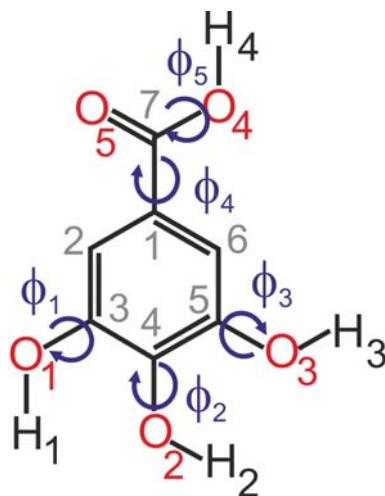# The complex polymorphic system of gallic acid - five monohydrates, three anhydrates and over 20 solvates

Doris E. Braun, Rajni M. Bhardwaj, Alastair J. Florence, Derek A. Tocher and Sarah L. Price

Electronic Supporting Information

# 1   COMPUTATIONAL

Figure S1. Molecular diagram of gallic acid. The torsions and angles that were optimized within the crystal energy minimization[1] were: ϕ_1 (H1-O1-C3-C2), ϕ_2 (H2-O2-C4-C3), ϕ_3 (H3-O3-C5-C4), ϕ_4 (C2-C1-C7-O5), ϕ_5 (H4-O4-C7-C1), H1-O1-C3, H2-O2-C4, H3-O3-C5, and H4-O4-C7.

1.1 <u>Conformational analysis of gallic acid</u>

   Potential energy surface scans (Figure S2) revealed that the acid exhibits four planar conformational minima (Table S1), separated by significant barriers, arising from the rotation of the hydroxyl (C-C-O-H) and the carboxylic acid group (C-C-C-O), if the rotation of the carboxylic acid proton (C1-C7-O4-H4) is ignored. The OH-group can rotate significantly for minimal energy cost, in agreement with the experimental conformations. There is a significant energy difference for swapping the position of the COOH proton. The relative energies of the four fully optimized conformations depend on the level of theory, but there is a clear pattern of having pairs of conformations that are close in energy varying in the position of the COOH proton, i.e. the pair of conformations forming two intramolecular hydrogen bonds and the pair of conformations forming one intramolecular hydrogen bond (Table S1). The PCM (ε=3, see section 1.2) model narrows the energy gap between the structures forming two intramolecular hydrogen bonds and structures forming one intramolecular hydrogen bond.
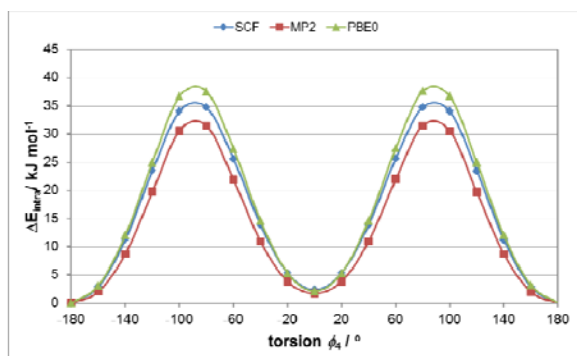
(a) φ4 (C2-C1-C7-O5): conf2-conf1-conf2
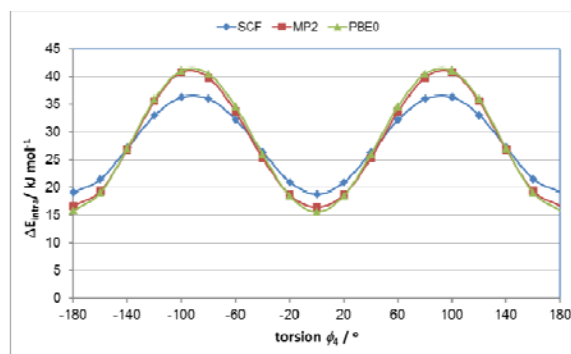
(b) φ4 (C2-C1-C7-O5): conf3-conf4-conf3
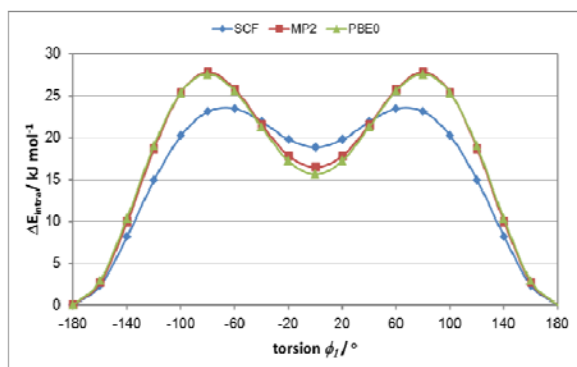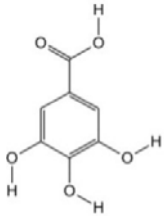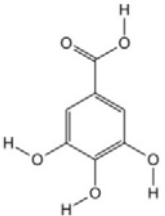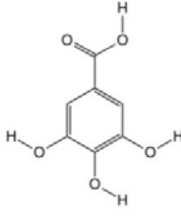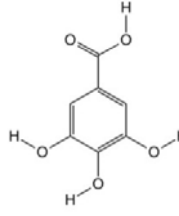
(c) φ1 (H1-O1-C3-C2): conf2-conf4-conf2
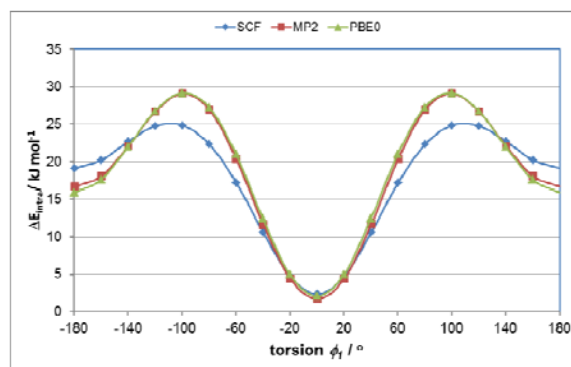
(d) φ3 (H3-O3-C5-C4): conf3-conf1-conf3

Figure S2. Potential energy surface scans for gallic acid with respect to (a) φ4 (C2-C1-C7-O5) starting with conf2 (-180 ° and 180 °), (b) φ4 (C2-C1-C7-O5) starting with conf3 (-180 ° and 180°), (c) φ1 (H1-O1-C3-C2) starting with conf2 (-180 ° and 180 °), and (d) φ3 (H3-O3-C5-C4) starting with conf3 (-180 ° and 180 °) at the SCF, MP2 and PBE0 level of theory with the 6-31G(d,p) basis set. The rest of the molecule was optimized at each value of φ, which was calculated every 10°. For conformational diagrams of the minima see Table S1.

Table S1. Conformational energy minima for gallic acid and their energy (ΔE_{intra}) values with respect to the global conformational minimum (conf2), for different wave functions.

$+ \Delta E_{intra}$ [kJ mol^{-1}]	conf1	conf2	conf3	conf4
SCF/6-31G(d,p)	2.36	0	19.15	18.82
MP2/6-31G(d,p)	1.69	0	16.68	16.49
PBE0/6-31G(d,p)	2.07	0	15.89	15.63
MP2/6-31G(d,p), PCM ε=3	1.51	0	8.73	8.56
PBE0/aug-cc-pVTz, PCM ε=3	1.47	0	7.17	6.97
Conformation present in experimental forms:	AH-II°, AH-III, MH-III	AH-I, MH-I°	MH-III, MH-IV	MH-II, MH-V

1.2 Computational generation of the anhydrate and monohydrate crystal energy landscapes

Hypothetical crystal structures of gallic acid anhydrates (Z'=1 and Z'=2) and monohydrates (Z'=1), for each of the four planar conformations, were generated with the program CrystalPredictor, which uses a low-discrepancy sequence to search the crystal packing space with quasi-random values for unit cell dimensions, molecular orientations and positions, followed by rigid molecule lattice energy minimization.[2] This resulted in four searches for Z'=1 and ten for Z'=2 anhydrates and four searches for the Z'=1 monohydrates. For each of the hydrate searches crystal structures were generated in 24 common space groups ($P1$, P-1, $P2_1$, $P2_1/c$, $P2_12_12$, $P2_12_12_1$, $Pna2_1$, $Pca2_1$, $Pbca$, $Pbcn$, $C2/c$, Cc, $C2$, $P2_1/m$, $P2/c$, $C222_1$, $Pnma$, $Fdd2$, $Fddd$, $P4_1$, $P4_12_12$, $P4_32_12$, $P3_1$, R-3). The anhydrate searches covered 32 common space groups (with the addition of Pc, Cm, $C2/m$, $Pmn2_1$, $Pmmn$, $Pbcm$, $Pccn$, $Pnna$). In all searches the molecules were in a general position and held rigid. Searches were continued until 230,000 structures had been energy minimized for each search for the hydrates and Z'=2 anhydrates and 50,000 for the Z'=1 anhydrates. The model for the intermolecular forces was an isotropic atom-atom potential (FIT) using atomic charges fitted to the PBE0/6-31G(d,p) electrostatic potential using the CHELPG scheme.[3]

Following each CrystalPredictor search, the 10,000 lowest energy structures, or all those within 20 kJ mol^{-1} range with respect to that search minimum, were used as starting points for lattice energy minimizations (DMACRYS[4]) using distributed multipoles[5] derived from the PBE0/aug-cc-pVTZ charge densities and the FIT parameters for the repulsion-dispersion energy. Conformations were kept rigid.

The most stable structures were refined allowing the conformational flexibility to optimize the proton positions and COOH torsion (ϕ4), using CrystalOptimizer.[1] About 400 (126 Z'=1 and 280 Z'=2) anhydrate and 300 monohydrate structures had the acid and water conformation within the crystal structures refined, varying the 5 non-aromatic torsion angles, the 4 C-O-H and the water bond angles (Figure S1). All molecular structures and energies were calculated at the PBE0/6-31G(d,p) level of theory, with the intermolecular lattice energies evaluated form the PBE0/aug-cc-pvTZ wave function distributed multipoles and the FIT potential parameters for the repulsion-dispersion models. The choice of structures to be refined by CrystalOptimizer took into account the known problem that the intramolecular energy penalty was likely to be overestimated for conf3 and conf4 (see Table S1).

To approximate the polarization of the molecule charge distribution in the crystal, as has been found necessary in CSP studies of peptides[6,7] and dihydroxybenzoic acids,[8] the charge density used in the evaluation of E_{latt} was generated in a dielectric continuum, with a dielectric constant ε=3, a value typical of organic crystals. Thus, the final PBE0/aug-cc-pVTz electron density calculations used the polarizable continuum model (PCM)[9] implemented in Gaussian03. The intramolecular energy penalty ΔE_{intra} was calculated from the electron density calculations derived in the CrystalOptimizer step.

The relationships between crystal structures were examined using the *XPac* program[10] and differences quantified using the overlay[11] of a 15 molecule cluster, rmsd$_{15}$, as calculated using the Molecular Similarity Module in Mercury.[12]

1.3 Periodic ab initio calculations (DFT-D): relative energy differences

Periodic electronic structure calculations have the advantage of not requiring separation into inter- and intramolecular interactions and that polarization is being automatically modelled well.

Therefore periodic electronic structure calculations were carried out on all experimental and closely related structures (which differ in proton positions, Table S2) with the CASTEP plane wave code[13] using the Perdew-Burke-Ernzerhof (PBE) generalized gradient approximation (GGA) exchange-correlation density functional[14] and ultrasoft pseudopotentials,[15] with the addition of a semi-empirical dispersion correction (Tkatchenko and Scheffler[16]). The results reported were obtained using a plane wave cut of energy of 780 eV and a Monkhorst-Pack[17] Brillouin zone sampling grid of spacing 2π x 0.07Å^{-1}. The required force tolerance for a successful geometry optimization in each run was 0.05 eV Å^{-1}. The enthalpy differences between the polymorphs changed by < 0.2 kJ mol^{-1} on doubling the cut off energy to 1560 eV, while the enthalpy change on a sampling grid of spacing 2π x 0.03Å^{-1} was smaller (<0.1 kJ mol^{-1}).

Table S2. Periodic ab initio calculations (DFT-D) and polarizable continuum model (PCM (ε=3), see 1.2) calculations for selected anhydrate and monohydrate structures. The structures with the protons position either derived from single crystal X-ray diffraction experiments or used as input for the structure solutions from powder X-ray diffraction data are in bold.

| Structure ID	(related to) "exptl."phase	Molecule Ψ [PCM (ε=3)]/ ΔE_{latt} / kJ mol^{-1}	Crystal Ψ [DFT-D]/ ΔE_{latt} / kJ mol^{-1}
A1_1	**AH-II° ($C2/c$, Z'=1)**	**0**	**0**
A13_2	AH-II° ($P2_1/c$, Z'=2)	+1.10	+0.76
A14_119	AH-II° (Cc, Z'=2)	+1.60	+8.86
A22_473	**AH-I (P-1, Z'=2)**	**-1.16**	**+1.91**
A12_38	AH-I (P-1, Z'=2)	+0.55	+4.41
A13_20	AH-I (P-1, Z'=2)	+0.06	+6.73
A12_1180	AH-I (P-1, Z'=2)	-1.19	+7.08
A11_4369	AH-I (P-1, Z'=2)	+1.16	+8.12
A1_20	**AH-III ($P2_1/c$, Z'=1)**	**+4.77**	**+3.16**
A3_86	AH-III ($P2_1/c$, Z'=1)	+2.66	+4.96
A4_20	AH-III ($P2_1/c$, Z'=1)	+12.49	+10.94
H2_312	**MH-I° ($P2/n$, Z'=1)**	**+1.99**	**+1.51**
H4_1532	**MH-II ($P2_1/c$, Z'=1)**	**+3.49**	**+5.07**
H2_104	MH-II ($P2_1/c$, Z'=1)	-4.06	+6.99
MH-III	**MH-III (P-1, Z'=4)**	**+1.45**	**+1.00**
H3_933	**MH-IV ($P2_1/c$, Z'=1)**	**0**	**0**
H4_1904	**MH-V ($P2_1/c$, Z'=1)**	**-2.80**	**+11.62**

1.4 Modelling of the experimental forms

The computational models were successful in reproducing the experimental anhydrate and hydrate structures of gallic acid (Table S3, Figures S3 to S10).

Table S3. Quality of representation of the experimental anhydrate and monohydrate structures.

	Lattice parameters (cell vectors/Å, angles/°)						cell density (g cm^{-3})	rmsd$_{15}$[a] (Å)
	a	b	C	α	β	γ		
AH-I (exp.)	7.318	8.253	11.715	100.47	90.23	90.99	1.624	−0.24
CryOpt (calc.)	7.501	8.117	11.534	98.41	91.25	92.56	1.628	0.33
PCM (ε=3) (calc.)	7.590	8.048	11.413	98.17	90.84	93.41	1.641	0.20
DFT-D (calc.)	7.170	8.337	11.554	98.99	92.99	91.24	1.668	
AH-II° (exp.)	25.685	4.927	11.251	90	106.23	90	1.653	–
CryOpt (calc.)	25.052	5.061	10.949	90	103.48	90	1.674	0.21
PCM (ε=3) (calc.)	26.276	4.762	11.455	90	108.65	90	1.664	0.20
DFT-D (calc.)	26.499	4.715	11.337	90	108.33	90	1.681	0.22
AH-III (exp.)	5.230	5.265	24.793	90	102.11	90		–
CryOpt (calc.)	5.427	5.158	24.326	90	101.74	90	1.695	0.27
PCM (ε=3) (calc.)	5.461	5.072	24.491	90	102.86	90	1.707	0.30
DFT-D (calc.)	5.524	4.769	25.230	90	99.05	90	1.721	0.50
MH-I (exp.)[b]	14.150	3.622	19.246	90	129.27	90	1.636	–
CryOpt (calc.)	14.241	3.489	19.142	90	129.43	90	1.701	0.19
PCM (ε=3) (calc.)	14.125	3.488	18.923	90	129.08	90	1.727	0.19
DFT-D (calc.)	14.068	3.587	18.920	90	129.60	90	1.699	0.08
MH-II (exp.)[c]	9.794	3.612	21.591	90	91.27	90	1.636	–
CryOpt (calc.)	9.636	3.536	21.560	90	92.23	90	1.702	0.11
PCM (ε=3) (calc.)	9.542	3.542	21.416	90	91.11	90	1.728	0.11
DFT-D (calc.)	9.618	3.575	21.403	90	91.19	90	1.698	0.09
MH-III (exp.)d	7.117	7.506	30.679	86.56	83.81	68.70	1.649	–
CryOpt (calc.)	7.076	7.579	30.452	87.35	82.79	68.41	1.659	0.14
PCM (ε=3) (calc.)	7.043	7.543	30.487	87.20	83.26	68.86	1.687	0.15
DFT-D (calc.)	7.161	7.411	30.627	79.57	89.64	66.93	1.688	0.13
MH-IV (exp.)e	5.794	4.719	28.688	90	95.08	90	1.599	–
CryOpt (calc.)	6.478	3.989	30.370	90	101.95	90	1.633	0.84
PCM (ε=3) (calc.)	6.378	3.972	30.312	90	102.12	90	1.664	0.73
DFT-D (calc.)	6.209	4.034	30.501	90	101.05	90	1.667	0.65
MH-V (exp.)	7.607	3.641	26.792	90	98.42	90	1.702	–
CryOpt (calc.)	7.631	3.814	25.092	90	98.79	90	1.732	0.32
PCM (ε=3) (calc.)	7.569	3.815	25.019	90	98.64	90	1.750	0.33
DFT-D (calc.)	7.567	3.854	25.458	90	98.94	90	1.704	0.32

