IUCrJ

Volume 10 (2023)

Supporting information for article:

Leucopterin, the white pigment in butterfly wings: structural analysis by PDF fit, FIDEL fit, Rietveld refinement, solid-state NMR and DFT-D

Federica Bravetti, Lukas Tapmeyer, Kathrin Skorodumov, Edith Alig, Stefan Habermehl, Robert Hühn, Simone Bordignon, Angelo Gallo, Carlo Nervi, Michele R. Chierotti and Martin U. Schmidt

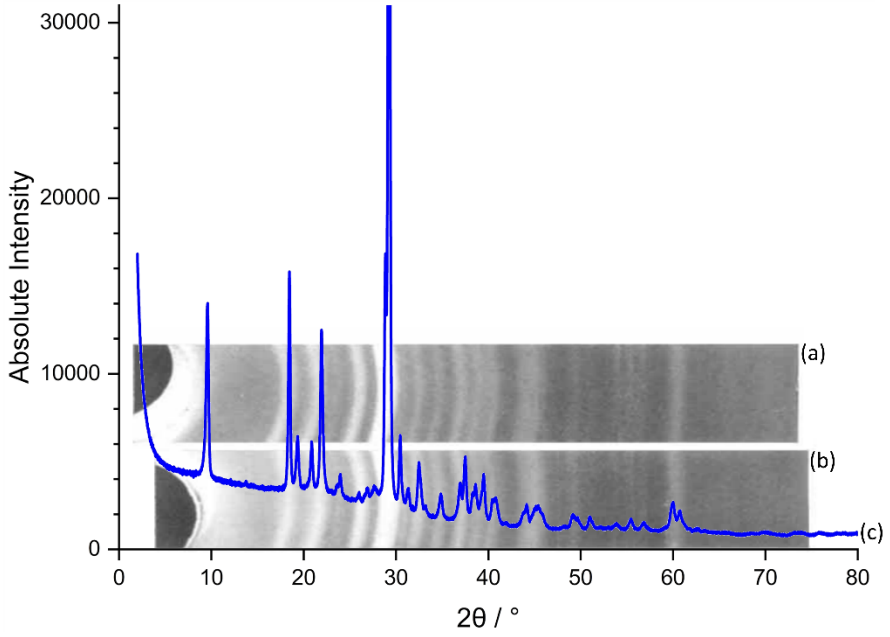

Figure S1 Comparison of the Debye-Scherrer patterns of (a) synthetic leucopterin from Purrmann (1940), (b) natural leucopterin isolated from butterflies, and (c) our synthetic leucopterin (blue curve). The images of the Debye-Scherrer films in (a) and (b) were taken from Purrmann (1940).

