Published in final edited form as:
Cryst Growth Des. 2016 June 1; 16(6): 3480–3496. doi:10.1021/acs.cgd.6b00459.

Computational and Experimental Characterization of Five Crystal Forms of Thymine: Packing Polymorphism, Polytypism/Disorder and Stoichiometric 0.8-Hydrate

Doris E. Braun[†,*], **Thomas Gelbrich**[†], **Klaus Wurst**[§], and **Ulrich J. Griesser**[†]

[†]Institute of Pharmacy, University of Innsbruck, Innrain 52c, 6020 Innsbruck, Austria

[§]Institute of General, Inorganic and Theoretical Chemistry, University of Innsbruck, Innrain 80/82, 6020 Innsbruck, Austria

Abstract

New polymorphs of thymine emerged in an experimental search for solid forms, which was guided by the computationally generated crystal energy landscape. Three of the four anhydrates (AH) are homeoenergetic (**A°** – **C**) and their packing modes differ only in the location of oxygen and hydrogen atoms. **AHs A°** and **B** are ordered phases, whereas **AH C** shows disorder (X-ray diffuse scattering). Anhydrates **AHs A°** and **B** are ordered phases, whereas **AH C** shows disorder (X-ray diffuse scattering). Analysis of the crystal energy landscape for alternative **AH C** hydrogen bonded ribbon motifs identified a number of different packing modes, whose 3D structures were calculated to deviate by less than 0.24 kJ mol^{-1} in lattice energy. These structures provide models for stacking faults. The three anhydrates **A°** – **C** show strong similarity in their powder X-ray diffraction, thermoanalytical and spectroscopic (IR and Raman) characteristics. The already known anhydrate **AH A°** was identified as the thermodynamically most stable form at ambient conditions; **AH B** and **AH C** are metastable but show high kinetic stability. The hydrate of thymine is stable only at water activities (a_w) > 0.95 at temperatures ≤25 °C. It was found to be a stoichiometric hydrate despite being a channel hydrate with an unusual water:thymine ratio of 0.8:1. Depending on the dehydration conditions, either **AH C** or **AH D** is obtained. The hydrate is the only known precursor to **AH D**. This study highlights the value and complementarity of simultaneous explorations of computationally and experimentally generated solid form landscapes of a small molecule anhydrate ↔ hydrate system.

Keywords

Crystal structure prediction; X-ray diffraction; thermal analysis; isothermal calorimetry; vibrational spectroscopy; gravimetric moisture sorption/desorption; water activity; dehydration

[*]Author to whom correspondence should be addressed: Dr. Doris E. Braun, Institute of Pharmacy, University of Innsbruck, Innrain 52c, 6020 Innsbruck, Austria, Tel: +43(0)512 507 58653, doris.braun@uibk.ac.at.

Associated Content

Computational and experimental details (PDF). This information is available free of charge via the Internet at http://pubs.acs.org/.

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Europe PMC Funders Author Manuscripts

1 Introduction

Crystal polymorphism in general is the ability of a material, with a given chemical composition and chemical bonding situation, to adopt more than one crystal structure. The phenomenon poses a significant technological challenge owing to its influence on solid-state properties that depend on the crystal structure, such as solubility, stability, bioavailability, etc.1–3 As a consequence, polymorphism has been the subject of much research as it is associated with quality control issues for manufacturing fine chemicals, such as pharmaceuticals, dyes or explosives. Understanding the factors that determine the formation of one or more modification(s) is of fundamental importance.4,5 However, in the absence of a universally-accepted definition for crystal polymorphism, with various definitions not specifying the degree of difference between the crystalline modifications,6,7 a differentiation between true polymorphs and disorder,8 i.e. variation in the crystalline samples due to growth conditions (disorder through defect), can sometimes be very challenging.

Polytypism has been defined as follows, "*An element or compound is polytypic if it occurs in several structural modifications, each of which can be regarded as built up by stacking layers of (nearly) identical structure and composition, and if the modifications differ only in their stacking sequence.*"9 Thus, polytypism is seen as a special case of polymorphism, i.e. two-dimensional translations within layers of molecules are essentially preserved.

Organic fine chemicals, in particular active pharmaceutical ingredients (APIs), frequently co-crystallize with water, resulting in a new crystalline phase, termed hydrate. Hydrates are a subset of a larger class of crystalline solids, named solvates, which are characterized by the inclusion of the solvent of crystallization in the crystal structure of the compound.10 Water molecules are small in size and feature hydrogen bonding donor and acceptor groups which can interact with a host molecule or with one another.11,12 As it is virtually impossible to exclude water from products and manufacturing processes, hydrates should be identified, extensively studied and characterized.13–19 Similar to polymorphism, hydrate formation or dehydration can affect the manufacturability, stability, as well as performance (bioavailability) of a (drug) product. The dehydration and rehydration processes of a system forming hydrate(s) can be complex and difficult to control.19–31 Numerous classifications have been proposed for hydrates. Morris and Rodriguez-Hornedo32,33 classified the hydrate structures into three categories that are discernible by common analytical techniques: isolated site hydrates (water molecules are located on isolated sites in the crystal), channel hydrates (chains of water molecules) and ion associated hydrates (metal ions are coordinated to water). The channel hydrates have water molecules located in tunnels or connected pockets, which are exposed to the atmosphere. The water molecules in these hydrates tend to be mobile and can readily escape the tunnels on modest increase in temperature or decrease in relative humidity (RH).34 The knowledge of the thermodynamic parameters related to the hydration/dehydration of organic solid phases is paramount to exert control over this phase transition.35

Thymine (5-methyluracil, Figure 1), a pyrimidine derivative, is an essential component of the deoxyribonucleic acid (DNA) molecule. It is a simple rigid molecule which would have

C_{2v} symmetry if C2=O2 and C4-H were not different functional groups. The compound was first isolated and synthesized in 1903.36,37 The crystal structure of hydrated thymine was first reported in 1961,38 and the reaction mechanism of the dehydration of thymine single crystals to anhydrous thymine has been studied with hot-stage thermal microscopy and related to the channel characteristics of the hydrate.39 A second phase of thymine was reported in 1974 and significant differences in IR bands have been found for the two phases.40 The structure of one of the thymine anhydrates (denoted as **AH A**° in this study) has been reported so far.41,42 Furthermore, a combined experimental and computational screen for thymine solid forms has been performed by Barnett et al.43 The crystallization experiments reproduced the known forms, the hydrate and anhydrate (**AH A**°). The computationally generated anhydrate crystal energy landscape revealed that thymine is likely to be polymorphic. **AH A**° was found in the computationally generated anhydrate crystal energy landscape, albeit not as the global energy minimum structure.43

In the present work we reinvestigate and unravel the solid state behavior of this over 100 years old compound. The structural and thermodynamic relationships between four anhydrous polymorphs and hydrated thymine is explored through a combination of experiment and computation. A broad range of analytical techniques were applied to characterize the thymine solid forms emerging from a polymorph screen and resolve the complexity of the thymine forms. The applied techniques included hot-stage microscopy (HTM), differential scanning calorimetry (DSC), thermogravimetric analysis (TGA), RH-perfusion isothermal calorimetry, powder and single crystal X-ray diffractometry, Infrared and Raman spectroscopy, gravimetric moisture sorption/desorption analysis and water activity measurements. Computational work, crystal structure prediction and lattice energy modelling, aimed at elucidating the system, its dehydration mechanisms, structure models for two anhydrate structures (**AH C** and **AH D**), and the stability of the thymine anhydrate crystal forms relative to other possible structures and ultimately to give a detailed picture of the solid state for thymine.

2 Materials

2.1 Materials and Preparation of Thymine Solid Forms

Thymine was purchased from Sigma (Lot # BCBM3353V, anhydrate **A**°). The reagents and solvents used were also purchased from Sigma and were of analytical grade.

Thymine hydrate (**Hy**) was prepared either by producing a saturated solution of thymine in water at 40 °C and subsequent cooling of the solution to 8 °C to allow crystallization or by slurring thymine in water at 10 °C for five hours. The resulting hydrate crystals were filtered and dried for 10 days at 92% RH. All experiments performed on **Hy** were done with samples that had been stored over a saturated KNO_3 solution (92% RH).

Anhydrate A° (**AH A**°) was obtained in any slurry experiment (temperature range from 10 to 60 °C) using organic solvents (e.g., methanol, 1-butanol, *etc.*) and with any thymine solid form as the starting material. Sublimation experiments at temperatures > 160 °C resulted in **AH B**. The experiments were carried out at ambient pressure on a Kofler hot bench at 160–240 °C; Thymine (**AH A**°) was heated between two glass slides separated by a spacer ring

of 50 mm height. Dehydration of **Hy** at 0% RH (over P_2O_5, under vacuum or N_2 purge) at room temperature or lower temperatures yielded anhydrate D (**AH D**). Dehydration at higher RH (43% – 62.5 %) resulted directly in anhydrate C (**AH C**).

2.2 X-ray Diffraction

2.2.1 Powder X-ray Diffraction—PXRD patterns were obtained using an X'Pert PRO diffractometer (PANalytical, Almelo, NL) equipped with a theta/theta coupled goniometer in transmission geometry, programmable XYZ stage with a well plate holder, a Cu-K$\alpha_{1,2}$ radiation source with a focussing mirror, a 0.5° divergence slit and a 0.02° Soller slit collimator on the incident beam side, a 2 mm antiscattering slit and a 0.02° Soller slit collimator on the diffracted beam side and a solid state PIXcel detector. The patterns were recorded at a tube voltage of 40 kV and tube current of 40 mA, applying a step size of $2\theta =$ 0.013° with 80s or 200s per step in the 2θ range between 2° and 40°. VGI stage (VGI 2000M, Middlesex, UK) was used for non-ambient RH measurements.

The diffraction patterns were indexed with DICVOL04 using the first twenty peaks. The space group was determined on the basis of a statistical assessment of systematic absences44 as implemented in the DASH structure solution package.45 Pawley fits46 were performed with Topas Academic V5.47 The background was modelled with Chebyshev polynomials and the modified Thompson-Cox-Hastings pseudo-Voigt function was used for peak shape fitting.

2.2.2 Single Crystal X-ray Diffraction—Crystals of **AH B**48 suitable for a single-crystal structure determination were obtained by sublimation. The data collection was carried out, using Mo radiation, on a Bruker APEX-II CCD diffractometer controlled by the APEX2 software.49 Data integration and reduction were performed using the SAINT software.49 The crystal structure was solved and refined with SHELXT and SHELXL, respectively.50,51 The twin law was identified using the TwinRotMat routine in PLATON. 52 The two twin components are related by a two-fold rotation about the *a* axis, with refined major/minor occupancy fractions of 0.943(1) : 0.057(1). All non-hydrogen atoms were refined anisotropically. Hydrogen atoms were located in difference maps and those bonded to N atoms were refined with a distance restraint of N–H = 0.86(1) Å and a free U_{iso} parameter. The positions of hydrogen atoms in the CH and CH_3 groups were refined using riding models, with U_{iso} parameters set at $1.2U_{eq}$ and $1.5U_{eq}$, respectively, of the parent C atom.

2.3 Calorimetry/Thermal Analysis

2.3.1 Hot-stage Microscopy (HSM)—A Reichert Thermovar polarization microscope, equipped with a Kofler hot-stage (Reichert, A), was used for HSM investigations. Photographs were taken with an Olympus DP71 digital camera (Olympus, D).

2.3.2 Differential Scanning Calorimetry (DSC)—DSC thermograms were recorded on a DSC 7 or Diamond DSC (Perkin-Elmer Norwalk, Ct., USA), controlled by the Pyris 7.0 software. Using a UM3 ultramicrobalance (Mettler, Greifensee, CH), samples of approximately 2 - 25 mg were weighted into open/closed aluminium pans or hermetically

sealed (high-pressure) pans. For the construction of the temperature/composition phase diagram, **Hy/AH A°** mixtures were prepared by mixing the two phases. The **Hy**/water mixtures were prepared by placing precisely weighted amounts of **Hy** and pure water (with the aid of a Hamiltion syringe) and a second accurate weight measurement into high-pressure DSC pans. The sealed sample pans were stored for 24 hours before the DSC runs were started to equilibrate the mixtures. For low temperature measurements the samples were frozen (liquid N_2) before the DSC runs. The samples were heated using rates ranging from 2 to 10 °C min^{-1}, with dry nitrogen as the purge gas (purge: 20 ml min^{-1}). The two instruments were calibrated for temperature with pure benzophenone (mp 48.0 °C) and caffeine (236.2 °C), and the energy calibration was performed with indium (mp 156.6 °C, heat of fusion 28.45 Jg^{-1}). The errors on the stated dehydration temperatures (extrapolated onset temperatures) and enthalpy values were calculated at the 95% confidence intervals (CI) and are based on at least three measurements.

2.3.3 Thermogravimetric Analysis (TGA)—TGA was carried out with a TGA7 system (Perkin-Elmer, Norwalk, CT, USA) using the Pyris 2.0 software. Approximately 2 - 8 mg of sample was weighted into a platinum pan. Two-point calibration of the temperature was performed with ferromagnetic materials (Alumel and Ni, Curie-point standards, Perkin-Elmer). Heating rates ranging from 2 to 10 °C min^{-1} were applied and dry nitrogen was used as purge gas (sample purge: 20 mL min^{-1}, balance purge: 40 mL min^{-1}).

2.3.4 Isothermal Calorimetry—RH perfusion calorimetry experiments were performed using a TAM III nanocalorimeter and a 4 mL stainless steel RH perfusion ampoule (TA Instruments, USA). The relative humidity was controlled with two mass flow controllers and dry N_2 was used as the carrier gas at a constant flow rate of 100 mL h^{-1}. 19.5 to 28.8 mg of sample was accurately weighed and used. The humidity profile (% RH vs. time) was measured as follows: 80% to 0% in one step (starting material: **Hy**). The RH perfusion cell was calibrated with saturated solutions of NaCl (75.3% RH), $Mg(NO_3)_2$ (52.8% RH) and LiCl (11.3% RH). The heat flow of the empty RH perfusion ampoule (derived from baseline runs with the same humidity steps) was subtracted from the heat flow of the sample measurement. The errors on the stated dehydration enthalpy value are calculated at the 95% CI based on three measurements.

2.4 Coulometric Karl Fischer Titration

Coulometric water determinations were performed using a coulometric titrator C20 instrument (Mettler Toledo, Switzerland).

2.5 Vibrational Spectroscopy

2.5.1 Infrared Spectroscopy—FT-IR spectra were recorded with a Bruker Vertex 70 spectrometer (Bruker Analytische Messtechnik GmbH, D). Potassium bromide (KBr) discs were prepared by gently grinding and mixing <0.5 mg thymine with 230 mg KBr in a mortar with pestle, evacuating of the mixture in the pressing tool for 30 seconds at ca. 10 mbar and applying a pressure of about 800 MPa for about 60 seconds using a hydraulic press. The spectra were recorded in the range of 4000 to 400 cm^{-1} at an instrument resolution of 2 cm^{-1}

(64 scans per spectrum). Temperature conditions were adjusted with a SPECAC (Grasebury Specac Limited, Orpington, UK) variable temperature cell and a temperature control unit.

