Electronic Supplementary Material (ESI) for ChemComm.

Supporting information

A crystal structure prediction enigma solved: The gallic acid monohydrate system – surprises at 10 K

Anna A. Hoser, Ioana Sovago, Arianna Lanza and Anders Ø. Madsen*

Experimental methods

Crystallization

GAM was purchased from Sigma-Aldrich and used without further purification. It was crystallized according to the procedure described by Clarke *et al*.[1] in order to produce the desired form IV (space group $P2_1/c$, unit cell dimensions: 9.7573(4), 3.5579(1), 21.5170(9), 91.338(1)).

10 K X-ray diffraction

A suitable crystal of dimensions 0.2×0.31×0.19 mm^3 was mounted on a Bruker D8 Venture diffractometer equipped with a Mo tube and a Photon 100 CMOS detector. Initially the unit cell was determined at room temperature and found to be in accordance with the previous structure determination.[1] The crystal was then rapidly cooled to 10(2) K using an open-flow liquid Helium cryostat (CIA LHe) and the temperature was maintained at 10(2) K during a unit cell determination and data collection. At this temperature, a new larger unit cell was found, and a full dataset for this cell was collected to a resolution of 0.6 Å. Subsequently, a redetermination of the unit cell at room temperature was performed and at this temperature, no sign of the larger unit cell was found. The structure, corresponding to the larger cell at 10 K, was unambiguously found using direct methods, as implemented in the Shelxs program, as well as by using the charge flipping method implemented in the Olex2 program.

The diffraction images collected at 10 K do not show sign of diffuse scattering or peak broadening. This indicates that the structure found at 10 K corresponds to long-range rather than short-range ordering with respect to the high-temperature phase.

