemistry, University College London, 20 Gordon Street, London, UK WC1H 0AJ. E-mail: s.l.price@ucl.ac.uk; Fax: +44 (0)20 7679 7463; Tel: +44 (0)20 7679 4622

‡ Current address: Diamond Light Source Ltd., Harwell Science and Innovation Campus, Didcot, Oxfordshire, UK OX11 0DE.
§ Current address: Phlexglobal Ltd, Mandeville House, 62 The Broadway, Amersham, Bucks, UK HP7 0HJ.

appear in the polymorphs and diverse solvates. This suggested performing a search of 5-substituted uracil structures, including solvates, in the CSD to examine the prevalence of the different hydrogen bonding motifs for 5-substituted uracils.

In order to interpret the results of this CSD search, twelve uracils with simple substituents were studied by a combination of crystal structure prediction, all using the same atom–atom potentials and methodology (R = H, CH_3, CH_2CH_3, $CHCH_2$, CN, OH, NH_2, NO_2, F and Cl), with experimental crystallization screening (R = H, CH_3, CH_2CH_3, $CHCH_2$, CN, NH_2, Cl, Br and I). Since a large amount of experimental work has been carried out in the past on the 5-substituted uracils where R = H,[24] OH,[25] F[16,26] and NO_2,[27] further crystallization experiments focused on the other 8 molecules which were subjected to varying degrees of screening in order to search for polymorphs or obtain a first crystal structure, and ensure that no readily crystallized forms were missed. Where growing single crystals suitable for single crystal X-ray diffraction proved impossible (R = $CHCH_2$ and NH_2), the compounds were subjected to powder X-ray diffraction and the patterns compared with the simulated powder patterns of the theoretical structures obtained from the computational search. The results from the 10 searches for low energy crystal structures of each compound were compared with each other and the known experimental structures in order to determine whether any strong preferences for particular hydrogen bonding motifs were observed and to ascertain if variations could be rationalized in terms of substituent interactions.

2. Methods and experimental

2.1 Analysis of 5-substituted uracil crystal structures in the Cambridge Structural Database

A search of the CSD (Version 5.29, Nov 2007)[9] was performed to find all of the crystal structures that involved 5-substituted uracils (Scheme 1), including polymorphs, solvates and co-crystals. The resulting 109 hits (ESI Table S1.1†) were first reduced to just those with full structural coordinates available (92) and then the 36 structures which included metallic elements were removed. Additionally, where multiple versions of the same structure were available, just the "best" one was kept, i.e., the one with hydrogen atoms present, the one where data was collected at a lower temperature or the one with the lowest R factor. The remaining 49 structures (ESI Table S1.2†) were analysed visually using Mercury[28] with hydrogen bonds being considered present only if the N–H···O hydrogen bond angle was ≥ 125° as well as the default setting for N···O intermolecular interaction distances of ≤3.07 Å.
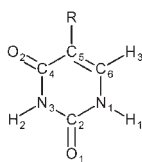
Scheme 1 5-Substituted uracil showing the numbering scheme used, where R = H, CH_3, CH_2CH_3, $CHCH_2$, CN, OH, NH_2, NO_2, F, Cl, Br and I.

2.2 Experimental

Thymine, uracil, 5-bromouracil and 5-aminouracil were purchased from Aldrich Chemicals, 5-iodouracil and 5-cyanouracil were purchased from Alfa Aesar, 5-ethyluracil and 5-vinyluracil were purchased from Avocado Research Chemicals Ltd. and 5-chlororuracil was purchased from Fisher Scientific; all were used without further purification. The 9 compounds were recrystallized by sublimation and from a number of solvents under various conditions as part of a manual experimental polymorph screen. Further details and results can be found in ESI Table S3.1.†

5-Ethyluracil 1. Single crystals were grown by cooling a saturated 50 : 50 water : acetone solution from 70 °C to room temperature over a period of 90 min.

5-Cyanouracil 2. Crystals suitable for single crystal X-ray diffraction were grown by slow evaporation of a saturated 1,4-dioxane solution.

5-Fluorouracil benzonitrile solvate (1 : 1) 3. Crystals suitable for single crystal X-ray diffraction were grown by slow evaporation of a saturated benzonitrile solution.

5-Fluorouracil formamide solvate (1 : 1) 4. Crystals suitable for single crystal X-ray diffraction were grown by slow evaporation of a saturated formamide solution.

5-Chlorouracil 5. Single crystals were grown by slow evaporation of a saturated 1,4-dioxane solution.

5-Bromouracil 6. Single crystals were grown by slow evaporation of a saturated 1,4-dioxane solution.

2.3 Single crystal X-ray diffraction

Single crystal X-ray experiments were performed on a Bruker AXS SMART APEX CCD detector diffractometer equipped with a Bruker AXS Kryoflex open flow cryostat (graphite monochromated Mo-Kα radiation λ = 0.71073 Å; ω scans). The single crystal structures were solved by direct methods using SHELXS-97[29] and all non-hydrogen atoms were located using subsequent difference-Fourier methods in SHELXL-97.[30] Other details of crystal data, data collection and processing are given in Table 1. For structures **2–6**, absorption corrections were applied by a semi-empirical approach using SADABS.[31] For structures **1–4**, hydrogen atoms were located from the difference Fourier map and freely refined whilst for **5** and **6** the hydrogen atoms were placed in geometrically calculated positions and refined using a riding model. The isomorphic structures **5** and **6** are, as previously established, disordered although no attempt was made at that time to model it.[32] The molecule in the asymmetric unit is disordered between two orientations corresponding to a 180° rotation about the O1 C2 C5 R axis (R = Cl or Br), i.e., there is effectively an interchange of the oxygen and hydrogen atoms on C4 and C6. The disorder was modelled over two sites, corresponding to the two orientations, for all of the non-hydrogen atoms and the site occupancy factor for each component refined. The occupancy ratios were fixed at 85 : 15 and 73 : 27 for structures **5** and **6**, respectively, and the co-ordinates and anisotropic displacement parameters

Table 1 Crystallographic data summary for compounds **1–6**

	5-Ethyluracil (**1**)	5-Cyanouracil (**2**)	5-Fluorouracil benzonitrile (**3**)	5-Fluorouracil formamide (**4**)	5-Chlorouracil (**5**)	5-Bromouracil (**6**)
Empirical formula	$C_6H_8N_2O_2$	$C_5H_3N_3O_2$	$C_{11}H_8N_3O_2F$	$C_5H_6N_3O_3F$	$C_4H_3N_2O_2Cl$	$C_4H_3N_2O_2Br$
M	140.14	137.10	233.20	175.13	146.53	190.99
Crystal system	Triclinic	Monoclinic	Monoclinic	Monoclinic	Monoclinic	Monoclinic
Space group	$P\bar{1}$	$P2_1/n$	$P2_1/c$	$P2_1/m$	$P2_1/n$	$P2_1/n$
a/Å	3.9193(18)	9.0098(18)	7.0460(7)	6.827(4)	8.4393(10)	8.5680(11)
b/Å	5.754(3)	6.6035(13)	24.035(2)	6.111(3)	6.8412(8)	6.8823(9)
c/Å	14.366(7)	9.1805(18)	6.8640(7)	8.424(4)	9.3679(12)	9.5715(12)
α/°	100.027(7)	90	90	90	90	90
β/°	96.109(7)	98.124(3)	116.554(2)	90.313(8)	104.201(2)	103.440(2)
γ/°	92.374(8)	90	90	90	90	90
U/Å^3	316.6(3)	540.72(19)	1039.80(17)	351.4(3)	524.33(11)	548.95(12)
Z	2	4	4	2	4	4
T/K	150(2)	150(2)	150(2)	150(2)	150(2)	150(2)
μ/mm^{-1}	0.113	0.136	0.118	0.153	0.633	7.397
Reflections collected	4168	4294	6296	2892	4328	4554
Unique reflections (R_{int})	1346 (N/A)	1270 (0.0199)	2458 (0.0158)	866 (0.0243)	1234 (0.0163)	1305 (0.0212)
Final R_1 [$F > 4\sigma(F)$]	0.0609	0.0398	0.0366	0.0407	0.0330	0.0254
wR_2 (all data)	0.1453	0.1121	0.1119	0.1108	0.0941	0.0630

were then refined in conjunction with geometric and vibrational restraints. The crystal of **1** was a non-merohedral twin so GEMINI[33] was used to index the data which showed that there were two, approximately equal, components. The output p4p files for each component were read back into SMART[34] and run through the BRAVAIS and L.S. routines. The data were combined to give a single p4p file and integrated simultaneously using SAINT+.[35] The twin law was found to be 1 −0.01 −0.02/0 1 0.01/0.25 −0.06 1. The dataset had 1731 data from component 1 only, 1711 data from component 2 only and 725 data belonging to both. I/σ for overlapping reflections was 19.5. The dataset was corrected for absorption using TWINABS[36] which was also used to produce an HKLF 4 file (non-overlapping reflections for component 1 only) for structure solution and refinement. The use of an HKLF 5 file, including reflections for both components, for refinement gave no improvement to the structure, parameters or data completeness.

