Electronic Supplementary Material (ESI) for ChemComm.

Electronic Supporting Information

Theoretically derived thermodynamic properties can be improved by the refinement of low-frequency modes against X-ray diffraction data

Anna A. Hoser, Marcin Sztylko, Damian Trzybiński, Anders Ø. Madsen

1. Datasets and single-crystal X-ray diffraction measurement details

Single-crystal X-ray diffraction analysis of glycine polymorphs

High-quality single crystals of **α-gly** and **β-gly** were selected for X-ray diffraction (XRD) experiments at T = 100(2) K. The crystals were mounted with paratone-N oil using a glass capillary (**α-gly**) and MiTeGen micromount (**β-gly**) (Fig. S1). Diffraction data were collected using an Agilent Technologies SuperNova single-source diffractometer with Mo $K\alpha$ radiation (λ = 0.71073 Å) using CrysAlis RED software (CrysAlis CCD and CrysAlis RED, Oxford Diffraction, Oxford Diffraction Ltd, Yarnton, 2008). Multi-scan empirical absorption correction using spherical harmonics, as implemented in the SCALE3 ABSPACK scaling algorithm, was applied (CrysAlis CCD and CrysAlis RED, Oxford Diffraction, Oxford Diffraction Ltd, Yarnton, 2008). The structure was determined using the SHELXT program[1]. The structures were solved by direct methods, and then successive least-squares refinement by the full-matrix least-squares method was performed on F^2 using the SHELXL program[1]. All the H atoms linked to N atoms were located on the Fourier difference electron density map and refined using U_{iso}(H) = 1.2U_{eq}(N). The N–H bond lengths were maintained at 0.87 Å. Remaining H atoms were positioned geometrically with a C–H bond length of 0.93 Å and constrained to ride on their parent atoms with U_{iso}(H) = 1.2U_{eq}(C).

Supplementary figures and tables

Table S1. Crystal data and structure refinement details for **α-gly** and **β-gly**.

Identification code	**α-gly**	**β-gly**
Empirical formula	$C_2H_5NO_2$	$C_2H_5NO_2$
Formula weight	75.07	75.07
Temperature/K	100(2)	100(2)
Crystal system	monoclinic	monoclinic
Space group	$P2_1/n$	$P2_1$
a/Å	5.0867(2)	5.0757(2)
b/Å	11.7982(4)	6.1737(2)
c/Å	5.4598(2)	5.3882(2)
α/°	90	90
β/°	111.983(5)	113.422(4)
γ/°	90	90
Volume/Å^3	303.84(2)	154.931(11)
Z	4	2
ρ_{calc}g/cm^3	1.641	1.609
μ/mm^{-1}	0.146	0.143
F(000)	160.0	80.0
Crystal size/mm^3	0.26 × 0.12 × 0.09	0.69 × 0.10 × 0.06
Radiation	Mo $K\alpha$ (λ = 0.71073)	Mo $K\alpha$ (λ = 0.71073)
2Θ range for data collection/°	6.908 to 61.01	8.244 to 60.978
Index ranges	-7 ≤ h ≤ 7, -16 ≤ k ≤ 16, -7 ≤ l ≤ 7	-7 ≤ h ≤ 7, -8 ≤ k ≤ 8, -7 ≤ l ≤ 7
Reflections collected	4778	9425
Independent reflections	932 [R_{int} = 0.0265,	942 [R_{int} = 0.0314,

	R_{sigma} = 0.0198]	R_{sigma} = 0.0131]
Data/restraints/parameters	932/3/55	942/4/55
Goodness-of-fit on F^2	1.078	1.100
Final *R* indexes [$I \geq 2\sigma(I)$]	R_1 = 0.0317, wR_2 = 0.0779	R_1 = 0.0229, wR_2 = 0.0638
Final *R* indexes [all data]	R_1 = 0.0386, wR_2 = 0.0841	R_1 = 0.0233, wR_2 = 0.0642
Largest diff. peak/hole / e $Å^{-3}$	0.50/-0.31	0.41/-0.19
Flack parameter	–	0.3(4)