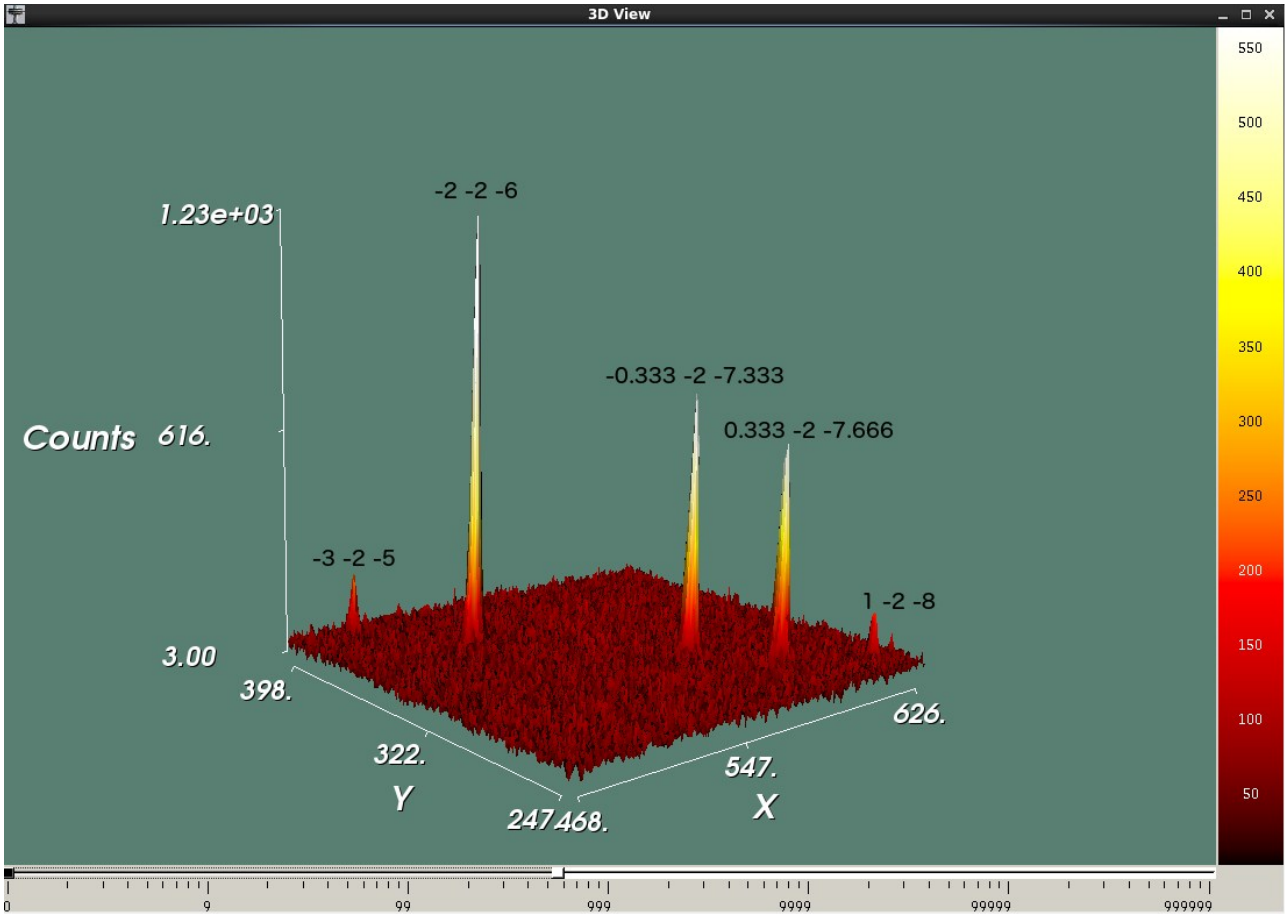


Figure S1 A 3D plot of intensities in a frame taken from the 10 K data collection. The peaks have been indexed according to the small cell. The peaks indexed with non-integral values correspond to the super-cell. They have similar widths as the peaks belonging to the smaller cell. The peak widths do not change going from 123 K to 10 K.

123 K high-resolution X-ray diffraction

Data collection was performed at 123 K using the same diffractometer as for the 10 K data collection. Here the temperature was maintained using an open-flow nitrogen cooling device (CIA LN2). In order to ensure the highest possible resolution, a larger crystal of dimensions 0.45×0.32×0.18 mm^3 was used. Based on an initial scan to confirm the unit cell dimensions and find the orientation matrix, a strategy was set up to ensure a complete dataset without detector overload. Three identical dataset were collected with varying exposure times of 2, 10 and 100 seconds per degree. Subsequently the data were integrated using SAINT incorporated into Brukers Apex package.[2] The data were scaled to a common dataset using the SORTAV program[3] in the DREADD package by Blessing.[4]Spherical atom refinement was performed initially using SHELXL-2013,[5] with full-matrix least-squares on F^2 and using all the unique data. The non-H atoms were refined by allowing anisotropic thermal motion. All calculations were carried out using the WinGX package[6] of crystallographic programs. A summary and statistics of the collected data are given in Table S1.

Multipole refinement -123K data

The refinement using aspherical form factors against the 123 K data of GAM IV was performed using the formalism of Hansens and Coppens[7]as implemented in the XD-2006 software[8]:

$$\rho(\mathbf{r}) = P_{\text{core}}\rho_{\text{core}}(r) + P_{\text{val}}\kappa^3\rho_{\text{val}}(\kappa r) + \sum_{l=0}^{l_{\max}} \kappa'^3 R_l(\kappa' r) \sum_{m=0}^{l} P_{lm\pm} d_{lm\pm}(\Omega)$$

The multipole expansion was truncated at the octupole level for the O and C atoms and the quadrupole level for the H atoms. The ADPs for H atoms in the multipole refinements were estimated using SHADE3.[9]The core and spherical-valence scattering factor for each pseudoatom were derived from the relativistic Dirac-Fock wavefunctions of Su and Coppens[10] expanded in terms of the single-ζ functions of Bunge, Barrientos and Bunge.[11] The valence deformation functions for the atoms used a single-ζ Slater-type radial function multiplied by the density-normalized spherical harmonics. The radial fits for each atoms were optimized by refinement of their expansion-contraction parameters κ, κ'.

Similarity of packing in the 10 K structures

The Z'=3 structure was subjected to a structural similarity check against the Z'=1 structure. A similarity check of a 15 molecules cluster gives an RMSD value of 0.081. The overlay of this cluster is shown in Figure 3a. The mean differences in intermolecular angles *Δa* and planes *Δp* for the Z'=1 and Z'=3 models equals *Δa* = 0.9° and *Δp*= 2.6° (XPacprogram[12]). The dissimilarity index[13] between both structures equals 2.8. The differences between the three molecules in the asymmetric unit of Z'=3 are illustrated in Figure 3b.