[a]Reproduction of the experimental crystal structures was evaluated by the optimal root-mean square overlay of all non-hydrogen atoms in a 15 molecule coordination cluster (rmsd$_{15}$).[11] Experimental structures (CSD refcodes): [b]KONTIQ01, [c]KONTIQ03, [d]KONTIQ04 and [e]KONTIQ.
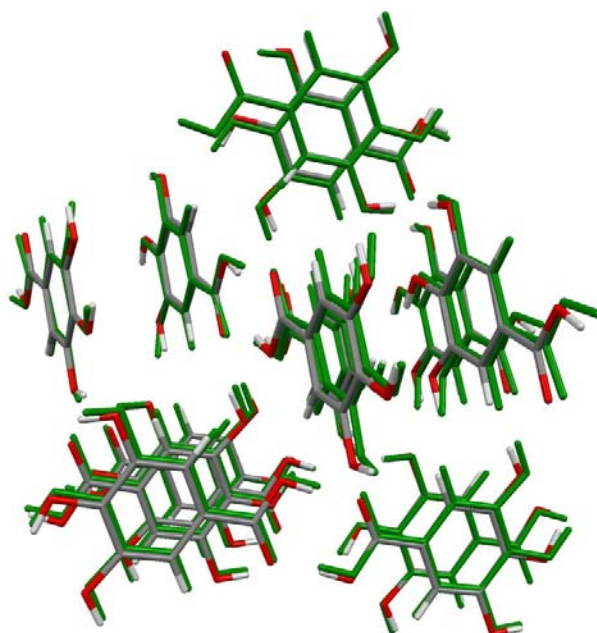
Figure S3. Overlay of the 15 molecule cluster of the observed structure of **AH-I** (colored by element and calculated DFT-D structure (green), rmsd$_{15}$=0.20 Å.
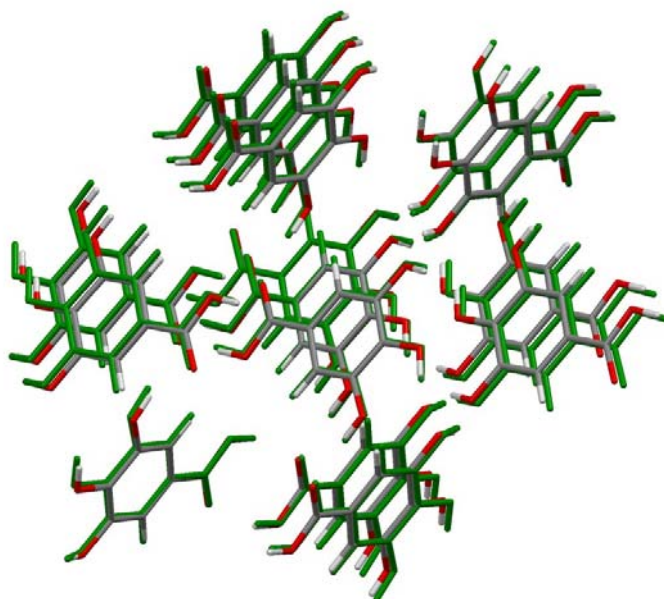
Figure S4. Overlay of the 15 molecule cluster of the observed structure of **AH-II°** (colored by element and calculated DFT-D structure (green), rmsd$_{15}$=0.22 Å.
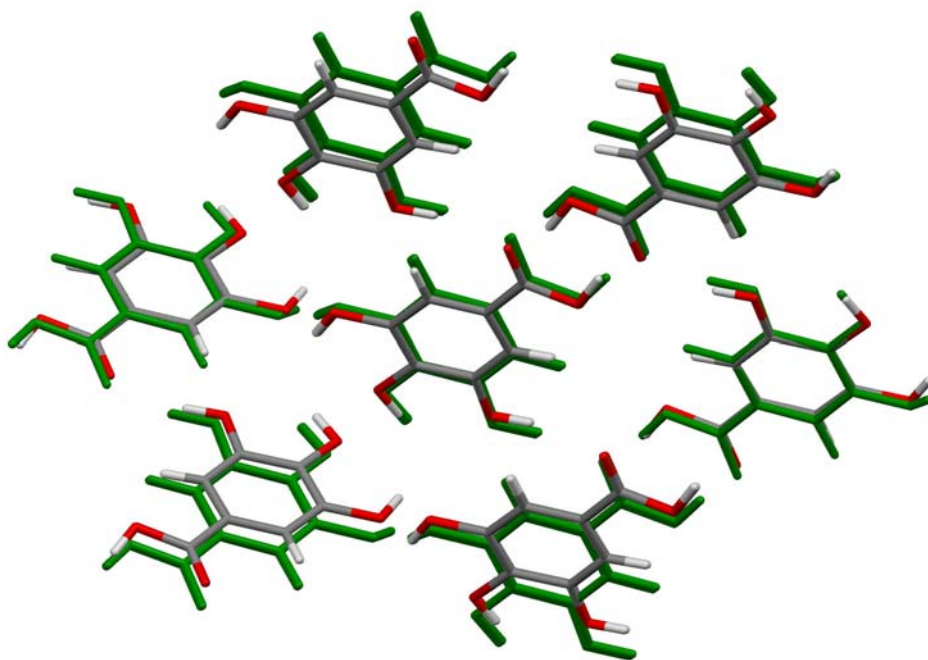
Figure S5. Overlay of the 15 molecule cluster of the observed structure of **AH-III** (colored by element and calculated DFT-D structure (green), $rmsd_{15}$=0.50 Å.
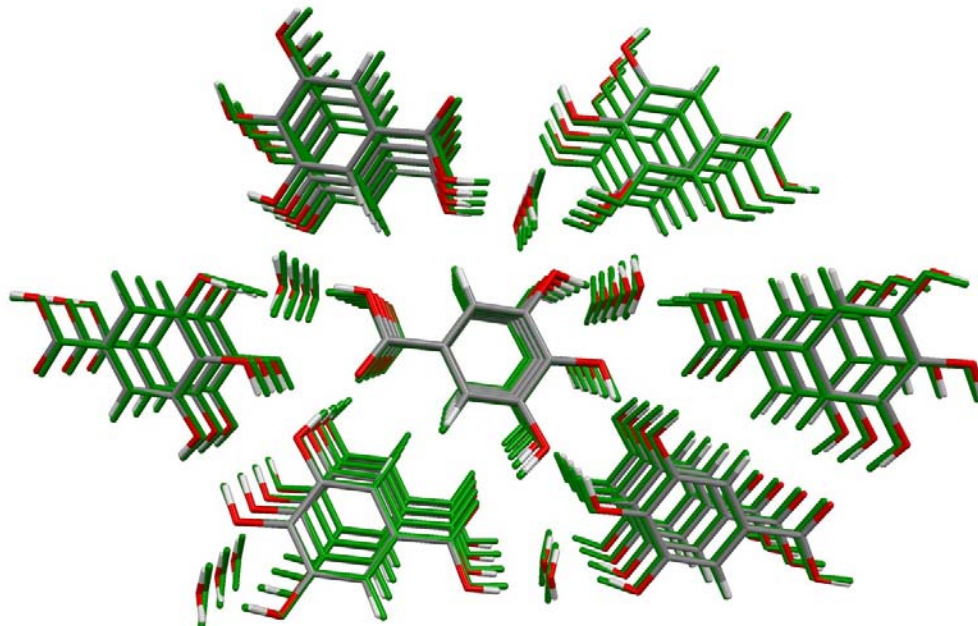
Figure S6. Overlay of the observed **MH-I°** structure of gallic acid (colored by element and calculated DFT-D structure (green), $rmsd_{15}$=0.08 Å.
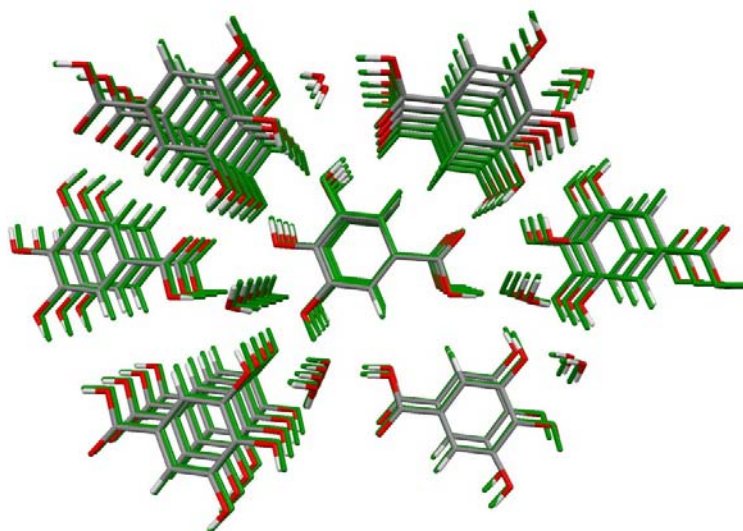
Figure S7. Overlay of the observed **MH-II** structure of gallic acid (colored by element and calculated DFT-D structure (green), $rmsd_{15}$=0.09 Å.
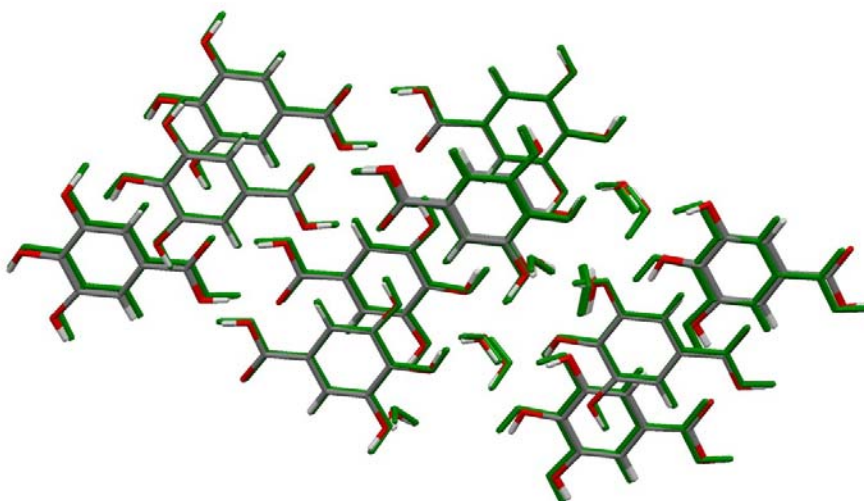
Figure S8. Overlay of the observed **MH-III** structure of gallic acid (colored by element and calculated DFT-D structure (green), $rmsd_{15}$=0.13 Å.

Figure S9. Overlay of the observed **MH-IV** structure of gallic acid (colored by element and calculated DFT-D structure (green), $rmsd_{15}$=0.65 Å.
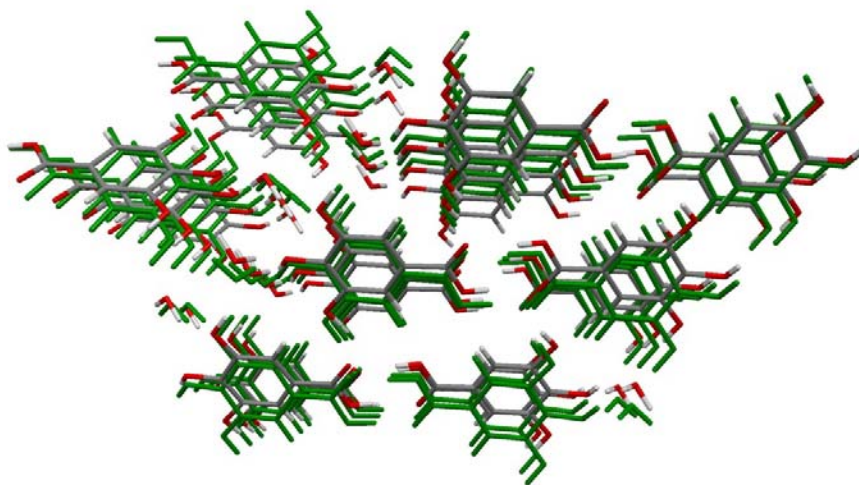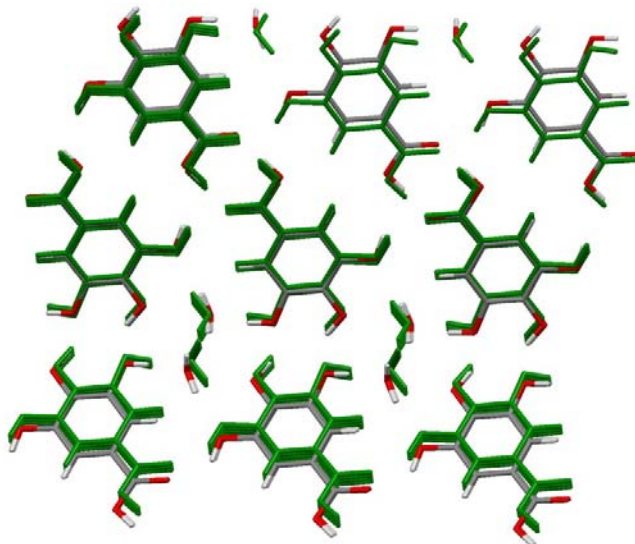

Figure S10. Overlay of the observed **MH-V** structure of gallic acid (colored by element and calculated DFT-D structure (green), $rmsd_{15}$=0.32 Å.

1.5 The crystal energy landscapes

Table S4. The low-energy crystal structures of gallic acid anhydrates shown on Figure 2b. Experimental structures (incl. structures differing in proton positions) are indicated in bold (Table S2 identifies the experimental structure).

Structure	Space group	Cell parameters a/Å	b/Å	c/Å	α/°	β/°	γ/°	E_{latt}/ kJ mol^{-1}	density/ g cm^{-3}	H-bonding[a]
A12_1180	***P*-1**	**11.373**	**8.079**	**7.571**	**87.01**	**86.01**	**82.22**	**-155.69**	**1.650**	**N2**
A22_473	***P*-1**	**7.590**	**11.413**	**8.048**	**81.83**	**93.41**	**90.84**	**-155.67**	**1.647**	**N2**
A23_34	$P2_1/c$	31.828	3.579	13.857	90	56.85	90	-154.96	1.710	D2*
A1_1	***C2/c***	**26.276**	**4.762**	**11.455**	**90**	**108.65**	**90**	**-154.51**	**1.664**	**N1**
A13_20	***P*-1**	**8.425**	**7.283**	**11.366**	**99.31**	**79.22**	**102.54**	**-154.45**	**1.703**	**N2**
A11_12	Pbca	50.310	4.763	11.319	90	90	90	-154.43	1.667	N1
A1_589	Pbcn	25.456	4.735	11.250	90	90	90	-154.29	1.667	N1
A11_7	C2/c	30.013	4.996	21.794	90	54.95	90	-154.13	1.695	Nx
A11_126	$P2_1/c$	14.960	12.182	7.325	90	89.78	90	-153.97	1.702	D1
A12_38	***P*-1**	**8.077**	**7.446**	**11.835**	**89.95**	**79.33**	**87.85**	**-153.96**	**1.628**	**N2**
A13_7	$P2_1/c$	7.383	10.031	17.519	90	105.71	90	-153.65	1.809	C2
A13_2	***P2₁/c***	**11.219**	**4.909**	**24.713**	**90**	**80.79**	**90**	**-153.41**	**1.682**	**N1**
A11_4369	***P*-1**	**8.060**	**14.915**	**7.471**	**60.96**	**87.68**	**65.66**	**-153.35**	**1.630**	**N2**
A1_21	P-1	8.451	7.110	6.445	64.31	100.97	90.47	-152.98	1.655	D1
A14_119	***Cc***	**26.943**	**4.277**	**12.857**	**90**	**65.42**	**90**	**-152.90**	**1.677**	**N1**
A12_2	$P2_1/c$	3.784	11.009	35.955	90	60.33	90	-152.63	1.737	other
A22_152	Pca2₁	16.597	3.619	21.279	90	90	90	-152.57	1.768	C1
A24_21	$P2_1/c$	22.496	3.631	16.419	90	106.06	90	-152.50	1.754	C1
A14_154	Cc	14.246	3.656	25.779	90	80.60	90	-152.31	1.706	D1
A12_261	P-1	6.440	17.388	7.161	104.82	63.52	89.83	-152.14	1.644	D1
A1_25	$P2_1/c$	5.356	25.962	4.803	90	85.38	90	-152.01	1.697	D1
A3_86	***P2₁/c***	**4.991**	**5.621**	**24.228**	**90**	**74.26**	**90**	**-151.85**	**1.727**	**D1**
A33_6	Pc	13.687	3.548	13.838	90	73.00	90	-151.63	1.758	C2
A14_132	***P2₁/c***	**25.147**	**4.345**	**12.719**	**90**	**78.49**	**90**	**-151.56**	**1.660**	**N1**
A13_31	$P2_1/c$	7.303	10.287	16.850	90	82.95	90	-151.44	1.799	C2
A2_18	$P2_1/c$	8.665	3.677	19.953	90	86.02	90	-151.40	1.782	C1
A2_583	P-1	12.012	7.110	6.441	115.74	67.17	135.29	-151.40	1.633	D1
A22_1335	P-1	8.665	3.677	21.197	89.99	69.91	90.00	-151.39	1.782	C1
A11_94	$P2_1/c$	3.745	16.090	22.638	90	98.52	90	-151.34	1.675	C1*,dx
A12_5	$P2_1/c$	3.700	29.444	11.884	90	84.70	90	-151.15	1.753	other
A4_30	P-1	7.223	7.375	7.407	108.63	113.55	62.38	-150.93	1.787	C1
A12_26	$P2_12_12_1$	3.729	33.285	10.143	90	90	90	-150.91	1.795	C12
A12_7	$P2_1/c$	3.556	30.277	12.173	90	86.96	90	-150.88	1.727	other
A23_5	P-1	24.428	7.390	3.660	87.12	92.21	83.43	-150.87	1.725	dx
A13_98	C2/c	52.531	3.773	13.537	90	86.38	90	-150.85	1.688	dx
A24_7	$P2_12_12_1$	3.594	46.506	7.827	90	90	90	-150.83	1.727	C1
A4_1	Pc	6.767	3.611	13.838	90	73.18	90	-150.78	1.745	C1
A24_14	$P2_12_12_1$	3.590	7.831	46.531	90	90	90	-150.70	1.727	C1
A24_141	P-1	7.535	12.000	7.496	90.07	82.34	77.53	-150.67	1.724	other
A33_2	C2/c	20.977	5.383	23.980	90	75.39	90	-150.60	1.725	Nx
A44_5023	$P2_1/c$	24.961	7.943	7.698	90	122.21	90	-150.56	1.750	dx
A4_3	$P2_1/c$	3.849	7.940	21.851	90	104.89	90	-150.54	1.751	dx
A33_53	$P2_1/c$	7.240	17.241	10.149	90	85.79	90	-150.47	1.789	C2
A13_3	***P2₁/c***	**10.742**	**5.191**	**24.236**	**90**	**102.14**	**90**	**-150.43**	**1.711**	**D1**