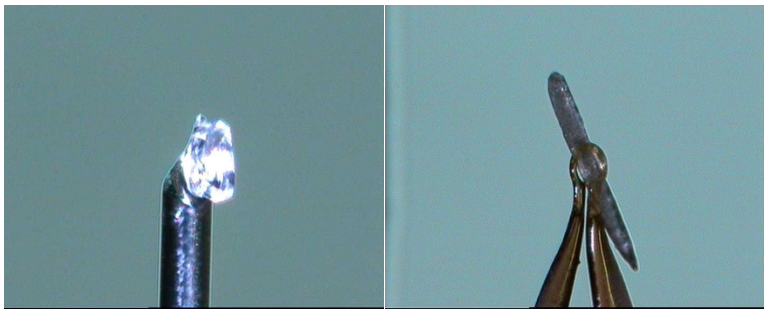

Figure S1. Single crystals of **α-gly** (left) and **β-gly** (right) selected for XRD analysis.

Table S2. Fractional atomic coordinates (×10^4) and equivalent isotropic displacement parameters ($Å^2$ × 10^3) of **α-gly**. U_{eq} is defined as one-third of of the trace of the orthogonalised U_{IJ} tensor.

Atom	*x*	*y*	*z*	U_{eq}
C1	5696(2)	6250.2(8)	5658.6(18)	8.18(19)
C2	5584(2)	6458.0(8)	2859.6(18)	9.05(19)
N1	7955.7(18)	5887.1(7)	2404.8(16)	9.49(18)
O1	8024.9(15)	5937.4(6)	7364.0(13)	10.85(18)
O2	3451.0(15)	6423.6(6)	6064.7(14)	11.71(18)

Table S3. Anisotropic displacement parameters ($Å^2$ × 10^3) of **α-gly**. Anisotropic displacement factor exponent takes the form $-2\pi^2[h^2a^{*2}U_{11} + 2hka^*b^*U_{12}+\ldots]$.

Atom	U_{11}	U_{22}	U_{33}	U_{23}	U_{13}	U_{12}
C1	10.5(4)	6.4(4)	8.3(4)	-1.0(3)	4.3(3)	-1.1(3)
C2	8.4(4)	11.5(4)	8.1(4)	1.0(3)	4.0(3)	1.9(3)
N1	9.7(4)	12.1(4)	7.4(4)	0.6(3)	4.1(3)	1.5(3)
O1	10.5(3)	13.9(3)	8.1(3)	1.1(2)	3.5(3)	1.7(3)
O2	9.4(3)	16.1(4)	11.2(3)	-1.7(3)	5.7(3)	-0.1(3)

Table S4. Bond lengths for **α-gly**.

Atom	Atom	Length/Å	Atom	Atom	Length/Å
C1	C2	1.5275(13)	C1	O2	1.2590(12)
C1	O1	1.2572(11)	C2	N1	1.4817(12)

Table S5. Valence angles of **α-gly**.

Atom	Atom	Atom	Angle/˚	Atom	Atom	Atom	Angle/˚
O1	C1	C2	117.40(8)	O2	C1	C2	116.96(8)
O1	C1	O2	125.63(9)	N1	C2	C1	111.52(7)

Table S6. Torsion angles of **α-gly**.

A	B	C	D	Angle/˚	A	B	C	D	Angle/˚
O1	C1	C2	N1	19.48(12)	O2	C1	C2	N1	-161.51(8)

Table S7. Hydrogen atom coordinates (×10^4 Å) and isotropic displacement parameters (×10^3 Å^2) of **α-gly**.

Atom	*x*	*y*	*z*	U_{eq}
H2B	3797.72	6178.13	1596.67	11
H2A	5672.51	7266.45	2576.16	11
H1A	7890(30)	5160(8)	2720(30)	14
H1B	9630(20)	6148(11)	3460(20)	14
H1C	7890(30)	5980(12)	757(19)	14

Table S8. Fractional atomic coordinates (×10^4) and equivalent isotropic displacement parameters (×10^3 Å^2) of **β-gly**. U_{eq} is defined as one-third of the trace of the orthogonalised U_{IJ} tensor.

Atom	*x*	*y*	*z*	U_{eq}
C1	6365(2)	5374(2)	5601(2)	8.6(2)
C2	6131(2)	5500(2)	2684(2)	9.6(2)
N1	8531(2)	4355(2)	2350(2)	9.1(2)
O1	8829(2)	5081.7(18)	7421.5(18)	12.0(2)
O2	4095.9(19)	5650.5(17)	5971.1(18)	10.8(2)

Table S9. Anisotropic displacement parameters (Å^2 × 10^3) of **β-gly**. The anisotropic displacement factor exponent takes the form $-2\pi^2[h^2a^{*2}U_{11} + 2hka^*b^*U_{12}+\ldots]$.