Table S1. Crystallographic data for gallic acid Monohydrate form IV at 10 and 123 K. Notice that the two refinements at 10 K (Z'=1 and Z'=3 models) were conducted against the same measurement, but in the case of Z'=1 the reflections which are only in accordance with the larger cell (Z'=3) had to be disregarded, leaving only 1/3 of the data for the refinement.

Compound formula	$C_7H_8O_6$	$C_7H_8O_6$	$C_7H_8O_6$
Polymorphic form	IV'	IV	IV
Z'	3	1	1
M_r	188.1	188.1	188.1
Space group	$P2_1/c$	$P2_1/c$	$P2_1/c$
Crystal system	Monoclinic	Monoclinic	Monoclinic
a/ Å	23.7685(14)	9.7824(5)	9.7573(4)
b/Å	3.5428(2)	3.5428(2)	3.5579(1)
c/Å	28.7556(19)	21.4669(10)	21.5170(9)
β/deg	112.831(2)	91.149(5)	91.338(1)
V/Å^3	2231.71(21)	743.83 (1)	746.77(1)
Z	12	4	4
D_{calc}/g cm^{-1}	1.68	1.68	1.66
λ/Å	0.71073	0.71073	0.71073
μ/mm^{-1}	0.151	0.151	0.15
Temperature/K	10	10	123
Crystal size/mm^3	0.2×0.31×0.19	0.2×0.31×0.19	0.45×0.32×0.18
θrange/deg	2.6-35.0	2.8-34.9	2.8-46.6
Max (sin θ)/λ /Å^{-1}	0.807	0.805	1.02
No. of data measured	53894	16090	125942
No. of unique data	9228	2659	6528
	-37 ≤h≤ 34	-10 ≤h≤ 10	-41 ≤h≤ 41
hkl range	-5 ≤k≤ 5	-5 ≤k≤ 5	-20 ≤k≤ 20
	-43 ≤l≤ 46	-30 ≤l≤ 34	-36 ≤l≤ 36
R_{int}	0.035	0.0258	0.063
R_σ	0.031	0.0194	0.031
No. of data in refinement	9228	2659	6528
No. of refined parameters	426	142	150
Final $R[F^2>2\sigma(F^2)]$	0.051	0.052	0.058
$wR(F^2)$	0.159	0.140	0.136
Goodness of fit S	1.019	1.091	1.093
Extrema in residual map/ e Å^{-3}	1.100 → -0.524	1.09 → -1.01	0.78 → -0.32
Max shift/esd in last cycle	0.002	0.000	0.001
Multipole refinement			
No. of data in refinement			6528
No. of refined parameters			346
Final $R[F>3\sigma(F)]$			0.0354
$wR(F)$			0.0394
Goodness of fit S			1.9943
Extrema in residual map/ eÅ^{-3} (all data)			0.467→ -0.323
Max shift/esd in last cycle			5.2×10^{-6}