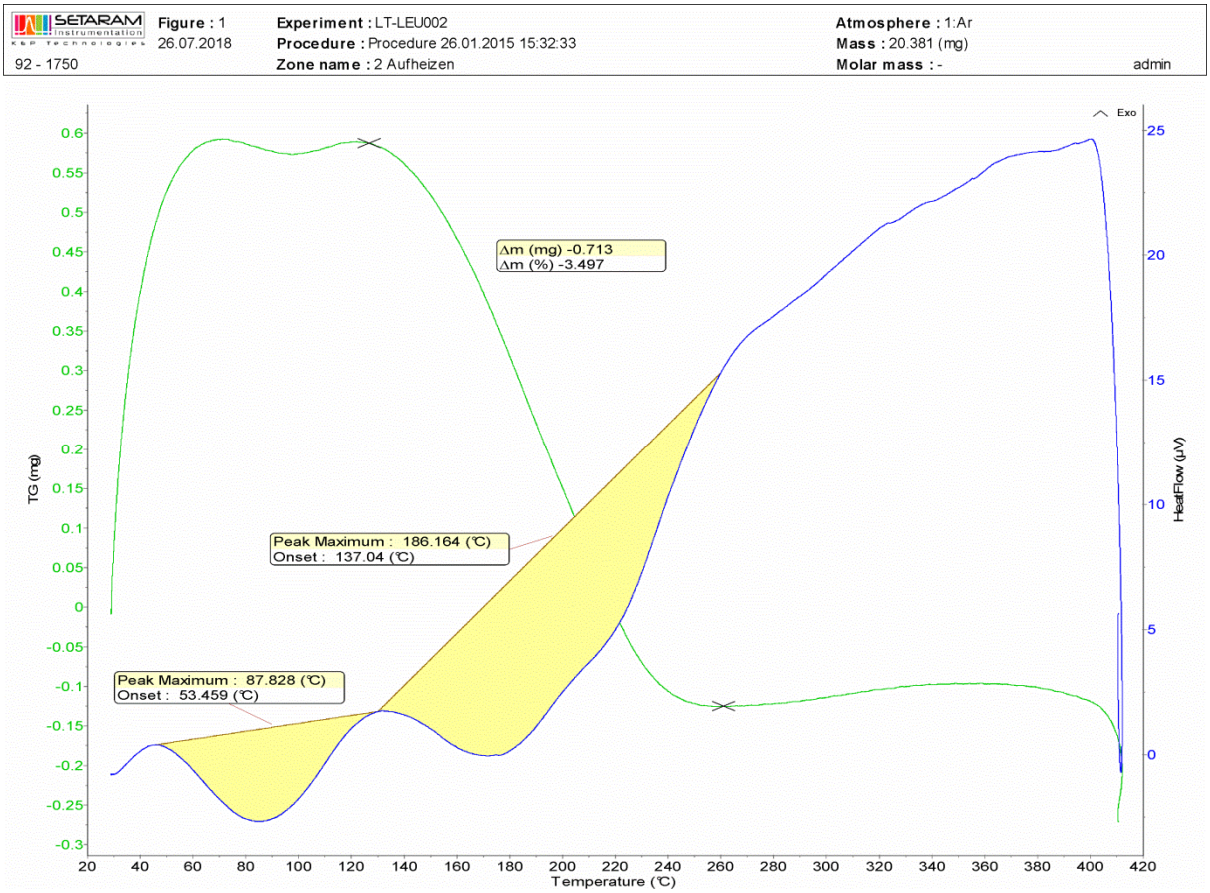

Figure S2 DTA-TG of leucopterin hemihydrate.

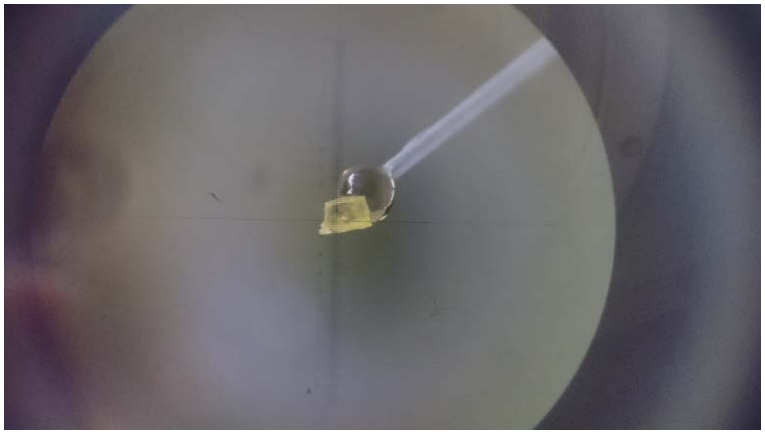

Figure S3 Single crystal of leucopterin hemihydrate on a glass pin.

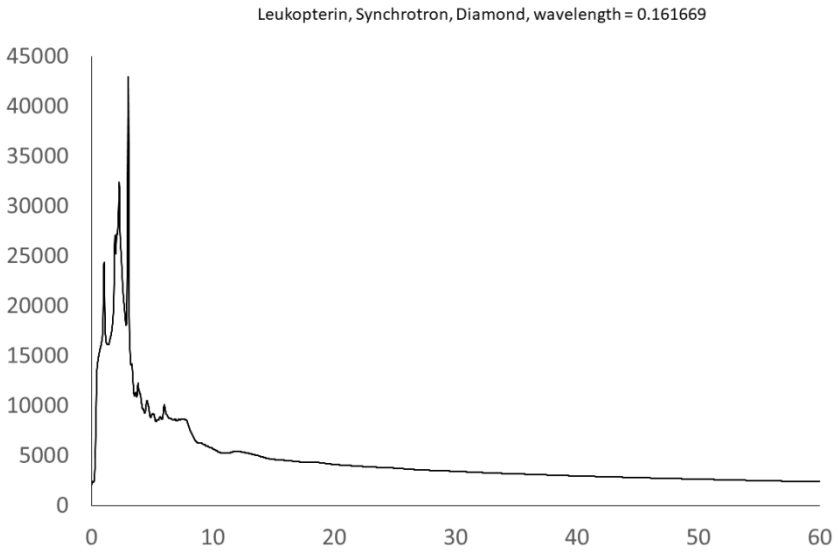

Figure S4 Synchrotron powder diffraction data.