Structure	Space group	Cell parameters						E_{latt}/ kJ mol^{-1}	density/ g cm^{-3}	H-bonding[a]
		a/Å	b/Å	c/Å	α/°	β/°	γ/°			
A13_4	***P2₁/c***	**10.777**	**5.176**	**28.565**	**90**	**123.87**	**90**	**-150.41**	**1.708**	**D1**
A2_80	***C2/c***	**27.924**	**4.313**	**12.435**	**90**	**67.14**	**90**	**-150.32**	**1.637**	**N1**
A2_43	*P2₁/c*	8.644	10.257	7.095	90	89.43	90	-150.21	1.796	C1
A2_61	*P3₁*	13.230	13.230	3.526	90	90	120	-150.16	1.586	other
A23_17	***P2₁/c***	**11.271**	**4.867**	**30.635**	**90**	**126.54**	**90**	**-150.12**	**1.674**	**N1**
A2_49	*P2₁/c*	12.906	3.625	14.144	90	76.41	90	-150.05	1.757	C1
A13_166	*P*-1	3.657	7.292	25.810	82.88	91.47	77.17	-150.04	1.699	D1
A13_35	***P2₁***	**23.572**	**5.311**	**5.265**	**90**	**92.13**	**90**	**-150.01**	**1.715**	**D1**
A33_5	*Pc*	3.573	11.733	15.564	90	89.95	90	-149.93	1.732	C2
A2_1	C2/c	25.019	3.481	15.493	90	111.89	90	-149.91	1.805	other
A2_192	*P2₁/c*	5.384	26.416	4.742	90	95.07	90	-149.90	1.682	D1
A44_17	*Pc*	13.520	3.589	13.836	90	73.63	90	-149.83	1.754	L1
A11_53	*P2₁/c*	3.748	22.272	16.254	90	93.90	90	-149.82	1.669	L1*,dx
A33_19	*P2₁/c*	6.576	3.715	53.569	90	91.02	90	-149.80	1.727	dx
A14_13	*P2₁/c*	26.063	3.760	13.591	90	87.27	90	-149.78	1.699	C6*
A4_167	*P2₁/c*	6.319	3.886	27.718	90	104.91	90	-149.75	1.718	D1
A1_20	***P2₁/c***	**5.461**	**5.075**	**28.963**	**90**	**55.53**	**90**	**-149.74**	**1.707**	**D1**
A22_95	*P2₁2₁2₁*	14.280	3.563	25.108	90	90	90	-149.72	1.769	other
A1_60	*P2₁/c*	14.676	3.766	13.424	90	71.22	90	-149.71	1.609	Nx
A24_1311	P2/c	14.554	3.699	24.434	90	95.82	90	-149.71	1.727	Nx
A3_5	*P2₁2₁2₁*	10.872	3.615	16.724	90	90	90	-149.70	1.719	C2
A3_2	***C2/c***	**23.518**	**5.354**	**10.738**	**90**	**101.98**	**90**	**-149.70**	**1.709**	**N1**
A14_142	*P*-1	14.053	7.295	7.307	63.95	89.33	77.22	-149.66	1.730	D1
A33_21	Pc	3.582	10.926	16.982	90	75.08	90	-149.63	1.760	C2
A1_171	*P2₁/c*	6.597	9.232	11.290	90	92.55	90	-149.60	1.645	dx
A34_92	*P2₁/c*	3.847	16.414	21.251	90	82.26	90	-149.59	1.700	dx*
A12_90	*P2₁2₁2₁*	13.536	3.539	28.404	90	90	90	-149.59	1.661	other
A2_2	C2/c	23.866	3.500	15.497	90	74.58	90	-149.58	1.811	C3
A1_320	*P2₁/c*	16.802	10.128	3.727	90	84.28	90	-149.55	1.790	C2
A34_1	***P*-1**	**5.072**	**5.539**	**23.522**	**90.40**	**85.70**	**90.12**	**-149.53**	**1.715**	**D1**
A44_32	*Pc*	10.939	3.585	16.423	90	92.37	90	-149.53	1.756	C1
A2_3	*P2₁/c*	12.617	3.536	15.568	90	112.20	90	-149.51	1.757	C3
A22_10	C2/c	46.924	3.540	15.284	90	85.29	90	-149.48	1.786	C3
A12_10	*P2₁/c*	3.749	27.464	12.885	90	97.28	90	-149.46	1.717	other
A1_15	*Pna2₁*	15.560	11.667	3.720	90	90	90	-149.43	1.673	C2
A22_7	C2/c	46.432	3.507	15.533	90	94.05	90	-149.41	1.791	C3
A14_59	*P1*	14.287	8.299	3.636	61.01	61.43	72.35	-149.41	1.713	D1
A22_3	*C2/c*	46.848	3.482	15.547	90	94.58	90	-149.34	1.788	C3
A1_13	*Pca2₁*	13.434	3.866	12.864	90	90	90	-149.33	1.691	C4
A4_10	*Pc*	7.779	3.569	14.050	90	123.34	90	-149.33	1.734	C1
A14_125	*P2₁/c*	14.585	12.340	7.288	90	86.15	90	-149.32	1.727	D1
A11_184	C2/c	12.692	7.749	27.227	90	94.14	90	-149.31	1.692	D1
A4_9	*P2₁*	3.576	7.770	11.741	90	86.74	90	-149.30	1.735	C1
A2_105	*P*-1	10.240	5.887	6.781	69.90	65.25	87.92	-149.28	1.634	D1
A3_4	*Pna2₁*	10.739	16.903	3.658	90	90	90	-149.19	1.702	C2
A1_49	*P2₁/c*	6.110	4.200	36.182	90	46.79	90	-149.14	1.670	D1
A24_20	*P2₁/c*	22.783	3.610	16.444	90	106.40	90	-149.13	1.742	C1
A22_5	*C2/c*	54.108	3.513	15.302	90	60.71	90	-149.12	1.782	C3

| Structure | Space group | Cell parameters | | | | | | E_{latt}/ kJ mol^{-1} | density/ g cm^{-3} | H-bonding[a] |
		a/Å	b/Å	c/Å	α/°	β/°	γ/°			
A4_73	*Cc*	10.156	8.626	7.221	90	85.45	90	-149.07	1.792	C1
A4_44	*P2₁/c*	7.312	8.617	10.073	90	86.25	90	-149.05	1.784	C1
A2_6	*P2₁/c*	12.071	3.498	15.649	90	75.93	90	-149.04	1.763	C3
A23_47	*P2₁/c*	27.399	3.609	13.980	90	76.68	90	-148.95	1.680	dx
A22_22	C2/c	30.388	3.539	25.483	90	67.23	90	-148.85	1.789	C3
A4_107	C2/c	23.698	4.508	12.816	90	96.11	90	-148.84	1.660	Nx
A23_40	*P2₁/c*	26.717	3.604	13.752	90	90.29	90	-148.72	1.707	dx
A44_175	*P*-1	8.616	7.390	14.979	70.34	124.39	125.50	-148.65	1.764	C1
A23_294	*P2₁/c*	5.398	51.910	4.733	90	84.78	90	-148.61	1.711	D1*
A1_327	*P*-1	7.474	7.260	7.055	110.64	93.88	100.00	-148.60	1.617	dx
A2_15	*P2₁/c*	3.572	24.307	8.098	90	66.52	90	-148.59	1.752	C2
A2_155	*P2₁/c*	11.958	6.517	8.469	90	82.28	90	-148.58	1.728	C2
A4_52	*P2₁/c*	7.225	8.626	10.199	90	85.42	90	-148.56	1.783	C1
A22_11	C2/c	25.227	3.506	30.577	90	68.10	90	-148.55	1.801	C3
A23_22	***P2₁/c***	**24.211**	**5.216**	**10.714**	**90**	**79.05**	**90**	**-148.51**	**1.701**	**D1**
A23_927	***Cc***	**24.927**	**4.266**	**12.852**	**90**	**90.71**	**90**	**-148.44**	**1.654**	**N1**
A13_10	*P*-1	10.740	3.821	16.611	83.09	90.71	108.74	-148.44	1.764	C2
A44_6055	*Pc*	10.863	3.592	16.511	90	88.32	90	-148.42	1.755	C1
A24_228	*P2₁/c*	30.068	3.651	13.779	90	116.61	90	-148.42	1.671	D2
A3_11	*P2₁/c*	4.021	10.110	17.553	90	113.30	90	-148.39	1.724	C1
A2_11	*P*-1	11.213	3.730	8.357	103.57	106.96	88.06	-148.36	1.739	dx
A2_9	*P2₁/c*	24.669	3.516	7.589	90	79.32	90	-148.31	1.747	C3
A23_20	*P2₁/c*	26.493	3.609	13.853	90	89.70	90	-148.19	1.706	C6*

[a]Classification according to Figure 2d. Structures marked with an asterisk (*): only part of the structure forms the specified H-bonding motif.

Table S5. The low-energy crystal structures of gallic acid monohydrate shown on Figure 2a. Experimental structures (incl. structures differing in proton positions) are indicated in bold (Table S2 identifies the experimental structure).

Structure	Space group	a/Å	b/Å	c/Å	α/°	β/°	γ/°	E_{latt}/ kJ mol^{-1}	density/ g cm^{-3}	H-bonding[a]
H3_314	*P*-1	3.777	6.343	15.300	98.57	84.31	86.16	-236.55	1.738	D1
H2_246	*P*2$_1$/*c*	10.676	3.487	19.402	90	89.01	90	-236.17	1.730	C1
H2_104	***P*2$_1$/*c***	**9.598**	**3.461**	**21.363**	**90**	**88.61**	**90**	**-235.75**	**1.761**	**C1**
H2_1607	C2/c	19.433	3.563	20.866	90	86.98	90	-235.13	1.732	C1
H3_324	*P*2$_1$/*c*	3.700	30.334	6.453	90	87.37	90	-234.86	1.727	D1
H4_1904	***P*2$_1$/*c***	**7.569**	**3.815**	**25.019**	**90**	**98.64**	**90**	**-234.49**	**1.750**	**D1**
H4_1510	*P*-1	6.399	15.441	3.733	94.66	93.52	96.09	-232.67	1.713	D1
H1_1563	*P*2$_1$/*c*	5.049	14.618	10.064	90	87.45	90	-232.57	1.684	dx
H3_914	*P*2/*n*	30.296	3.642	6.644	90	84.60	90	-232.48	1.712	D2
H1_40	*P*-1	3.541	6.830	15.818	76.08	76.76	86.31	-232.35	1.729	D1
H3_933	***P*2$_1$/*c***	**6.378**	**3.972**	**35.269**	**90**	**57.17**	**90**	**-231.69**	**1.664**	**D1**
H4_423	*P*-1	10.924	6.926	16.136	115.02	41.35	78.34	-231.58	1.722	D1
H4_707	*P*2$_1$/*n*	6.429	29.835	3.705	90	86.89	90	-231.57	1.761	C5
H3_1046	*P*2$_1$/*c*	4.098	29.537	7.599	90	126.80	90	-231.51	1.697	D1
H4_458	*P*2$_1$/*c*	3.707	30.573	6.471	90	87.48	90	-231.26	1.706	D1
H1_2944	*P*2$_1$/*c*	15.437	3.483	14.122	90	102.04	90	-231.21	1.683	D2
H3_49	*P*-1	7.019	6.388	8.958	87.58	75.20	68.74	-230.98	1.729	D1
H2_465	***P*2$_1$/*c***	**9.376**	**3.584**	**21.748**	**90**	**99.45**	**90**	**-230.97**	**1.733**	**C1**
H1_2249	***P*2$_1$/*n***	**30.773**	**3.704**	**6.513**	**90**	**88.06**	**90**	**-230.94**	**1.684**	**D1**
H4_280	*P*-1	3.649	15.021	6.750	87.22	87.39	95.63	-230.89	1.701	D1
H2_538	I2/a	18.728	3.509	21.893	90	87.28	90	-230.86	1.739	C1
H1_1850	C2/c	10.730	7.532	22.699	90	127.05	90	-230.80	1.707	D1
H2_3283	*P*2$_1$/*c*	13.331	3.575	21.392	90	45.47	90	-230.74	1.719	C1
H2_4956	A2/n	20.294	3.644	19.580	90	87.18	90	-230.72	1.728	C1
H2_3508	*P*2$_1$/*c*	5.065	14.668	10.057	90	87.63	90	-230.67	1.674	dx
H4_1519	*Cc*	3.673	29.906	6.483	90	87.51	90	-230.34	1.757	C5
H1_1373	C2/c	31.389	3.465	14.113	90	75.26	90	-230.34	1.684	D1
H2_1987	***P*2$_1$/*c***	**9.292**	**3.591**	**21.909**	**90**	**93.06**	**90**	**-230.28**	**1.712**	**C1**
MH4	***P*-1**	**7.043**	**7.543**	**30.487**	**87.20**	**83.26**	**68.86**	**-230.24**	**1.687**	**D1**
H1_325	*P*2$_1$/*c*	3.878	30.613	7.139	90	60.97	90	-230.01	1.686	D1
H3_1273	*P*2$_1$/*c*	4.102	6.324	33.599	90	119.24	90	-229.76	1.643	D1
H1_21	Fdd2	30.603	27.050	3.614	90	90	90	-229.75	1.671	dx
H2_312	***P*2/*n***	**14.853**	**3.488**	**14.125**	**90**	**81.49**	**90**	**-229.70**	**1.727**	**C1**
H4_2142	*P*2$_1$/*c*	3.625	6.662	30.108	90	81.60	90	-229.44	1.737	C5
H2_239	*P*-1	6.882	16.018	3.531	105.45	94.00	77.59	-229.42	1.705	D1
H2_2871	*P*2$_1$/*n*	14.596	3.482	14.094	90	94.62	90	-229.36	1.750	C1
H3_66	*P*2$_1$2$_1$2$_1$	7.555	25.157	3.746	90	90	90	-229.23	1.755	C5
H2_76	*P*-1	3.491	10.680	9.844	86.80	96.79	97.93	-229.22	1.732	C1

Structure	Space group	Cell parameters						E_{latt}/ kJ mol^{-1}	density/ g cm^{-3}	H-bonding[a]
		a/Å	b/Å	c/Å	α/°	β/°	γ/°			
H4_321	*C*2/*c*	30.538	3.830	14.223	90	115.40	90	-229.19	1.663	D1
H1_592	*P*2₁/*c*	14.461	7.485	7.060	90	86.49	90	-229.14	1.638	D1
H1_1876	*A*2/*a*	14.481	3.487	30.395	90	78.82	90	-228.97	1.660	D1
H3_255	*P*2₁/*c*	3.768	6.526	29.673	90	92.22	90	-228.86	1.714	D1
H3_1687	*P*2₁/*n*	4.218	29.218	6.123	90	87.04	90	-228.72	1.658	D1
H2_6823	*C*2/*c*	31.885	3.471	14.087	90	74.06	90	-228.65	1.667	D1
H2_326	*P*2₁/*c*	9.957	3.697	19.638	90	95.19	90	-228.64	1.736	C1
H1_4516	*P*2₁/*c*	4.600	14.723	10.994	90	80.97	90	-228.62	1.670	dx
H3_115	*C*2/*c*	30.637	3.843	13.908	90	64.97	90	-228.57	1.685	D1
H2_413	*A*2/*n*	19.142	3.569	21.183	90	86.11	90	-228.54	1.731	C1
H2_3215	*P*2₁/*c*	15.782	3.597	13.347	90	99.03	90	-228.43	1.670	D2
H1_478	*P*2₁/*c*	15.660	3.576	13.511	90	79.75	90	-228.35	1.678	D2
H4_877	*P*2₁/*c*	6.599	3.690	30.329	90	79.69	90	-228.32	1.720	C6
H2_2237	C2/c	30.553	3.483	14.324	90	79.70	90	-228.32	1.666	D1
H4_19	*P*2₁/*c*	13.733	3.834	14.016	90	89.22	90	-228.32	1.693	C6
H4_948	*Pca*2₁	29.883	3.607	6.675	90	90	90	-228.31	1.737	C5
H4_1532	**_P_2₁/_c_**	**9.542**	**3.542**	**21.416**	**90**	**92.11**	**90**	**-228.20**	**1.728**	**C1**
H4_213	C2/c	25.663	4.005	15.498	90	111.58	90	-228.20	1.687	N1
H3_284	*P*-1	4.328	11.157	7.682	82.47	79.46	78.49	-228.17	1.757	D1
H1_667	*P*2₁2₁2₁	14.690	13.837	3.700	90	90	90	-227.98	1.661	C6
H4_836	*P*2/*c*	6.664	3.594	30.604	90	96.39	90	-227.98	1.715	D2
H1_259	*P*2₁/*c*	15.349	3.474	14.304	90	74.96	90	-227.96	1.697	C6
H3_610	*P*-1	3.648	43.065	7.538	119.11	63.75	59.78	-227.77	1.690	D1
H4_1468	*P*-1	7.608	11.490	4.324	79.29	78.13	92.80	-227.64	1.726	D1
H1_2304	*P*2₁2₁2₁	18.647	3.647	10.744	90	90	90	-227.63	1.710	C2
H3_1377	*P*-1	9.581	6.256	7.500	58.46	99.88	84.20	-227.58	1.670	D1
H2_1812	*Fdd*2	30.293	9.452	10.099	90	90	90	-227.54	1.728	dx
H3_4227	*P*2₁/*a*	8.621	11.799	7.294	90	90.09	90	-227.06	1.684	dx
H3_128	*P*2₁/*c*	3.605	29.486	7.849	90	120.33	90	-226.78	1.735	C5
H2_841	P2/c	15.333	3.580	13.815	90	74.43	90	-226.57	1.711	D1

[a]Classification according to Figure 2c.