2.5.2 Raman Spectroscopy—Raman spectra were recorded with a Bruker RFS 100 Raman-spectrometer (Bruker Analytische Messtechnik GmbH, D), equipped with a Nd:YAG Laser (1064 nm) as the excitation source and a liquid-nitrogen-cooled, high sensitivity Ge-detector. The spectra (128 scans per spectrum) were recorded in aluminum sample holders with a laser power of 300 mW and a resolution of 2 cm^{-1}. For investigating H_2O/D_2O exchange, the number of scans and laser power were increased to 1064 and 400 mW, respectively. Samples were stored and measured in hygrostats as detailed in ref. 53.

2.6 Gravimetric Moisture Sorption/Desorption Experiments

Moisture sorption and desorption studies were performed with the automatic multisample gravimetric moisture sorption analyzer SPS23-20μ (ProUmid, Ulm, D). 100 - 600 mg of sample was used for each analysis. The measurement cycles were started at either 40% or 90% RH with an initial stepwise desorption (decreasing humidity) to 0%, followed by a sorption cycle (increasing humidity) up to 90% (95%) and back to 0% relative humidity. RH changes were set to 5% for all cycles. The equilibrium condition for each step was set to a mass constancy of ± 0.001 % over 30 (60) minutes and a maximum time limit of 48 hours for each step.

2.7 Determination of the Critical Water Activity (Slurry Method)

Excess of thymine **AH D** was stirred (500 r.p.m.) in ≥0.5 mL of methanol/water mixtures, each containing a different mole fraction of water corresponding to a defined water activity54,55 (Section 1.3 of the Supporting Information) at 10.0 ± 0.1 °C and 25.0 ± 0.1 °C for seven days. Samples were withdrawn and the resulting phase (wet cake) was determined using PXRD (measured between to mylar foils to avoid a phase transformation during measurement).

2.8 Computational Modelling

2.8.1 Computational Generation of the Anhydrate Crystal Energy Landscape —The molecular geometries of the six tautomers of thymine (1 diketo, 4 keto-enol and 1 dienol) were obtained by an ab initio optimization at the PBE0/6-31G(d,p) level of theory using GAUSSIAN09.56 There was no indication for the presence of tautomers for thymine other than the diketo tautomer.57 Hypothetical anhydrate crystal structures were generated with the program CrystalPredictor.58–60 150,000 structures were randomly generated in 48 space groups, keeping the molecular geometry rigid (diketo tautomer). The structures were relaxed to a local minimum in intermolecular lattice energy, calculated from the FIT61 exp-6 repulsion-dispersion potential and atomic charges, fitted to electrostatic potential around the PBE0/6-31G(d,p) charge density using the CHELPG scheme.62 The energies of all structures within 30 kJ mol^{-1} of the global energy minimum were refined (5,910 structures), using DMACRYS63 with a more realistic, distributed multipole model64 for the electrostatic forces which had been derived using GDMA265 to analyse the PBE0/6-31G(d,p) charge density. The orientation of the methyl group (Figure 1) in the 262 most stable structures (10 kJ mol^{-1} range with respect to the global minimum structure) was

Europe PMC Funders Author Manuscripts

minimized with the program CrystalOptimizer.66 Conformational energy penalties and isolated molecule charge densities were computed both at the PBE0/aug-cc-pVTz and MP2/6-31G(d,p) levels.

For selected low energy structures additional computationally more demanding calculations were performed, to further investigate the sensitivity of the crystal energies to theoretical methods. CASTEP plane wave code67 calculations employed the Perdew-Burke-Ernzerhof (PBE) generalized gradient approximation (GGA) exchange-correlation density functional68 and ultrasoft pseudopotentials,69 with the addition of either the Tkatchenko and Scheffler (TS)70 or the Grimme06 (G06)71 model. PIXEL calculations72–74 were performed on the low energy PBE-TS structures to estimate the repulsive (E_R), dispersion (E_D), electrostatic (Coulombic, E_C) and polarization (also called induction, E_P) contributions from individual pairs of molecules within a crystal. For more details see sections 2.3 and 2.5 of the Supporting Information.

The relationships between crystal structures were examined using the XPac program.75,76 The results described were obtained using all non-hydrogen atoms and routine medium cutoff parameters (δ_{ang} = 10°, δ_{tor} and δ_{dhd} = 18°).

2.8.2 Modelling of the Hydrate Structure—CASTEP (PBE-TS) was applied to derive structure models for thymine hydrate. Four models were generated, covering water:thymine ratios of 1:1, 0.875:1, 0.8125:1 and 0.75:1 (for more details, see section 2.7 of the Supporting Information). The structures were optimized, keeping the lattice parameters fixed at the RT values (derived from indexing and Pawley fitting, see section 1.7 of the Supporting Information).

3 Results

3.1 Experimental Solid Form Screen

The relatively poor solubility of thymine in organic solvents limits the experimental search space for solid forms. In a previous study, all evaporation and vapor diffusion experiments resulted exclusively in the known solid forms, either **AH A**° or **Hy**.43 We extended the experimental screen by solvent slurry, sublimation and systematic dehydration experiments. The slurry experiments, using 27 organic solvents or water and slurring in the temperature range in between 10 and 20 °C, resulted either in the known anhydrous form **AH A**° (all organic solvents) or the hydrate (using water). Sublimation experiments, carried out on a Kofler hot bench at temperatures > 160 °C, resulted in the anhydrate **AH B**.

Systematic dehydration studies of **Hy** were performed and the resulting product was analysed with PXRD and TGA (see section 1.2 of the Supporting Information). Three parameters, the dehydration temperature, relative humidity, and duration were varied. No phase transformation was observed within 72 hours if the hydrate was stored at RHs ≥72% RH at 25 °C. A third anhydrate, **AH C**, was observed in dehydration experiments performed at RHs between 43% and 62.5% at 25 °C. A lower RH (11% to 31%) resulted in a mixture of **AH C** and a fourth polymorph, **AH D**. At the driest conditions (over P_2O_5), phase pure **AH D** was obtained. **AH D** could also be produced by N_2 purge (TGA) or vacuum drying at

temperatures ≤30 °C. Prolonging the dehydration time, but not increasing the temperature, resulted in an **AH D** to **AH C** phase transformation. Furthermore, a slow transformation of **Hy** to anhydrous thymine (**AH C**) was observed at RHs ≤84%. Increasing the dehydration temperature resulted in an increase in the amount of **AH C** and a decrease in **AH D**. Anhydrate **A°** was observed in dehydration experiments, but only if **Hy** was seeded with **AH A°** or if dehydration occurred under isochoric conditions (peritectic decomposition in high-pressure DSC capsules, see section 3.5.1).

Cooling crystallization and solvent evaporation experiments using either methanol, 1-butanol, 1,4-dioxane, acetonitrile, ethyl acetate, nitromethane, 2-butanone or water, with the addition of AH **B**-**D** seed crystals, were performed with the aim to grow single crystals of the new polymorphs (see section 1.5 of the Supporting Information). The majority of the evaporation and crystallization experiments yielded mixtures of anhydrates, with **AH A°** being the major phase. It was possible to induce nucleation and growth of **AH B** and/or **AH C**, albeit in the case of **AH C** only concomitantly with other forms. Anhydrate **D** was never obtained in the solvent crystallization/evaporation experiments. “High quality” single crystals, suitable for laboratory single crystal diffraction experiments, were obtained only for phases **AH A°** and **AH B**.

3.2 Single Crystal Structures

The Cambridge Structural Database (CSD)77 contains entries for two solid form of thymine, an anhydrate (**AH A°**, Refcode family: THYMIN41,42) and a monohydrate (THYMMH).38 Herein we report the single crystal structure of a second anhydrate and two other anhydrate structures deduced from a combination of experimental and computational techniques (see section 3.4.1).

3.2.1 Thymine Hydrate—The single crystal structure of thymine *monohydrate* was reported by Gerdil in 1961.38 The hydrate crystallizes in the monoclinic space group $P2_1/c$, with $Z'=1$. The thymine molecule adopts the planar diketo conformation (Figure 1) and forms four strong N–H···O/O···H–N hydrogen bonds across two different centers of symmetry to adjacent thymine molecules, i.e. two $R_2^2(8)$ ring motifs78. These interactions result in a one-periodic H-bonded structure based on a non-polar ribbon motif43 (Figure 2). The ribbons are stacked in the *a* and *c* directions. Adjacent ribbon stacks are related by the glide planes and are inclined about 155° relative to one another, forming an undulating sheet of anti-parallel, non-polar ribbons. Thymine forms a fifth strong hydrogen bond, O2···H–O_{water} to water. The water molecules are located in channels parallel the *c* crystallographic axis and form a hydrogen bonded *zig-zag* chain propagating in the same direction (Figure 2b). The water molecules serve as bridges between two ribbon stacks.

3.2.2 Thymine Anhydrates A° and B—The thymine **AH A°** crystal structure was first reported, without proton positions, in 1969 (THYMIN).41 A second structure solution reports the proton positions (THYMIN01).42 This anhydrate crystallizes in the monoclinic space group $P2_1/c$, with $Z' = 1$ and displays the planar diketo conformation (Figure 1). As in **Hy**, the thymine molecule forms four strong N–H···O/O···H–N hydrogen bonds resulting in two distinct $R_2^2(8)$ ring motifs, so that thymine molecules are linked into a ribbon. This

ribbon has a 2_1 symmetry and is therefore polar and differs from the analogous chain in **Hy**, which displays inversion symmetry (Figure 3). In **AH A°** adjacent ribbons are related by 2_1 screw axes, resulting in polar parallel-ribbon planes in **AH A°** (Figure 3a) and adjacent planes are related by inversion symmetry.

The second thymine anhydrous polymorph (**AH B**) crystallizes in the monoclinic space group *C*2/*c*, with $Z' = 1$ and again the thymine molecule adopts the planar diketo conformation. As in **Hy**, thymine forms four strong N–H···O/O···H–N hydrogen bonds to adjacent thymine molecules across two crystallographically distinct centers of symmetry. This interaction results in two $R_2^2(8)$ rings and leads to the non-polar ribbon motif. Adjacent non-polar ribbons are related by a 2_1 screw axis and are therefore anti-parallel (Figure 3c). As in **AH A°**, adjacent ribbon planes are related by inversion symmetry. **AH A°** and **AH B** are indistinguishable, except for their C2=O2 and C4–H (Figure 1) groups.

3.3 Computational Screen for Thymine Anhydrous Forms

The experimental screen for thymine solid forms was extended by calculating the anhydrate ($Z'=1$) crystal energy landscape, updating the computational search from 2008,[79] which already suggested alternative thermodynamically feasible packings of thymine. A large number of distinct thymine structures were found in the computational search with over 70 unique structures within 5 kJ mol^{-1} and 22 structures within 2 kJ mol^{-1} of the global minimum (Figure 4). All of the lowest energy structures form either the polar or anti-polar ribbon motif (Figure 3). It has to be noted that the computational screen was performed parallel to the experimental screen and finished at a time when only the structure of **AH A°** was known.

The experimental structures **AH A°** and **AH B** were found amongst the lowest energy structures and are 0.57 and 0.68 kJ mol^{-1}, respectively, less stable than the global minimum structure. Two additional structures, 14_260 (**AH C**, *P*2/*c*) and 12_790 (*I*2/*c*) were identified to be indistinguishable from **AH A°** and **AH B** if the relative positions of the C2=O2 and C4–H groups (Figure 1) are ignored. These four structures were found to be within an energy range of only 0.24 kJ mol^{-1}. The isomorphic dehydrate structure (**Hy** framework structure) was found among the low energy - low density structures. The structure was calculated to be 5.46 kJ mol^{-1} less stable than **AH A°**.

In all the low energy structures the ribbon motif is either planar (as in **AH A°, AH B** and **Hy)** or is inclined, leading to either planar (as in **AH A°** and **AH B**) or undulating layers (as in **Hy**). Layers are then either stacked in a parallel fashion (**Hy**) or “anti-parallel”, i.e. inversion symmetry between adjacent layers (**AH A°** and **AH B**). An assignment of the low energy structures to different packing motifs is given in Table S11 of the Supporting Information. Overall, stacks of parallel ribbons (Figure 4b, structures I-IVb) are more dense than the structures where layers are stacked in an “anti-parallel” fashion (Figure 4b, structures I-IVa). The lowest energy structure, 1_1727, is a non-polar parallel-ribbon structure, whose layers are stacked in a parallel fashion. The next group of structures, rank 2 to 15, “anti-parallel” layer stacks, includes the experimental structures. Computed structures showing the same type of layer motif (e.g. polar parallel, denoted with the same symbol in

Figure 4b) could interconvert easily, without breaking any of the strong intermolecular interactions and stacking. Thus, the latter group (rank 2 to 15 structures) consists of four distinct packing types as illustrated in Figure 3. On the other hand, structures belonging to different packing types cannot interconvert without a break-up of strong hydrogen bond interactions or requires the inversion of adjacent ribbon motifs. Taking the uncertainty of the modelling and the negligible energy differences between the lowest energy structures into account may allow to conclude that the structures 11_168 (**AH B**), 14_260 (**AH C**) and 12_790 are the most promising candidates for alternate thymine polymorphs/polytypes. The lowest energy structure should be considered as a possible high(er) pressure polymorph.

An alternative search, using a different level of theory for calculating the conformational energy penalties and isolated molecule charge densities (PBE0/aug-cc-pVTz instead of MP2/6-31G(d,p)) still had the latter structures as the most stable ones, albeit the stability order changed slightly (Table S10 of the Supporting Information), supporting the assumption that the calculated energy differences between the lowest energy structures are likely to be similar in magnitude to the modelling errors. The equivalence of the lattice energy (corresponds approximately to intermolecular energy, as thymine can be classified as a rigid molecule) for a range of coordination environments may imply that there could be a large number of structures showing variations of the ribbons, i.e. variations within the layers, or variations in the stacking.

3.4 Comparison of Computational and Experimental Results

3.4.1 Crystal Structures of AH C and AH D—

The crystal structures of **AHs C** and **D** were derived by comparing the experimental PXRD patterns with patterns simulated from the computationally generated structures (Figure 4).

AH C was obtained directly *via* desolvation experiments of **Hy** at RH ≥43 % and indirectly *via* the **AH D** to **AH C** transformation at room temperature. Furthermore, **AH C** crystallized concomitantly in solvent crystallization/evaporation experiments (see Section 1.5 of the Supporting Information). The experimental powder X-ray diffraction data (section 1.7 of the Supporting Information) indexed to a monoclinic unit cell (a = 12.8955(30) Å, b= 6.8299(4) Å, c = 6.7721(14) Å, β = 104.984(23)°, V = 573.62(19) Å^3) and space group $P2/c$. Cell volume and spectroscopic data (section 1.6 of the Supporting Information) indicated the presence of a $Z' = 1$ structure. By comparing the experimental PXRD pattern to the simulated powder patterns of the computationally generated lowest-energy structures (Figure 4), it was possible to identify structure 14_260 as the best match for **AH C** (Figure S44 of the Supporting Information). The latter structure is one of the four structures that are indistinguishable from **AH A**° and **AH B** if the relative positions of the C2=O2 and C4–H groups (Figure 1) are ignored and was previously identified as one of the likely alternative thymine polymorphs. It has to be noted that the PXRD Pawley fit improved if modelled as a binary mixture using **AH C** and **AH B** or 12_790. The second phase accounted for a peak broadening, in particular in the 2theta range in between 12.7 and 14.7°. Peak broadening may be a result of diffuse scattering and thus, that **AH C** may show variations in the structure (section 4.3). The main structural component of **AH C** (14_260) has the non-polar parallel layer motif (Figure 3d): The thymine molecule forms four strong N–H···O/O···H–N

hydrogen bonds to adjacent thymine molecules across two different centers of symmetry. Adjacent non-polar ribbons are related via inversion symmetry. For the computed **AH C** *.res file see section 2.10 of the Supporting Information.