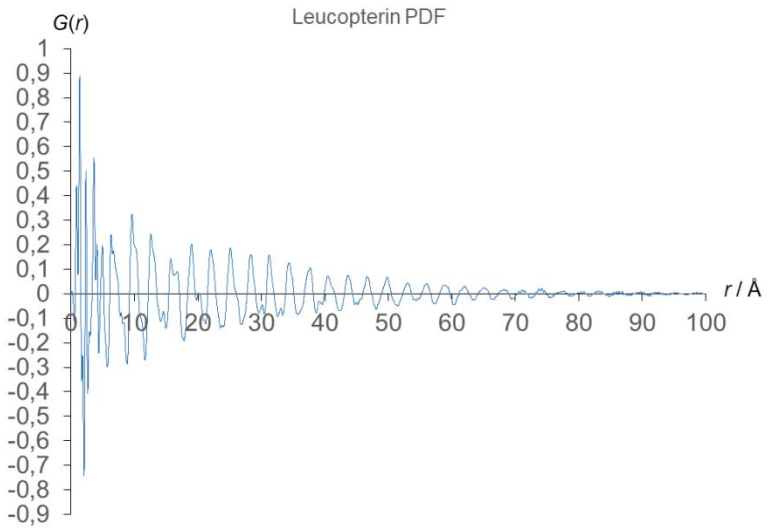

Figure S5 Experimental PDF.

PDF global fits

PDF global fits were performed with the aim to solve the crystal structure of the hemihydrate. All PDF fits were performed without the water molecule. At first, PDF fits were run in various space groups, not including *P*2/*c*. The best resulting fit is shown in Figure S6a. The corresponding crystal structure was wrong. After *P*2/*c* turned out to be the correct space group, additional PDF global fits were run in *P*2/*c*. The R_{wp}^{PDF} value dropped, and the best fit (Figure 4 in the main text) corresponds to the correct structure. Figure S6b gives an overlay of the structures from PDF fit and single-crystal data.