The hypothetical anhydrate and monohydrate crystal structures are available in *.res from the authors on request.

2 EXPERIMENTAL

2.1 <u>Materials</u>

Gallic acid (purity $\geq$ 97.5%) was purchased from Sigma. The compound was recrystallized for purification from a hot saturated ethanol solution at 25 °C and desolvated in a drying oven at 80 °C. For the solvent screens a set of 29 solvents was chosen, which were all of analytical quality and all organic solvents were purchased from Aldrich or Fluka. The set of solvents was methanol, ethanol, 1-propanol, 2-propanol, 1-butanol, 2-butanol, toluene, xylene, hexane, cyclohexane, dichloromethane, dichloroethane, chloroform, ethyl methyl ketone, acetone, diethyl ether, diisopropyl ether, cyclohexanone, 1,4-dioxane, tetrahydrofurane, ethyl acetate, acetic acid, formic acid, acetonitrile, nitromethane, pyridine, dimethyl formamide, dimethyl sulfoxide and water.

2.2 Solid form screen

2.2.1 Solvent Screen

Crystallization experiments included fast and slow solvent evaporation, cooling crystallization and solvent-mediated transformation. Mixed solvents (2nd solvent) contained a 1:1 ratio of the two solvents. In total more than 165 solvent crystallization conditions were explored mainly at room temperature (RT). The crystallization products were identified with hot-stage microscopy, IR spectroscopy and powder X-ray diffractometry.

Table S6. Results of gallic acid solid form screen: solvent mediated transformation experiments[a] (II° - anhydrate form II°, MH-I° – monohydrate form I°, S-Dx2 – dioxane disolvate, S-Dx-H dioxane-water solvate).

Solvent	Solid Form	Solvent (mixture)	Solid Form
1-Butanol	II°		
1-Propanol	II°	1-propanol+water	MH-I°
2-Butanol	II°		
2-Propanol	II°	2-propanol+water	MH-I°
Acetic acid	II°	acetic acid+Water	MH-I°
Acetone	II°	acetone+water	MH-I°
Acetonitrile	II°	acetonitrile+water	MH-I°
Cyclohexane	II°		
Chloroform	II°		
Dichloroethane	II°		
Dichloromethane	II°		
Diethyl ether	II°		
Dimethyl formamide	Solvates[b] (DMF1 > DMF2, DMF4)	dimethyl formamide+water	Solvates[b] (DMF2, DMF3, DMF4)
Dimethyl sulfoxide	Solvates[b] (DMSO1, DMSO3 DMSO4)	dimethyl sulfoxide+water	Solvates[b] (DMSO3, DMSO4)
Diisopropyl ether	II°		
1,4-Dioxane	S-Dx2	1,4-dioxane+water	S-Dx-H
Ethyl methyl ketone	II°		
Ethyl acetate	II°		
Ethanol	II°	ethanol+water	MH-I°
Formic acid	II°		
Hexane	II°		
Methanol	II°	methanol+water	MH-I°
Nitromethane	II°		
Pyridine	II°		
Tetrahydrofurane (THF)	II°	THF+water	MH-I°
Toluene	II°		
Water	MH-I°		
Xylene	II°		

[a]Anhydrous gallic acid and few drops of solvent were ground in a Retsch grinding mill MM 200 for 10 minutes. [b]See Table S10.

Table S7. Results of gallic acid solid form screen: fast evaporation experiments[a] (S$_{X-H}$ – isostructural heterosolvates containing one mole water per mole gallic acid, MH – monohydrates, S-DMF and S-DMSO – solvates containing dimethyl formamide or dimethylsulfoxide, II° – anhydrate form II°, III – anhydrate form III, THF – tetrahydrofurane).

Solvent	Solid Form[b]	2nd solvent water Solid Form	2nd solvent THF Solid Form	2nd solvent toluene Solid Form
1-Butanol	S$_{X-H}$ (MH-III) MH-IV	n.a.	MH-I°	n.a.
1-Propanol	S$_{X-H}$ (MH-III) MH-IV MH-I°	MH-II	n.a.	MH-III MH-V
2-Butanol	MH-V S$_{X-H}$ (MH-III) MH-IV	n.a.	MH-III MH-V	n.a.
2-Propanol	S$_{X-H}$ (MH-III) MH-IV MH-I°	MH-I°	n.a.	n.a.
Acetic acid	S$_{X-H}$ (MH-III) MH-I°	S$_{X-H}$ MH-I°	n.a.	n.a.
Acetone	S$_{X-H}$ (MH-III) MH-II MH-I°	MH-II MH-I°	n.a.	n.a.
Acetonitrile	S$_{X-H}$ MH-I°	MH-II MH-I°	n.a.	n.a.
Dimethyl formamide	S-DMF1 S-DMF2	S-DMF1 S-DMF2 S-DMF1 (85)[b] S-DMF4 (85)[b]	n.a.	n.a.
Dimethyl sulfoxide	S-DMSO2	S-DMSO4 S-DMSO2	n.a.	n.a.
1,4-Dioxane	S$_{X-H}$	S$_{Dx-0.5}$ MH-I°	n.a.	n.a.
Ethyl methyl ketone	MH-IV S$_{X-H}$	n.a.	S$_{X-H}$	n.a.
Ethyl acetate	MH-I° S$_{X-H}$	n.a.	n.a.	n.a.
Ethanol	S$_{X-H}$ (MH-III)	S$_{X-H}$ MH-I° MH-II MH-III	MH-IV	S$_{X-H}$ MH-IV II° (85)[b] III (85)[b]
Formic acid	MH-I°	n.a.	n.a.	n.a.
Methanol	S$_{X-H}$ (MH-III) MH-I° MH-II MH-III	MH-I° MH-II	n.a.	S$_{X-H}$ MH-IV MH-II MH-I°
Tetrahydrofurane	S$_{X-H}$ MH-IV	n.a.	n.a.	n.a.
Water	MH-II (MH-I°)	MH-II (MH-I°)	MH-I° MH-II	n.a.
Toluene	Not soluble	n.a.	II°	n.a.

[a]A saturated solution (at RT) of anhydrous gallic acid was filtered and the solvent was evaporated from a watch glass at RT. [b]Evaporation temperature (° C) as stated in parenthesis, desolvation/dehydration product; [b]see Table S10; n.a. – not attempted.

Table S8. Results of gallic acid solid form screen: slow evaporation experiments[a] (S_{X-H} – isostructural heterosolvates containing one mole water per mole gallic acid, MH – monohydrates, S-DMF and S-DMSO – solvates containing dimethyl formamide or dimethylsulfoxide, II° – anhydrate form II°, III – anhydrate form III, THF – tetrahydrofurane).

Solvent	Solid Form[b]	2nd solvent water Solid Form	2nd solvent THF Solid Form	2nd solvent toluene Solid Form
1-Butanol	MH-IV	n.a.	MH-IV	II° MH-IV, S_{X-H}
1-Propanol	S_{X-H}	MH-II	n.a.	MH-I°, MH-IV
2-Butanol	MH-III (S_{X-H})	n.a.	II°	MH-III II°
2-Propanol	MH-I°, MH-IV	MH-I°	n.a.	II°
Acetic acid	MH-I°	II°	MH-I°, MH-II	MH-I°, MH-II
Acetone	S_{X-H} (MH-III) MH-I°	n.a.	n.a.	II° S_{X-H}
Acetonitrile	MH-I° MH-IV	MH-II MH-I°	II° S_{Dx-2}, MH-I°	II° $S_{X-H,}$ MH-III
Chloroform	n.a.	n.a.	II°	II° S_{X-H}
Cyclohexane	n.a.	n.a.	II°	
Dichloroethane	n.a.	n.a.	II°	II°
Dichloromethane	n.a.	n.a.	II° (III)	MH-I° S_{X-H}
Diethyl ether	n.a.	n.a.		II°
Diisopropyl ether	n.a.	n.a.	II° S_{Dx-2}	II°
Dimethyl formamide	S-DMF 4	n.a.	n.a.	n.a.
Dimethyl sulfoxide	DMSO1, DMSO2	n.a.	n.a.	n.a.
1,4-Dioxane	S_{Dx-2}, $SD_{x-0.5}$ S_{X-H}	MH-I° MH-III	n.a.	S_{Dx-2}, $S_{Dx-0.5}$ S_{X-H}
Ethyl methyl ketone	S_{X-H} (MH-III) II° MH-IV	n.a.	II° S_{Dx-2}	II° S_{Dx-2}
Ethyl acetate	MH-IV MH-II	n.a.	II°	II° MH-III
Ethanol	S_{X-H} (MH-III) MH-I°	MH-I° MH-II	n.a.	II°, S_{X-H} MH-III, MH-I°
Hexane	n.a.	n.a.	S_{Dx-2} (MH-III)	II°
Methanol	MH-I, MH-II MH-IV, S_{X-H} II°	n.a.	MH-I°	MH-I°, MH-II S_{X-H} II°
Nitromethane	n.a.	n.a.	II°	n.a.
Tetrahydrofurane	S_{X-H}	n.a.	n.a.	MH-I°, MH-IV
Pyridine	salt	n.a.	n.a.	MH-III
Water	MH-I°, MH-II	n.a.	n.a.	n.a.
Toluene	n.a.	n.a.	II° MH-IV	n.a.
Xylene	n.a.	n.a.	II° MH-I°	II°

[a]A saturated solution (at RT) of anhydrous gallic acid was filtered and the solvent was evaporated from an open vial at RT; [b]See Table S10; n.a. – not attempted.

Table S9. Results of gallic acid solid form screen: crystallization experiments[a] (II° – anhydrate form II°, MH-I° – monohydrate form I°, S_{AA} – acetic acid solvate, S_{FA} – formic acid solvate, S_{DX} – dioxane solvate, S_{X-H} – isostructural heterosolvates containing 1 mole water per mole gallic acid).

Solvent	Method	Solid form[b]
1-Butanol	F/S	II°
2-Butanol	F/S	II°
1-Propanol	F	S_{X-H} (MH-III)
	S	MH-I°
2-Propanol	F/S	II°
Ethanol	F/S	MH-I°
Methanol	F/S	MH-I°
Acetic acid	F/S	S_{AA}
Acetone	F	S_{X-H} (MH-III), MH-I°
	S	MH-I°
Acetonitrile	F/S	II°
1,4-Dioxane	F/S	S_{DX2} ($S_{DX0.5}$)
Ethyl methyl ketone	F/S	II°
Ethyl acetate	F/S	II°
Fomic acid	F	S_{FA}
Tetrahydrofurane	F/S	MH-I°
Water	F/S	MH-I°

[a]A hot saturated solution (close to the boiling point of each solvent used) was either cooled fast (F, in ice) or slow (S, test tube wrapped in aluminium foil) to 0° or RT, respectively. [b]See Table S10.

2.2.2   Sublimation experiments

Initial experiments were carried out under a Reichert polarization microscope (Vienna, A) equipped with a Kofler hot-stage. Anhydrous gallic acid was placed on a microscopic glass slide, covered with a cover slip and heated. Strong sublimation of the acid was observed at 210 °C. All three anhydrous polymorphs were found. **AH-I** seed crystals (in addition to **AH-II°**) could be harvested from sublimation experiments. For details see section 2.4.

Based on the initial results additional sublimation experiments were conducted. Anhydrous gallic acid was placed between two microscopic slides, which were separated by a glass ring (height approx. 2 mm) and stored on a preheated metal block (240 °C). Anhydrates **AH-I** and **AH-II°** were obtained. The sample had no contact with the top microscopic slide. For more details see section 2.4.

2.2.3 Desolvation studies (preparation of anhydrates I and III):

Monohydrates and solvates were desolvated at different conditions:

a. 0% RH at 25 °C

Selected solvate/hydrate samples were stored in a desiccator over silica-gel for six months, resulting predominantly in **AH-II°** (traces of **AH-I** were identified in some samples).

b. Moisture sorption/desorption analysis (see section 2.6)

c. Thermal desolvation at higher temperatures (at ambient moisture conditions)

Solvate/hydrate phases were stored at temperatures ranging 110 to 125 °C. Desolvation resulted predominantly in **AH-II°**.

d. Thermal desolvation upon slowly heating (at ambient moisture conditions)

Solvate/hydrate samples were slowly heated (on a Reichert polarized microscope equipped with a Kofler hot-stage, heating rate ca. 5 °C min^{-1}) until a desolvation process was observed (i.e. darkening of the crystals. Temperature was then held until desolvation was completed. Several of the solvates/hydrates resulted in products with **AH-I** and **AH-III** being the major phase: Desolvating **MH-I°**, **MH-II**, **S-DMF4** (dimethyl formamide monosovlate) or **S-AA** (acetic acid monosolvate) were found to be the most promising for producing **AH-I**, and desolvating **S-DMF2** (dimethyl formamide water heterosolvate) the most promising for producing **AH-III**.

For the dioxane di-solvate (**S-Dx2**) a two-step desolvation process was observed, i.e. desolvation to a lower stoichiometric solvate (dioxane hemisolvate, **S-Dx0.5**) and in a second step desolvation to **AH-I** and **AH-II°**.

The desolvation experiments were repeated using a transmission infrared microscope (Bruker IFS 25 spectrometer connected to a Bruker microscope I with a 15x-Cassegrain-objective, Bruker Analytische Messtechnik GmbH, Ettlingen, D), equipped with a heatable accessory holder. Freshly prepared crystals of the solvates/hydrates were pressed onto a ZnSe disc and heated in 2 degree steps until desolvation was observed (< 80 °C). Every 2 degrees a spectrum was recorded, i.e. temperature held for 20 seconds (spectral range 4000 to 600 cm^{-1}, resolution 4 cm^{-1}, 32 interferograms per spectrum). Selected spectra are given in Figure S11.
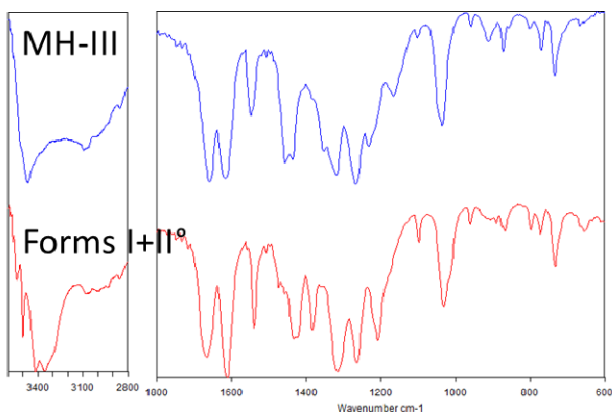
MH-I° → AH-I
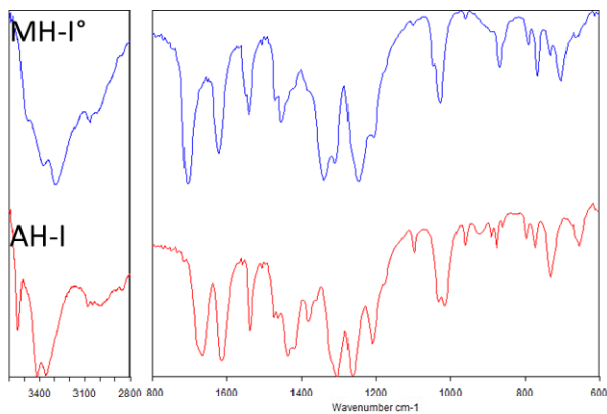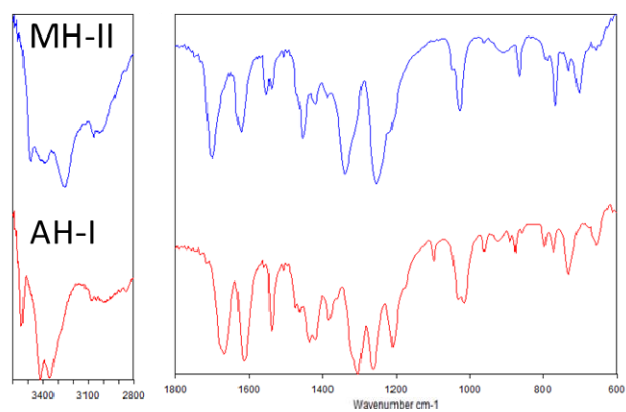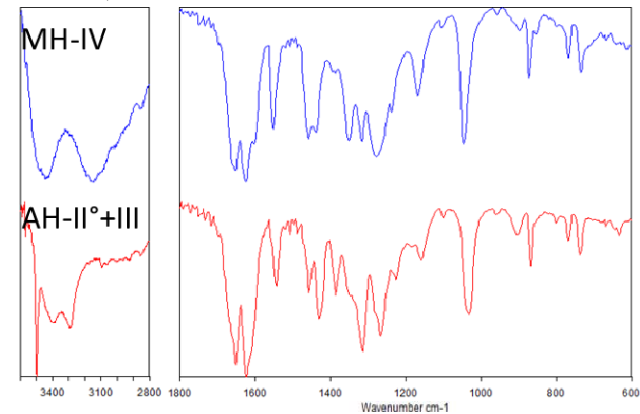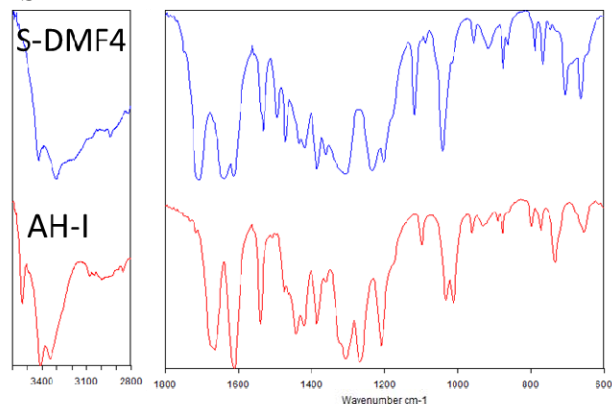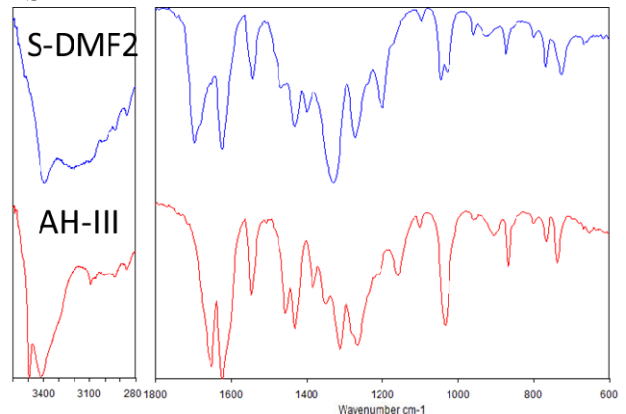

MH-II → AH-I

MH-III → AH-I + AH-II°

MH-IV → AH-II° + AH-III

S-DMF4 → AH-I

S-DMF2 → AH-III

Figure S11. Fourier Transform – Infrared spectra of heating cycles showing the starting (in blue) and desolvation product (in red) for selected gallic acid solvates (S-DMF2 and S-DMF4)/hydrates (**MH-I° - IV**).