Crystal structure diagrams were produced using SHELXTL[37] and Mercury[28] with interactions calculated as being contacts shorter than the sum of the van der Waals radii.

2.4 Crystal structure prediction

A gas phase *ab initio* model for each molecule was obtained by optimization of the MP2/6-31G(d,p) energy using the program GAUSSIAN03.[45] For R = CH$_2$CH$_3$, CHCH$_2$ and OH, minimizations were also performed with the functional group in the opposite orientation to establish whether this conformation was sufficiently low in energy to possibly be found in the crystal. For R = NO$_2$, a conformation constrained to have the NO$_2$ in the same plane as the uracil (as this was closer to the torsion angles observed for the three polymorphs than that obtained by the free optimization) was also tested. A distributed multipole analysis[46] (using GDMA1[47]) of the *ab initio* charge density of the molecule was performed to provide an accurate description of the electrostatic contribution to the lattice energy in the rigid molecule crystal structure modelling.

This atomic multipole electrostatic model automatically represents the electrostatic effects of lone pair and π-electron density.[48] All other intermolecular contributions to the lattice energy were represented by an empirical repulsion–dispersion model of the form:

$$U = \sum_{i \in 1, k \in 2} (A_{\iota\iota}A_{\kappa\kappa})^{1/2} \exp(-(B_{\iota\iota} + B_{\kappa\kappa})R_{ik}/2) - \frac{(C_{\iota\iota}C_{\kappa\kappa})^{1/2}}{R_{ik}^6}$$

where atom i in molecule 1 of type ι and atom k in molecule 2 of type κ are separated by a distance R_{ik}. The parameters used had been empirically fitted to a range of crystal structures.[49–51] No computational work was done for 5-iodouracil and 5-bromouracil because of a lack of suitable parameters. Previous searches on 5-hydroxyuracil[25] and thymine[52] with slightly different C parameters were compared.

The hypothetical crystal structures for all 5-substituted uracils were generated by MOLPAK,[53] which performs a systematic grid search on orientations of the rigid central molecule in 39 common co-ordination geometries of organic molecules, belonging to the space groups Cc, $C2$, $C2/c$, $P1$, $P\bar{1}$, $P2/c$, $P2_1$, $P2_1/c$, $P2_12_12$, $P2_12_12_1$, $Pba2$, Pc, $Pca2_1$, $Pna2_1$, $Pbcn$, $Pbca$, $Pma2$, $Pmn2_1$ and $Pnn2$ with one molecule in the asymmetric unit. Approximately 3000 of the densest packings were then used as starting points for lattice energy minimization by DMAREL[54,55] using the atom–atom based model potential described above. Thus, the search only produces structures with one entire molecule in the asymmetric unit of the space groups considered. The distinct low energy minima within 10–20 kJ mol^{-1} of the global minimum were established by considering the reduced cell parameters[56] using PLATON.[57] The second derivative properties of each unique lattice energy minimum were examined[58] and those that were mechanically unstable were eliminated.

The unique structures within 5 kJ mol^{-1} of the global minimum were analyzed visually using Mercury[28] in the same way as the experimental crystal structures.

3. Results and discussion

3.1 Analysis of 5-substituted uracil crystal structures in the Cambridge Structural Database

The 49 structures obtained from the CSD search after various criteria had been applied were analyzed in terms of their hydrogen bonding motifs (Table 2; further information is available in ESI Tables S1.1 and S1.2†). Although quite a few of the structures exploited the extra functionality available with the more complex R groups, many did not and neither did the presence of additional co-crystallizing agents necessarily affect the hydrogen bonding motif formed.

Structural analysis of these 49 structures showed quite a range of different motifs but dimers (13 hits) and the symmetric ribbon (11 hits) (Scheme 2) stand out as being particularly prevalent. Additionally, the symmetric ribbon formation appears to transcend the use of functionalized substituents[25,32,38,40–43] and the presence of molecules of co-crystallization or solvation.[20,39,41,44] Although all 11 of these structures can be classified as having the symmetric ribbon motif, they show quite a diversity on closer inspection, as they cover a range of degrees of undulation of the chain. The observation that the 5-substituted uracils (Scheme 1) could have C_{2v} symmetry about the O1 C2 C5 R axis if it were not for the major difference between C4=O2 and C6–H3, means that the relative orientations of the molecules within these ribbons becomes significant.

Thus, the symmetric ribbon can be classified as polar, with all the C4=O2 carbonyls pointing in the same direction within

Table 2 Evaluation of the hydrogen bonding motifs of the 49 5-substituted uracils in the CSD. The designated motif is based solely on the hydrogen bond donors and acceptors available on the uracil

Hydrogen bonding motif utilizing only the uracil functionality		Number of occurrences
0-D (19)	No interactions	6
	Dimers	13
1-D (19)	Chains	5
	Symmetric ribbon	11
	Asymmetric ribbon	3
2-D (9)	4,4-grid	4
	6,3-net	4
	4.8^2-net	1
3-D (2)	All	2

a ribbon (Schemes 2a and c), or non-polar, with the C4=O2 carbonyls pointing in opposite directions for the two halves of the ribbon (Schemes 2b and d). On consideration of the adjacent ribbon, the sheets formed can be termed parallel (Schemes 2a and b) or anti-parallel (Schemes 2c and d).

Other common one-dimensional motifs include the asymmetric ribbon[16,59,60] (Scheme 3a), which utilizes all the available hydrogen bonding acceptors and donors, and chains composed of single N–H···O interactions[17,27,43,61,62] (Scheme 3b). This second type tends to be formed when the remaining uracil donors and acceptors participate in hydrogen bonding interactions with the solvent molecule,[17,27] or the substituent in the 5-position.[43,61,62] There are 9 cases which form two-dimensional nets. Uracil forms a brick wall type 6,3-net (Scheme 3c) through
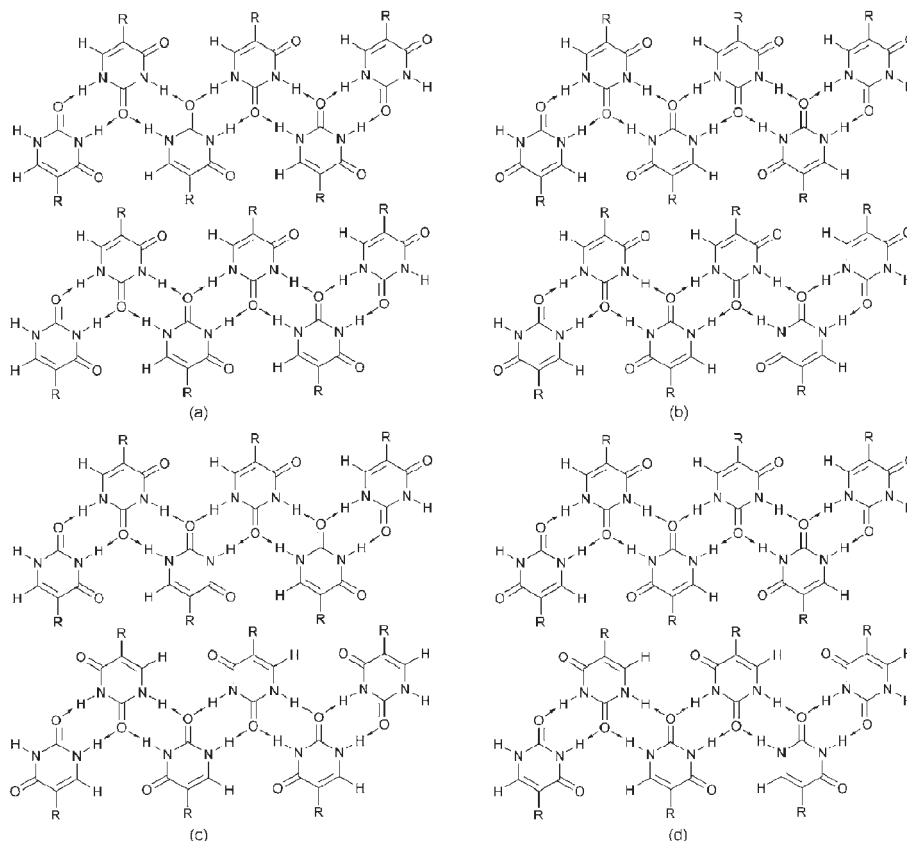
Scheme 2 The variations observed in the symmetric ribbon motif (a) polar parallel,[20,38–41] (b) non-polar parallel,[25,42] (c) polar anti-parallel[43] and (d) non-polar anti-parallel.[44] Hydrogen bonds are denoted by arrows.