The **AH D** polymorph was obtained exclusively in fast dehydration experiments at lowest RH (over P_2O_5, vacuum drying or N_2 purge) at temperatures ≤30 °C. The experimental PXRD data (section 1.7 of the Supporting Information) indexed to a monoclinic unit cell (a = 4.5568(8) Å, b = 23.0159(58) Å, c = 5.3940(6) Å, β = 93.928 (13)°, V = 564.39(18) Å^3) and space group $P2_1/n$. Based on the cell volume and spectroscopic data it was concluded that **AH D** is also a $Z' = 1$ structure. The experimental powder pattern matches the simulated powder pattern of the computationally generated structure 32_113 (Figure S45). **AH D** and **Hy** show structural similarity. The two structures have non-polar ribbon stacks and adjacent ribbons are aligned in anti-parallel fashion (Figure 3c). In **AH D** the adjacent ribbon stacks, related by the glide planes, are inclined about 119° and shifted with respect to each other by approx. half of the stacking distance in contrast to the **Hy** packing (Figure 5). The transformation of **Hy** to **AH D** seems to be facile. Water egress is expected to occur along the water channels (Figure 2b, along the crystallographic c axis) and the solvent accessible space should facilitate the rearrangement of the ribbon stacks. No breaking of strong intermolecular hydrogen bonds and stacking interactions (close contacts) is required for the **Hy** to **AH D** transformation. For the *.res file of the computed **AH D** structure see section 2.10 of the Supporting Information.

3.4.2 Hydrate Structure – Void Structures—A characteristic feature of the thymine anhydrate crystal energy landscape (Figure 4a) is that low energy-low density structures were calculated to be thermodynamically feasible, c.f. 4-aminoquinaldine AH II80. The Mercury Hydrate Analyzer tool was used to visualize the water/void space in the computed low energy and low density structures. (For details see section 2.9 of the Supporting Information.) Based on the position in the crystal energy landscape, which is correlated with the shape of the void space, it was possible to subdivided the low density structures into two groups (Figure 4a). 'Group 1' structures, within 3.59 to 4.41 kJ mol^{-1} with respect to the global energy minimum structure, show interconnected voids in close proximity to the C2=O2 group (Figure S42 of the Supporting Information), but no open channel void space. In contrast, the four 'Group 2' structures contain open channels. The void space is also in close proximity to the C2=O2 group (Figure 1) in these structures. The latter hydrogen bonding donor group is not involved in a strong hydrogen bonding interaction in any of the computed structures. The hydrate framework (hypothetical isomorphic dehydrate) belongs to the 'Group 2' structures. Despite their lower packing efficiency, these structures were calculated to be only 5.91 to 6.32 kJ mol^{-1} less stable than the global minimum structure. The four 'Group 2' structures may interconvert easily into each other. Furthermore, the 'Group 2' structures are related to three other, lower energy and high(er) density structures (52_112, 27_119 and **AH D,** Figure S40 of the Supporting Information). The facile transformation to more stable packings may explain why it was not possible to dehydrate **Hy** to an isomorphic dehydrate, despite the packing being a minimum in Figure 4b.

3.5 Characterization of Thymine Solid Forms

3.5.1 Thermal Dehydration Reactions—Crystallization of thymine from water at temperatures ≤25 °C results in hexagonal elongated plates. As previously reported by Perrier and Byrn,39 removal of the plates from the mother liquor leads to complete loss of transparency, indicating dehydration, within two hours. Dehydration occurs from end faces and proceeds toward the centre of the crystal parallel to the *c*-axis, i.e. along the water channels. In silicon oil preparations, dehydration is inhibited at RT. Upon increasing the temperature, segmentation of the plates is observed at 55 °C (Figure 6) and a fast loss of birefringence is observed, which starts from the cracks on the crystal surface. This indicates that the water molecules are released by a diffusion controlled process. Dehydration at temperatures below 60 °C results in aggregates, with the original shape of the **Hy** crystals being maintained, consisting of small anhydrate crystals ("pseudomorphosis"81), which is characteristic of the desolvation of stoichiometric solvates (hydrates).82 At higher heating rates (~10 °C min^{-1}, crystals embedded in silicon oil), a peritectic transformation, (dehydration and simultaneous fast nucleation and growth of anhydrate crystals) is observed at approx. 60 °C. The anhydrate crystals decompose at temperatures above 315 °C.

The TGA curves of **Hy** show a mass loss of 10.31 ± 0.01%, which corresponds to 0.81 ± <0.01 mol of water per mol thymine (Figure 7a). Under a nitrogen purge dehydration starts immediately.83 The water content was therefore also determined using Karl Fischer titration. Coulometric water determination indicated a molar ratio of 0.82 ± 0.01 mol of water per mol thymine and confirmed that the thymine hydrate is not a monohydrate.

Slow-heating DSC experiments of **Hy** (Figure 7a) using pinhole lids show that the dehydration of **Hy** to **AH C** (with **AH D** impurities) takes place below 60 °C. By increasing the heating rate (to 10 °C min^{-1}) and using sealed or high-pressure DSC pans, the peritectic transformation of **Hy** to **AH A°** was observed at 61.0 ± 1.5 °C (n=12).

The temperature/composition phase diagram was constructed using **Hy** and **AH A°** (the thermodynamically most stable anhydrate in the investigated temperature range, see section 3.5.6). Within the hydrate/water region (>0 to 54.95 mol % thymine), no reaction between free (unbound) water and **Hy** is observed. The DSC curves showed the melting of pure ice, the peritectic **Hy** to **AH A°** transition (Figure 7b), and at a higher temperature dissolution and decomposition of the anhydrate and/or the release of water vapor from the DSC pan after the transformation experiment. Both, the melting temperature of ice and the pertitectic transition temperature were found to be independent of the **Hy**/water composition (Figure 7d). This situation indicates only minor interactions between water and thymine, which is also demonstrated by the low water solubility of the compound (< 4 mg mL^{-1} at 25 °C, see Section 1.4 of the Supporting Information). The peritectic transformation reaction confines also the upper phase boundary in the **AH A°/Hy** region of the phase diagram (54.95 to < 100 mol % thymine). Because of the weak interaction between thymine and water and the thermal decomposition of this high melting compound, it was not possible to record the solid-liquid phase transition (liquidus line) above the peritectic temperature. At heating rates of ≤10 °C min^{-1} the decomposition of **AH A°** occurs at temperature above 315 °C.

3.5.2 Moisture Dependent Stability of Thymine Hydrate and Anhydrates—The hydration and dehydration behavior of the thymine solid forms was investigated in gravimetric moisture sorption/desorption studies at 25 °C, long-time storage studies over saturated salt solutions at 25 °C, and slurry experiments using water/methanol mixtures with different water activities (a_w) at 10 and 25 °C.

In the automated gravimetric desorption studies, dehydration of **Hy** started at RH values < 70% (Figure 8a). The dehydration rate between 65% and 60 % RH is slow (months) and accelerates significantly at RHs < 60% (< 2days). Systematic storage studies at different RH values (98, 92, 84, 75, 62.5, 52, 43, 31, 24, 11 and 0%, see section 1.2 of the Supporting Information) revealed that different dehydration products (polymorphs) may result at different humidity conditions. **AH C** is formed at higher RHs (≥43% and ≤84%), whereas at lower RHs (< 43%) **AH D** is concomitantly formed with **AH C**. Phase-pure **AH D** was only obtained over P_2O_5, but never in the stepwise desorption experiments. The desorption studies confirm that the hydrate stoichiometry is 0.82 ± <0.01 molequivalent of water (n=3).

The four thymine anhydrates show no pronounced water uptake (**AH A°** < 0.20%, **AH B**: < 0.13%, **AH C**: < 0.16% and **AH D**: < 0.10%) up to 90% RH (Figure 8b) and do not transform to the hydrate under the tested conditions. The three anhydrates **AH A° – C** were stable over the entire investigated RH range, while **AH D** transformed to **AH C**. The kinetics of this moisture induced transformation was found to increase with increasing RH.

To determine the critical water activity (a_w) of the thymine anhydrate ↔ hydrate system **AH D** was added to methanol/water mixtures of various compositions (section 1.3 of the Supporting Information) and equilibrated for one week under stirring. Samples were withdrawn from time to time and analyzed with PXRD. These slurry experiments revealed that the hydrate is only stable at the highest RH/a_w conditions (at a_w = 1). All experiments at lower water activities (a_w ≤0.95) resulted in **AH A°** indicating that this polymorph is the thermodynamically stable one at ambient temperature. From these studies we can conclude that the critical a_w for the thymine anhydrate ↔ hydrate system is > 0.95 at 10 and 25 °C.

3.5.3 Modelling of the Hydrate Structure—Three independent methods, TGA, Karl Fischer titration and gravimetric desorption studies, indicate that the **Hy** is not a monohydrate but contains 0.82 mol water per mol thymine. Structural anomalies associated with the water of crystallization have been reported in 1961, and were based on experimental diffraction data, and short water···water interactions were postulated in the case of an assumed monohydrate stoichiometry.38 It has been concluded that only five out of six water positions are occupied (0.83 moles of water per mol thymine). There seems to be no periodic packing fault in the arrangement of the water molecules, but rather that the water molecules are statistically distributed over the available sites.38 We used ab initio calculations (DFT-D) on thymine monohydrate to achieve a better understanding of the hydrate structure on the molecular level. Starting from the monohydrate structure, four RT hydrate models were generated, which covered water:thymine ratios of 1:1, 0.875:1, 0.8125:1 and 0.75:1 (see section 2.7 of the Supporting Information).

Figure 9 contrasts exemplarily the 1:1 (a) and 0.8125:1 (b) models. Close contacts are observed between the water molecules (O···O distance: 2.642 Å) in the monohydrate model. Removing water molecules located at isolated sites from the monohydrate or (1 x 1 x n) supercell structures thereof and optimizing the atomic positions of the thymine and water molecules while keeping the lattice parameters fixed did not alter the thymine positions. Atomic positions of the water molecules of the four models differ in their *z* coordinates, i.e. in their position within the water channels (parallel to the *c* axis). In all hydrate models each water molecule is still engaged in three strong hydrogen bonding interactions, two to adjacent water molecules and one to thymine (C2=O2). Thus, the hydrogen bonded water chains are maintained, but the *zig-zag* arrangement is broken (Figure 9b). The removal of water molecules leads to unused C2=O2 donor groups (Figure 1), the functional group which is also unused in all experimental and computed low energy anhydrate structures (see section 4.1). More importantly, the hydrate models with stoichiometric water:thymine ratios of <1:1 resolved the problem of short water O···O distances. The computed DFT-D models confirm the simple model suggested by Gerdil.38

3.5.4 Water Diffusion in Thymine Hydrate Monitored Using D/H Exchange—

To investigate the mobility of the water molecules in **Hy**, D/H exchange experiments were performed. Thymine hydrate was exposed to deuterium oxide vapor (~100% RH) and characterized by Raman spectroscopy at different time points (Figure 10a) to monitor the water dynamics in **Hy**. The Raman spectrum of **Hy** (t=0) shows strong ν(C–H) vibrations at 3076 cm^{-1} and 3020 – 2840 cm^{-1} for the C4–H and group CH-H_3 groups, respectively. On exposure to D_2O, the spectrum changes within less than 1.5 hours, the ν(O–D) peaks of D_2O molecules, replacing H_2O molecules in **Hy**, evolve with time. Thus, water diffusion in and out of the crystal structure is rapid. Since all of the water molecules are located in the same channel a complete exchange can be expected, similar to the caffeine 4/5 hydrate34.

Thymine also has N–H protons that could exchange a hydrogen for a deuterium. Therefore, IR and Raman spectra of thymine hydrates, containing either H_2O or D_2O from the H/D exchange studies, were dehydrated and compared. The comparison showed that within 112 h of D_2O vapor exposure the N–H was not affected. However, slurring thymine in D_2O for 12 hours resulted not only in and O–D but also N–D exchange (Figure 10b).

3.5.5 Enthalpy of the Hydrate/Anhydrate Transition—

The enthalpy of the **Hy/AH** transition was estimated from DSC and isothermal calorimetry (with the aid of an RH perfusion cell) experiments. The dehydration enthalpy, $\Delta_{dehy}H_{Hy\text{-}AH\_C/D}$, measured in open DSC pans (pinhole lid, Figure 7a), can be subdivided (application of Hess's law) into the enthalpy of hydrate to anhydrate transformation, $\Delta_{trs}H_{Hy\text{-}AH\_C/D}$, and that of the vaporization of the expelled 0.82 mol of water.

$$\Delta_{\mathrm{trs}}H_{Hy-AH\_C/D}=\Delta_{dehy}H_{Hy-AH\_C/D}-0.82\Delta_{vap}H_{H2O} \quad (1)$$

If we subtract the known enthalpy value for the vaporization of water at the dehydration temperature (T_{dehy} ~ 55 °C, $\Delta_{vap}H_{H2O}$ = 42.699 kJ mol^{-1})84 from the measured enthalpy of

dehydration (41.6 ± 0.6 kJ mol^{-1}) according to eq. (1), we can estimate the enthalpy of the hydrate to anhydrate (**AH C** > **AH D**) phase change to be 6.6 ± 0.6 kJ mol^{-1}.85

With the RH-perfusion cell (25 °C) an enthalpy of dehydration ($\Delta_{dehy}H_{Hy\text{-}AH\_C/D}$) of −41.4 ± 0.5 kJ mol^{-1} was obtained. Using eq. (1) and the vaporization of water at 25 °C (43.99 kJ mol^{-1}), the transition enthalpy of **Hy** to a mixture of **AH C** and **AH D** ($\Delta_{trs}H_{Hy\text{-}AH\_C/D}$) was calculated to be 6.2 ± 0.5 kJ mol^{-1}. The two estimated transformation enthalpies are in good agreement with one another, considering the instability of **Hy** at RHs < 70 % (Figure 8b) and variability in the dehydration products.

DSC experiments in hermetically sealed pans can provide transformation temperatures and transformation enthalpies.86 In the case of two hydrates with the same (or almost the same) stoichiometry, the transformation enthalpy corresponds to the enthalpy difference between these two phases. However, in the case of a change in stoichiometry, the measured enthalpy value also includes an unknown contribution from the enthalpy of solution of a fraction of the dehydration product in the liberated water. At the peritectic temperature and ambient pressure, only 0.2% of thymine would dissolve in the liberated water of the **Hy** structure. This rare situation allows to determine the $\Delta_{trs}H_{Hy\text{-}AH\_A}$ directly in a hermetically sealed DSC pan. Indeed, the measured enthalpy value of 6.9 ± 0.1 kJ mol^{-1} is only slightly higher than those determined in the open DSC pan and via the RH perfusion method estimated for the **Hy** to **AH** phase transformation.

The measured transformation energy between the hydrate and the anhydrate lies at the lower end of the range of reported monohydrate to anhydrate transformation energies (5.7 to 22.9 kJ mol^{-1}),87–89 indicating the presence of a hydrate with lower thermodynamic stability than for example the morphine monohydrate,87 which is an exceptionally stable hydrate.