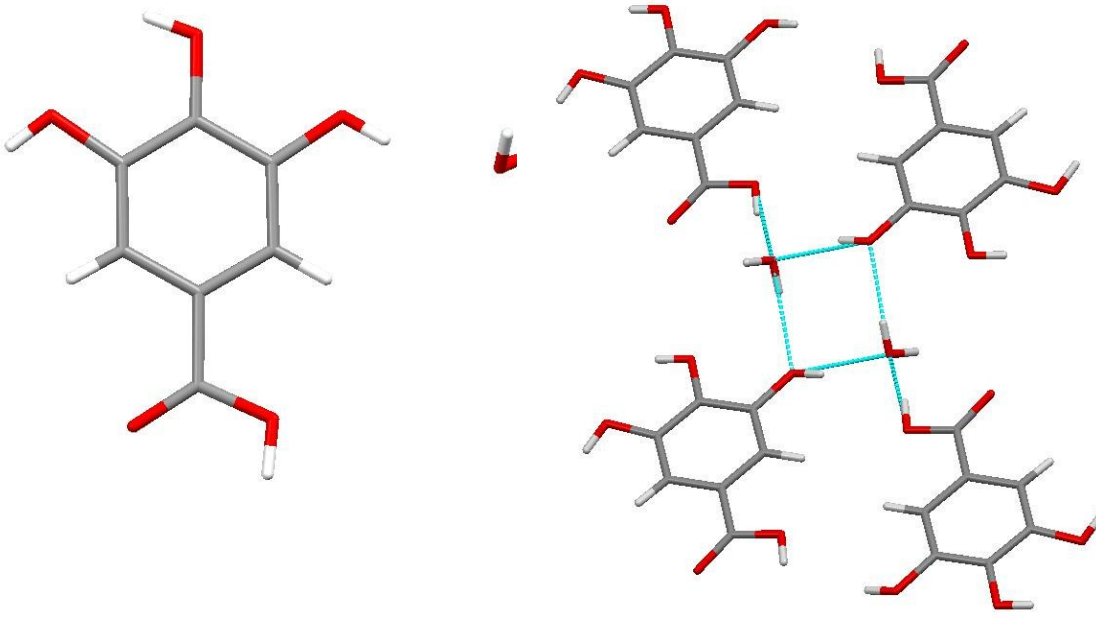
Transformation matrix applied

Theoretical methods

Ab-initio periodic calculations were performed using the CRYSTAL09 software[14] at the B3LYP-d[15]/TZP[16] level of theory. An empirical dispersion correction was applied, as proposed by Grimme[17] and modified for crystals.[18] A counterpoise correction was used to account for basis superposition error. Before the energy calculations were performed we geometry-optimized the hydrogen atom positions.

Table S2. Theoretical Estimates of cohesive Energy (kJmol^{-1}) for gallic acid monohydrate 10 K (Z'=1 and Z'=3). Method: B3LYP-d/TZP. The row named 'Z'=1 disorder part 1' computes the energy with the dominating hydrogen bond pattern found in the Z'=1 structure, while the row named 'Z'=1 disorder part 2' related to the reverse hydrogen bond pattern (20 % of the crystal). The third row is the calculation for the Z'=3 structure. See Fig. S2 below for an illustration.

	Periodic-DFT-d CRYSTAL09	
Z'=1 disorder part 1	-124.72	**+0.25**
Z'=1 disorder part 2	-118.01	**+6.69**
Z'=3	-124.97	**0.0**