This journal is © The Royal Society of Chemistry and the Centre National de la Recherche Scientifique 2008
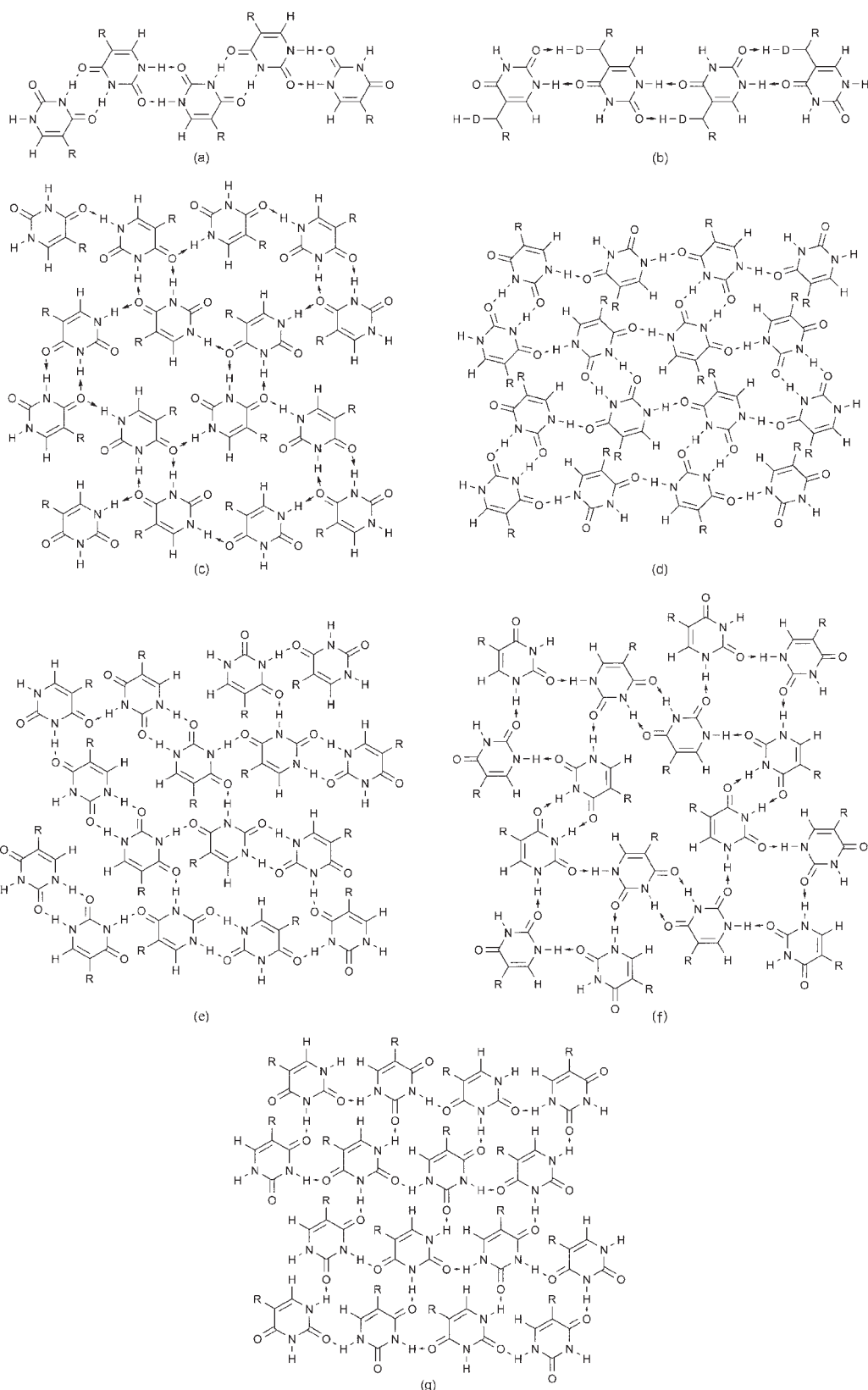
Scheme 3 Some simple motifs observed for 5-substituted uracils (hydrogen bonds are denoted by arrows); (a) asymmetric ribbon observed for form **II** of 5-fluorouracil, (b) chains formed by single N–H···O interactions with secondary hydrogen bonds formed by the substituent in the 5-position, (c) 6,3-net (brick wall) observed for uracil, (d) alternative 6,3-net found for 5-(diphenylphosphino)uracil, (e) 6,3-net (herringbone) observed for the 1,4-dioxane solvate of 5-nitrouracil, (f) 4.8²-net observed for form **I** of 5-fluorouracil and (g) 4,4-grid observed for form **III** of 5-nitrouracil.

Fig. 1 Plots of lattice energy against density of the unique structures within 5 kJ mol^{-1} of the global minimum denoted by the motif adopted when just considering the hydrogen bonding interactions formed by the uracil for (a) uracil, (b) thymine, (c) 5-ethyluracil, (d) 5-cyanouracil, (e) 5-hydroxyuracil and (f) 5-chlorouracil. The structures corresponding to the known forms are indicated by red open shapes.

than the global minimum and known experimental form. An alternative search,[52] with a reduced carbon repulsion, confirmed the small energetic preference for the symmetric ribbon motif.

Table 3 Summary of the results from the computational polymorph screen for each of the 5-substituted uracils studied
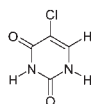
Molecule	Structure	REFCODE / Motif of experimentally observed forms	(Rank) ΔE^a/kJ mol^{-1}		Molecular overlay^b/Å (Packing overlay^c/Å)		Motif of global minimum structure	ΔE/kJ mol^{-1} to symmetric ribbon^d / Motif^e
Uracil		URACIL 6,3-Net	(20th) −1.22		0.0277 0.239		Symmetric ribbon (Scheme 2a)	0.16 6,3-Net
Thymine		THYMIN01 Symmetric ribbon	(3rd) −0.29		0.0214 0.172		Symmetric ribbon (Scheme 2a)	2.31 Asymmetric ribbon
5-Ethyluracil		1 Symmetric ribbon	(5th) −0.70		0.068 0.357		Symmetric ribbon (Scheme 2b)	2.01 6,3-Net
5-Vinyluracil		No structure Powder data suggests symmetric ribbon	—		—		Symmetric ribbon (Scheme 2a)	0.27 Asymmetric ribbon
5-Cyanouracil		2 Ladder	(4th) −1.76		0.0485 0.179		Dimers (ESI Fig. S2.5a)	−4.51
5-Hydroxy-uracil		QARWAI Symmetric ribbon	(5th) −2.70		0.0476 1.797		Dimers (ESI Fig. S2.7)	−2.70
5-Aminouracil		No structure Powder data suggests symmetric ribbon	—		—		Dimers (other interactions make this 3-D)	−3.02
5-Nitrouracil		NIMFOE Asymmetric ribbon	Opt (Not found) −4.07 / Planar (Not found) −4.57		Opt — / Planar —		Opt Dimers / Planar 4,4-Grid (Scheme 3g)	−6.96
		NIMFOE01 Asymmetric ribbon	(37th) −8.29 / (Not found) −4.41		0.227 1.546 / —		(ESI Fig. S2.8)	
		NIMFOE02 4,4-Grid	(6th) −3.72 / (1st) 2.68		0.232 0.749 / 0.0432 0.193			
5-Fluorouracil		FURACL01 4.8²-Net	(Not found) −5.91		—		Asymmetric ribbon (Scheme 3a)	−1.97
		FURACL03 Asymmetric ribbon	(1st) 0.09		0.0324 0.220			

Table 3 (*continued*)

Molecule	Structure	REFCODE Motif of experimentally observed forms	(Rank) ΔE^a/kJ mol^{-1}	*Molecular overlay*^{b}/Å Packing overlay^c/Å	Motif of global minimum structure	ΔE/kJ mol^{-1} to symmetric ribbon^d Motif^e
5-Chlorouracil		CLURAC10, **5** Symmetric ribbon	(4th) 1.36	*0.0341* 0.160	Symmetric ribbon (Scheme 2d)	3.52 Uracil-like 6,3-Net

^{a} ΔE negative values give the energy between the known form and the global minimum structure and positive values give the energy between the known form at the global minimum and the next lowest energy hypothetical structure. ^{b} Molecular overlay is the RMS value as calculated by Mercury for all non-hydrogen atoms between the experimental and *ab initio* optimized molecules. ^{c} Packing overlay is the RMS value calculated for a 15 molecule co-ordination sphere using the crystal packing similarity function in Mercury between the experimental and computed crystal structure using the *ab initio* optimized molecular structure. The tolerance level in the atom–atom distances for finding the overlay was the 20% default, except for 5-hydroxyuracil, where a value of 50% was required, and 5-nitrouracil (optimized) overlaid with NIMFOE01, where a value of 60% was required. ^{d} ΔE of symmetric ribbon—positive numbers give the energy difference between the lowest ranked structure not based on the symmetric ribbon and the symmetric ribbon at the global minimum; negative numbers give the energy gap from the global minimum to the lowest energy symmetric ribbon. ^{e} Motif—where the global minimum is a symmetric ribbon, the motif of the lowest ranked structure not based on the symmetric ribbon is given.