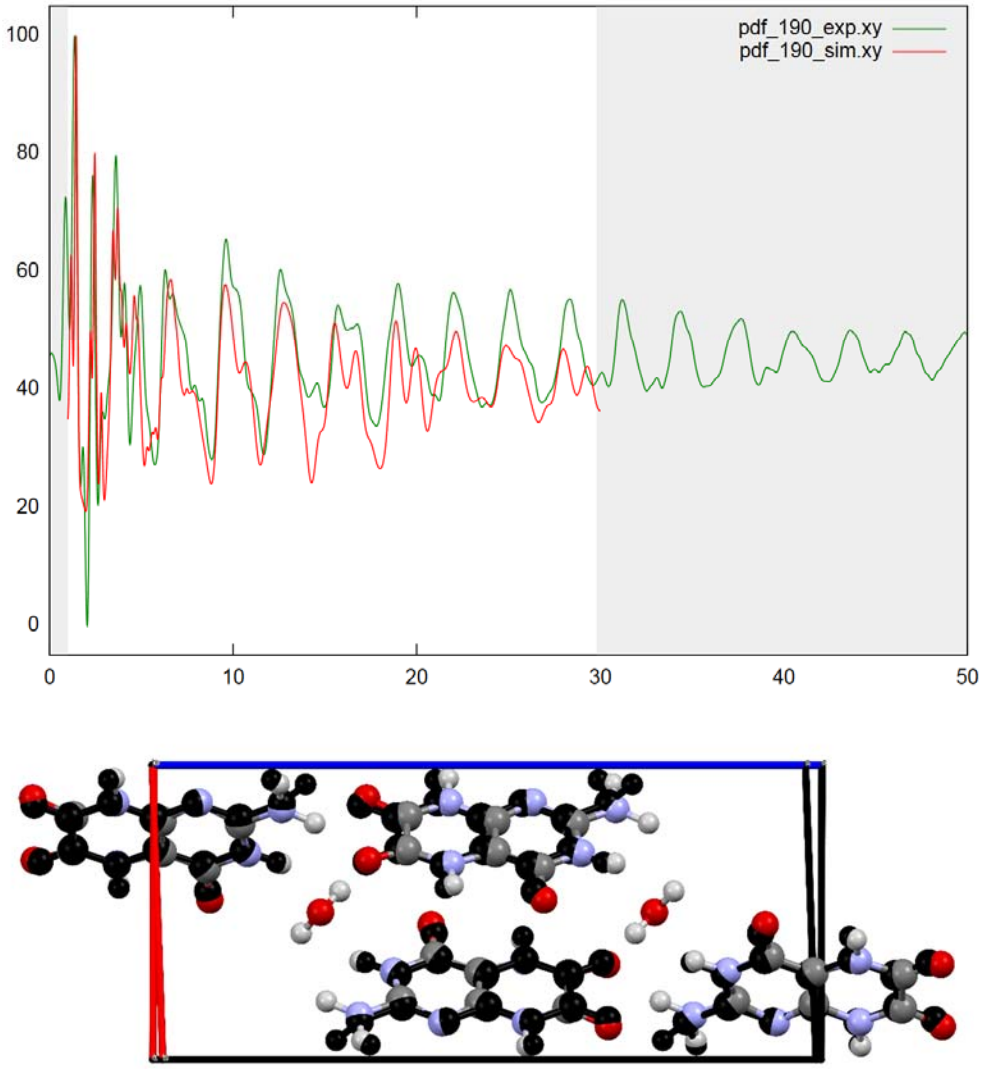

Figure S6 PDF global fit for structure solution of the hemihydrate. (a) Best PDF fit obtained in other space groups except *P*2/*c*: space group P2$_1$/c, a = 4.82 Å , b = 3.92 Å, c = 41.22 Å, β = 108.4°, V = 738.82 Å^3. This crystal structure is wrong. (b) Structure from the best PDF fit obtained in *P*2/*c*: a = 8.08 Å, b = 4.82 Å, c = 17.94 Å, β = 88.0°, V = 700.60 Å^3. R_{wp}^{PDF} = 38.05 %. The structure from PDF fit is drawn in black, and overlayed with the structure from single-crystal data (in colour).

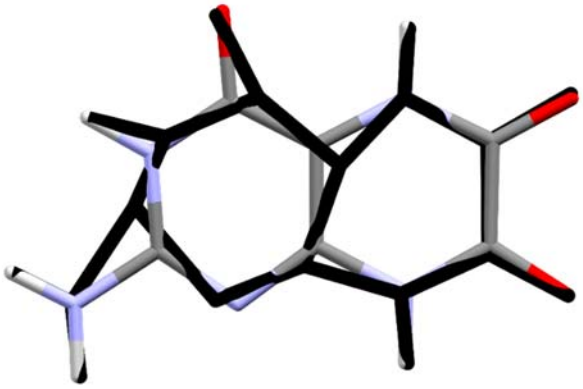

Figure S7 Overlay of the distorted molecule of leucopterin 0.2-hydrate after the unrestrained refinement (in black) with the correct one (coloured). During the free refinement, the occupation of the oxygen atom representing the water molecule dropped from 0.42 to 0.212(11), corresponding to a 0.08-hydrate.

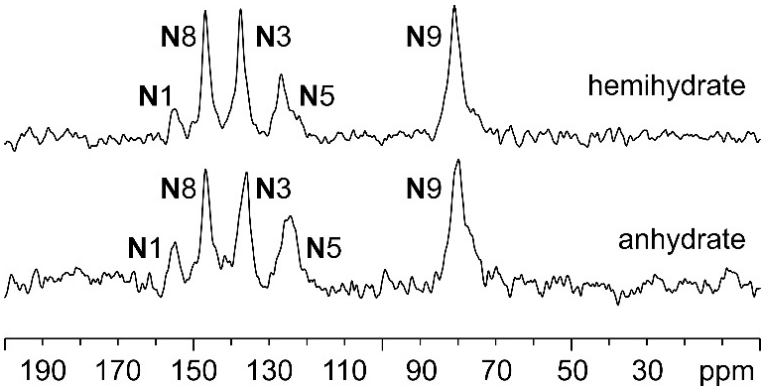

Figure S8 Assigned ^{15}N CPMAS spectra (contact time 4 ms) of leucopterin hemihydrate and anhydrate.