3.5.6 Thermodynamic and Kinetic Stability of Thymine Anhydrates—The different anhydrous polymorphs were examined by DSC (Figure 11). Heating rates of 50 °C min^{-1} had to be applied to separate the melting process from decomposition of thymine. Each of the DSC curves of **AH A°**, **AH B** and **AH C** show melting as the only thermal event. Their very similar melting point onset temperatures and heat of fusion values (Table 1) indicate that the latter three polymorphs are homeoenergetic. To confirm that no undetected phase transformation had taken place upon heating, the DSC runs were stopped at 290 °C and the samples were analyzed with PXRD. **AH D** shows an exothermic phase transformation to **AH B** at a temperature > 120 °C. The energy difference between **AH D** and **AH B** was measured to be 0.8 ± 0.1 kJ mol^{-1}. The exothermic nature of the phase transformation indicates that the polymorphic pair **AH D/AH B** is monotropically90,91 related, with **AH D** being metastable in the entire temperature range. Furthermore, all four polymorphs were cooled, in sealed DSC pans, to –196 °C (liquid nitrogen) to investigate for possible low temperature phase transformation(s), but no transformations were observed.

The lattice energies of the structures were estimated using ab initio electronic structure calculations on either the molecule or crystal and PIXEL energy calculations (Table S13 of the Supporting Information). With the exception of the PIXEL energy calculations, **AH A°–C** were calculated to be homeoenergetic, which is in agreement with the DSC results. The

relative stability order of **AH D** and **AH A°** (ΔE_{latt}: –0.38 to +2.45 kJ mol^{-1}) varied slightly, depending on the used wavefunction for deriving the electrostatic and conformational energy penalty contributions as well as the used dispersion correction. PIXEL energy calculations overestimated ΔE_{latt}.

Liquid assisted grinding experiments of the four polymorphs in water and organic solvents (section 1.1 of the Supporting Information) were employed to determine the thermodynamically most stable anhydrate at ambient temperature. Slurry experiments in organic solvents resulted exclusively in **AH A°**, as did slurry experiments in water at temperatures > 25 °C. Thus, **AH A°** is the thermodynamically most stable thymine anhydrous form in the temperature range of 10 to 60 °C. **AH D** transformed to **AH C** at room temperature and to **AH B** at higher temperatures (Figure 11). By contrast, no solid state transformations were observed for forms **AH A°**, **AH B**, **AH C** or mixtures thereof within four months (end of investigation time). Thus, **AH B** and **AH C** show a high kinetic stability which can be related to the fact that the polymorphs are homeoenergetic and a transformation of **AH B** to **AH A°** or **AH C** to **AH A°** would require the break-up of strong hydrogen bonded ribbon motifs (Figure 3).

4 Discussion

4.1 Thymine – Packing Possibilities in the Solid State

Crystal structure prediction methods are based on searching for the thermodynamically most stable crystal structures,92 and provide information about the packing arrangements, hydrogen bonding motifs, and in case of flexible molecules the conformation, which can be expected in the solid state. The computationally generated crystal energy landscape for Z′=1 anhydrate structures of thymine (Figure 4) revealed that all the low-energy structures are based on one out of two possible strong hydrogen bonded ribbon motifs (polar or apolor ribbons, Figure 3). The two ribbons are based on a combination of the same two $R_2^2(8)$ motifs. In the polar ribbon motif, as seen in **AH A°**, each of the two $R_2^2(8)$ rings was calculated to have an underlying pairwise interaction energy contribution of approx. –65 kJ mol^{-1} (PIXEL energy calculations, Figure S31a of the Supporting Information). The two $R_2^2(8)$ rings in the apolar ribbon chains, as in **AH B,** have pairwise energies of approx. –50 and –80 kJ mol^{-1} (Figure S31c of the Supporting Information). Thus, based on the strong hydrogen bonding interactions there is no preference for either the polar or apolar ribbon motif. As C1=O1 (Figure 1) acts as the acceptor for both N1/2–H···O1 hydrogen bonds, the second C2=O2 group is left without a strong hydrogen bond interaction. A C4–H···O2 close contact is present in all calculated structures (pairwise energy contribution approx. –13 kJ mol^{-1} in planar ribbon layers). The only possibility to satisfy the unused hydrogen bonding donor group is by introducing additional hydrogen bonding acceptor groups, i.e. solvate (hydrate) formation.

Starting from one of the two ribbon motifs, adjacent ribbon chains can be arranged in a parallel or anti-parallel fashion, so that planar or undulating layers are formed. Stacking of these layers was again found to be possible in a parallel or anti-parallel mode, leading to numerous options for the arrangement of ribbon motifs in a periodic way. But, as seen for

other “symmetric” small organic molecules, eniluracil8 or phloglucinol82, order-disorder phenomena are possible, such as interdigitating parallel/anti-parallel ribbons or stacking faults of ribbon chains or layers. Thus, the crystal energy landscape (Figure 4) may warn us of the possibility of stacking faults.

4.2 Why does Thymine Form an Unstable Stoichiometric Hydrate?

The mismatch between the numbers of available hydrogen bond donor and acceptor groups may be seen as a driving force for thymine hydrate formation. The computational search for hydrate structures has one difficulty in that different hydrate stoichiometries have to be considered. This leads to exponential computational expenses in terms of resources and time required to generate hydrate crystal energy landscapes if compared to the computing of the crystal energy landscape of an anhydrate. Exploring the anhydrate crystal energy landscape for void structures can be an alternative, cheaper way to gain an insight into the possibility of solvate (hydrate93) or inclusion compound94 formation. However, this may require the generation of higher energy structures and only a subset of the possible solvate types, the so-called Class 2: channel solvates33, may be detected within the anhydrate crystal energy landscape. In the case of thymine, this kind of analysis resulted in two groups of void structures, one with interdigitating and one with open channel void space in close proximity to the unused C2=O2 group (Figure 1).

Addition of water molecules to a member of the open void structure family and optimizing their position in the crystal (monohydrate stoichiometry used) resulted in the **Hy** structure. The pairwise intermolecular PIXEL energy for the O_{water}–H···O2 interaction was calculated to be –19.3 kJ mol^{-1} and the strongest pairwise water···water interaction, present twice, –10.5 kJ mol^{-1} (Figure S36 of the Supporting Information). Even water, the smallest possible solvent molecule, is too big to be incorporated in the hydrate framework structure in the 1:1 stoichiometry (monohydrate). It has been shown (section 3.5.3) that only a sub-stoichiometric ratio of water molecules, 5/6 mol water per mol thymine, is present, but this means that not all of the thymine C2=O2 groups are used and that the water···water interactions are longer (normal range) and therefore weaker as estimated by the PIXEL energy calculations. An isostructural solvate, i.e. with water replaced by an organic solvent molecule, is not expected to form, as this would require a widening of the channels, decreasing the stability of the framework structure.

The sub-stoichiometric water ratio in the thymine hydrate and the fact that the water molecules are located in open channels may incorrectly imply the presence of a non-stoichiometric hydrate. The most important feature of a non-stoichiometric hydrate is that the structure (framework) is retained,10 while the water content changes in accordance with the RH/a_w conditions of its environment. This is in contrast to the fact that thymine crystallizes in an uncommon, but well-defined water:compound ratio, similar to the hydrate of caffeine.95,96 The moisture dependent PXRD dehydration (Figures S4 & S5 of the Supporting Information) and RH dependent storage experiments of **Hy** confirmed that this phase behaves like a stoichiometric hydrate and that it is not possible to reduce the water content of the **Hy** structure without inducing a transformation to either **AH C** or **AH D**. Thus, the presence of 5/6 mol of water is essential for the **Hy** structure to exist. This result is

Europe PMC Funders Author Manuscripts

supported by the calculations showing that the isostructural dehydrate structure is energetically less favourable than other very closely related and more stable low-energy structures, and it is therefore not expected to be observed experimentally.

The observations that the same hydrogen bond ribbon motifs are formed regardless of the presence of water molecules in the crystal lattice, that the water molecules are hydrogen bonded to H-bond acceptor groups which are not engaged in strong hydrogen bond interactions in the anhydrates and that anhydrate structures lacking C2=O2 hydrogen bonds are stable may explain why the hydrate phase of thymine is unstable. This hydrate is the thermodynamically most stable form only in water and at temperatures ≤25 °C. With this knowledge, hydrate formation can be avoided easily in production or during storage. In contrast, it is practically impossible to avoid dehydration to anhydrous thymine once hydrate formation has occurred. In general, processing, handling and storage of a phase with a stability behaviour of thymine hydrate would pose an extreme challenge.

4.3 Packing Polymorphism and Order-Disorder Phenomena in Thymine

The PXRD pattern (Figure 12), IR and Raman spectra (section 1.6 of the Supporting Information) of three of the four thymine anhydrates, AHs **A°** – **C**, show close similarity. In contrast, the PXRD and spectroscopic characteristic of **Hy** and **AH D** clearly differ from those of the other thymine solid forms. A closer analysis of the PXRD data of AHs **A°** – **C** reveals that the three phases can be distinguished in the 2theta region from 12.5 to 17°. **AH A°** and **AH B** are highly crystalline, whereas the PXRD pattern of **AH C** shows peak broadening, which could be caused by an order-disorder phenomenon (diffuse scattering). These observations, together with the fact that the computationally generated anhydrate crystal energy landscape (Figure 4) is populated by a number of homeoenergetic 3D packing modes, with very similar lattice parameters (Table S11 of the Supporting Information, Table 2) raises the question whether the forms are true polymorphs, or represent different degrees of ordering/domain structures (such as aspirin,97 orotic acid,98 eniluracil8).

The single crystal data of **AH A°** and **AH B** did not indicate of the presence of disorder (diffuse scattering), which is consistent with the presence of sharp diffraction peaks and the absence of any peak broadening in their respective PXRD diffractograms. Furthermore, the corresponding Pawley fits did not indicate peak positions that were unaccounted for by the structure model established form single crystal data. Thus, the bulk material corresponds to the crystal chosen for the single crystal XRD experiment, and **AH A°** and **AH B** are true polymorphs that may be described as packing polymorphs or polytypes.

In contrast, the PXRD pattern of **AH C** contains disorder streaks (Figure 12) and gives only an imperfect fit with the PXRD pattern calculated on the basis of structure 14_260 (section 1.7 of the Supporting Information). Nevertheless, 14_260 was identified as the main crystal domain of **AH C**. Analysis of the low energy computed structures readily showed that there are possibilities of mutual substitution of the strong hydrogen bonded ribbon motifs (apolar *vs.* polar), layer stacks, etc., thus, ‘mixed’ structures are probable. Lattice parameters and packing diagrams of ordered components that may represent the domains of possible disordered structures are summarized in Table 2 and Figure 13, respectively. The four

structures are indistinguishable from one another if their relative positions of the C2=O2 and C4–H groups (Figure 1) are ignored.

Thus, **AH C** seems not to be an ordered structure. Nevertheless, **AH C** has to be seen as a thymine polymorph, as the packing arrangement of the main domain is distinct from the packing observed in **AH A°** and **AH B**. The nature of the disorder determines whether the rearrangement to the most stable ordered domain/form is facile or so difficult that the ordered phase is practically unobtainable, as recently reported for the anhydrate of a related compound, orotic acid.98 In the case of **AH C**, the disorder cannot be a local disorder of the orientation of molecules around the C5-C3-C1-O1 axis, as this would bring two C=O oxygens into close contact. Thus, as indicated by the calculations, the disorder is likely to involve the non-polar parallel ribbons which are occasionally replaced by polar ribbons, or the translation symmetry between parallel ribbons is broken. Once one particular type of ribbon motif/layer is formed, a considerable rearrangement would be required, which is unlikely in the solid state. Interestingly, the degree of disorder showed hardly any variation between individual **AH C** batches prepared by applying different dehydration/evaporative crystallization conditions. Furthermore, ageing of the sample (storing at ambient or elevated temperature) did also not induce an ordering process, which confirms that starting from **AH C**, as observed in this study, it is practically impossible to produce an ordered **AH C** phase.

The fourth packing polymorph of thymine, **AH D**, is easy to distinguish from the other thymine solid forms by PXRD or vibrational spectroscopy. Its low crystallinity (broad PXRD peaks compared to all other forms, Figure 12) can be related to the fact that this form was exclusively obtained *via* dehydration. Procedures that may increase the crystallinity of a phase, i.e. storage and increased in temperature resulted in a phase transformation. **AH D** has to be seen as a metastable intermediate phase. This latter phase can be stabilized for two to six weeks if stored at low RHs.

The four characterized polymorphs of thymine are packing polymorphs, with the crystallization method determining the degree of crystallinity and order-disorder. Cooling crystallization and slow evaporation experiments predominantly lead to ordered phases (**AH A°** and **AH B**), whereas **AH C** is formed by evaporation and dehydration, **AH D** exclusively *via* dehydration. Thus, the time factor for nucleation and growth seems to determine the degree of order. The structural similarity and stability of the forms may be the reason why different thymine polymorphs crystallizes concomitantly.99

4.4 Do we expect to find more polymorphs?

The analysis of the computed crystal energy landscape, based on similarity and differences in the more likely (lower energy) structures, requires human intervention.100 Experience is needed to interpret and correlate the experimental screening results and computed crystal energy landscape(s) to estimate whether (i) alternate computed structures are possible candidates for additional polymorphs,88,89,101 (ii) closely related structures are likely to disordered,8,102–104 or (iii) whether a structure is too similar to an experimental one so that it unlikely to even survive as metastable polymorph.

The computationally generated anhydrate crystal energy landscape of thymine (Figure 4) clearly indicates that thymine can be expected to be polymorphic and that various types of disorder and/or modulated structures are possible. The computed structures 11_168 (**AH B**), 12_790 and 14_260 (**AH C**) had to been seen as alternative polymorphs and/or represent possibilities for disorder. Two of the three structures, 11_168 (**AH B**) and 14_260 (**AH C**), were indeed found to be experimental polymorphs, **AH B** an ordered phase and **AH C** presumed to show order-disorder phenomena. Experimental evidence for 12_790 has not been provided, but it may be that the appropriate experiment has not been performed yet. The plethora of possible experiments, which extends beyond conventional (solvent) crystallization methods, e.g. crystallization under confinement,105 heteronuclear screening using polymers,106 or crystallization in electric107 or ultrasound108 fields, makes it impossible to cover the entire experimental screening space. If it does indeed exist, the physico-chemical properties of 12_790 can be expected to be similar to those of the three polymorphs **AH A°** – **C**, i.e. it would be metastable but showing a high kinetic stability.

The fact that **AH A°** (confirmed to be the thermodynamically most stable anhydrate in the temperature range from 10 to 60 °C) was not found as the global minimum structure and that the crystal energy landscape shows structures of higher density leads to the assumption that high(er) pressure polymorphs of thymine exist.104,109

Despite the ‘obvious’ candidate structures, this combined experimental and computational study has demonstrated that there is still no straightforward way to identify putative experimental forms, c.f. **AH D**, let alone propose routes to the predicted structures. Numerous alternative packings of thymine are plausible. In the case of **AH D** the route from **Hy** to **AH D** has been rationalized based on the structural characteristics of the two phases (Figure S40 of the Supporting Information).