3.5 5-Ethyluracil

The crystal structure of 5-ethyluracil (R = CH$_2$CH$_3$), **1**, consists of non-polar symmetric ribbons which are formed through N–H···O hydrogen bonds (ESI Fig. S4.1 and S4.2†). These ribbons then stack such that flat parallel sheets (Scheme 2b) are produced similar to 5-hydroxyuracil[25] and 5-[1-(3-methoxycarbonyl-*O*-ethylpseudoureido)]uracil.[42]

The result from the computational search (Fig. 1c; ESI Table S2.3†) for crystal structures of 5-ethyluracil was similar to that of thymine, in that the majority of the low energy structures were based on the symmetric ribbon (29/39 within 5 kJ mol^{-1} of the global minimum). The structure corresponding to the experimental form (Fig. 2) was found 0.7 kJ mol^{-1} above the global minimum just slightly less stable than

Fig. 2 Overlay of the experimental crystal structure of 5-ethyluracil (black), with the structure found in the computational search (grey) showing (left) the sheets and (right) viewed perpendicular to the sheets.

structures based on sheets of non-polar parallel symmetric ribbons which were stacked differently.

The *ab initio* optimized molecular conformer, with the ethyl group *syn*-periplanar to the uracil C–H, is very similar to that seen in the crystal structure, and was approximately 20 kJ mol^{-1} more stable than the *anti*-periplanar conformation. Hence, it seems unlikely that other molecular conformations would be found in any condensed phase.

3.6 5-Vinyluracil

In contrast to 5-ethyluracil, 5-vinyluracil (R = CHCH$_2$) is calculated to be conformationally flexible, with two minima on the conformational energy surface which differ by only 2.21 kJ mol^{-1}. The most stable conformation has the vinyl group almost coplanar with the uracil and a =C–H···O distance of 2.28 Å. However this conformer can pack slightly less favourably than the markedly non-planar minimum with the vinyl group on the opposite side (ESI Fig. S2.1†), with the net result that there are a range of structures of both conformers that are virtually equi-energetic (ESI Tables S2.4 and S2.5, Fig. S2.10†). Both conformations show predominantly symmetric ribbons in the low energy range. There are many closely related structures with differing powder patterns, close in energy, giving the possibility of stacking disorder in real crystals.

Despite our best efforts, we were unable to grow crystals of 5-vinyluracil suitable for single crystal X-ray diffraction. A low-resolution powder X-ray diffraction pattern (ESI Fig. S5.1†) shows some similarities to several of the symmetric ribbon structures and one other low energy structure found in the search. Thus we have to leave 5-vinyluracil as a problem for either crystal growth or expert solution from powder data, but note that it is plausible that the structure will be related to one or more of the low energy structures found in the search.

3.7 5-Cyanouracil

Similarly to the majority of 5-substituted uracils, the crystal structure of 5-cyanouracil (R = CN), **2**, (ESI Fig. S4.3†)
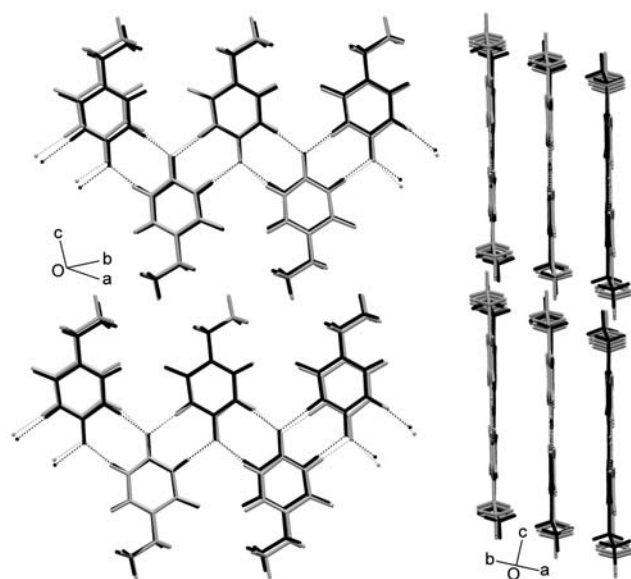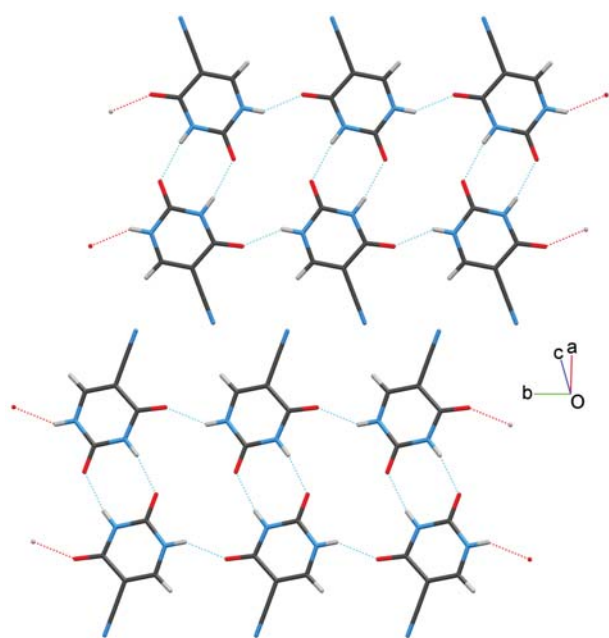
Fig. 3 Unique ladder motif observed for 5-cyanouracil (C—dark grey; H—light grey; N—blue; O—red).
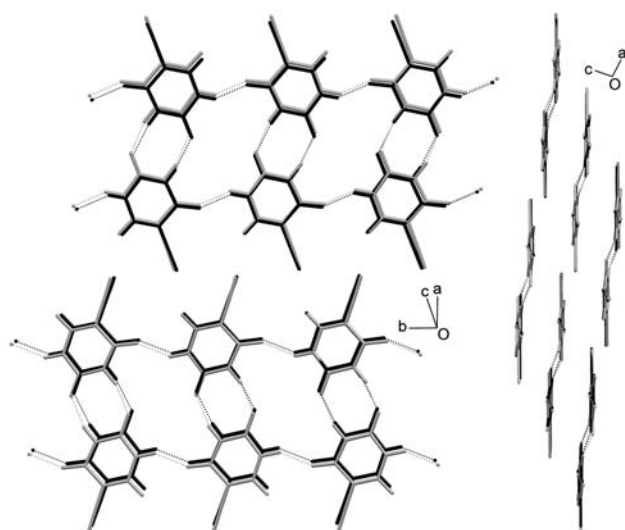
Fig. 4 Overlay of the experimental crystal structure of 5-cyanouracil (black), with the structure found in the computational search (grey) showing (left) the sheets and (right) viewed perpendicular to the sheets.

exhibits N–H⋯O double hydrogen bonds, formed between two adjacent molecules. However, unlike all of the previously published 5-substituted uracil structures, these dimers are linked by single N–H⋯O hydrogen bonds to give a novel ladder motif (Fig. 3 and 4).