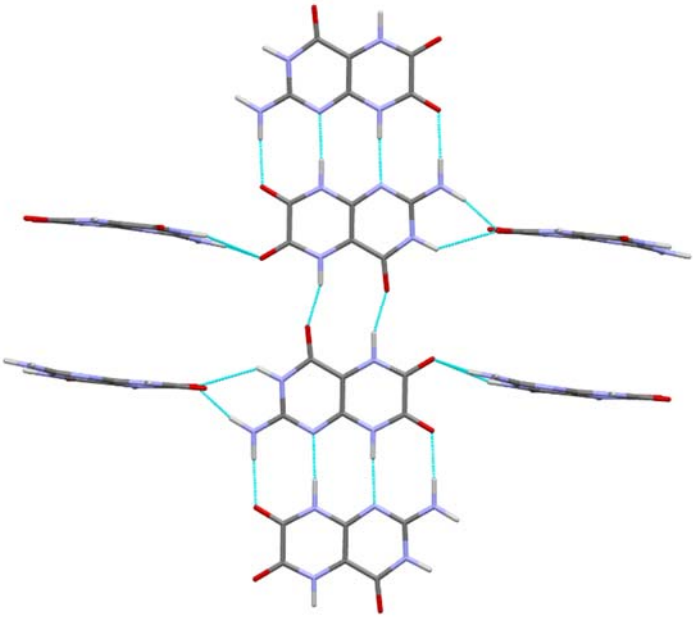

Figure S9 Molecular chains in the crystal structure of leucopterin anhydrate, after DFT-D optimisation with fixed lattice parameters. Colour code in all drawings: C = grey, O = red, N = blue, H = white, hydrogen bonds = turquoise. View direction [120].

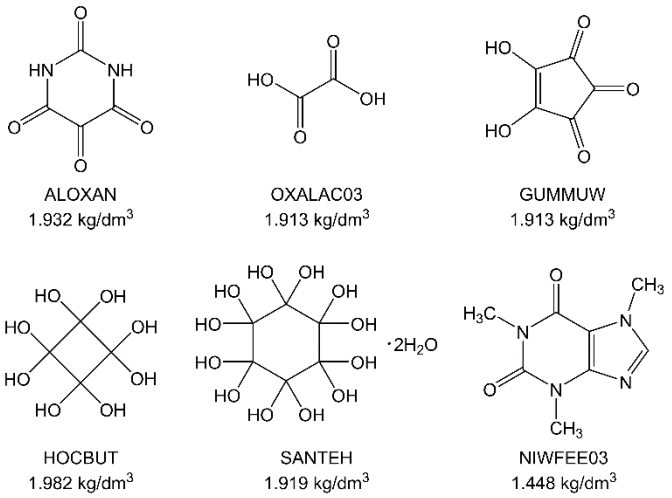

Figure S10 Molecular structures, CCDC refcodes and densities of the crystal structures of some non-nitro compounds with a density higher than 1.909 kg/dm^3 at ambient conditions, and caffeine.