5 Conclusions

Thymine, first isolated and synthesized in 1903,36,37 has been found to exist in at least four polymorphs. Beside the common form **A°** three new metastable anhydrates have been characterized for the first time. Additionally, thymine forms a hydrate with an unusual but well-defined stoichiometric water to thymine ratio of 5:6 (0.8 hydrate).The fact that the water molecules can escape easily from the channels of the hydrate complicates the handling of hydrated thymine. The hydrate can be seen as an intermediate phase for providing access to different anhydrous polymorphs of thymine (**AH C** and **AH D**) by means of varying the dehydration conditions. **AH D** is a further example80,102,110,111 of a polymorph which can only be produced by desolvating a solvate/hydrate under specific conditions and not any other crystallization route.

The solid state of thymine appears to be dominated by two types of homeoenergetic hydrogen bonded ribbons, which can be arranged in numerous crystallographically distinct packing modes. The computationally generated anhydrate crystal energy landscape not only indicated that thymine is polymorphic but also revealed that several closely related structures are so similar in energy that disorder/polytypism and concomitant nucleation/growth of the

anhydrates **A**° – **C** are possible. These three polymorphs were found to differ by ≤0.3 kJ mol^{-1} in heat of fusion and lattice energy.

Only the combination of experimental and computational characterization methods, i.e. X-ray diffraction, vibrational spectroscopy, thermal analysis, crystal structure prediction and lattice energy modelling, allowed to successfully distinguish between polymorphism and disorder in the solid state forms of thymine. Overall, this study demonstrates that the computationally generated crystal energy landscape can indicate the possibility and the nature of a disorder phenomenon and that it is an invaluable complementary technique for characterizing and understanding the organic solid state at a molecular level.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgment

The authors are grateful to Prof. S.L. Price (University College London) for the use of the DMACRYS, to Profs. C. C. Pantelides and C. S. Adjiman (Imperial College London) for the use of the CrystalPredictor and CrystalOptimizer programs, and to Moritz Walter for experimental assistance. DEB gratefully acknowledges funding by the Elise Richter (FWF, project V436-N34) programme of the Austrian Science Fund. Calculations were supported by the Austrian Ministry of Science BMWF as part of the *UniInfrastrukturprogramm* of the Research Platform Scientific Computing at the University of Innsbruck.

References

1. Yu L, Reutzel SM, Stephenson GA. Pharmaceut Res. 1998; 1:118–127.
2. Brittain, HG. Polymorph in Pharmaceutical Solids. Second Edition. Brittain, HG., editor. Informa Healthcare; 2009. p. 1-23.
3. Chemburkar SR, Bauer J, Deming K, Spiwek H, Patel K, Morris J, Henry R, Spanton S, Dziki W, Porter W, Quick J, et al. Org Process Res Dev. 2000; 4:413–417.
4. Dunitz JD. Acta Crystallogr., Sect. B. 1995; 51:619–631.
5. Davey RJ. Chem Commun. 2003:1463–1467.
6. Bernstein, J. Polymorphism in Molecular Crystals. Clarendon Press; Oxford: 2002.
7. Threlfall TL. Analyst. 1995; 120:2435–2460.
8. Copley RCB, Barnett SA, Karamertzanis PG, Harris KDM, Kariuki BM, Xu MC, Nickels EA, Lancaster RW, Price SL. Cryst Growth Des. 2008; 8:3474–3481.
9. Guinier A, Bokii GB, Boll-Dornberger K, Cowley JM, Durovic S, Jagodzinski H, Krishna P, De Wolff PM, Zvyagin BB. Acta Crystallogr., Sect. A: Found. Crystallogr. 1984; A40:399–404.
10. Griesser, UJ. Polymorphism: In the Pharmaceutical Industry. Hilfiker, R., editor. Wiley-VCH; Germany: 2006. p. 211-233.
11. Rodriguez-Hornedo N, Murphy D. J Pharm Sci. 2004; 93:449–460. [PubMed: 14705201]
12. Vippagunta SR, Brittain HG, Grant DJW. Adv Drug Deliver Rev. 2001; 48:3–26.
13. Campeta AM, Chekal BP, Abramov YA, Meenan PA, Henson MJ, Shi B, Singer RA, Horspool KR. J Pharm Sci. 2010; 99:3874–3886. [PubMed: 20575000]
14. Zhao XS, Siepmann JI, Xu W, Kiang YH, Sheth AR, Karaborni S. J Phys Chem B. 2009; 113:5929–5937. [PubMed: 19351127]
15. Morissette SL, Almarsson O, Peterson ML, Remenar JF, Read MJ, Lemmo AV, Ellis S, Cima MJ, Gardner CR. Adv Drug Deliver Rev. 2004; 56:275–300.
16. Aaltonen J, Alleso M, Mirza S, Koradia V, Gordon KC, Rantanen J. Eur J Pharm Biopharm. 2009; 71:23–37. [PubMed: 18715549]
17. Allesoe M, Tian F, Cornett C, Rantanen J. J Pharm Sci. 2010; 99:3711–3718. [PubMed: 19798763]

18. Braun DE, Koztecki LH, McMahon JA, Price SL, Reutzel-Edens SM. Mol Pharm. 2015; 12:3069–3088. [PubMed: 26075319]
19. Raijada D, Bond AD, Larsen FH, Cornett C, Qu H, Rantanen J. Pharm Res. 2013; 30:280–289. [PubMed: 22996567]
20. Khankari RK, Grant DJW. Thermochimica Acta. 1995; 248:61–79.
21. Malaj L, Censi R, Gashi Z, Di MP. Int J Pharm. 2010; 390:142–149. [PubMed: 20117196]
22. Zhang GGZ, Law D, Schmitt EA, Qiu Y. Adv Drug Delivery Rev. 2004; 56:371–390.
23. Phadnis NV, Suryanarayanan R. J Pharm Sci. 1997; 86:1256–1263. [PubMed: 9383736]
24. Braga D, Grepioni F, Chelazzi L, Campana M, Confortini D, Viscomi GC. CrystEngComm. 2012; 14:6404–6411.
25. Bernardes CES, da Piedade MEM. Cryst Growth Des. 2012; 12:2932–2941.
26. Pina MF, Pinto JF, Sousa JJ, Fabian L, Zhao M, Craig DQM. Mol Pharm. 2012; 9:3515–3525. [PubMed: 23051151]
27. Stephenson GA, Diseroad BA. International Journal of Pharmaceutics. 2000; 198:167–177. [PubMed: 10767566]
28. Berzins A, Skarbulis E, Rekis T, Actins A. Cryst Growth Des. 2014; 14:2654–2664.
29. Rager T, Geoffroy A, Hilfiker R, Storey JMD. Phys Chem Chem Phys. 2012; 14:8074–8082. [PubMed: 22481217]
30. Kumar SS, Nangia A. Cryst Growth Des. 2014; 14:1865–1881.
31. Braun DE, Bhardwaj RM, Florence AJ, Tocher DA, Price SL. Cryst Growth Des. 2013; 13:19–23.
32. Morris, KR., Rodriguez-Hornedo, N. Encyclopedia of Pharmaceutical Technology. Swarbrick, J., Boylan, J., editors. Marcel Dekker; New York: 1993. p. 393-440.
33. Brittain HG, Morris KR, Boerrigter SXM. 2009:233–281. 192 ed.
34. Ahlqvist MUA, Taylor LS. Int J Pharm. 2002; 241:253–261. [PubMed: 12100853]
35. Lafontaine A, Sanselme M, Cartigny Y, Cardinael P, Coquerel G. J Therm Anal Calorim. 2013; 112:307–315.
36. Levene PA. Ztschr physiol Ch. 1903; 39:4–8.
37. Wheeler HL, Merriam HF. Amer Chem J. 1903; 29:478–492.
38. Gerdil R. Acta Crystallogr. 1961; 14:333–344.
39. Perrier PR, Byrn SR. J Org Chem. 1982; 47:4671–4676.
40. Kaliszan R, Janicki H. Acta Pol Pharm. 1974; 31:121–122. [PubMed: 4832098]
41. Ozeki K, Sakabe N, Tanaka J. Acta Crystallogr., Sect. B. 1969; 25:1038–1045.
42. Portalone G, Bencivenni L, Colapietro M, Pieretti A, Ramondo F. Acta Chem Scand. 1999; 53:57–68.
43. Barnett SA, Hulme AT, Issa N, Lewis TC, Price LS, Tocher DA, Price SL. New J Chem. 2008; 32:1761–1775.
44. Markvardsen AJ, David WIF, Johnson JC, Shankland K. Acta Crystallogr., Sect. A. 2001; 57:47–54. [PubMed: 11124502]
45. David WIF, Shankland K, van de Streek J, Pidcock E, Motherwell WDS, Cole JC. J Appl Crystallogr. 2006; 39:910–915.
46. Pawley GS. J Appl Crystallogr. 1981; 14:357–361.
47. Coelho, AA. Topas Academic V5. Coelho Software; Brisbane: 2012.
48. Crystal data for AH B: C5H6N2O2, M = 126.12, monoclinic, a = 24.919(3) Å, b = 6.8385(7) Å, c = 6.6132(7) Å, β = 90.081(4)°, V = 1127.0(2) Å3, T = 183(2) K, space group C2/c, Z =8, 1028 reflections measured; 3.1° ≤θ ≤50.5° (λ = 0.71073 Å).The final R1 value was 0.0424 (I > 2σ(I)). The final wR(F2) values were 0.1057 (I > 2σ(I)) and 0.1166 (all data). The max. and min. residual densities were 0.30 and −0.28 e– Å3, respectively.CCDC No. 1470238
49. Bruker APEX2 and SAINT. Bruker AXS Inc; Madison, Wisconsin, USA: 2009.
50. Sheldrick GM. Acta Crystallogr., Sect. C: Struct. Chem. 2015; 71:3–8. [PubMed: 25567568]
51. Sheldrick GM. Acta Crystallogr., Sect. A: Found. Adv. 2015; 71:3–8. [PubMed: 25537383]
52. Spek AL. Acta Crystallographica Section D. 2009; 65:148–155.

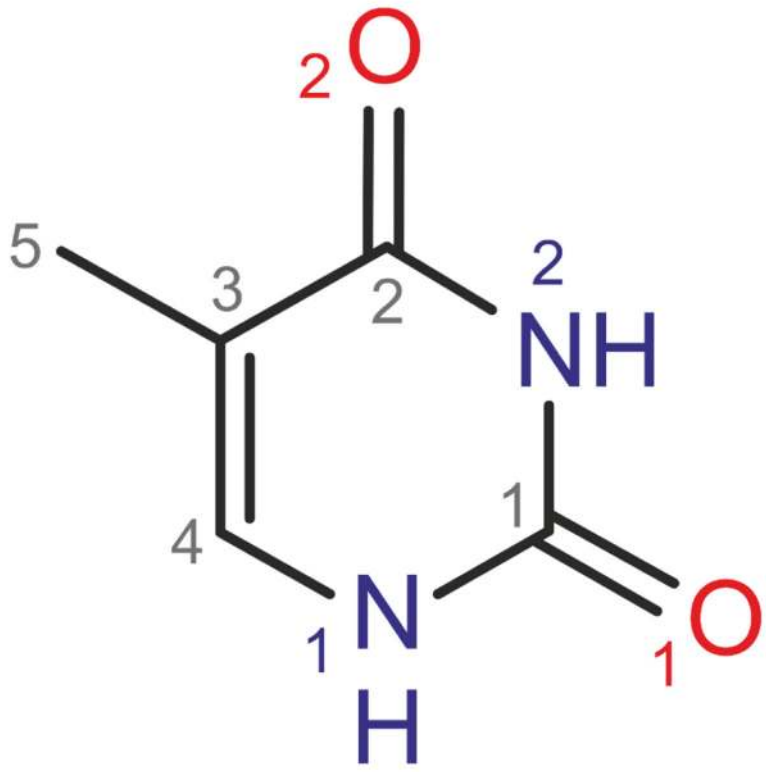
53. Ahlqvist MUA, Taylor LS. J Pharm Sci. 2002; 91:690–698. [PubMed: 11920754]
54. Goelles F. Monatsh Chem. 1961; 92:981–991.
55. Zhu H, Yuen C, Grant DJW. Int J Pharm. 1996; 135:151–160.
56. Frisch, MJ., Trucks, GW., Schlegel, HB., Scuseria, GE., Robb, JMA., Cheeseman, R., Scalmani, G., Barone, V., Mennucci, B., Petersson, GA., Nakatsuji, H., et al. Gaussian 09. Gaussian Inc; Wallingford CT: 2009.
57. Tsuchiya Y, Tamura T, Fujii M, Ito M. J Phys Chem. 1988; 92:1760–1765.
58. Karamertzanis PG, Pantelides CC. J Comput Chem. 2005; 26:304–324. [PubMed: 15622548]
59. Karamertzanis PG, Pantelides CC. Molecular Physics. 2007; 105:273–291.
60. Habgood M, Sugden IJ, Kazantsev AV, Adjiman CS, Pantelides CC. J Chem Theory Comput. 2015; 11:1957–1969. [PubMed: 26574397]
61. Coombes DS, Price SL, Willock DJ, Leslie M. J Phys Chem. 1996; 100:7352–7360.
62. Breneman CM, Wiberg KB. J Comput Chem. 1990; 11:361–373.
63. Price SL, Leslie M, Welch GWA, Habgood M, Price LS, Karamertzanis PG, Day GM. Phys Chem Chem Phys. 2010; 12:8478–8490. [PubMed: 20607186]
64. Stone AJ. J Chem Theory Comput. 2005; 1:1128–1132. [PubMed: 26631656]
65. Stone, AJ. GDMA: A Program for Performing Distributed Multipole Analysis of Wave Functions Calculated Using the Gaussian Program System. University of Cambridge; Cambridge, United Kingdom: 2010. version 2.2
66. Kazantsev AV, Karamertzanis PG, Adjiman CS, Pantelides CC. J Chem Theory Comput. 2011; 7:1998–2016. [PubMed: 26596459]
67. Clark SJ, Segall MD, Pickard CJ, Hasnip PJ, Probert MJ, Refson K, Payne MC. Z Kristallogr. 2005; 220:567–570.
68. Perdew JP, Burke K, Ernzerhof M. Phys Rev Lett. 1996; 77:3865–3868. [PubMed: 10062328]
69. Vanderbilt D. Phys Rev B. 1990; 41:7892–7895.
70. Tkatchenko A, Scheffler M. Phys Rev Lett. 2009; 102 073005-1-073005/4.
71. Grimme S. J Comput Chem. 2006; 27:1787–1799. [PubMed: 16955487]
72. Gavezzotti A. New J Chem. 2011; 35:1360–1368.
73. Gavezzotti A. J Phys Chem B. 2002; 106:4145–4154.
74. Gavezzotti A. J Phys Chem B. 2003; 107:2344–2353.
75. Gelbrich T, Hursthouse MB. CrystEngComm. 2006; 8:448–460.
76. Gelbrich T, Threlfall TL, Hursthouse MB. CrystEngComm. 2012; 14:5454–5464.
77. Allen FH. Acta Crystallogr., Sect. B. 2002; 58:380–388. [PubMed: 12037359]
78. Etter MC, MacDonald JC, Bernstein J. Acta Crystallographica, Section B: Structural Science. 1990; B46:256–262.
79. Barnett SA, Johnson A, Florence AJ, Price SL, Tocher DA. Cryst Growth Des. 2008; 8:24–36.
80. Braun DE, Gelbrich T, Kahlenberg V, Griesser UJ. CrystEngComm. 2015; 17:2504–2516. [PubMed: 26726294]
81. Kuhnert-Brandstaetter M, Proell F. Mikrochimica Acta. 1983; 3:287–300.
82. Braun DE, Tocher DA, Price SL, Griesser UJ. J Phys Chem B. 2012; 116:3961–3972. [PubMed: 22390190]
83. The measurements, were performed in a pan that was covered with a one pinholed lid.
84. Riddick, JA., Bunger, WB. Techniques of Chemistry, Vol. 2: Organic Solvents: Physical Properties and Methods of Purification. 4th ed. Wiley-Interscience; New York: 1986.
85. The heat capacity term (difference between the heat capacities of water in vapor and liquid states) was ignored, as we used the heat of vaporization of water at the dehdyration temperature.
86. In DSC experiments of hydrates, performed with hermetically sealed pans, any water that is released from the hydrate will be kept in the system. An equilibrium water vapor pressure will build up in the small volume above the sample. Thus, the composition of the condensed phase will remain relatively constant.