The search for possible crystal structures (Fig. 1d; ESI Table S2.6†) of 5-cyanouracil using the *ab initio* optimized geometry found that the ability of the cyano substituent to form inter-actions is exploited and the majority of the low energy structures are based on either simple dimers that are linked by N–H⋯N≡C interactions into 6,3-nets (ESI Fig. S2.5†) or chains which are linked by N–H⋯N≡C interactions into

4,4-grids (ESI Fig. S2.6†). The known structure was found to be 1.76 kJ mol^{-1} less stable than the global minimum. In contrast to the searches performed on the hydrocarbon sub-stituted uracils (CH$_3$, CH$_2$CH$_3$, CHCH$_2$), this search pro-duced the lowest energy symmetric ribbon 4.5 kJ mol^{-1} above the global minimum. A simplistic view of molecular packing might indicate that the cyano groups ought to be able to interdigitate to form the symmetric ribbon motif in a similar fashion to thymine, and this is observed in the computed lowest symmetric ribbon structure. However, this interdigita-tion of the cyano groups is clearly not competitive with the observed structure and other alternatives.

3.8 5-Hydroxyuracil

The previously reported crystallization screen[25] yielded a symmetric ribbon structure of 5-hydroxyuracil (R = OH). The non-polar parallel symmetric ribbons are joined by addi-tional O–H⋯O2 double hydrogen bonds (O⋯O2 = 2.67 Å; O–H⋯O2 = 132.4°) to form slightly stepped sheets.

The energy difference between the experimental structure and the global minimum of the computational search (Fig. 1e; ESI Table S2.7†) is 2.7 kJ mol^{-1}, with the few lower energy structures being based on 4,4-grids of 5-hydroxyuracil dimers. The relative energies are clearly sensitive to the potential model, as a previous search[25] with a reduced carbon repulsion gave a qualitatively similar picture, but with the known structure being further (3.1 kJ mol^{-1}) from the global mini-mum. Similarly, although the conformational barrier for rotation of the hydroxyl group is around 30 kJ mol^{-1}, small changes of up to 30° are possible with such a small energy penalty that a more sophisticated calculation,[70] optimizing this torsion angle within the crystal packing, might well reorder the structures. Hence, the hydroxyl functionality appears to be capable of both stabilizing the symmetric ribbon motif, and forming competitive structures with the hydroxyl hydrogen bonding to uracil's carbonyl.

3.9 5-Aminouracil

The low energy structures predicted for 5-aminouracil (R = NH$_2$) show some particularly favoured 3-D structures in which the NH$_2$ group is involved as both a hydrogen bond donor and acceptor (ESI Table S2.8; Fig. S2.10†). The sym-metric ribbon motif only occurs 3.02 kJ mol^{-1} above the global minimum, and these ribbons are arranged in a criss-cross fashion. Interdigitation of the amine groups to form symmetric ribbon sheets results in structures almost 7 kJ mol^{-1} above the global minimum. However, these relative energies are very sensitive to the amine geometry which is quite pyramidal in the *ab initio* optimized structure. An alternative search with a conformation where the amine group was constrained to be planar (not reported here) yields various symmetric ribbon motifs as the most stable structures, though these were significantly less stable than the crystal structures found using the optimized pyramidal amine geometry. The approximately 10 kJ mol^{-1} difference between the global minima obtained using the two geometries will be increased by the small planarization barrier for the amine group, which

has been shown to be very demanding of *ab initio* methodology for nucleic acid bases.[71]

Despite our best efforts, we were unable to grow crystals of 5-aminouracil suitable for single crystal X-ray diffraction. A low-resolution powder X-ray diffraction pattern shows some similarities to those simulated for the symmetric ribbon structures found in the search (ESI Fig. S5.2†), with a qualitative comparison favouring fc35 over bd74. Given the sensitivity of the computed crystal structures to the amine conformation, it would probably be necessary to refine the amine conformation within low energy crystal structures, using DMAflex,[70] in order to provide simulated powder patterns that might be useable as a starting point for Rietveld refinement. This might prove worthwhile if high resolution powder data could not be indexed and used to solve the structure.

3.10 5-Nitrouracil

5-Nitrouracil (R = NO₂) has 3 polymorphic forms,[59,60] one monohydrate[72] and 9 other co-crystals.[27,60] Two of the polymorphs[59,60] adopt the asymmetric ribbon (Scheme 3a) whilst the third[59] forms a 4,4-grid (Scheme 3g). The majority of the co-crystallized forms exhibit interactions between the different species, resulting in 6 of the structures forming 5-nitrouracil dimers,[27,72] two having no uracil–uracil interactions at all[27] and one chain structure.[27] Only the 1,4-dioxane solvate[27] does not form any interactions between the 5-nitrouracil and the solvate molecule.

The gas phase *ab initio* optimized molecule has a torsion angle for C4 C5 N_{NO2} O_{NO2} of 27.8° which is very different from the corresponding angles of 5.9° (NIMFOE), 3.5° (NIMFOE01) and 3.5° (NIMFOE02) for the three polymorphs. Since this angle is very different, a second *ab initio* optimization of the gas phase molecule was performed such that the NO₂ group was constrained to lie within the same plane the rest of the uracil molecule (ESI Fig. S2.2†). The conformational energy difference is calculated to be 1.4 kJ mol⁻¹ explaining why there is such a high degree of variation in the observed experimental 5-nitrouracil crystal structures (ESI Table S1.3 for measured nitro torsion angles†), with torsion angles varying between the virtually planar 1.0° for the 1,4-dioxane solvate[27] to 19.2° for the *N,N'*-dimethyl-piperazine solvate.[27]

Both the predicted crystal structures and the experimental crystal structures exhibit a variety of different motifs, which are dependent on the nitro conformation (ESI Tables S2.9, S2.10; Fig. S2.10†), and all three polymorphs and the complexes with ethanol or DABCO and water involve the nitro group in certain key interactions. 59% of the 68 hypothetical structures found in the search with the gas phase optimized conformation exhibit close nitro–nitro interactions, with 69% exhibiting close nitro–uracil interactions. For the search carried out with the planar conformation, 55% of the 44 structures have nitro–nitro interactions and 57% have nitro–uracil close contacts. Only 9% and 14% of structures found in the searches with the optimized and planar conformations, respectively, have no interactions involving the nitro group. In accord with the observation that none of the experimental forms adopt the symmetric ribbon, the search shows that this

motif is energetically disfavoured with a 3.24 kJ mol⁻¹ energy gap between the lowest energy experimental structure and the lowest energy symmetric ribbon structure in the search with the *ab initio* optimized conformation, with the NO₂ groups interdigitating to form sheets. Using the planar conformation, form **III** (NIMFOE02) was found at the global minimum, approximately 4.5 kJ mol⁻¹ more stable than forms **I** and **II** (which were not found in the search). The lowest energy structure showing the symmetric ribbon motif in this search was 9.4 kJ mol⁻¹ less stable than the lowest energy experimentally known form.

3.11 5-Fluorouracil

Previous studies on 5-fluorouracil (R = F) have shown it to be polymorphic[16] with form **II** having sheets of asymmetric ribbons (Scheme 3a) whilst form **I** is a $Z' = 4$ structure that adopts a unique 4.8²-net (Scheme 3f). 5-Fluorouracil also forms 9 co-crystals and solvates (ESI Table S1.2†) which adopt a wide variety of motifs from the theophylline monohydrate co-crystal[73] and DMSO solvate[19] which have no interactions between the uracil moieties, through dimers (for cytosine,[74] 9-ethylhypoxanthine[75] and 1-methylcytosine[76]) and chains (1,4-dioxane solvate[17]) to two and three-dimensional structures observed for the DMF solvate[18] and thymine solid solution,[21] respectively. Only one structure, the trifluoroethanol solvate,[20] adopts the symmetric ribbon motif.

The computational search (ESI Table S2.11; Fig. S2.10†) found the asymmetric ribbon of form **II** as the global minimum in energy, and two asymmetric ribbon structures are slightly more stable than a variety of symmetric ribbon structures. These, and some 6,3-nets are calculated to be more stable than the $Z' = 4$ structure of the well known form **I**. The heat of fusion rule and melting points of the two polymorphs of 5-fluorouracil suggest that the calculated relative lattice energies are the wrong way around.[16]

As part of our polymorph screening studies on 5-fluorouracil a number of solvates were isolated: 1,4-dioxane,[17] DMF,[18] DMSO[19] and TFE[20] have already been reported and were included in the CSD analysis. Since then, two further solvates have been isolated and structurally characterized; 5-fluorouracil benzonitrile (1 : 1), **3**, and 5-fluorouracil formamide (1 : 1), **4**. The benzonitrile solvate structure (Fig. 5; ESI Fig. S4.4 and S4.5†) is similar to that observed for the TFE solvate, as this also forms symmetric ribbons (polar, parallel) separated by regions of interdigitating solvate molecules. It is noteworthy that in contrast to 5-cyanouracil, the benzonitrile molecules form ribbons with interdigitating cyano groups.