2.2.4 Summary of experimental screen

The experimental solid form screen resulted in three anhydrate polymorphs (**AH-I**, **AH-II°** and **AH-III**), five monohydrates (**MH-I°**, **MH-II**, **MH-III**, **MH-IV** and **MH-V**) and 22 novel solvates. Solvates (Table S10) were obtained from acetic acid, formic acid, dioxane, dimethyl formamide, dimethyl sulfoxide, lower alcohols, acetone, ethyl methyl ketone and tetrahydrofurane. The majority of the solvates were mixed solvates with water (heterosolvates), with ten being isostructural heterosolvates.

Table S10. Gallic acid solvates emerging from the solvent screen.

Type	Solvate	GallicAcid:Water:Solvent ratio[a]
Homo-solvates	Acetic acid (S_{AA})	1:0:1
	Dioxane 0.5 ($S_{Dx-0.5}$)	1:0:0.5
	Dioxane 2 (S_{Dx-2})	1:0:2
	Dimethyl formamide (S_{DMF4})	1:0:1
	Formic acid (S_{FA})	1: 0: n.d.
Hetero-solvates with Water	Dioxane (S_{Dx-H})	1:1:~0.25
	Dimethyl formamide (S_{DMF3})	1:1:~0.75
	Dimethyl sulfoxide (S_{DMSO3})	1:1:0.5
	Dimethyl sulfoxide (S_{DMSO4})	1:0.5:0.5
Isostructural Solvates[b] (Hetero-solvates with water).	Dimethyl formamide (S_{DMF2}), Dimethyl sulfoxide (S_{DMSO1}), Methanol, 1-Propanol, 2-Propanol, 1-Butanol, Acetone, Ethyl methyl ketone, Tetrahydrofurane, Dioxane (S_{X-H})	1:1:x (depending on solvent)
Additional Solvates (nor characterized in detail)	Dimethyl sulfoxide (S_{DMSO2})	n.d.
	Dimethyl sulfoxide (S_{DMSO5})	n.d.
	Dimethyl formamide (S_{DMF})	n.d. (contains water)

[a]Ratio determined by Coulometric Karl-Fischer-Titration and Thermogravimetric Analysis.
[b]Classified as isostructural based on IR and PXRD measurements. n.d. – not determined.

2.3 Identification of gallic acid solid forms

2.3.1 Vibrational Spectroscopy

Infrared spectra were recorded with a diamond ATR crystal on a Perkin Elmer Spectrum One Fourier Transform spectrophotometer (Perkin Elmer, Norwalk Ct., USA) over a range of 4000 to 650 cm^{-1} with a resolution of 2 cm^{-1} (24 scans). The spectra were analysed with the Opus v 5.5 software.

IR spectroscopy (Figures S12 – S18) was used to identify the gallic acid solid forms and to get some insight into the structural features of the forms. The absence of bands in the spectral range of 1730 to 1705 cm^{-1} indicates that the carboxylic acid is dimerized.[18] The most characteristic region for discriminating the forms is the region of the ν(O-H) vibrations.
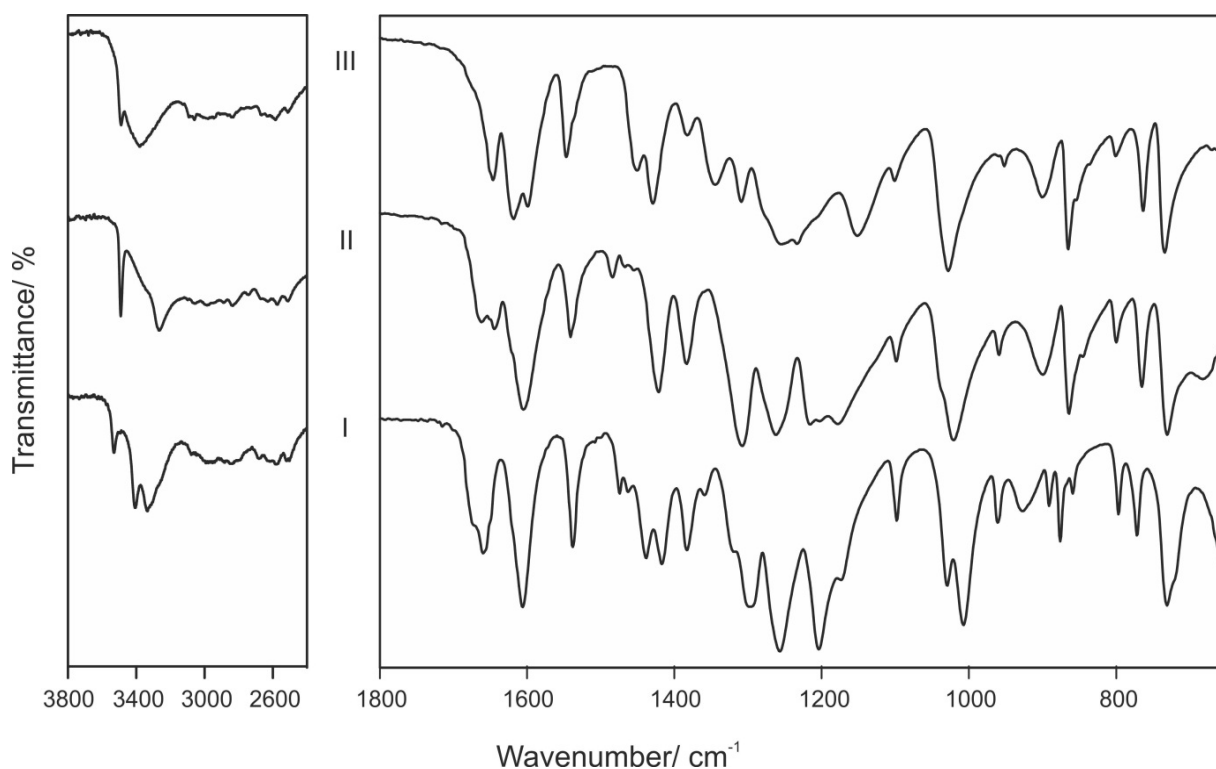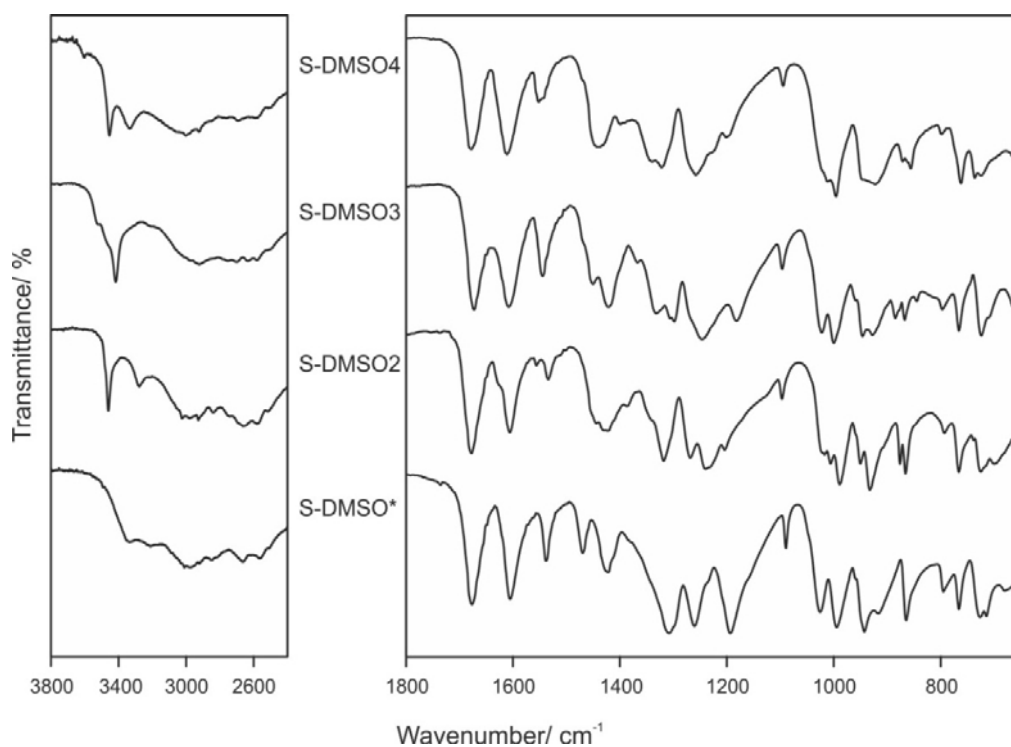

Figure S12. FT-IR spectra of gallic acid anhydrates (I – **AH-I**, II – **AH-II°** and III – **AH-III**).
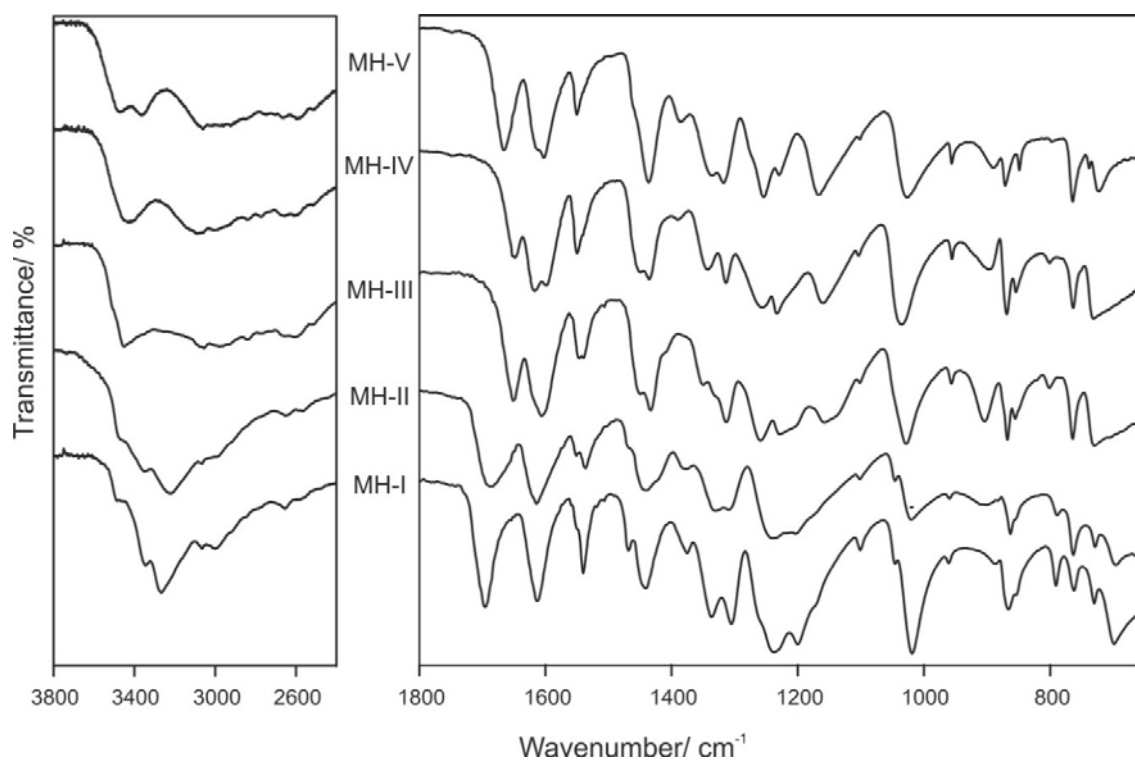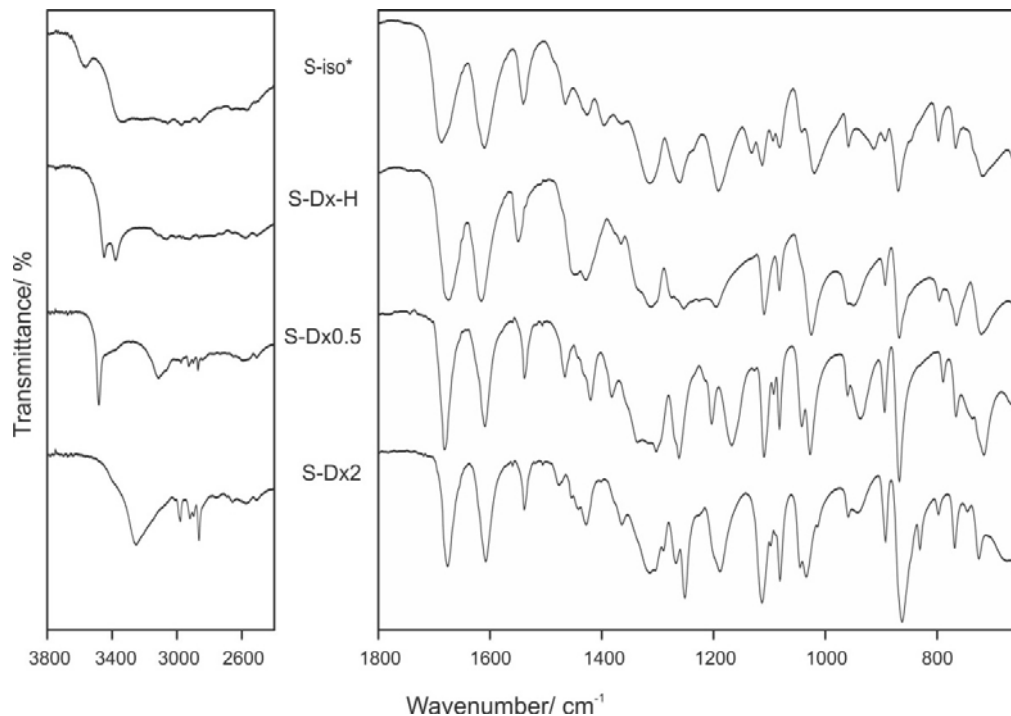
**Figure S13.** FT-IR spectra of gallic aicd monohydrates.
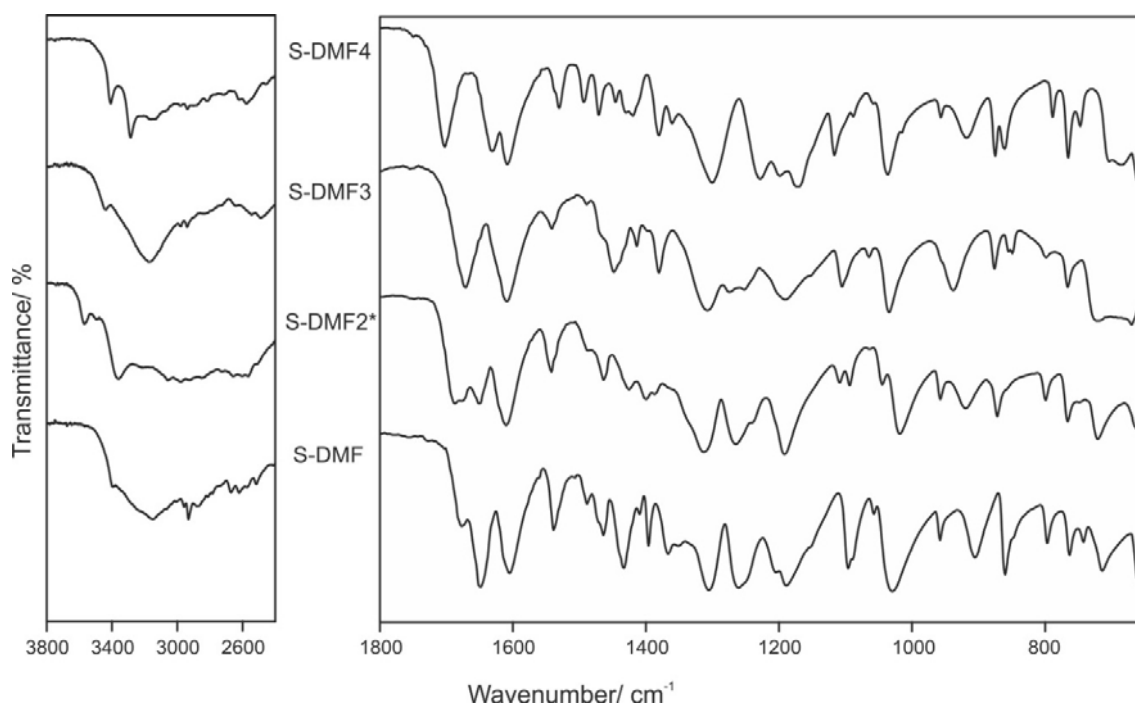
Figure S14. FT-IR spectra of gallic acid solvates containing the solvent dimethyl formamide (see Table S10). Spectrum marked with asterisk (*): isostructural to solvates presented in Figure S18.
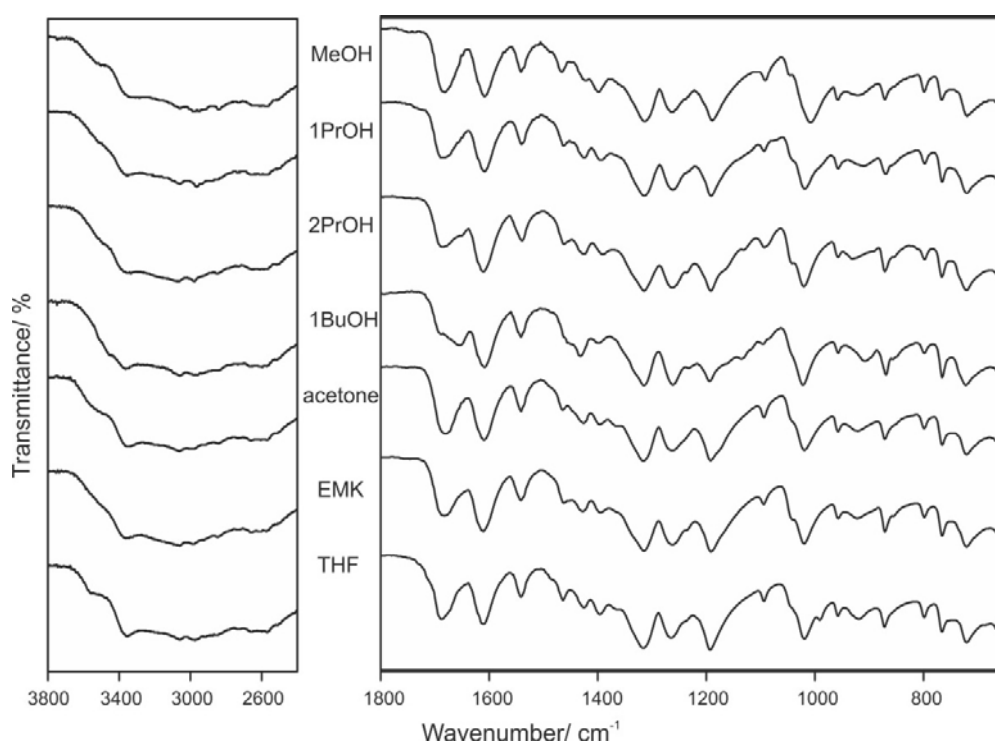
Figure S15. FT-IR spectra of gallic acid solvates containing the solvent dimethyl sulfoxide (see Table S10). Spectrum marked with asterisk (*): isostructural to solvates presented in Figure S18.