(a) **(b)**

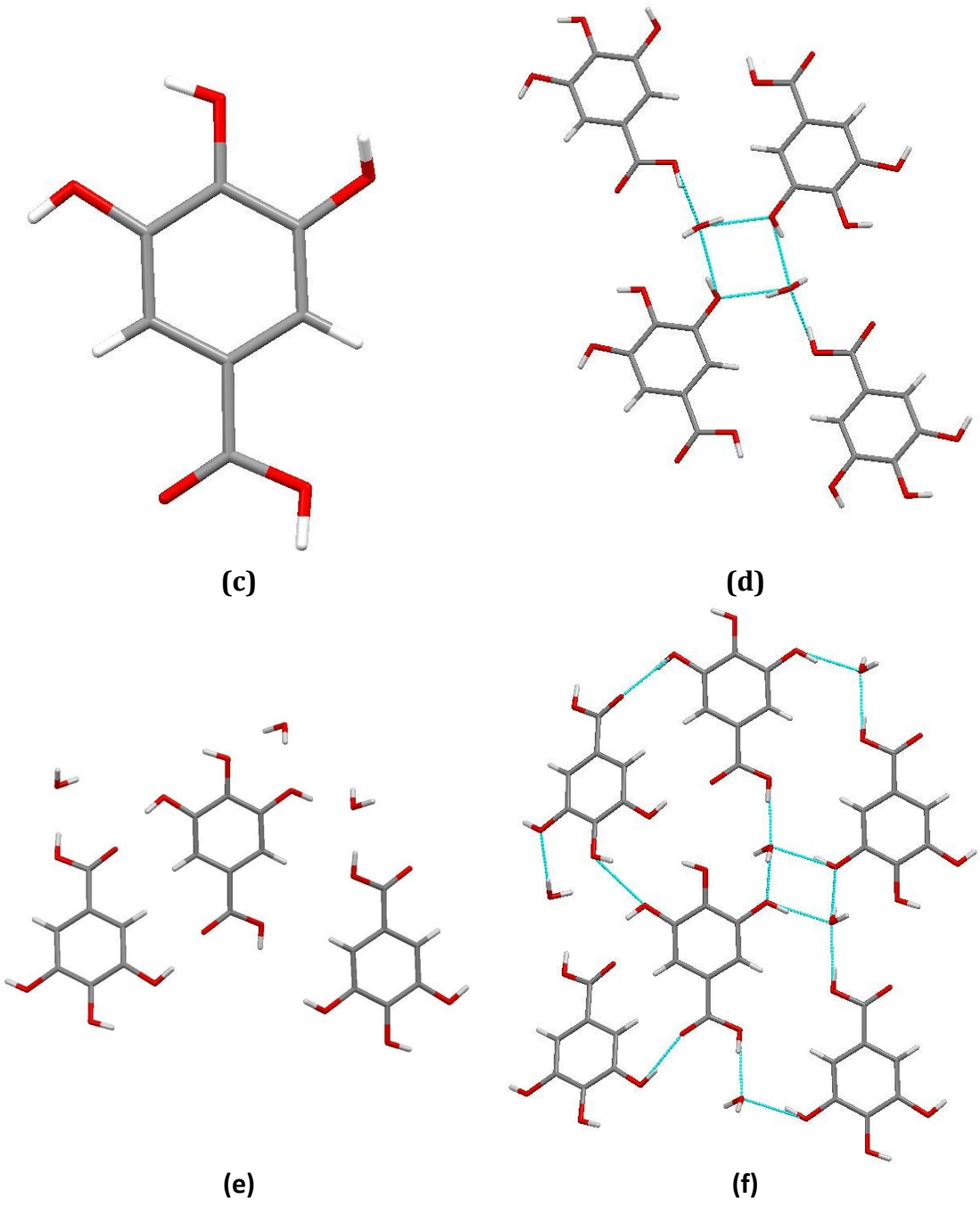


Figure S2. Structures used for theoretical computations: **(a)** and **(b)** disorder part 1, **(c)** and **(d)** disorder part 2, **(e)** and **(f)** Z'=3 structure

Geometry-optimization of Z'=3 structures.
Starting from the experimental Z'=3 structure (10 K) and the corresponding Z'=1 structure expanded to Z'=3, the structures were optimized at the B3LYP-d/TZP level of theory. Geometry convergence criteria were the default values in the Crystal14 program. When the hydrogen positions were initiated at the main positions found in the experimental structures, these structures optimizes to a geometry closely resembling the experimental Z'=3 structure. An illustration is given below, Fig. S3.

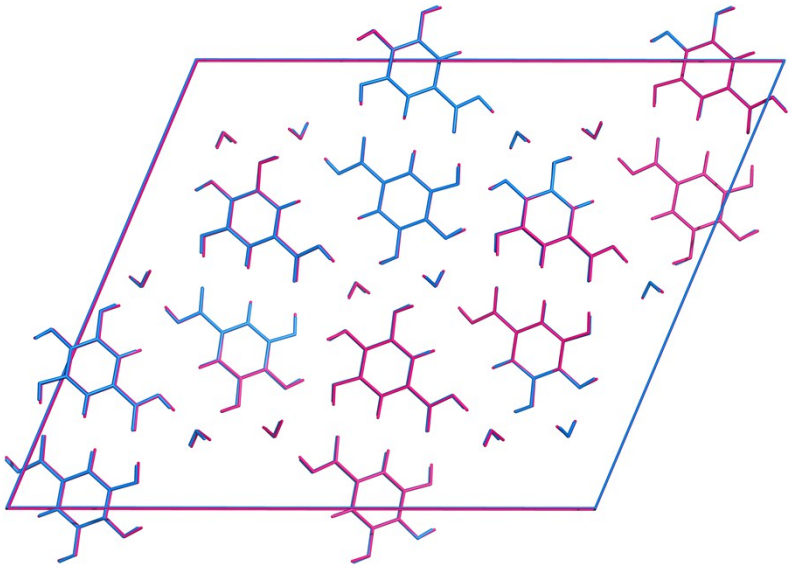


Figure S3: A comparison of the experimental Z'=3 structure (Blue) and the geometry-optimized structure (Pink) in the plane spanned by the a and c axes. The geometry optimization was initiated from the experimental Z'=3 structure.

Table S3. Unit cells parameters after optimization compared with the experimental T = 10 K parameters. Method: B3LYP-D*/TZP.

	Z1 expanded to Z3 optimized	Original Z3 optimized	Experimental Z3
a [Å]	23.932	23.770	23.7685
b [Å]	3.527	3.517	3.5428
c [Å]	28.78	28.672	28.7556
B[°]	113.60	112.809	112.831
Energy [*au*] Per unit cell	-8675.171914	-8675.169812	

Calculation of entropy due to disorder

Two types of disorder stabilize the GAM structure by increasing the residual entropy.