The 5-fluorouracil formamide solvate, **4**, forms a novel 6,3-net with brick wall type topology (Fig. 6; ESI Fig. S4.6 and S4.7†). The formamide molecules are vital to the hydrogen bonded motif as there are no interactions between 5-fluorouracil molecules, instead N–H···O (N1···O = 2.71 Å; N1–H···O = 176.5° and N···O8 = 2.91 Å; N–H···O8 = 165.5°) hydrogen bonds give chains of alternating 5-fluorouracil and formamide molecules which are cross-linked by further double N–H···O hydrogen bonds (N3···O = 2.81 Å; N3–H···O = 175.6° and N···O7 = 2.92 Å; N–H···O7 = 164.1°).
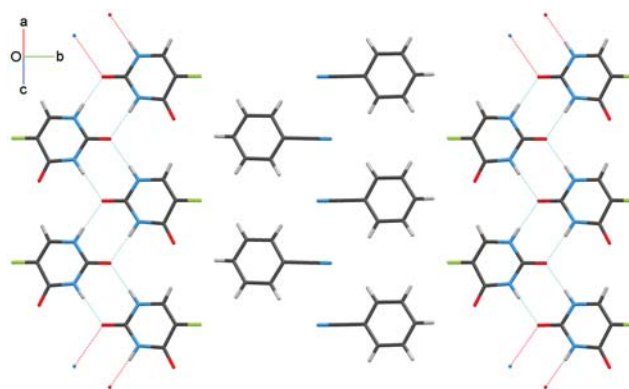
Fig. 5 The sheets formed by alternating rows of benzonitrile and 5-fluorouracil symmetric ribbons (C—dark grey; H—light grey; N—blue; O—red).
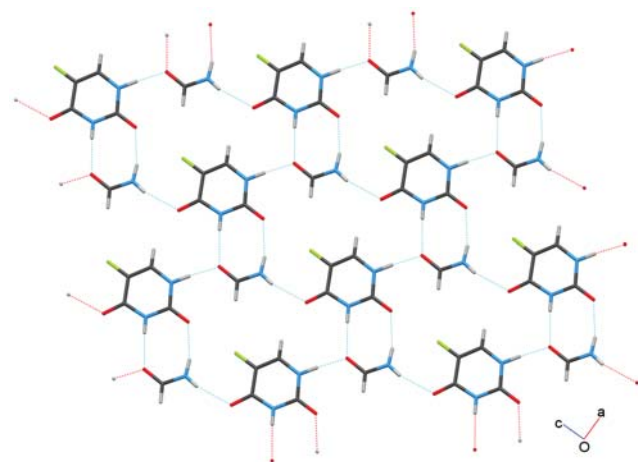

Fig. 6 The sheets formed by 5-fluorouracil formamide solvate (C—dark grey; H—light grey; N—blue; O—red).

3.12   5-Chlorouracil

5-Chlorouracil (R = Cl), **5**, consists of sheets of non-polar anti-parallel symmetric ribbons (Scheme 2d; ESI Fig. S4.8 and S4.9†). The crystal structure was found to be disordered with the 5-chlorouracil molecule disordered between two orientations corresponding to a 180° rotation about the O1 C2 C5 Cl axis corresponding effectively to an interchange of the oxygen atom on C4 and hydrogen atom on C6. By considering the two components separately and the ribbons formed, it can be seen that both parts form the same structure. This structure, and the disorder observed, is very similar to that found for 5-ethynyluracil (eniluracil, R = C≡C–H).[77]

The crystal energy landscape is dominated by symmetric ribbon structures (Fig. 1f; ESI Table S2.12†), with all four sheet variants occurring within 1.5 kJ mol^{-1}. Since these structures are only distinguishable by the orientation of the C6–H3 and C4=O2 groups, there is little energy penalty for growth errors involving stacking of the sheets, interdigitation of the ribbons or even different types of ribbon being incorporated as defects in the structure. Both the major and minor components of the disordered crystal structure of 5-chlorouracil correspond to ak108, the anti-parallel non-polar symmetric ribbon structure. However, the clear occurrence of

disorder in the structure can be attributed to the equi-energetic interchangeable motifs. This relationship between the crystal energy landscape showing this set of closely related symmetric ribbons and disorder in the single crystals has recently been demonstrated by the variable disorder in four single crystal determinations of eniluracil.[77]

3.13   5-Bromouracil

The crystal structure of 5-bromouracil (R = Br), **6**, was found to exhibit an identical structure to 5-chlorouracil and is also composed of disordered sheets of non-polar anti-parallel symmetric ribbons (ESI Fig. S4.10 and S4.11†). The similarity in the anisotropic van der Waals radii[78] for Br···Br and Cl···Cl allows both Br and Cl to interdigitate in a quite similar, favourable,[79,80] geometry, despite the symmetric ribbon motif constraint. The similarities between the three disordered single crystal structures of 5-bromouracil, 5-chlorouracil and eniluracil, and in the crystal energy landscapes of 5-chlorouracil and eniluracil suggest the same cause of the disorder[77] and hence likely sample dependence.

3.14   5-Iodouracil

Attempts to crystallize 5-iodouracil from eleven solvents only found the previously reported structure. This is a 4,4-grid with close non-linear iodine–iodine interactions. It is not obvious that the observed iodine interactions would significantly stabilize this structure over a symmetric ribbon structure, although the larger van der Waals radius of iodine would increase the C–X···X angle from those observed for bromo and chlorouracil. The large anisotropic polarizability of organic iodine makes evaluating the relative energies of these structures too computationally demanding.

4.   Discussion

The CSD survey shows that the symmetric ribbon is a frequently observed hydrogen bonding motif for the 5-substituted uracils, where there is an extended network of uracil···uracil hydrogen bonding. Many substituents on the uracil or other components in the crystal competed successfully to give structures in which the uracil did not hydrogen bond to itself or only produced dimers. However, the symmetric ribbon motif was able to accommodate quite a wide

range of substituents and solvates, which generally interdigitate to form sheets, although with the flexibility to undulate. Most of the alternative motifs observed for 5-substituted uracils (Scheme 3) are far more limited in the size and type of substituent that can be accommodated without disrupting the uracil···uracil hydrogen bonded network.

We have attempted to understand the preference for the symmetric ribbon motif, by studying the experimental and computed low energy structures for a range of simple 5-substituted uracils. It appears that two simple uracil hydrogen bonding motifs, the symmetric ribbon and the asymmetric ribbon, are amongst the energetically feasible structures for all 10 5-substituted uracils studied here. The nature of the substituent plays a key role in determining which other hydrogen bonding motifs are also energetically feasible, for example, R = H is the only substituent small enough for the brick wall 6,3-net (Scheme 3c), although much larger substituents can be accommodated within the alternative herringbone 6,3-net (Scheme 3d) arrangement.

The nature of the substituent is vital in determining the packing of the symmetric and asymmetric ribbons. For example, R = NO_2, ensures that the asymmetric ribbons pack in a herringbone motif, whereas R = F gives a planar sheet structure. Many of the substituents can interdigitate to form nearly planar sheets. The substituents with conformational flexibility adopt the lowest energy nearly-planar conformations, and the packing of the ethyl and hydroxyl groups dictates the adoption of the non-polar, parallel (Scheme 2b) version of the symmetric ribbon motif. However, the simple substituents R = Cl and C≡C–H do not produce a significant energy difference between the parallel and anti-parallel interdigitation of either the polar or non-polar ribbons in the layer crystal structures, resulting in the disorder found in these crystal structures.

The crystal energy landscapes have been calculated with a relatively standard method which neglects all thermal effects, assumes that the molecule adopts the gas phase *ab initio* optimized conformation, uses an empirically based model intermolecular potential and only searches for $Z' = 1$ structures in the most common space groups. Undoubtedly, allowing small changes in conformation in response to crystal packing forces would change the relative energies of the structures, particularly in the case of NH_2 pyramidalization, NO_2 flexibility and rotation of CH_2CH_3, $CHCH_2$ and OH. The relative energies of the different hydrogen bonding motifs might change somewhat with a more accurate intermolecular potential that included the induction energy explicitly. Our estimates of the harmonic $k = 0$ phonons suggest that there will be variations in zero-point energy and thermal energy that would also affect the relative order of stability of some structures at ambient conditions. However, the observation that all searches had a known experimental structure within 3 kJ mol^{-1} of the global minimum and all polymorphs were found within 6 kJ mol^{-1} of the global minimum supports the argument that this approach is accurate enough to conclude that the range of structures found within 5 kJ mol^{-1} can be considered thermodynamically feasible.