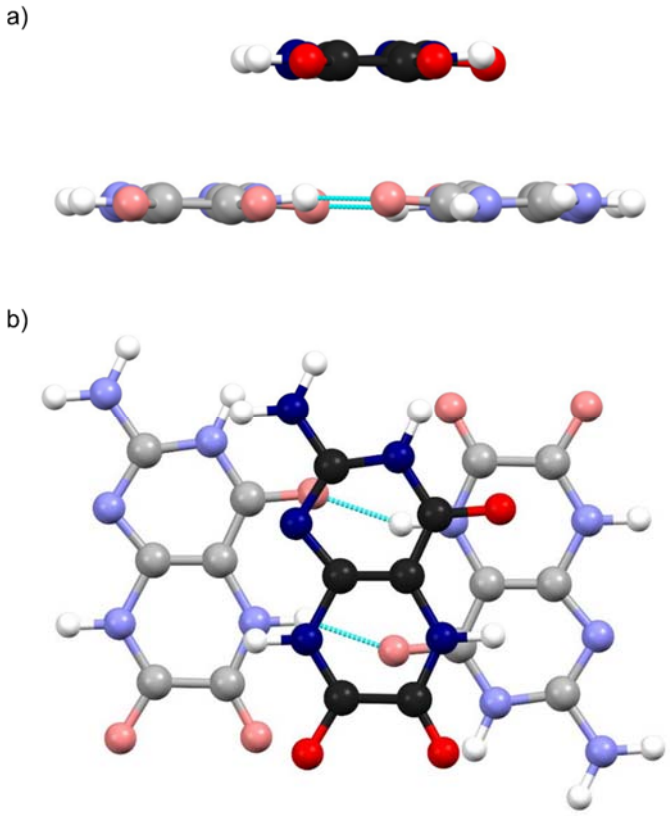

Figure S11 Molecular packing in leucopterin hemihydrate. (a) Stacking of molecules in neighbouring chains. (b) Perpendicular view. One molecule is highlighted.

Table S1 ^{1}H, ^{13}C and ^{15}N computed and experimental chemical shifts, peak assignments and RMSE values of leucopterin hemihydrate containing tautomer T1. Calc 1 and Calc 2 refer to the chemical shifts computed with the B86r or optB88 method, respectively. See Scheme 1 for atom numbering (#).

	^{1}H chemical shift (ppm)			**^{13}C chemical shift (ppm)**			**^{15}N chemical shift (ppm)**		
#	**Exp**	**Calc 1**	**Calc 2**	**Exp**	**Calc 1**	**Calc 2**	**Exp**	**Calc 1**	**Calc 2**
1							154.9	152.4	155.3
2				153.4	151.6	152.0			
3	10.2	9.9	10.0				137.5	133.0	133.5
4				156.2	154.5	154.4			
4a				99.3	102.6	102.6			
5	10.2	11.2	11.1				126.8	131.9	131.4
6				153.4	152.1	152.2			

7				157.6	158.3	157.9			
8	11.6	12.8	12.4				146.9	150.1	149.8
8a				142.4	143.2	143.4			
9	6.9 7.9	6.9 7.4	6.9 7.5				81.0	79.7	77.1
H_2O	3.4	2.3	2.5						

Table S2 ^{1}H, ^{13}C and ^{15}N computed and experimental chemical shifts, peak assignments and RMSE values of leucopterin anydrate containing tautomer T1. Calc 1 and Calc 2 refer to the chemical shifts computed with the B86r or optB88 method, respectively. See Scheme 1 for atom numbering (#).

	^{1}H chemical shift (ppm)			**^{13}C chemical shift (ppm)**			**^{15}N chemical shift (ppm)**		
#	**Exp**	**Calc 1**	**Calc 2**	**Exp**	**Calc 1**	**Calc 2**	**Exp**	**Calc 1**	**Calc 2**
1							154.9	152.4	155.0
2				154.1	152.0	152.2			
3	7.9	7.5	7.6				146.8	149.7	149.4
4				156.0	153.7	153.5			
4a				99.6	103.7	103.8			
5	9.8	11.3	11.1				135.9	131.0	131.5
6				154.1	152.1	152.2			
7				157.8	158.7	158.4			
8	11.8	13.6	13.2				124.4	130.4	130.5
8a				142.3	143.6	143.7			
9	6.8 7.9	7.2 7.2	7.4 7.6				79.9	78.3	75.5

Table S3 Experimental (SCXRD) and computed cell parameters ($P2/c$, $Z = 4$) for the 17 structural models of leucopterine (T1-T17) hemihydrate, each one containing a different tautomer. Relative energies with respect to T1. Gas: single molecule in the gas phase, by Gaussian 09; Solid: in the solid state, by Quantum Espresso with the two vdW-DF2 methods B86r (in black) and optB88 (in red). ^{1}H, ^{13}C and ^{15}N chemical shift RMSEs for the computed structures.