We estimate these entropy contributions by calculating the *entropy of mixing*.

Disorder between three sites corresponding to the low-temperature modulation

According to a refinement of occupancies of the three molecular sites constrained at their positions transferred from the Z'=3 structure by transformation to the Z'=1 cell, the sites are occupied in the ratios x1 = 0.20, x2 = 0.42 and x3 = 0.38;

$$\Delta S_{mix} = -nR\left(x_1 \cdot \ln\left(x_1\right) + x_2 \cdot \ln\left(x_2\right) + x_3 \cdot \ln\left(x_3\right)\right) = 8.31 \cdot (0.8 \cdot \ln(0.8) + 0.1 \cdot \ln(0.1) + = 8.8$$

J mol^{-1} K^{-1}.

This contribution gives a 2.6 kJ mol^{-1} stabilizing contribution to the room temperature free energy.

Hydrogen bond pattern disorder

The reverse hydrogen bond patterns are not equally likely.
Assuming that the patterns appear in a ratios of 20% (x_1=0.2) to 80% (x_2=0.8) ratio, the entropy of mixing becomes

$$\Delta S_{mix} = -nR\left(x_1 \cdot \ln\left(x_1\right) + x_2 \cdot \ln\left(x_2\right)\right) = -8.31 \cdot (0.8 \cdot \ln(0.8) + 0.2 \cdot \ln(0.2)) = 4.2 \text{ J mol}^{-1} \text{ K}^{-1}$$

At room temperature this becomes a difference in free energy of 4.2 J mol^{-1} K^{-1} × 300 K = 1.2 kJ mol^{-1}.

References

1. H. D. Clarke, K. K.Arora, Ł. Wojtas, M. J. Zaworotko, *Cryst. Grow. Des.,* 2011**, 11**, 964–966.
2. APEX2/SAINT software, *Bruker AXS Inc: Madison*, Wisconsin, 2012.
3. H. R. Blessing, *J. Appl. Cryst.*, 1997, **30**, 421.
4. H. R. Blessing, *Rev. Cryst.* 1987, **1**, 3–58.
5. M. G.Sheldrick, *Acta Cryst.,*2008, **A64**, 112.
6. L. J. Farrugia, *J. Appl. Cryst.*, 2012, **45**, 849-854.
7. N. K. Hansen, P. Coppens, *Acta Cryst.*, 1978, **A34**, 909.
8. A. Volkov, P. Macchi, L. J. Farrugia,C. Gatti, P. R. Mallinson, T. Richter, T. Koritsanszky, *XD-2006*, 2006.
9. A. Ø. Madsen., J. Appl. Cryst., 2006, 39, 757-758.
10. Z. Su, P. Coppens, *Acta Cryst.,*1998, **A54**, 646.
11. F. C .Bunge, A. J. Barrientos, V. A. Bunge, *At. Data Nucl. Data Tab.*,1993, **53**, 113.
12. T. Gelbrich, B. M. Hursthouse, *Cryst. Eng. Comm.*, 2005, **7**, 324-336.
13. F. P. A. Fabbiani, B. Dittrich, A. J. Florence, T. Gelbrich, M. B. Hursthouse, ,W. F. Kuhs, N. Shankland, H. Sowa, *Cryst. Eng. Comm.,*2009,**11**, 1396 – 1406.
14. R. Dovesi, V. R. Saunders, C. Roetti, R. Orlando, C. M. Zicovich-Wilson, F. Pascale, B. Civalleri, K. Doll, N. M. Harrison, I. Bush, P. D'Arco,; M. Llunell, *CRYSTAL09*, University of Torino: Torino, 2009.
15. (a) A. D. Becke, *Phys. Rev. A.*,1988, **38**, 3098. (b) A. D Becke, *J. Chem. Phys.,*1993, **98**, 5648. (c) C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B.,*1988, **37**, 785.

16. A. Schafer, H. Horn, R. Ahlrichs, *J. Chem. Phys*., 1992, **97**, 2571.
17. S. Grimme, *J. Comp. Chem.,* 2006*,* **27**, 1787–1799.
18. B. Civalleri, C. M. Zicovich-Wilson, L. Valenzano, P. Ugliengo, *Cryst. Eng. Comm.,* 2008, **10,** 405–410.