The number of structures predicted on this crystal energy landscape would undoubtedly be increased by a wider search, considering higher Z' values. It is uncertain whether this would find new hydrogen bonding motifs that are only possible within $Z' = 2$ (as observed, for example, in the most stable polymorph of 7-fluoroisatin[81]). However, undoubtedly in all cases where there are two different stackings of the same uracil sheet that are very close in energy, then increasing Z' would increase the number of variants in the repeat stacking, *e.g.*, mixing the two stacking modes.

Thus, we have many low energy crystal structures with alternative hydrogen bonding motifs for all our uracils. The strength and directionality of these hydrogen bonding motifs would suggest that there is a sufficient barrier to prevent the facile transformation from one motif to another, which is supported by no transformations being detected in the DSC studies on 5-fluorouracil forms **I** and **II**, below the melting temperature. This implies that if a metastable polymorph were formed with a different hydrogen bonding motif to the thermodynamically stable form, it would persist to be a practically relevant polymorph. (This would be in marked contrast to a nearly spherical imide,[82] where existence of a plastic phase and calculations showed that it was unlikely that a dimer hydrogen bonded motif could be trapped as a metastable polymorph). Thus, the computed energy landscapes immediately raise the question as to why polymorphs have only been found for 5-substituted uracils with R = NO_2 and F, although these are also the only two 5-substituted uracils which have also been found to exhibit a wide range of further co-crystals and solvates.

The degree of polymorph screening performed on the other uracils is fairly limited, and it can only be said to indicate that no other forms that grow as X-ray quality single crystals are readily achievable by manual solvent studies. The 5-substituted uracils as a group decompose on heating before they melt, and are extremely insoluble, and hence the range of crystallization conditions that can be explored is limited. The observation that even a trace of water in dry nitromethane can prevent the growth of form **II** of 5-fluorouracil suggests a high degree of solvent specificity.[16] Indeed, the molecular dynamics rationalization of this observation,[83] *i.e.*, that strong hydration of the uracil promotes the early association by F···F intermolecular contacts and provides a barrier to the formation of the double hydrogen bond $R_2^2(8)$ motif, suggests that such kinetic effects may play a major role in determining the crystallization outcome of all the uracils.

A more subtle kinetic effect has been observed when the crystal energy landscapes suggest that there is the possibility of growth mistakes which would not be readily corrected because of the strength of the hydrogen bonding. This has been observed as disorder in the symmetric ribbon structures observed for R = Cl and Br. Since we have been unable to grow suitable crystals of 5-aminouracil and 5-vinyluracil, and yet the powder patterns bear some resemblance to two or more low energy symmetric ribbon structures, it is tempting to relate the crystal growth problems to more extensive disorder for R = NH_2 and $CHCH_2$ than characterized for R = Br, Cl and C≡C–H. However, our overall conclusion is that these are systems where the nature of the molecule severely limits the range of crystallization conditions that can be applied, and new techniques[84] may well show a more diverse solid state behaviour.

5. Conclusions

The symmetric ribbon motif (Scheme 2) for 5-substituted uracils superficially appears to be a strongly hydrogen bonded motif that might be expected to be quite a robust supramolecular synthon. Although there are other uracil hydrogen bonding motifs that have all the carbonyl acceptors involved in conventional hydrogen bonds between the uracil groups, (Scheme 3a, d, e, f, g) some of these require specific acceptors and are less frequently observed in the CSD survey. Our comparisons of the computed crystal energy landscapes and observed structures for a dozen simply substituted uracils show that the relative energies of the different uracils motifs are very dependent on substituent group. A specific motif is rarely favoured significantly in energy, at least compared with the energy range of possible polymorphism and the errors in the calculations. Indeed, for simple substituents like R = Cl, Br and C$\equiv$C–H the closeness in energy of different variants on the same symmetric ribbon structure rationalizes the disorder we observed in the crystal structure. The observed preference for the symmetric ribbon motif appears to merely reflect its ability to accommodate a wider range of substituents without disrupting the uracil···uracil hydrogen bonding network. Most of the substituents that are capable of hydrogen bonding with the uracil core form particular structures utilizing this interaction, with the exception of 5-hydroxyuracil where a hydroxyl hydrogen bonding network compatible with the symmetric ribbon motif is formed.

The computed energy landscapes appear[12] to be predicting that most of the uracils studied are likely to be polymorphic, in some cases are likely to exhibit a range of disorder, and may form multi-component systems to solve the packing problem. Indeed, the most extensively studied uracils, 5-fluorouracil and 5-nitrouracil, are polymorphic. Whilst both solvent dependent and disorder kinetic factors have been identified, it seems that the predominant factor limiting the solid form diversity of the uracils is the difficulty in covering a wide range of crystallization conditions. This study clearly illustrates the problems[1] in engineering a crystal structure when there are a number of competing hydrogen bonding motifs and that all substituents play a role in determining the observed crystal structure. However, the calculation of the possible range of low energy crystal structures should automatically balance all the different types of intermolecular interactions, making it a powerful aid to the understanding of the experimental crystal structures and the supramolecular recognition of the uracil functional group.

References

1 G. R. Desiraju, *Angew. Chem., Int. Ed.*, 2007, **46**, 8342–8356.
2 C. B. Aakeroy, *Acta Crystallogr., Sect. B*, 1997, **53**, 569–586.
3 C. V. K. Sharma, *Cryst. Growth Des.*, 2002, **2**, 465–474.
4 P. Erk, H. Hengelsberg, M. F. Haddow and R. van Gelder, *CrystEngComm*, 2004, **6**, 474–483.
5 B. Moulton and M. J. Zaworotko, *Chem. Rev.*, 2001, **101**, 1629–1658.
6 G. R. Desiraju, *Crystal Engineering: the Design of Organic Solids*, Elsevier, Amsterdam, 1989.
7 J. Chisholm, E. Pidcock, J. van de Streek, L. Infantes, S. Motherwell and F. H. Allen, *CrystEngComm*, 2006, **8**, 11–28.
8 F. H. Allen, W. D. S. Motherwell, P. R. Raithby, G. P. Shields and R. Taylor, *New J. Chem.*, 1999, **23**, 25–34.
9 F. H. Allen, *Acta Crystallogr., Sect. B*, 2002, **58**, 380–388.
10 D. Braga, F. Grepioni and J. J. Novoa, *Chem. Commun.*, 1998, 1959–1960.
11 J. D. Dunitz and A. Gavezzotti, *Angew. Chem., Int. Ed.*, 2005, **44**, 1766–1787.
12 S. L. Price, *Phys. Chem. Chem. Phys.*, 2008, **10**, 1996–2009.
13 T. Beyer and S. L. Price, *J. Phys. Chem. B*, 2000, **104**, 2647–2655.
14 M. C. Etter, *Acc. Chem. Res.*, 1990, **23**, 120–126.
15 T. C. Lewis, D. A. Tocher and S. L. Price, *Cryst. Growth Des.*, 2005, **5**, 983–993.
16 A. T. Hulme, S. L. Price and D. A. Tocher, *J. Am. Chem. Soc.*, 2005, **127**, 1116–1117.
17 A. T. Hulme and D. A. Tocher, *Acta Crystallogr., Sect. E*, 2004, **60**, O1781–O1782.
18 A. T. Hulme and D. A. Tocher, *Acta Crystallogr., Sect. E*, 2004, **60**, O1783–O1785.
19 A. T. Hulme and D. A. Tocher, *Acta Crystallogr., Sect. E*, 2004, **60**, O1786–O1787.
20 A. T. Hulme and D. A. Tocher, *Acta Crystallogr., Sect. E*, 2005, **61**, O3661–O3663.
21 S. A. Barnett, A. T. Hulme and D. A. Tocher, *Acta Crystallogr., Sect. C*, 2006, **62**, O412–O415.
22 A. T. Hulme and D. A. Tocher, *Cryst. Growth Des.*, 2006, **6**, 481–487.
23 A. Johnston, A. J. Florence, N. Shankland, A. R. Kennedy, K. Shankland and S. L. Price, *Cryst. Growth Des.*, 2007, **7**, 705–712.
24 N. Blagden, 2008, private communication.
25 R. C. B. Copley, L. S. Deprez, T. C. Lewis and S. L. Price, *CrystEngComm*, 2005, **7**, 421–428.
26 A. T. Hulme, *Experimental and Computational Studies of Polymorphism of Small Organic Molecules*, PhD thesis, University College London, UK, 2007.
27 R. Thomas, R. S. Gopalan, G. U. Kulkarni and C. N. R. Rao, *Beilstein J. Org. Chem.*, 2005, **1**, 15.
28 C. F. Macrae, P. R. Edgington, P. McCabe, E. Pidcock, G. P. Shields, R. Taylor, M. Towler and J. De Streek, *J. Appl. Crystallogr.*, 2006, **39**, 453–457.
29 G. M. Sheldrick, *SHELXS97*, Universityof Göttingen, Göttingen, Germany, 1997.
30 G. M. Sheldrick, *SHELXL97*, University of Göttingen, Göttingen, Germany, 1997.
31 *SADABS [2.03]*, Bruker AXS Inc., Madison, Wisconsin, USA, 2001.
32 H. Sternglanz and C. E. Bugg, *Biochim. Biophys. Acta*, 1975, **378**, 1–11.
33 *GEMINI [1.02]*, Bruker AXS Inc., Madison, Wisconsin, USA, 2000.
34 *SMART [5.625]*, Bruker AXS Inc., Madison, Wisconsin, USA, 2001.
35 *SAINT+ [6.45]*, Bruker AXS Inc., Madison, Wisconsin, USA, 2003.
36 *TWINABS [1.02]*, Bruker AXS Inc., Madison, Wisconsin, USA, 2003.
37 *SHELXTL [6.12]*, Bruker AXS Inc., Madison, Wisconsin, USA, 2000.
38 G. Portalone, L. Bencivenni, M. Colapietro, A. Pieretti and F. Ramondo, *Acta Chem. Scand.*, 1999, **53**, 57–68.
39 T. Sakurai and M. Okunuki, *Acta Crystallogr., Sect. B*, 1971, **27**, 1445–1453.