Structure	**ΔE (kJ/mol)**		**^{1}H RMSE (ppm)**	**^{13}C RMSE (ppm)**	**^{15}N RMSE (ppm)**	**Volume (Å^3)**	***a*** **(Å)**	***b*** **(Å)**	***c*** **(Å)**	***β*** **(°)**
	Gas	**Solid**								
SCXRD	/	/	/	/	/	710.274	8.0781	4.7930	18.3452	90.2238
T1	0.00	0.00	0.8	1.8	3.6	692.263	7.945	4.749	18.346	90.528
		0.00	0.6	1.8	3.5	710.724	8.091	4.764	18.440	90.519
T2	2.24	40.11*	1.3	1.7	4.6	763.008	10.893	4.757	17.803	55.908
		62.29	2.8	1.6	10.2	780.738	10.609	4.922	18.303	54.769
T9	17.25	84.21	0.8	5.8	35.1	683.122	7.423	5.067	18.192	86.695
		86.15	0.7	6.0	36.9	703.755	7.546	5.106	18.296	86.575

T7	14.75	88.19	1.4	2.6	13.1	738.629	9.304	4.717	17.954	110.379
		85.04	1.2	2.6	12.6	759.854	9.433	4.727	18.036	109.125
T11	9.04	89.79	1.5	6.8	48.1	727.800	8.442	4.804	17.953	91.777
		87.43	1.5	6.8	49.7	750.855	8.625	4.815	18.088	91.702
T4	11.20	90.72	1.2	5.4	44.8	738.546	8.263	4.857	18.418	92.214
		86.90	1.2	5.4	46.3	763.247	8.519	4.836	18.544	92.385
T17	32.16	96.18	1.1	4.4	37.4	717.726	10.460	3.722	18.901	102.765
		97.57	1.4	4.3	39.0	746.747	10.162	3.950	19.012	101.914
T15	9.41	98.42	1.6	2.3	25.3	712.686	8.005	4.690	18.993	91.683
		99.34	1.5	2.4	27.2	738.174	8.233	4.683	19.150	91.229
T14	21.43	122.66	0.4	10.9	60.0	725.651	7.474	5.291	18.441	95.751
		121.62	0.4	10.9	52.7	751.741	7.578	5.374	18.576	96.444
T8	18.28	126.94	1.8	4.9	24.8	803.216	8.550	6.011	17.591	62.801
		126.64	0.6	5.2	27.6	859.248	7.796	6.385	18.534	68.644
T6	16.43	127.18	0.7	5.8	18.9	736.255	8.888	4.416	18.772	92.195
		123.70	0.7	5.8	20.4	758.263	9.056	4.427	18.923	91.920
T3	15.28	137.38	3.1	2.6	33.0	738.481	9.406	5.479	16.643	59.432
		136.86	2.1	2.1	36.1	770.354	9.533	5.563	16.788	59.917
T13	18.02	143.06	0.6	8.4	50.1	825.023	8.119	5.182	19.881	99.503
		136.43	0.7	8.2	51.4	860.129	8.330	5.222	20.114	100.556
T5	23.00	151.01	1.3	4.2	27.8	761.000	8.467	5.167	17.825	102.654
		116.25	1.8	4.9	37.3	824.300	10.835	4.483	18.395	112.687
T12	26.36	155.06	1.1	10.1	62.7	716.744	7.491	5.324	17.980	88.269
		154.76	1.1	10.01	64.1	742.323	7.595	5.404	18.097	88.083
T16	20.79	165.18	1.3	4.3	39.1	729.309	9.377	4.555	17.179	96.360
		164.44	1.2	4.3	41.1	758.472	9.629	4.608	17.220	96.888
T10	30.68	206.76	0.8	9.4	54.4	835.577	7.084	6.604	18.146	100.193
		200.30	0.9	9.4	56.5	877.585	7.213	6.792	18.259	101.173

*During the DFT-D optimisation process, the T2 tautomer was converted into T1, but the molecular arrangement is different. This process proceeded via a rearrangement of the molecules in the cell (there is a significant change in the β angle of the optimised cell) which however resulted in a higher energy. This rearrangement allowed for the transformation of the O–H···N intramolecular interaction of T2 into the O···H–N intramolecular interaction of T1. The same interaction is present in T1, but the optimised T2 cell is very different from the T1 one. The T1 optimised cell parameters fit very well with the experimental data.