Figure S16. FT-IR spectra of gallic acid solvates containing the solvent dioxane (see Table S10). Spectrum marked with asterisk (*): isostructural to solvates presented in Figure S18.
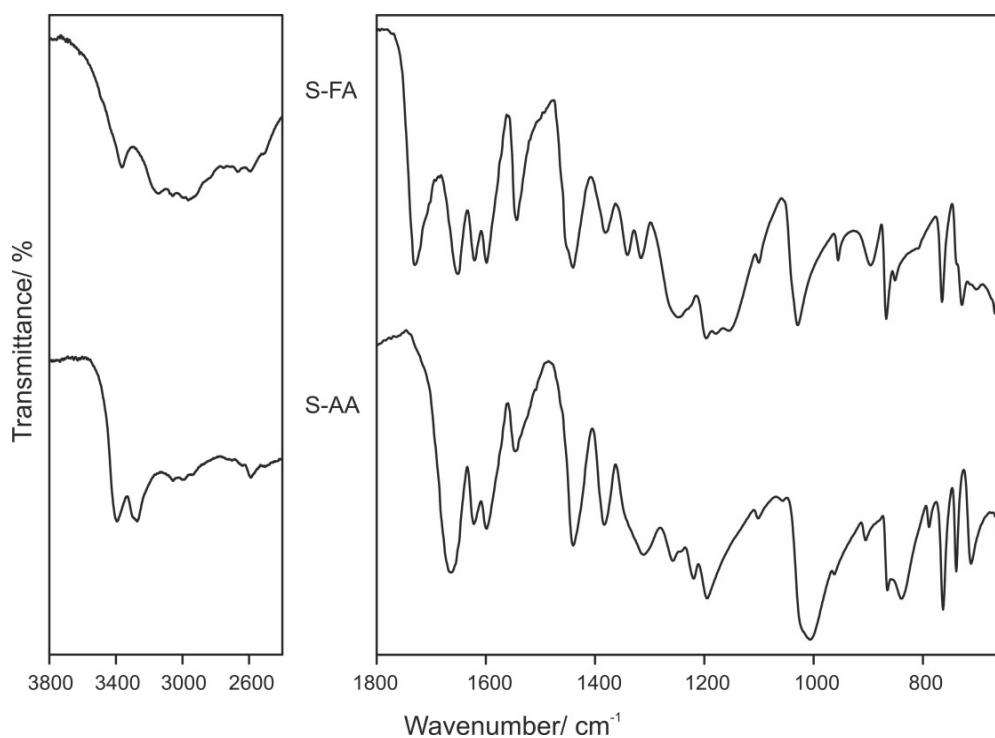
**Figure S17.** FT-IR spectra of gallic acid formic acid (FA) and acetic acid (AA) solvates.

Figure S18. FT-IR spectra of isostructural gallic acid hetero-solvates with water and organic solvent (MeOH − methanol, 1PrOH, 1-propanol, 2PrOH − 2-propanol, 1BuOH − 1-butanol, EMK − ethyl methyl ketone, THF − tetrahydrofurane).

2.3.2 Powder X-ray diffraction

The powder X-ray diffraction patterns used for phase identification were obtained using a X'Pert PRO diffractometer (PANalytical, Almelo, The Netherlands) equipped with a theta/theta coupled goniometer in transmission geometry, programmable XYZ stage with well plate holder, 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, tube current of 40 mA, applying a step size of $2\theta = 0.013°$ with 40 s per step in the 2θ range between 2° and 40°.

Powder pattern for the hydrates, contrasted to the simulated patterns from the experimental crystal structures are given in Figure S19.
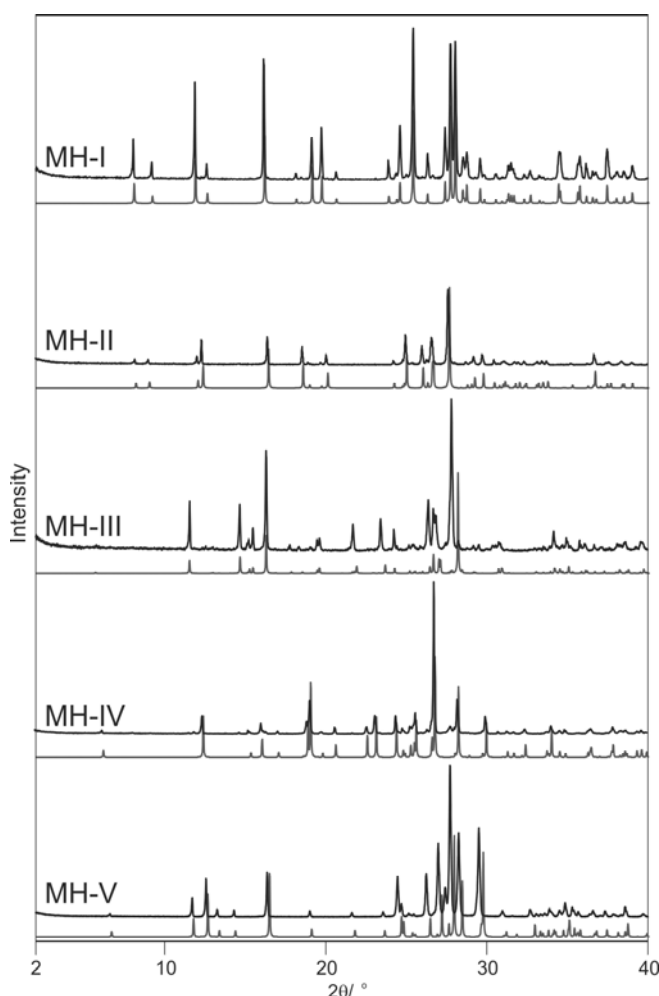
Figure S19. X-ray powder diffraction patterns of gallic acid monohydrates. The experimental diffractograms are contrasted with the calculated powder patterns (**MH-I**°: KONTIQ01, **MH-II**: KONTIQ03, **MH-III**: KONTIQ04, **MH-IV**: KONTIQ and **MH-V**: own structure solution from PXRD). Peak shifts arise from different measurement temperatures.

2.4 X-ray Crystallography (details for structure solutions)

Crystal structures were solved for all three anhydrates (**AH-II**° from single crystal data, **AH-I** from powder X-ray diffraction data and **AH-III** using the predicted structure as a starting point for Rietveld refinement) and the fifth monohydrate structure (from powder data).

2.4.1 Anhydrate form II° (**AH-II**°)

No structure for the anhydrate (AH-II°) had been published when we were working on the compound. After finishing the work and during preparing the manuscript three determinations of **AH-II**° have been reported (CSD: IJUMEG[19-21] ref-code family). The structural details of all four **AH-II**° structure solutions are in agreement.

Single crystal X-ray diffraction data were collected on a Bruker AXS SMART APEX CCD diffractometer using graphite monochromated Mo-Ka radiaton (λ = 0.71073 Å, ω scans). Unit cell and data sets were collected at 20 °C. Data integration and final unit cell parameters were obtained using the program SAINT+.[22] Absorption correction was applied using the program SADABS.[23] The structure was solved using SIR04[24] and refined with SHELXL-97.[25] All hydrogen atoms bonded to carbon atoms were generated by a riding model on idealized geometries with $U_{iso}(H)=1.2U_{eq}(C)$. The polar hydrogen atoms were identified form the difference map and refined isotropically, with the exception of H2 and H3, where the position was refined with a constrained O−H bond distance.

Crystal data of **AH-II**°: $C_7H_6O_5$, M_r = 170.12, monoclinic, space group $C2/c$, T = 293(2) K, size [mm]: 0.35 × 0.25 × 0.2, a = 25.685(3) Å, b = 4.9273(5) Å, c = 11.2508(12) Å, β = 106.234(2) °, V = 1367.1 (2) Å^3, Z = 8, ρ_{calc} = 1.653 g cm^{-3}, 4765 reflections measured, 1315 independent reflections, 1023 observed reflections, θ range for data collection: 3.3 − 26.0 h, k, l, range: −31<h<31, −6<k<6, −13<l <13, data: 732, restraints: 2, parameters: 124, $R[F^2> 2\sigma(F^2)]$= 0.036, $wR(F^2)$ = 0.095, R int = 0.028, Goodness of fit on F^2 = 1.11, $\Delta\rho_{max}$ = 0.18, $\Delta\rho_{min}$ (e Å^{-3}) = −0.21.
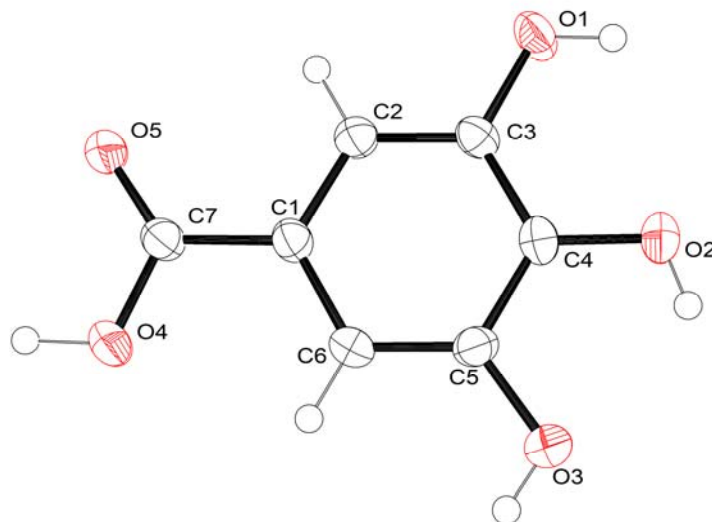
Figure S20. Thermal ellipsoid plot of gallic acid **AH-II**°. Displacement ellipsoids are drawn at the 50% probability level.
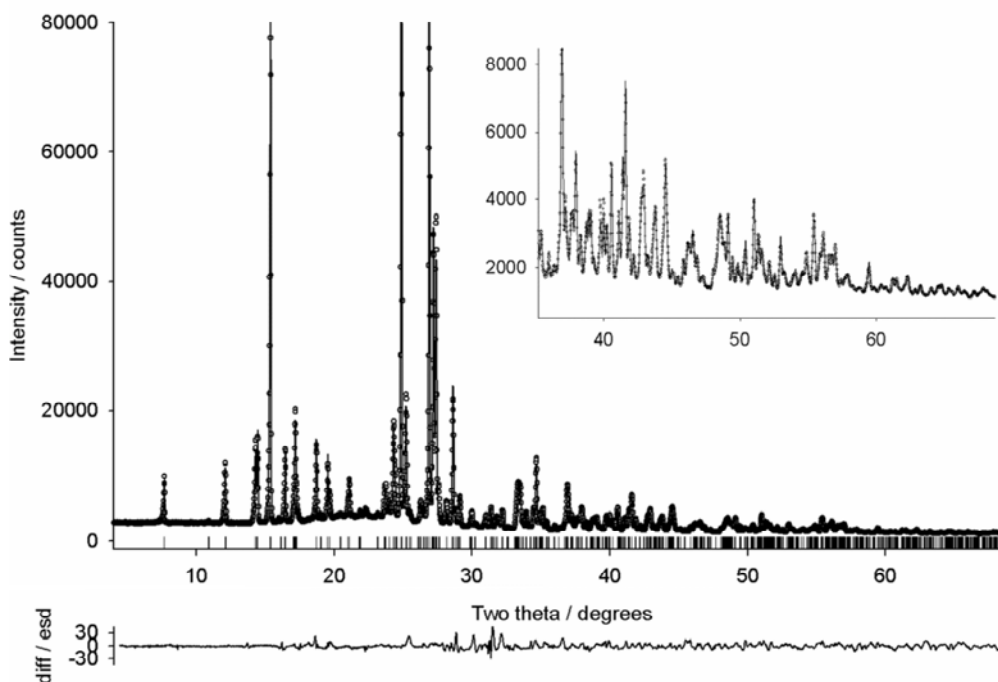
2.4.2 Anhydrate form I (**AH-I**)

Gallic acid **AH-I** was loaded in a rotating 0.7 mm borosilicate capillary and mounted on a Bruker AXS D8 powder X-ray diffractometer equipped with primary monochromator (CuKα_1, λ = 1.54056 Å) and Lynxeye position sensitive detector. Data was collected at room temperature using a variable count time scheme[26,27] (Table S11).

Table S11. Details of instrument and data collection parameters.

Instrument name	Bruker D8			
Radiation	Cu K$_{\alpha 1}$, λ=1.54056Å			
Generator settings	40kV/50mA			
Monochromator	Primary, focusing curved Ge 111			
Geometry	Transmission capillary geometry			
Detector	Bruker Lynxeye PSD			
Scan Type	Fixed		VCT	
Step size (°2θ)	0.016		0.016	
Data range (°2θ) and step time	2-40	8s	3-22	2s
			22-40	4s
			40-55	15s
			55-70	24s

The diffraction pattern was indexed to a triclinic unit cell using the first twenty peaks with DICVOL04 and the space group was determined to be *P*-1 based on statistical assessment of systematic absences,[28] as implemented in DASH structure solution package.[29] The data was background subtracted and Pawley refinement[30] was used to extract the intensities and their correlations. Simulated annealing was used to optimize the gallic acid model against the diffraction data set (93 reflections) in direct space. The internal coordinate (Z-matrix) description was derived from the PBE0/6-31G(d,p) gas phase global conformational minimum (conf2, Table S1), with O–H distances normalized to 0.9 Å and C–H distances to 0.95 Å. The choice of conformation (polar proton positions) was based on DFT-D calculations (Table S2). The structure was solved using 800 simulated annealing runs of 2.5 x 10^7 moves per run as implemented in DASH. Each gallic acid molecule was allowed seven (six external and one internal) degrees of freedom. The best solution returned a χ^2 ratio of ca. 2.79 (profile χ^2/pawley χ^2) and was used as the starting point for a rigid body Rietveld refinement[31] in TOPAS V4.1.[32] The rigid body description was derived from the Z-matrix used in the simulated annealing runs and the final refinement included a total of 87 parameters (53 profile, 14 preferred orientation, 6 cell, 1 scale, 1 isotropic temperature factor, 6 position and 6 rotation) yielding a final R$_{wp}$= 5.904 (Figure S21).