87. Braun DE, Gelbrich T, Kahlenberg V, Griesser UJ. Molecular Pharmaceutics. 2014; 11:3145–3163. [PubMed: 25036525]
88. Braun DE, Orlova M, Griesser UJ. Cryst Growth Des. 2014; 14:4895–4900.
89. Braun DE, Oberacher H, Arnhard K, Orlova M, Griesser UJ. CrystEngComm. 2016; doi: 10.1039/C5CE01758K
90. Burger A, Ramberger R. Mikrochimica Acta. 1979; 2:259–271.
91. Burger A, Ramberger R. Mikrochim Acta. 1979; 2:273–316.
92. Price SL. Chem Soc Rev. 2014; 43:2098–2111. [PubMed: 24263977]
93. Braun DE, Bhardwaj RM, Arlin JB, Florence AJ, Kahlenberg V, Griesser UJ, Tocher DA, Price SL. Cryst Growth Des. 2013; 13:4071–4083.
94. Cruz-Cabeza AJ, Day GM, Jones W. Chem Eur J. 2009; 15:13033–13040. [PubMed: 19876969]
95. Gerdil R, Marsh RE. Acta Crystallogr. 1960; 13:166–167.
96. Griesser UJ, Burger A. International Journal of Pharmaceuticals. 1995; 120:83–93.
97. Bond AD, Boese R, Desiraju GR. Angew Chem Int Ed. 2007; 46:618–622.
98. Braun DE, Nartowski KP, Khimyak YZ, Morris KR, Byrn SR, Griesser UJ. Mol Pharm. 2016; doi: 10.1021/acs.molpharmaceut.5b00856
99. Bernstein J, Davey RJ, Henck JO. Angew Chem Int Ed. 1999; 38:3441–3461.
100. Price SL, Reutzel-Edens SM. Drug Discov Today. 2016
101. Arlin JB, Price LS, Price SL, Florence AJ. Chem Commun. 2011; 47:7074–7076.
102. Braun DE, McMahon JA, Koztecki LH, Price SL, Reutzel-Edens SM. Cryst Growth Des. 2014; 14:2056–2072.
103. Habgood M. Cryst Growth Des. 2011; 11:3600–3608.
104. Neumann MA, van de Streek J, Fabbiani FPA, Hidber P, Grassmann O. Nat Commun. 2015; 6:7793. [PubMed: 26198974]
105. Lee AY, Lee IS, Dettet SS, Boerner J, Myerson AS. J Am Chem Soc. 2005; 127:14982–14983. [PubMed: 16248610]
106. Pfund LY, Matzger AJ. ACS Comb Sci. 2014; 16:309–313. [PubMed: 24933573]
107. Taleb M, Didierjean C, Jelsch C, Mangeot JP, Capelle B, Aubry A. J Cryst Growth. 1999; 200:575–582.
108. Ruecroft G, Hipkiss D, Ly T, Maxted N, Cains PW. Org Process Res Dev. 2005; 9:923–932.
109. Fabbiani FP, Buth G, Levendis DC, Cruz-Cabeza AJ. Chem Comm. 2014; 50:1817–1819. [PubMed: 24400322]
110. Braun DE, Kahlenberg V, Gelbrich T, Ludescher J, Griesser UJ. CrystEngComm. 2008; 10:1617–1625.
111. Schmidt AC, Niederwanger V, Griesser UJ. Journal of Thermal Analysis and Calorimetry. 2004; 77:639–652.

Europe PMC Funders Author Manuscripts

Synopsis

A joined computational and experimental study of thymine revealed four anhydrous polymorphs and a 0.8-hydrate. Three of the polymorphs are homeoenergetic due to their similar packing modes which differ only in the location of the oxygen and hydrogen atoms. Models for all solid forms could be proposed by combining crystal structure prediction with powder X-ray diffraction.



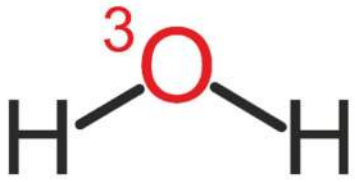


Figure 1.
Molecular diagram of thymine (5-methyluracil).

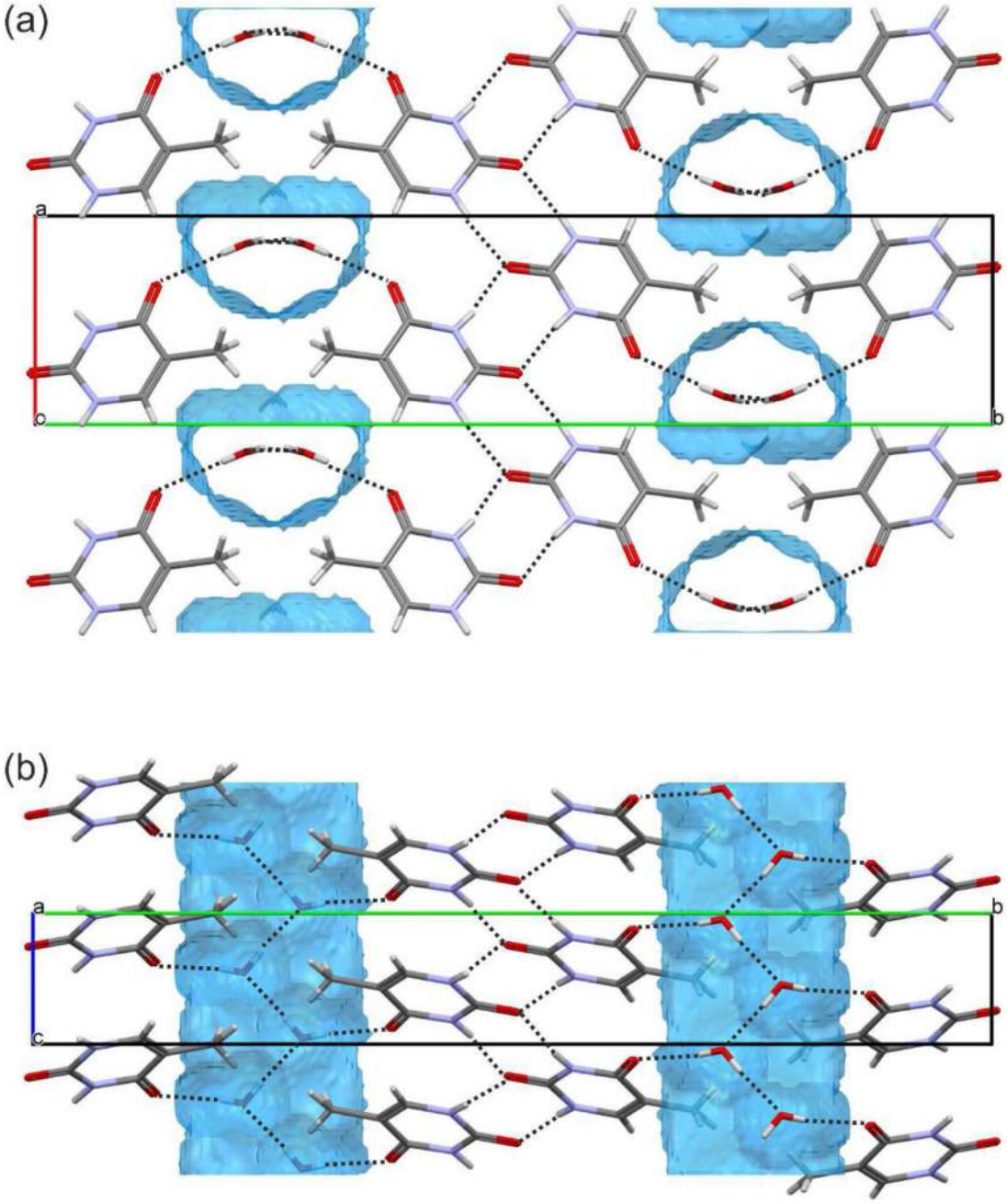


Figure 2.

Packing diagrams of thymine hydrate (THYMMH38) viewed along the crystallographic *c* axis (a) and *a* axis (b). Hydrogen bonds are denoted as black dotted lines and the water space (channels) in blue. Water space was calculated using the Hydrate Analyser tool in Mercury and a probe radius and approx. grid spacing of 1.0 and 0.15 Å, respectively.

Europe PMC Funders Author Manuscripts

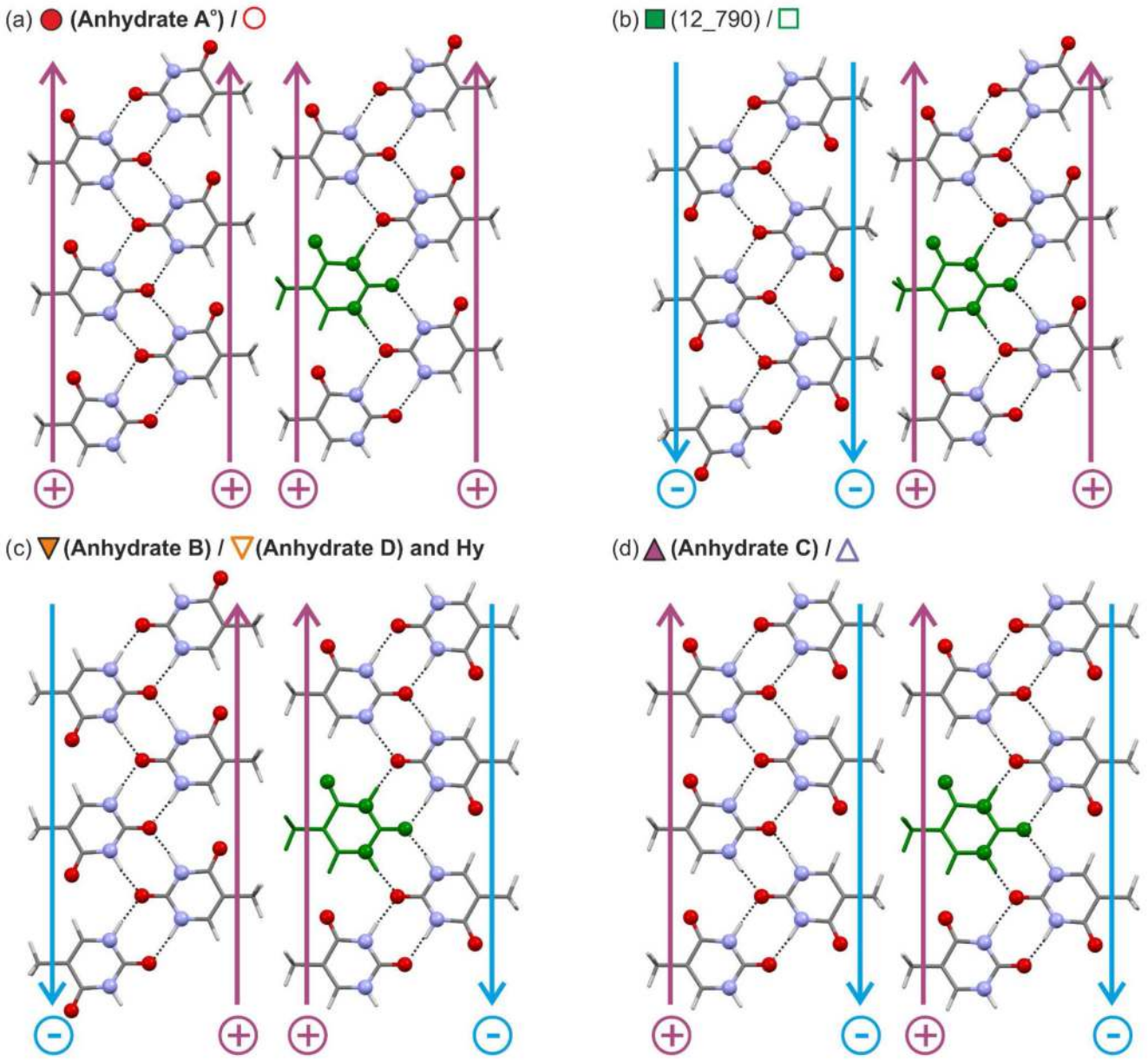


Figure 3.

Types of observed packing arrangements of ribbons in layers (planar or undulating): (a) polar parallel: I, (b) polar anti-parallel: II, (c) non-polar anti-parallel: IV, and (d) non-polar parallel: III. Roman numerals denote the classes of layer motifs. All layers are equivalent except for the relative positions of their C2=O2 and C4–H groups (Figure 1), but they differ in the polarity of the ribbons which is indicated by the arrows/"+" and "-" symbols.

Europe PMC Funders Author Manuscripts

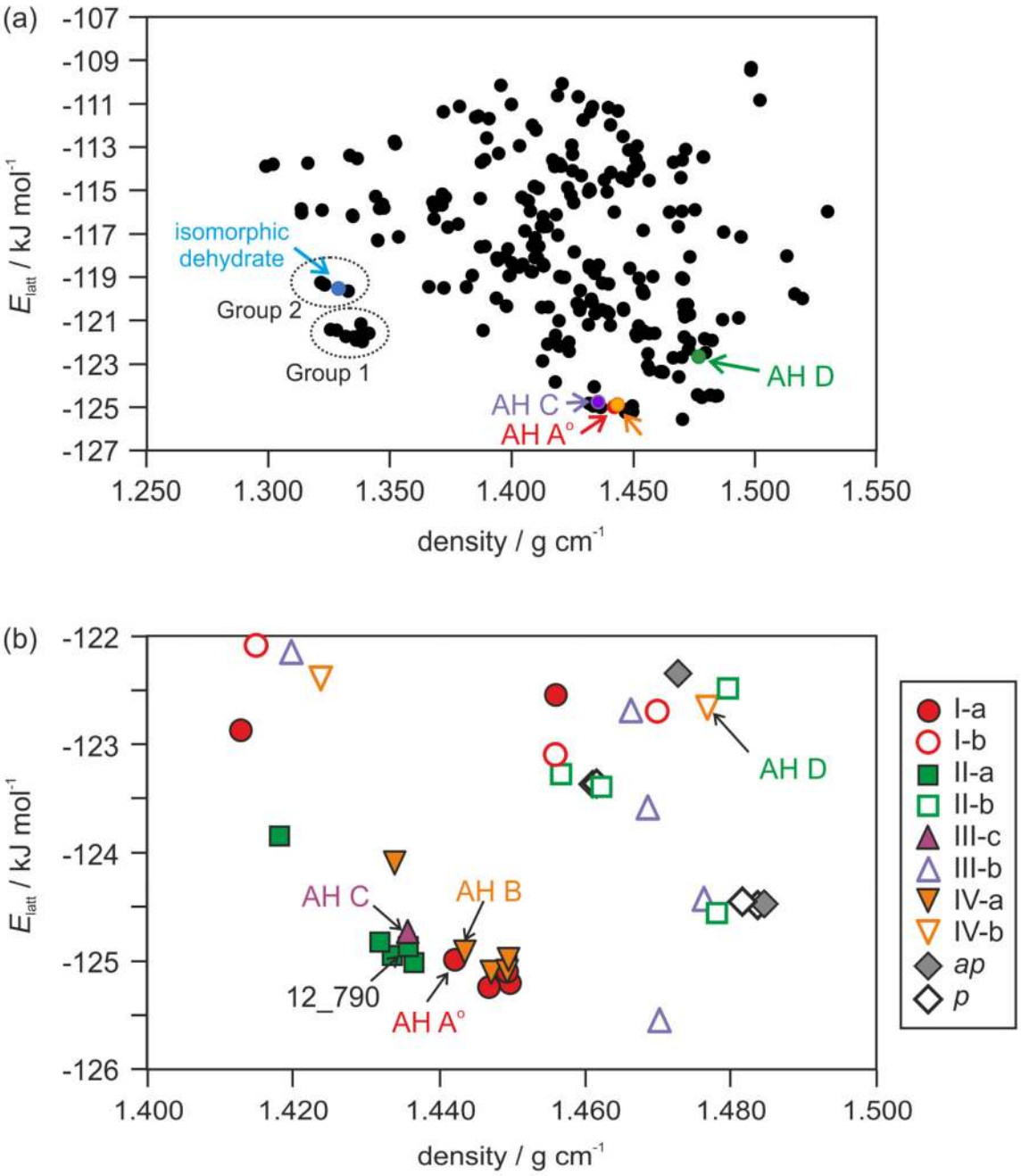


Figure 4.