40 A. Hempel, B. G. Lane and N. Camerman, *Acta Crystallogr., Sect. C*, 1997, **53**, 1707–1709.
41 G. J. B. Williams, A. J. Varghese and H. M. Berman, *J. Am. Chem. Soc.*, 1977, **99**, 3150–3154.
42 J. W. Chern, D. S. Wise, W. Butler and L. B. Townsend, *J. Org. Chem.*, 1988, **53**, 5622–5628.
43 M. Yokoyama, T. Ikeue, Y. Ochiai, A. Momotake, K. Yamaguchi and H. Togo, *J. Chem. Soc., Perkin Trans. 1*, 1998, 2185–2191.
44 R. Gerdil, *Acta Crystallogr.*, 1961, **14**, 333–344.
45 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, C. Gonzalez and J. A. Pople, *GAUSSIAN 03*, Gaussian, Inc., Wallingford, CT, 2004.
46 A. J. Stone and M. Alderton, *Mol. Phys.*, 1985, **56**, 1047–1064.
47 A. J. Stone, *GDMA: A Program for Performing Distributed Multipole Analysis of Wave Functions Calculated Using the Gaussian Program System [1.0]*, University of Cambridge, Cambridge, UK, 1999.
48 S. L. Price, *J. Chem. Soc., Faraday Trans.*, 1996, **92**, 2997–3008.
49 D. S. Coombes, S. L. Price, D. J. Willock and M. Leslie, *J. Phys. Chem.*, 1996, **100**, 7352–7360.
50 L. Y. Hsu and D. E. Williams, *Acta Crystallogr., Sect. A*, 1980, **36**, 277–281.
51 D. E. Williams and D. J. Houpt, *Acta Crystallogr., Sect. B*, 1986, **42**, 286–295.
52 T. C. Lewis, *Investigating Polymorphic Behaviour of Organic Molecules using Theoretical and Experimental Techniques*, PhD thesis, University College London, 2005.
53 J. R. Holden, Z. Y. Du and H. L. Ammon, *J. Comput. Chem.*, 1993, **14**, 422–437.
54 S. L. Price, D. J. Willock, M. Leslie and G. M. Day, *DMAREL 3.02*, 2001. http://www.ucl.ac.uk/~ucca17p/dmarelmanual/dmarel.html.
55 D. J. Willock, S. L. Price, M. Leslie and C. R. A. Catlow, *J. Comput. Chem.*, 1995, **16**, 628–647.
56 I. Krivy and B. Gruber, *Acta Crystallogr., Sect. A*, 1976, **32**, 297–298.
57 A. L. Spek, *PLATON, A Multipurpose Crystallographic Tool*, Utrecht University, Utrecht, The Netherlands, 2003.
58 G. M. Day, S. L. Price and M. Leslie, *Cryst. Growth Des.*, 2001, **1**, 13–26.
59 R. S. Gopalan, G. U. Kulkarni and C. N. R. Rao, *ChemPhysChem*, 2000, **1**, 127–135.
60 A. R. Kennedy, M. O. Okoth, D. B. Sheen, J. N. Sherwood and R. M. Vrcelj, *Acta Crystallogr., Sect. C*, 1998, **54**, 547–550.
61 A. Momotake, J. Mito, K. Yamaguchi, H. Togo and M. Yokoyama, *J. Org. Chem.*, 1998, **63**, 7207–7212.
62 H. M. Berman, D. E. Zacharias, H. L. Carrell and A. J. Varghese, *Biochemistry*, 1976, **15**, 463–467.
63 R. F. Stewart and L. H. Jensen, *Acta Crystallogr.*, 1967, **23**, 1102–1105.
64 H. Zimmermann, M. Gomm, J. Ellermann and E. Kock, *Acta Crystallogr., Sect. C*, 1987, **43**, 1798–1800.
65 S. Hanessian and D. J. Ritson, *J. Org. Chem.*, 2006, **71**, 9807–9817.
66 H. Sternglanz, G. R. Freeman and C. E. Bugg, *Acta Crystallogr., Sect. B*, 1975, **31**, 1393–1395.
67 B. Twamley, O. D. Gupta and J. M. Shreeve, *Acta Crystallogr., Sect. E*, 2002, **58**, O1040–O1042.
68 C. Ouvrard and J. B. O. Mitchell, *Acta Crystallogr., Sect. B*, 2003, **59**, 676–685.
69 S. L. Price and K. S. Wibley, *J. Phys. Chem. A*, 1997, **101**, 2198–2206.
70 P. G. Karamertzanis and S. L. Price, *J. Chem. Theory Comput.*, 2006, **2**, 1184–1199.
71 S. Y. Wang and H. F. Schaefer, *J. Chem. Phys.*, 2006, **124**, 044303.
72 B. M. Craven, *Acta Crystallogr.*, 1967, **23**, 376–383.
73 S. Zaitu, Y. Miwa and T. Taga, *Acta Crystallogr., Sect. C*, 1995, **51**, 1857–1859.
74 D. Voet and A. Rich, *J. Am. Chem. Soc.*, 1969, **91**, 3069–3075.
75 S. H. Kim and A. Rich, *Science*, 1967, **158**, 1046.
76 S. H. Kim and A. Rich, *J. Mol. Biol.*, 1969, **42**, 87.
77 R. C. B. Copley, S. A. Barnett, P. G. Karamertzanis, K. D. M. Harris, B. M. Kariuki, M. Xu, E. A. Nickels, R. W. Lancaster and S. L. Price, *Cryst. Growth Des.*, 2008, DOI: 10.1021/cg800517h.
78 S. C. Nyburg and C. H. Faerman, *Acta Crystallogr., Sect. B*, 1985, **41**, 274–279.
79 S. L. Price, A. J. Stone, J. Lucas, R. S. Rowland and A. E. Thornley, *J. Am. Chem. Soc.*, 1994, **116**, 4910–4918.
80 F. F. Awwadi, R. D. Willett, K. A. Peterson and B. Twamley, *J. Phys. Chem. A*, 2007, **111**, 2319–2328.
81 S. Mohamed, S. A. Barnett, D. A. Tocher, K. Shankland, C. K. Leech and S. L. Price, *CrystEngComm*, 2008, **10**, 399–404.
82 A. T. Hulme, A. Johnston, A. J. Florence, P. Fernandes, K. Shankland, C. T. Bedford, G. W. A. Welch, G. Sadiq, D. A. Haynes, W. D. S. Motherwell, D. A. Tocher and S. L. Price, *J. Am. Chem. Soc.*, 2007, **129**, 3649–3657.
83 S. Hamad, C. Moon, C. R. A. Catlow, A. T. Hulme and S. L. Price, *J. Phys. Chem. B*, 2006, **110**, 3323–3329.
84 A. Llinas and J. M. Goodman, *Drug Discovery Today*, 2008, **8**, 198–210.