Figure S21. Final observed (points), calculated (line) and difference [(yobs - ycalc)/ σ(yobs)] profiles for the Rietveld refinement of gallic acid anhydrate form I.
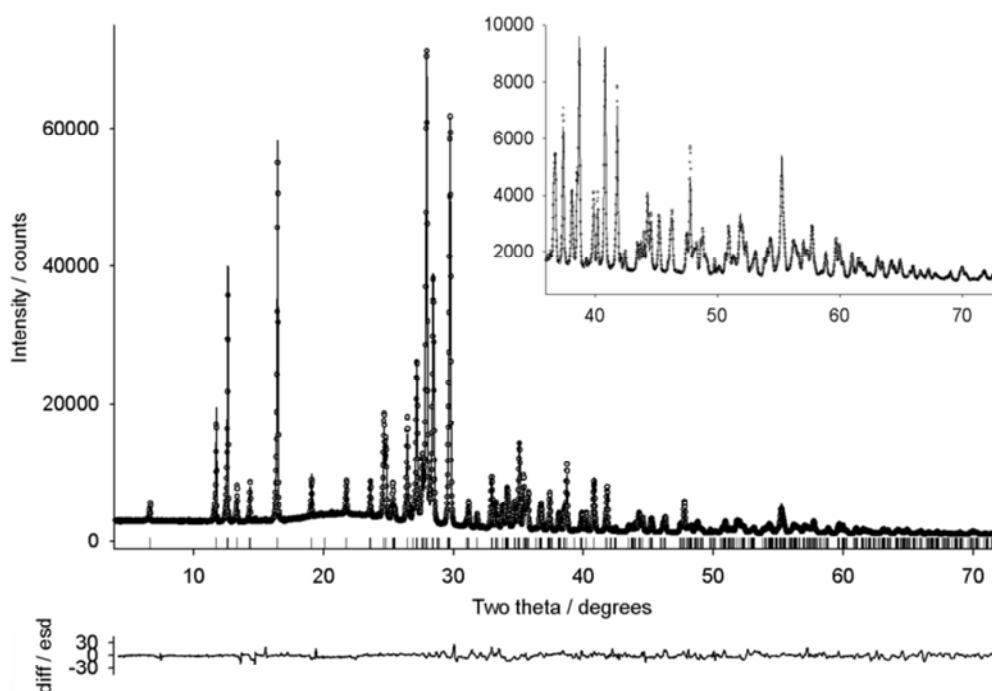
Crystal data of **AH-I**: $C_7H_6O_5$, Mr=170.12, triclinic, space group *P*-1, T = 25 °C, sample formulation: powder, specimen shape: 12 x 1.0 x 0.7 mm, wavelength: 1.54056 Å, a = 7.3183 (2) Å, b = 8.2536(2) Å, c = 11.7148(3) Å , α = 100.472(1) °, β = 90.234(2) °, γ = 90.984(2) ° V = 695.73(3) Å3, Z = 4, density = 1.624 g cm^{-3}, 2 theta range for data collection : 3 to 70°, background treatment: Chebyshev polynomial, No. of measured reflections: 580, Refinement method: Rietveld, data/parameter/restraints: 580/87/0, goodness of fit: 5.227 (on Y_{obs}), R_{wp} = 5.904, R_{exp} = 1.130 , R_p = 4.818.

The resulting structure from Rietveld refinement was further scrutinized by allowing all fractional coordinates to refine freely (168 parameters, R_{wp} = 4.792). As expected, the improve R_{wp} came at the expense of some chemical sense (e.g. slight distortion in planarity of benzene ring, movement of H atoms to nonsensical positions), but otherwise, the geometry of the independent molecules was well preserved, confirming the correctness of the constrained refined crystal structure.

2.4.3   Monohydrate form V (**MH-V**)

Gallic acid **MH-V** was loaded in a rotating 0.7mm borosilicate capillary and mounted on a Bruker AXS D8 powder X-ray diffractometer equipped with primary monochromator (CuKα_1, λ = 1.54056 Å) and Lynxeye position sensitive detector. Data was collected at 150K using a variable count time scheme[26,27] (Table S11). The diffraction pattern was indexed to a monoclinic unit cell using the first twenty peaks with DICVOL04 and the space group was determined to be P2$_1$/c based on statistical assessment of systematic absences,[28] as implemented in DASH structure solution package.[29] The data was background subtracted and Pawley refinement[30] was used to extract the intensities and their correlations. Simulated annealing was used to optimize the gallic acid and water model against the diffraction data set (61 reflections) in direct space. The internal coordinate (Z-matrix) description was derived from conf 4 (Table S1), with O–H distances normalized to 0.9 Å and C–H distances to 0.95 Å. The choice of conformation (polar proton positions) was based on the computed structures (section 1.3). The structure was solved using 800 simulated annealing runs of 2.5 x 10^7 moves per run as implemented in DASH. Each gallic acid and water molecule was allowed seven (6 external and 1 internal) and six (external) degrees of freedom respectively. The best solution returned a χ^2 ratio of ca. 3.26 (profile χ^2/pawley χ^2) and was used as the starting point for a rigid body Rietveld refinement[31] in TOPAS V4.1.[32] The rigid body description was derived from the Z-matrix used in the simulated annealing runs and the final refinement included a total of 71 parameters (44 profile, 9 preferred orientation, 4 cell, 1 scale, 1 isotropic temperature factor, 6 position and 6 rotation) yielding a final R$_{wp}$= 4.909 (Figure S22).

**Figure S22**. Final observed (points), calculated (line) and difference [(yobs - ycalc)/ σ(yobs)] profiles for the Rietveld refinement of gallic acid monohydrate V.

Crystal data of **MH-V**: $C_7H_6O_5\cdot H_2O$, Mr=188.13, monoclinic, space group: $P2_1$/c, T = 123 °C, sample formulation: powder, specimen shape: 12 x 1.0 x 0.7 mm, wavelength: 1.54056 Å, a = 7.60719(9) Å, b = 3.64133(4) Å, c = 26.7915(4) Å, β = 98.421(1) °, V = 734.13(2)Å^3, Z = 4, Z = 1, density = 1.702 g cm^{-3}, 2 theta range for data collection: 3 to 70°, background treatment: Chebyshev polynomial, No. of measured reflections: 358, Refinement method: Rietveld, data/parameter/restraints: 358/71/0, goodness of fit: 4.298 (on Y_{obs}), R_{wp} = 4.909, R_{exp} = 1.142, R_p = 4.179.

The resulting structure from Rietveld refinement was further scrutinized by allowing all fractional coordinates to refine freely (86 parameters, R_{wp} = 4.1). As expected, the improve R_{wp} came at the expense of some chemical sense (e.g. slight distortion in planarity of benzene ring, movement of H atoms to nonsensical positions), but otherwise, the geometry of the independent molecules was well preserved, confirming the correctness of the constrained refined crystal structure.

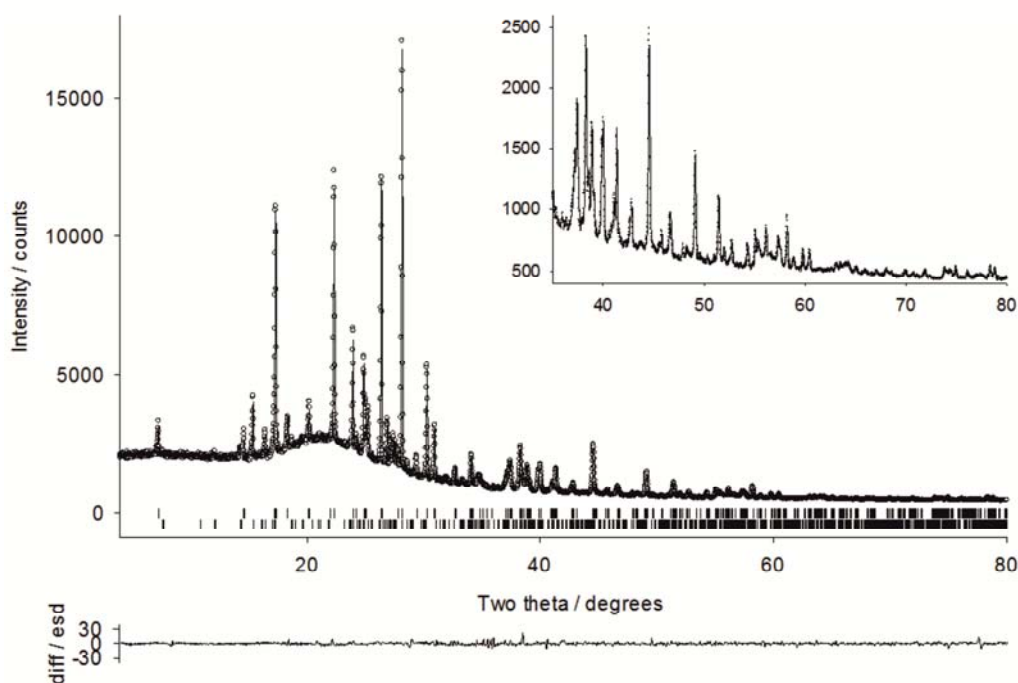### 2.4.4 Anhydrate form III (**AH-III**)

The IR spectrum of gallic acid **AH-III** exhibits a close resemblance to the IR spectra obtained for the three ladder based hydrate structures (D1 structures: **MH-III**, **MH-IV** and **MH-V**, Figure 2b), exemplarily shown in Figure S23 for **AH-III** and **MH-IV**. Based on this information the presence of a ladder based D1 structure was expected.
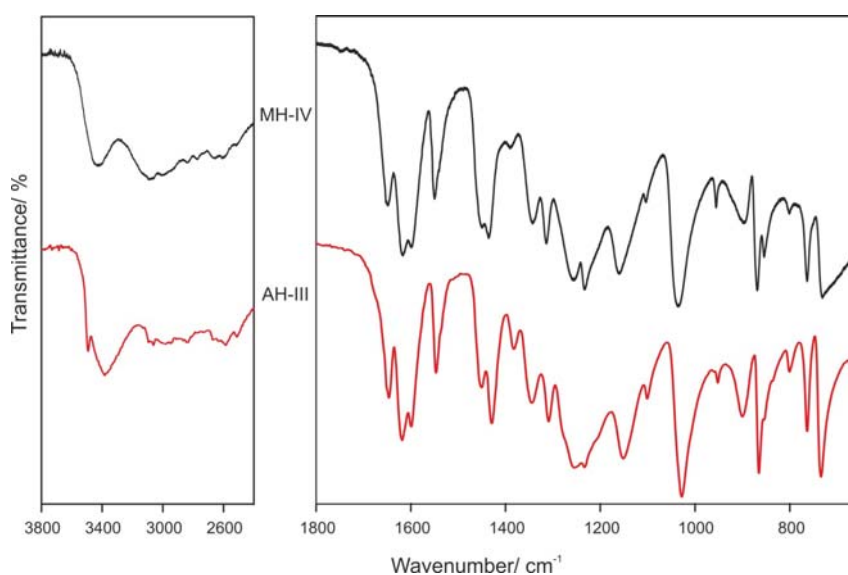
Figure S23. Comparison of the IR spectra of gallic acid **MH-IV** and **AH-III**.

Gallic acid **AH-III** was loaded in a rotating 0.7mmm borosilicate capillary and mounted on a Bruker AXS D8 powder X-ray diffractometer equipped with primary monochromator (CuKα_1, λ = 1.54056 Å) and Lynxeye position sensitive detector. Data were collected at room temperature using a variable count time scheme[26,27] (Table S11). Due to the presence of phase impurities, the

crystal structure prediction results were used to solve the structure. The simulated powder pattern of three of the computationally generated D1 ladder structures closely matched the reflections of the experimental **AH-III**. The three structures (A1_20, A3_86 and A4_20) differed only in proton position, but not in the packing. The most stable for the three structures (DFT-D calculations, with modified lattice parameters obtained after Pawley type refinement[30] against mixed phase data from **AH-III** – 72.5% and **AH-I** – 27.5%) was taken as a starting structure for a multi-phase restrained Rietveld refinement[31] in TOPAS Academic V4.1.[32] In the course of the refinement the **AH-I** unit-cell and peak-shape parameters were allowed to vary, whilst all atomic coordinates were fixed. All atomic positions (including H atoms) for **AH-III** structure were refined, subject to a series of restraints on bond lengths, bond angles and planarity. The final refinement included a total of 116 parameters (41 profile, 4 cell, 1 scale, 1 isotropic temperature factor, 15 preferred orientation and 54 positions yielding a final R_{wp}= 3.074 (Figure S24).

Figure S24. Final observed (points), calculated (line) and difference [(yobs - ycalc)/ σ(yobs)] profiles for the Rietveld refinement of gallic acid anhdyrate form III (major phase, tic marks: second row) and **AH-I** (byproduct, tic marks: first row).

Crystal data of gallic acid **AH-III**: $C_7H_6O_5$, M_r = 170.12, Monoclinic, space group; $P2_1/c$, T = 25 °C, sample formulation: powder, specimen shape: 12 x 1.0 x 0.7 mm, wavelength: 1.54056 Å, a = 5.2303(1) Å, b = 5.26495(12) Å, c = 24.7927(4) Å , β = 102.1116(17) °, Z = 4, density = 1.693 g cm^{-3}, 2 theta range for data collection: 3 to 80°, background treatment: Chebyshev polynomial, No. of measured reflections: 414, Refinement method: Rietveld, data/parameter/restraints: 414/116/44, goodness of fit: 2.065 (on Y_{obs}), R_{wp} = 3.074, R_{exp} = 1.460, R_p = 2.586.

2.5 Thermal Analysis, thermodynamic and kinetic stability

Thermogravimetric analysis (TGA) was carried out with a TGA7 system (Perkin-Elmer, Norwalk, CT, USA) using the Pyris 2.0 Software. Approximately 3-5 mg of sample was weighed into a platinum pan. Two-point calibration of the temperature was performed with ferromagnetic materials (Alumel and Ni, Curie-point standards, Perkin-Elmer). A heating rate of 5 K min^{-1} was applied and dry nitrogen was used as a purge gas (sample purge: 20 mL min^{-1}, balance purge: 40 mL min^{-1}).

Differential scanning calorimetry (DSC) thermograms were recorded on a DSC 7 (Perkin-Elmer) controlled by the Pyris 2.0 software. Using a UM3 ultramicrobalance (Mettler), samples of approximately 1 - 3 mg were weighed into sealed/perforated aluminium pans. The hydrate and solvate samples were heated from 20 to 120 °C using a rate of 10 °C min^{-1} and with dry nitrogen as the purge gas (purge: 20 ml min^{-1}). For faster heating rates a Diamond DSC (Perkin-Elmer), controlled by the Pyris 7.0 software, was used. Samples of approximately 0.5 mg were weighed with the same ultramicrobalance into sealed T_{zero} pans (Waters). The anhydrate samples were heated from 20 to > 320 °C with heating rates of 50 and 100 K 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 temperatures (extrapolated onset temperatures) and enthalpy values are calculated at the 95% confidence intervals (CI) based on at least three measurements.

Gallic acid **MH-I°** forms transparent block or needle like crystals, **MH-II** fine long needles, **MH-IV** small prismatic, and **MH-V** thin platy and needle like multi-layered crystals. **MH-III** batches were obtained via a transformation step (desolvation) and therefore the hydrate crystals already showed the typical signs of a desolvation product, i.e. white non-transparent powder appearing dark under polarized light microscopy. Single crystals of the hydrates turned opaque when slowly heated (heating rage < 5 °C min^{-1}), while the original shape of the crystals was more or less maintained ("pseudomorphosis"). The desolvation process was largely completed at about 75 °C for **MH-V**, 85 °C for **MH-IV**, 90 °C for **MH-III** and 115 °C for forms **MH-II** and **MH-I°** upon heating. Under these desolvation conditions **MH-I°** and **MH-II** dehydrated to forms **I** and/or **II°**, **MH-III** predominantly to form **II°** (+ **AH-I**), and **MH-IV** to **AH-II°** with **AH-III** impurities. At faster heating rates (> 10 °C min^{-1}) and embedding in high-viscosity silicon oil peritectic melting of the **MHs I°-IV** was observed (**MH-IV**: 111.5 – 113.5 °C, **MH-III**: 115 – 117 °C, **MH-II**: 119-120 °C, and **MH-I°**: 124 - 126 °C). The rather high peritectic temperatures (all above the boiling point of water) indicate that the water molecules are well accommodated in these four structures. For **MH-V** the melting point could not be determined. The **MH-V** differential scanning calorimetry (DSC) curve exhibited an exothermic event starting below 60 °C (closed pan, isochoric conditions), a phase transition to **MH-I°** (heat of transition -3.4 ± 0.2 kJ mol^{-1}).