(a, b) Lowest crystal energy structures for thymine anhydrate generated in CSP searches (CrystalOptimizer, MP2/6-31G(d,p)), with the experimental structures labelled. Encircled and marked as groups in (a) are low energy and low density structures. (b) The most stable computed structures are classified according to layer motifs (I-IV, Figure 3). Open/closed symbols denote distinct ribbon/layer stacks. For more details, see Table S11 of the Supporting Information. The structural data in *.res format are available from the authors on request.

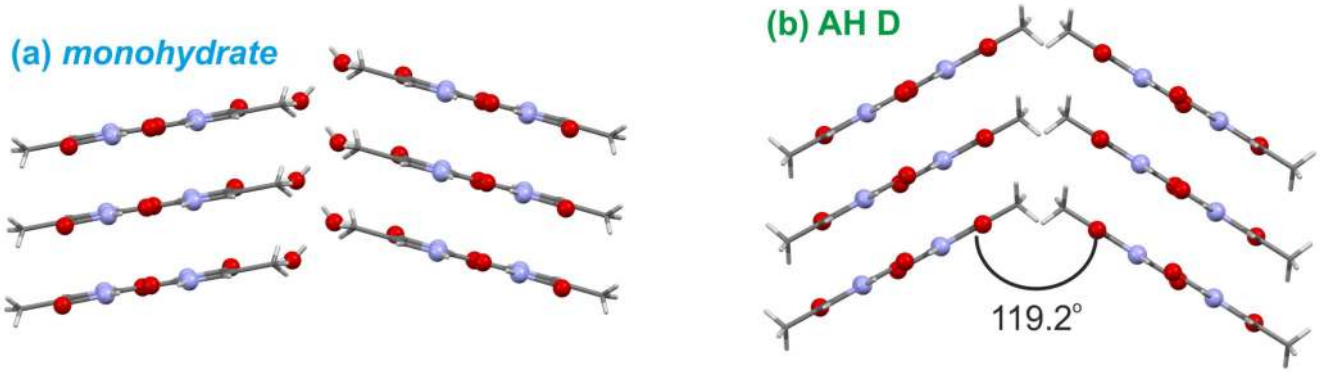


Figure 5.
Packing diagrams of (a) thymine hydrate and (b) **AH D**. Oxygen and nitrogen atoms are drawn as balls.

Europe PMC Funders Author Manuscripts

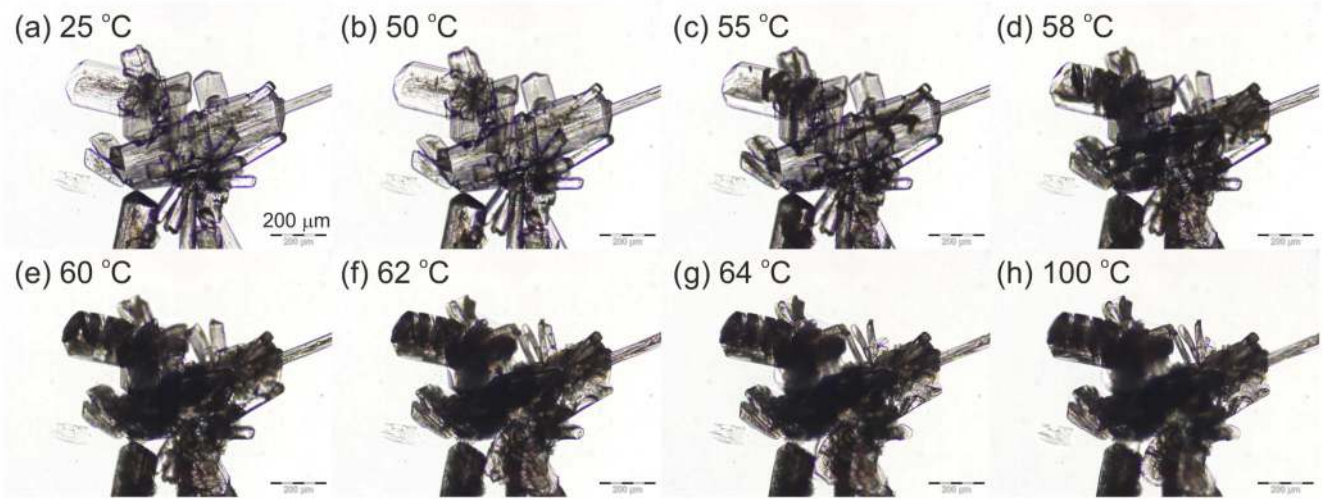


Figure 6.

Hot-stage thermal investigation of thymine **Hy** (embedded in silicon oil) showing the hydrate to anhydrate transition in the temperature range between 25 and 100 °C.

Europe PMC Funders Author Manuscripts

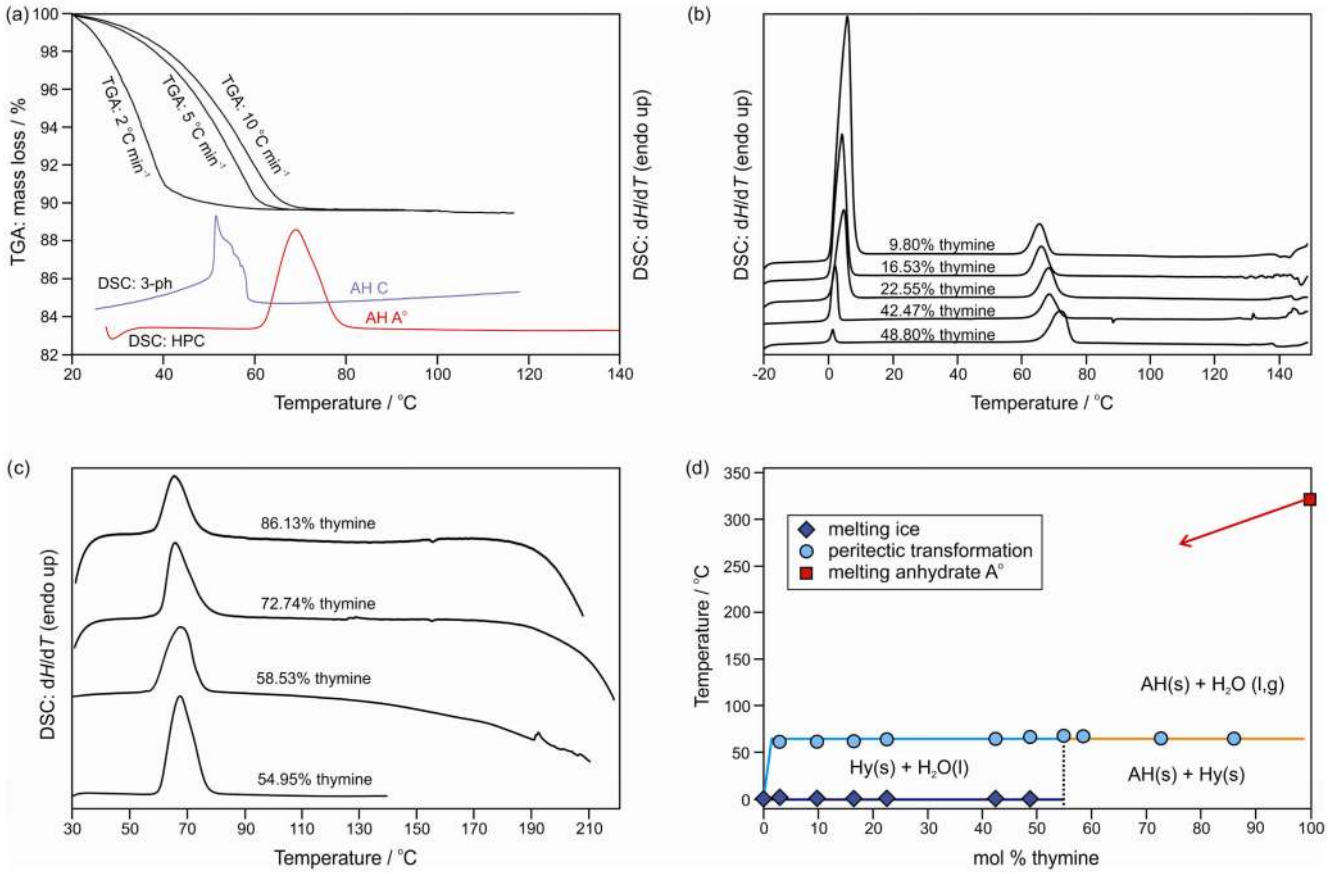


Figure 7.

(a) DSC and TGA thermograms of thymine hydrate (TGA curves: different heating rates were applied; DSC curves: HPC – high pressure capsule (hermetically sealed) and a heating rate of 10 °C min^{-1}, 3-ph: 3 pinhole capsule and a heating rate of 2 °C min^{-1}. (b),(c) Selected DSC thermograms of the thymine–water system (sealed capsules, heating rate 10 °C min^{-1}) used for constructing the (d) temperature/composition phase diagram (AH – anhydrate, Hy – hydrate, s – solid, l – liquid, g – gaseous).

Europe PMC Funders Author Manuscripts

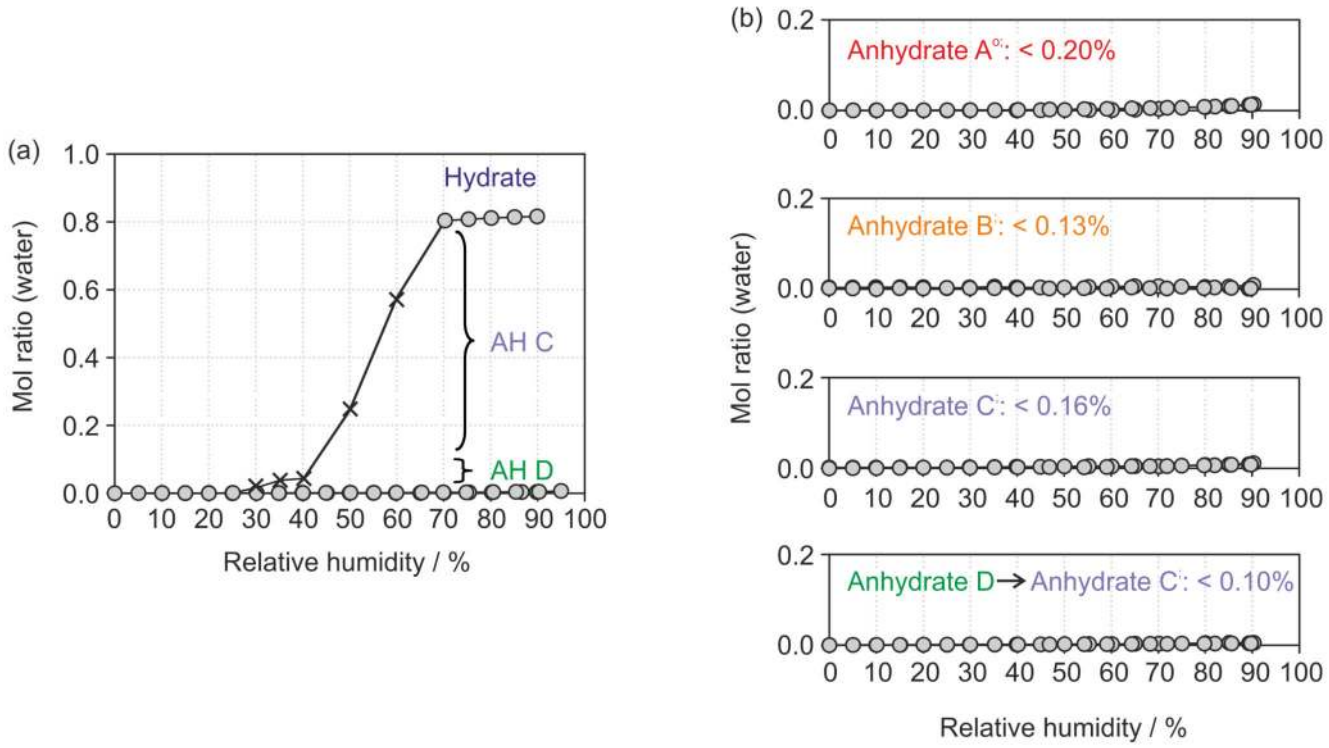


Figure 8.

Gravimetric moisture sorption and desorption curves of thymine solid forms at 25 °C: (a) hydrate and (b) anhydrates. The gray circles represent data points that satisfy the preset equilibria conditions (see section 2.6), whereas the crosses in (a) mark measurement values that did not reach equilibrium within the permitted time limit. Values for max. amounts of sorbed water are given in (b).

Europe PMC Funders Author Manuscripts

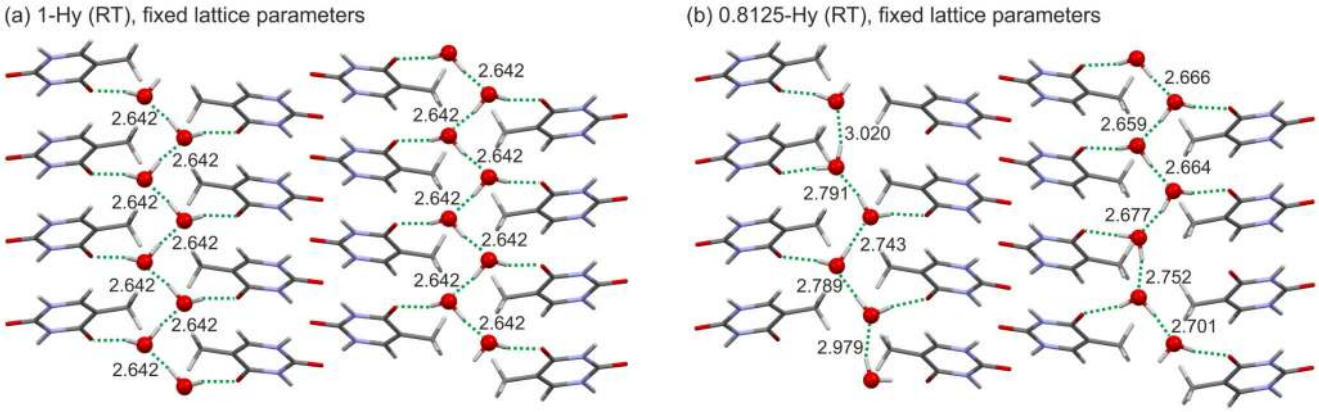


Figure 9.