The solvent:compound ratios of the five hydrate phases were determined with TGA (Figure S14). The TGA curves of all hydrates show one distinct step due to the loss of water. The mass loss of **MH-I** (9.53 %), **MH-II** (9.61 %), **MH-III** (9.55 %), **MH-IV** (9.70%), and **MH-V** (9.67 %) corresponds to 1.00 to 1.02 mole water per mole gallic acid, confirming the monohydrate stoichiometry. Furthermore, the TGA measurements indicated following thermal stability order:

MH-I° (most stable) > **MH-II** > **MH-III** > **MH-IV** > **MH-V** (least stable), essentially the same order as found for the order of the **MH-I°** - **MH-IV** peritectic melting temperatures. The samples were gently ground to similar particle size prior measurement to exclude particle size dependent influences on the dehydration temperatures.
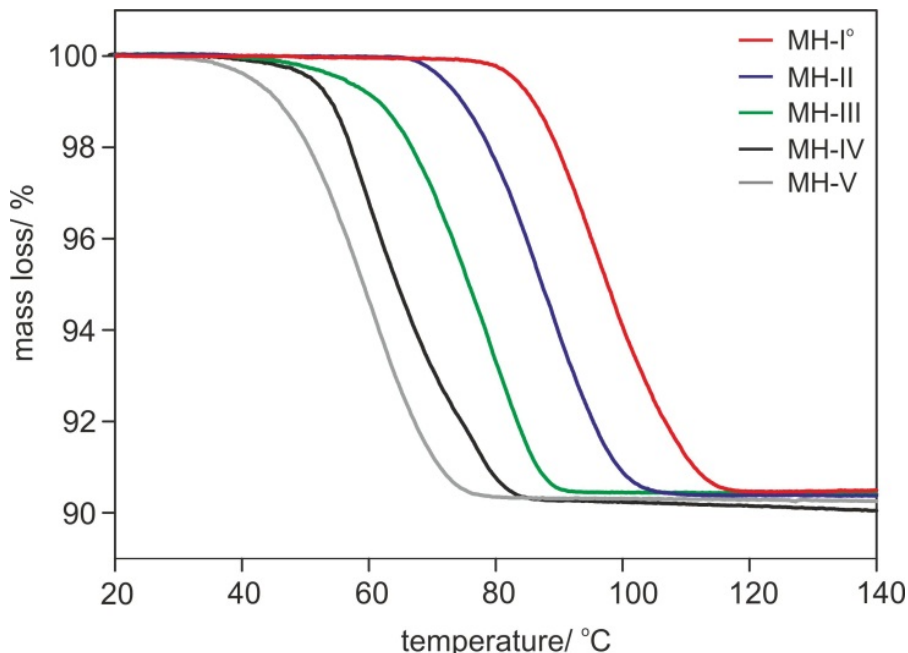
Figure S25. Overlay of the TGA curves of the GalAc MHs (temperature range: 20 – 140 °C, heating rate 5 °C min^{-1}).

Upon further heating (anhydrate) strong sublimation is observed at approximately 210 °C. Initial growth of small grainy crystals occurred (**AH-I** and few **AH-III** crystals), followed by the growth of spear-like and elongated prismatic crystals (**AH-II°**). The spear-like and prismatic crystals grew at expense of the grainy crystals (**AH-I**) which exhibit a higher vapor pressure. Simultaneously few of the originally grown crystals melted at 225 to 230 °C. The melting point corresponds to Lindpainter's "unstable" modification[33] (**AH-III** based on our nomenclature), but no transformation to the "stable" modification could be seen for the **AH-III** crystals. The presence of all three forms (and order of appearance) was confirmed with infrared spectroscopy. These observations could only be made if fast heating rates (> 10 °C min^{-1}) were applied. If slow heating rates were applied (< 5 °C min^{-1}) only **AH-II°** crystals were identified among the sublimed crystals. Above 235 °C strong thermolysis of the acid started. With the aid of differential scanning calorimetry, an endothermic **AH-II°** to **AH-I** phase transformation (> 245 °C, heating rate 50 °C min^{-1}, Figure S26) could be identified and the obtained form was confirmed with infrared spectroscopy and powder X-ray diffractometry. In the hot-stage microscopic investigations thermolysis of the acid was too strong to unambiguously identify this transformation (applying heating rates << 50 °C min^{-1}). An instant melting point of 263 to 265 °C could be determined on the hot-bench for the higher melting form (**AH-I**).
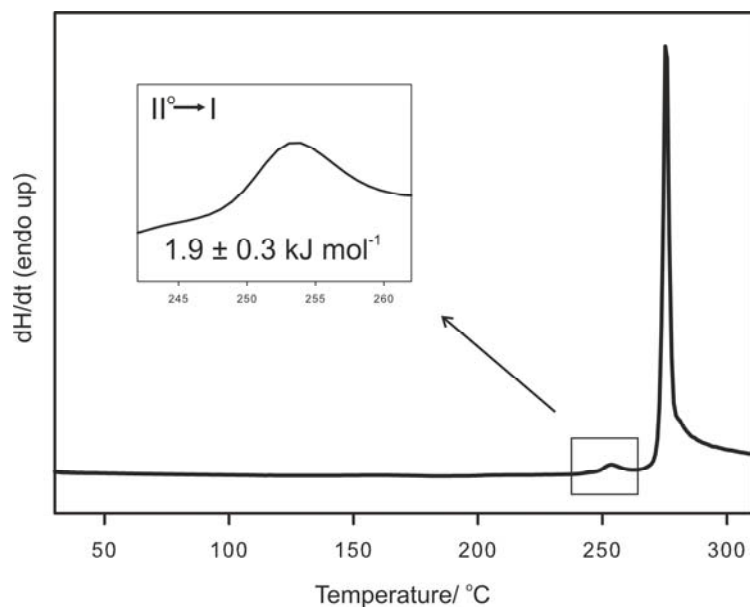
Figure S26. Differential scanning calorimetry curve that shows the transformation of AH-II° to AH-I (heating rate: 50 °C min⁻¹).

Additional sublimation experiments with the sample between two microscopic slides (no contact between the sample and the upper microscopic slide) were performed on top of a preheated metal block (metal surface temperature approx. 240 °C, glass surface on which crystal sublimed to approx. 190 °C at the end of the experiment). Grainy **AH-I** crystals were observed immediately on the glass surface. After few minutes the spear-like **AH-II°** crystals grew on top of the powder and grew further up to the upper microscopic slide, reaching the grainy **AH-I** crystals. Within 10 to 20 minutes all **AH-I** crystals "transformed" through the vapor phase to **AH-II°**. It has to be noted that all these thermal events were observed below the experimental **AH-II°** to **AH-I** transformation temperature ($T_{trs,expl.} > 245$ °C).

The combination of DSC and sublimation experiments suggested that the thermodynamic transition point (T_{trs}) is probably somewhere between 200 and 245 °C.

The thermodynamic relationship of the gallic acid anhydrate polymorphs is displayed in a semi-schematic energy/temperature diagram (Figure S27).
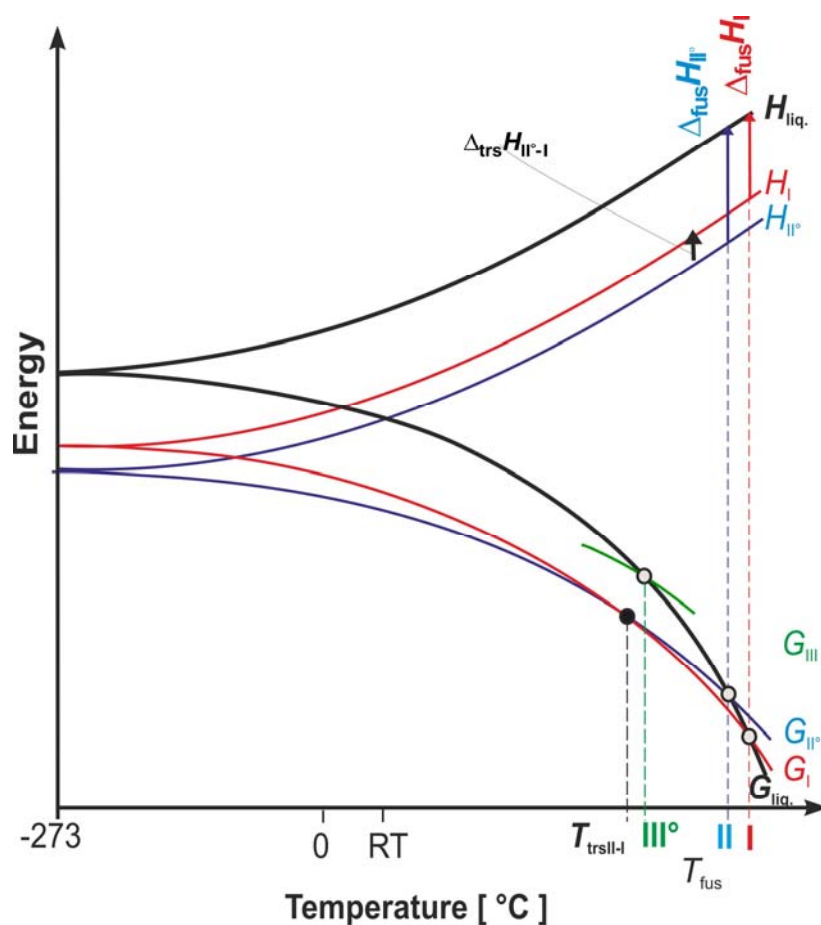
Figure S27. Semi-schematic energy/temperature diaram of gallic acid anhydrate polymorphs. T_{fus} = melting point, G = Gibbs free energy, H = enthalpy, $\Delta_{fus}H$ = enthalpy of fusion, T_{trs} – transition point, $\Delta_{trs}H$ = transition enthalpy, liq. $\Delta_{fus}H$ =liquid phase (melt).

2.6 Moisture Sorption/Desorption Analysis

The five monohydrate polymorphs and the thermodynamically most stable anhydrate form at room temperature (**AH-II**°) were subjected to gravimetric moisture sorption studies at 25 °C (Figure S28).

Dynamic moisture sorption and desorption studies were performed with the automatic multi-sample (gravimetric) moisture sorption analyzer (SPS11-10µ, Project-Messtechnik, Ulm, D). Approximately 100 to 600 mg of sample were used for the investigations. The samples were gently ground prior to measurement to reduce the influence of particle size and surface area. The measurement cycles were started at 40% relative humidity (RH). Sorption and desorption cycles covered the 10% to 90% RH range in 10% steps and the 0% to 10% range in 5% steps. The equilibrium condition for each step was set to a mass constancy of ± 0.01% over 15 minutes.

AH-II° absorbs water and transforms only to **MH-I°** at the highest RH value (90%, Figure S28). The measured mass gain of 10.39% corresponds to 0.98 mole water. **MH-I°**, **MH-II** and **MH-III** dehydrated only at driest conditions (below 5% RH) and **MH-IV** only under rather dry conditions (5% RH and below). In contrast to these high stability forms (moisture dependent) **MH-V** started dehydrating at ambient conditions (<40 % RH and below). Only **AH-II°** and **MH-I°** emerged from the sorption/desorption cycles as confirmed with PXRD. The profile of the moisture sorption-desorption isotherms (distinct steps and hysteresis between the sorption and desorption) are characteristic for stoichiometric hydrates.[34] From the hysteresis we can deduce that **AH-II°** as well as the hydrates **MH-I°**, **MH-II**, **MH-III** and **MH-IV** can be handled without problems and are expected to be storage stable provided extremely low (< 10% RH) or high (> 80 % RH) moisture conditions are avoided. However, a mixture of any hydrate and **AH-II°** started absorbing water already at lower RH (RH > 50%).
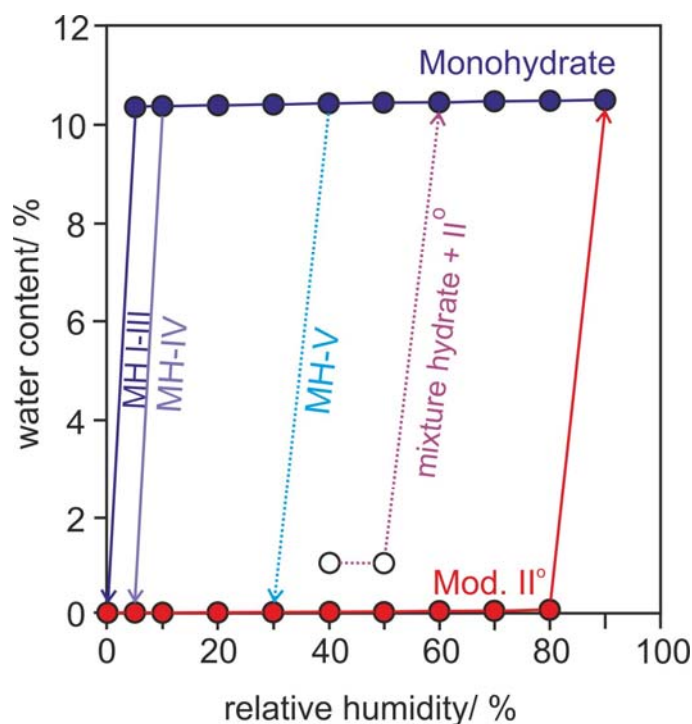
Figure S28. Moisture sorption/desorption isotherms of gallic acid hydrates, **AH-II°** and a hydrate/**AH-II°** mixture performed at 25 °C. Isotherms were constructed by the combination of the desorption cycles of the hydrates and sorption cycles of **AH-II°** and hydrate/**AH-II°** mixture.

Reference List

1. Kazantsev, A. V.; Karamertzanis, P. G.; Adjiman, C. S.; Pantelides, C. C. Efficient Handling of Molecular Flexibility in Lattice Energy Minimization of Organic Crystals. J. Chem. Theory Comput. 7[6], 1998-2016. 2011.

2. Karamertzanis, P. G.; Pantelides, C. C. *Mol. Phys.* **2007,** *105*, 273-291.

3. Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1990,** *11*, 361-373.

4. Price, S. L.; Leslie, M.; Welch, G. W. A.; Habgood, M.; Price, L. S.; Karamertzanis, P. G.; Day, G. M. *Phys. Chem. Chem. Phys.* **2010,** *12*, 8478-8490.

5. Stone, A. J. *J. Chem. Theory Comput.* **2005,** *1*, 1128-1132.

6. Cooper, T. G.; Hejczyk, K. E.; Jones, W.; Day, G. M. *J. Chem. Theory Comput.* **2008,** *4*, 1795-1805.

7. Day, G. M.; Cooper, T. G. *CrystEngComm* **2010,** *12*, 2443-2453.

8. Braun, D. E.; Karamertzanis, P. G.; Price, S. L. Which, if any, hydrates will crystallise? Predicting hydrate formation of two dihydroxybenzoic acids. Chem. Commun. (Cambridge, U. K.) 47[19], 5443-5445. 2011.

9. Cossi, M.; Scalmani, G.; Rega, N.; Barone, V. *J. Chem. Phys.* **2002,** *117*, 43-45.

10. Gelbrich, T.; Hursthouse, M. B. *CrystEngComm* **2005,** *7*, 324-336.

11. Chisholm, J. A.; Motherwell, S. *J. Appl. Crystallogr.* **2005,** *38*, 228-231.

12. Macrae, C. F.; Bruno, I. J.; Chisholm, J. A.; Edgington, P. R.; McCabe, P.; Pidcock, E.; Rodriguez-Monge, L.; Taylor, R.; van de Streek, J.; Wood, P. A. *J. Appl. Crystallogr.* **2008,** *41*, 466-470.

13. Clark, S. J.; Segall, M. D.; Pickard, C. J.; Hasnip, P. J.; Probert, M. J.; Refson, K.; Payne, M. C. *Z. Kristallogr.* **2005,** *220*, 567-570.

14. Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996,** *77*, 3865-3868.

15. Vanderbilt, D. *Phys. Rev. B* **1990,** *41*, 7892-7895.

16. Tkatchenko, A.; Scheffler, M. *Phys. Rev. Lett.* **2009,** *102*, 073005-1-073005/4.

17. Monkhorst, H. J.; Pack J.D. *Phys. Rev. B* **1976,** *13*, 5188-5192.

18. Colthup, N.; Daly, L. H.; Wiberley, S. E. *Introduction to Infrared and Raman Spectroscopy. 2nd Ed;* 1975.

19. Clarke, H. D.; Arora, K. K.; Wojtas, L.; Zaworotko, M. J. *Cryst. Growth Des.* **2011,** *11*, 964-966.

20. Zhao, J.; Khan, I. A.; Fronczek, F. R. *Acta Crystallogr Sect E Struct Rep Online* **2011,** *67*, o316-o317.

21. Hirun, N.; Saithong, S.; Pakawatchai, C.; Tantishaiyakul, V. *Acta Crystallogr. , Sect. E: Struct. Rep. Online* **2011,** *E67*, o787.

22. *Saint+: Area detector control and data integration and reduction software,* Brüker AXS: **2001**

23. *SADABS,* Sheldrick, G. M. Germany: University of Göttingen, **1997**

24. Burla, M. C.; Caliandro, R.; Camalli, M.; Carrozzini, B.; Cascarano, G. L.; De Caro, L.; Giacovazzo, C.; Polidori, G.; Spagna, R. *J. Appl. Crystallogr.* **2005,** *38*, 381-388.

25. Sheldrick, G. M. *Acta Crystallogr. , Sect. A.* **2008,** *A64*, 112-122.

26. Hill, R. J.; Madsen, I. C. In *Structure Determination from Powder Diffraction Data*, David, W. I. F., Shankland, K., McCusker, L. B., Baerlocher, Ch., Eds.; Oxford Science Publications: Oxford, **2002**; 98-117.

27. Shankland, K.; David, W. I. F.; Sivia, D. S. *J. Mater. Chem.* **1997,** *7*, 569-572.

28. Markvardsen, A. J.; David, W. I. F.; Johnson, J. C.; Shankland, K. *Acta Cryst. A* **2001,** *57*, 47-54.

29. David, W. I. F.; Shankland, K.; van de Streek, J.; Pidcock, E.; Motherwell, W. D. S.; Cole, J. C. *J. Appl. Crystallogr.* **2006,** *39*, 910-915.

30. Pawley, G. S. *J. Appl. Crystallogr.* **1981,** *14*, 357-361.

31. Rietveld, H. M. *J. Appl. Crystallogr.* **1969,** *2*, 65-71.

32. Coelho, A. A. *Topas user manual;* Version 4.1 ed.; Bruker AXS GmbH: Karlsruhe, Germany, 2004.

33. Lindpaintner, E. Microscopic study of polymorphic substances. II. Mikrochemie 27, 21-41. 1939.

34. Griesser, U. J. In *Polymorphism: In the Pharmaceutical Industry*, Hilfiker, R., Ed.; Wiley-VCH: Germany, **2006**; 211-233.