Packing diagrams of computationally generated thymine hydrate models differing in the water:compound ratio: (a) 1:1 and (b) 0.8125:1. Hydrogen bonding interactions involving water molecules are denoted as green dotted lines and numbers denote water O···O interactions in Å. For more details see section 2.7 of the Supporting Information.

Europe PMC Funders Author Manuscripts

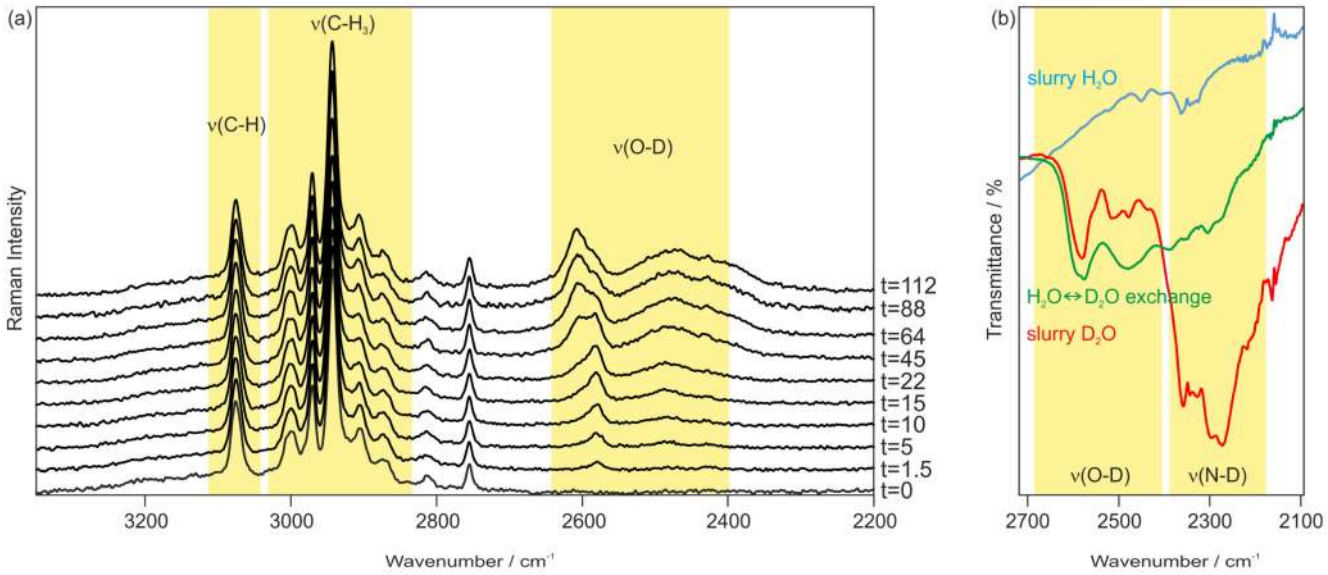


Figure 10.

(a) Raman spectra of thymine **Hy** as a function of the exposure time to D_2O vapor, ~100% RH. Peaks due to O–D stretching vibrations emerge after 1.5 hrs. (b) IR spectra of three thymine **Hy** samples prepared by slurring thymine in H_2O or D_2O or by exposing thymine **Hy** to D_2O vapor.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

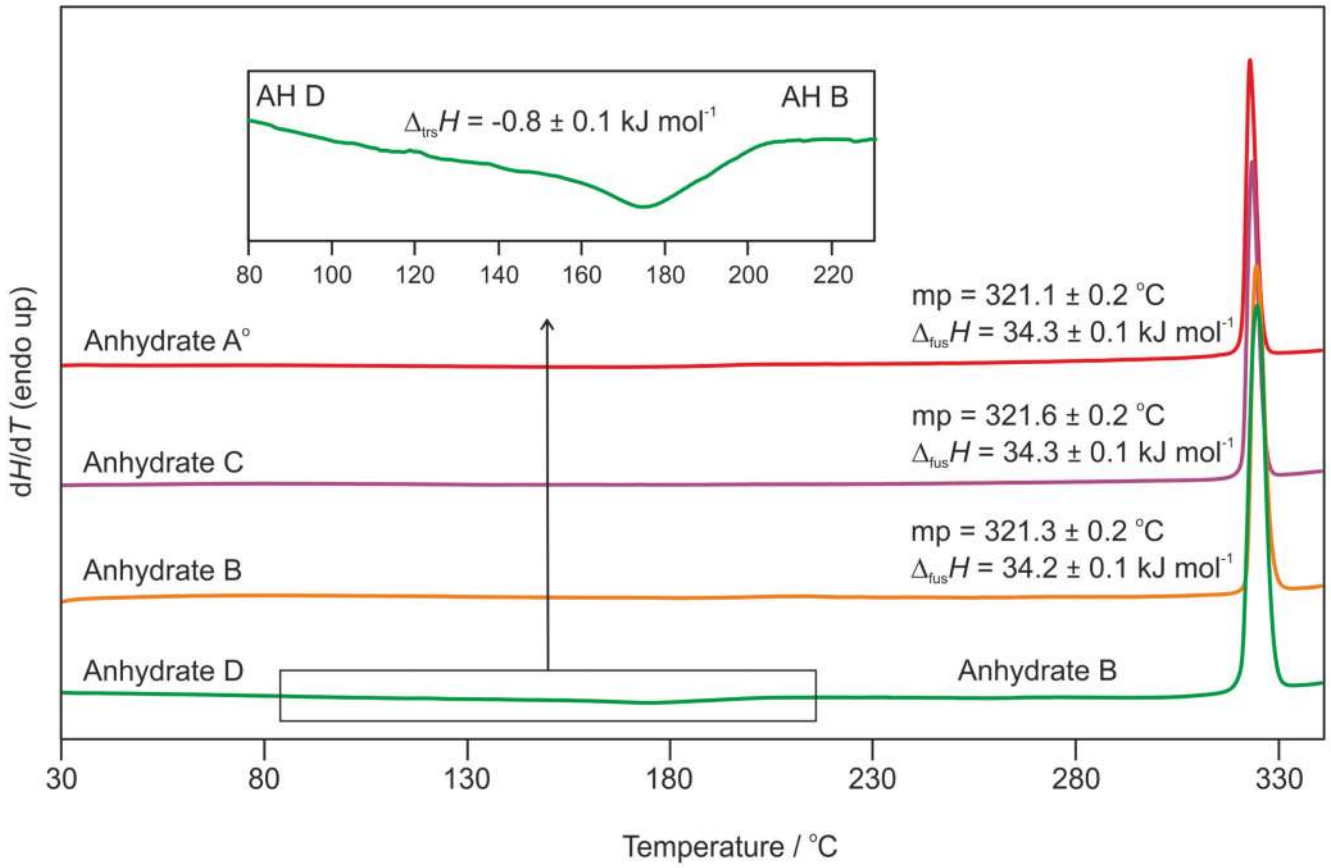


Figure 11.
DSC thermograms of thymine anhydrates, measured at a heating rate of 50 °C min^{-1}.

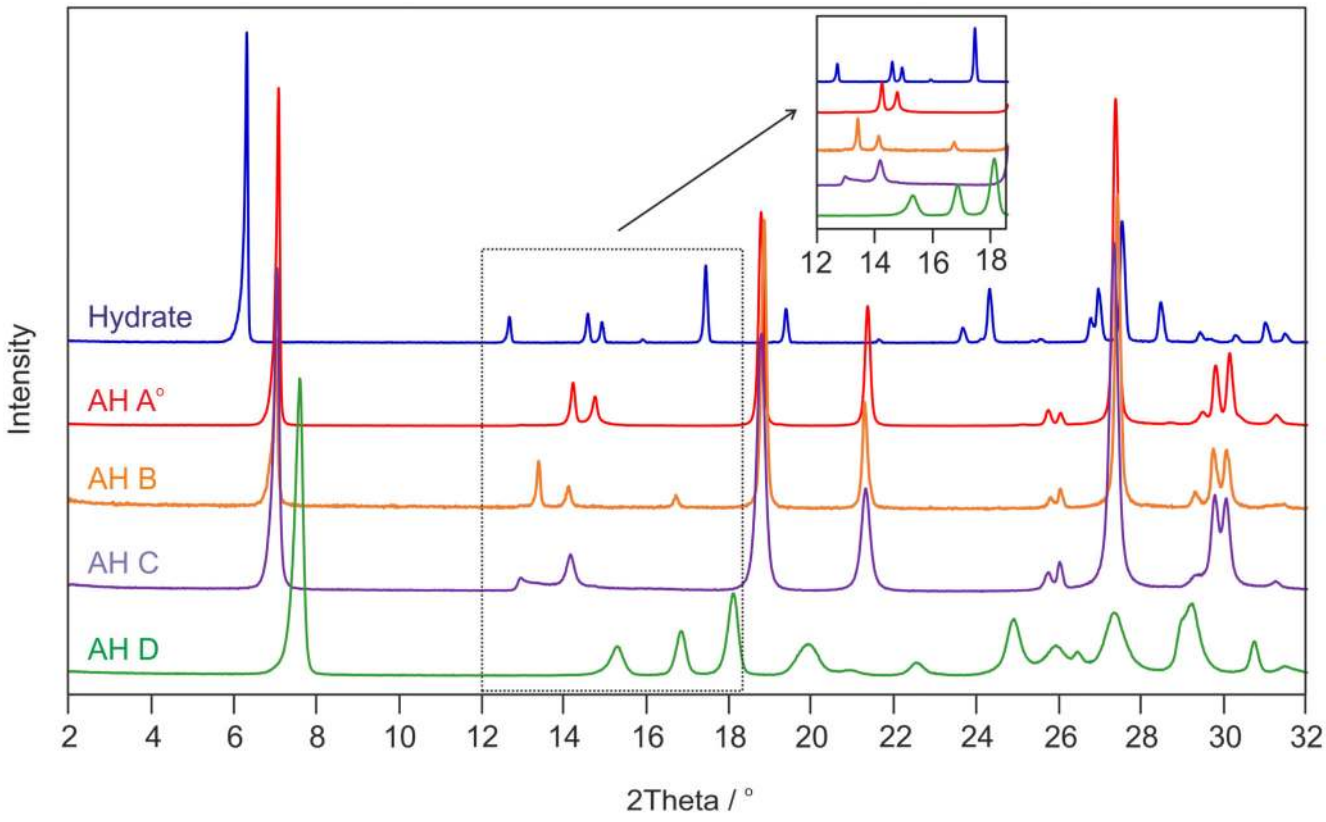


Figure 12.
Comparison of powder X-ray diffraction patterns of thymine solid forms.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

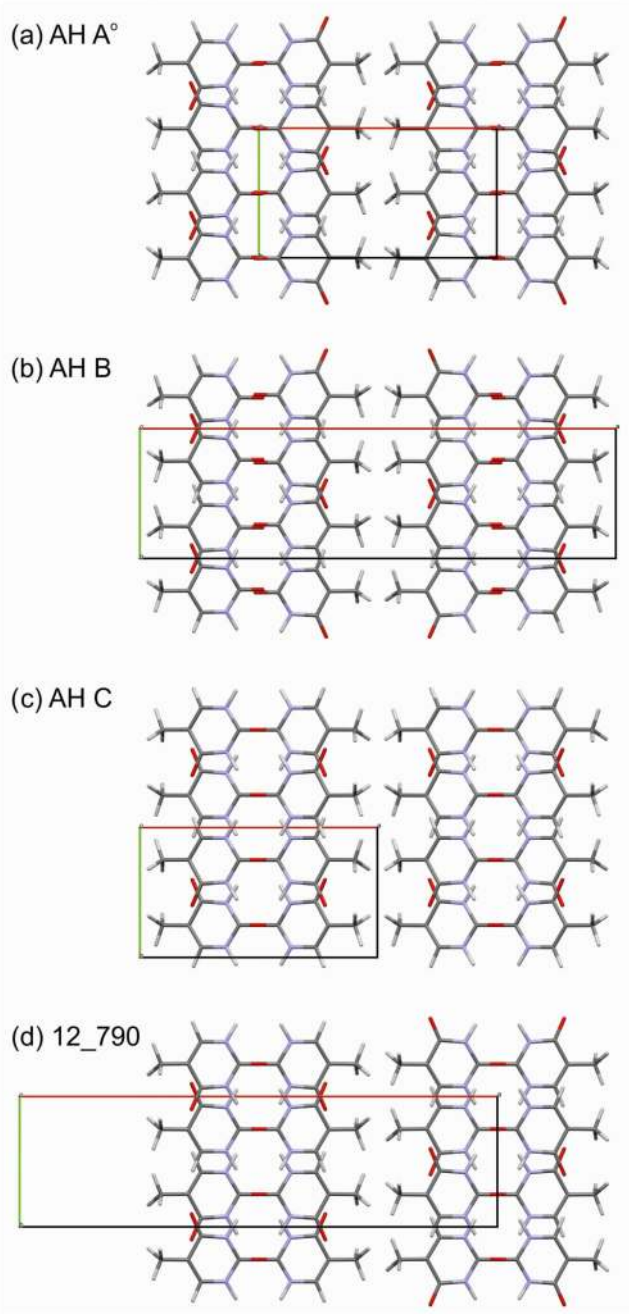


Figure 13.

Packing diagrams of **AH A°**, **AH B**, **AH C** and the computationally generated structure 12_790 viewed along the respective crystallographic *c* axis.

Europe PMC Funders Author Manuscripts

Table 1

Melting point (onset) temperatures, heat of fusion and transformation enthalpies for thymine anhydrates (measured at a heating rate of 50 °C min^{-1})

Polymorph	Melting point (onset) / °C	Heat of fusion / kJ mol^{-1}	Heat of transformation / kJ mol^{-1}
AH A°	321.1 ± 0.2	34.3 ± 0.1	–
AH B	321.3 ± 0.2	34.2 ± 0.1	–
AH C	321.6 ± 0.2	34.3 ± 0.1	–
AH D	–	–	–0.8 (to **AH B**)

Europe PMC Funders Author Manuscripts

Table 2

Lattice parameters derived from indexing the RT PXRD patterns of thymine anhydrates **A**° – **C** and computationally generated structure 12_790.

Mod.	Space Group	a / Å	b / Å	c / Å	β / °	Volume / Å^3
AH A°	$P2_1/c$	12.848(<1)	6.8296(2)	6.7439(3)	105.016(6)	571.52(6)
AH B	$C2/c$	25.014(2)	6.8321(2)	6.7061(2)	89.402(5)	1146.00(10)
AH C	$P2/c$	12.896(3)	6.8299(4)	6.7721(14)	104.984(23)	573.62(19)
12_790[a]	$C2/c$	25.014	6.8321	6.7061	90.598	1146.01

[a]Structure was optimized using the **AH B** lattice parameters.