Atom	U_{11}	U_{22}	U_{33}	U_{23}	U_{13}	U_{12}
C1	11.5(5)	7.5(6)	8.1(4)	-0.3(4)	5.2(4)	-1.1(5)
C2	9.6(5)	12.5(7)	7.3(5)	0.7(4)	4.0(4)	2.3(5)
N1	9.1(4)	11.4(5)	7.4(5)	-0.4(4)	3.8(4)	0.5(4)
O1	9.3(4)	19.1(6)	7.7(4)	-0.1(3)	3.5(3)	-1.3(3)
O2	10.2(4)	13.0(5)	11.3(4)	-1.3(4)	6.4(3)	0.1(4)

Table S10. Bond lengths for **β-gly**.

Atom	Atom	Length/Å	Atom	Atom	Length/Å
C1	C2	1.5303(16)	C1	O2	1.2565(14)
C1	O1	1.2582(15)	C2	N1	1.4801(16)

Table S11. Valence angles of **β-gly**.

Atom	Atom	Atom	Angle/˚	Atom	Atom	Atom	Angle/˚
O1	C1	C2	117.10(10)	O2	C1	O1	125.83(11)
O2	C1	C2	117.01(10)	N1	C2	C1	111.65(9)

Table S12. Torsion angles of **β-gly**.

A	B	C	D	Angle/˚	A	B	C	D	Angle/˚
O1	C1	C2	N1	24.55(16)	O2	C1	C2	N1	-158.00(11)

Table S13. Hydrogen atom coordinates (×10^4 Å) and isotropic displacement parameters (×10^3 Å^2) of **β-gly**.

Atom	*x*	*y*	*z*	U_{eq}
H2A	4325.86	4861.61	1485.73	12
H2B	6132.17	7006.8	2175.6	12
H1A	10190(30)	4800(30)	3490(30)	11
H1B	8590(40)	4540(30)	760(30)	11
H1C	8370(40)	2980(20)	2570(40)	11

2. Theoretical calculations

Periodic *ab-initio* density functional theory (DFT) calculations were performed for the selected systems using the *CRYSTAL17* program with the B3LYP[2] functional[3]. The standard 6-31G(d,p) basis set was applied. It has been used extensively and gives reasonable relative stabilities of related systems[4, 5]; it has been used for the estimation of vibrational frequencies. We optimised the geometry (not the cell parameters, only the coordinates, starting from the experimental geometry). Subsequently, we calculated the normal modes and their frequencies at the Γ point of the Brillouin zone. The convergence criteria for geometry optimisation were set to the default for frequency calculations using the *PREOPTGEOM* keyword. As we optimised only the coordinates, the frequencies were calculated using the cell parameters from XRD measurements. Input for *CRYSTAL17* frequency calculations can be easily created using a cif2crystal routine (www.shade.ki.ku.dk/docs/cif2crystal/cif2crystal.html).[6]

For the glycine polymorphs, we conducted 2 × 2 × 2 supercell DFT frequency calculations in CRYSTAL17. Before the calculations, we optimised the geometry using B3LYP and the standard 6-31G(d,p) basis set.

3. Normal-mode refinement

The idea of normal mode refinement was introduced in our previous work.[7] Here, we would like to briefly describe the model which we are using during normal mode refinement and focus on

differences between the initial approach[7] and current version implemented in the NoMoRe webserver.

In both approaches the normal-mode coordinates and frequencies are derived from the dynamical matrix, which is obtained from DFT calculations (in our case using the CRYSTAL17 program). The atomic mean square displacement matrix (MSD) can be calculated as [8]:

$$B_{atom}(k) = \frac{1}{Nm_k}\sum_{jq}\frac{E_j(q)}{\omega_j^2(q)}e(k\,|\,jq)e^*(k|jq))^T, \qquad [1]$$

where $e(k|jq)$ denotes the *k*th component of the eigenvector ***e***(*j****q***) and corresponds to atom *k* in normal mode *j* along the wavevector ***q***. $\omega_j(q)$ is the frequency of mode *j*, m_k is the mass of atom *k*, and E_j(***q***) is the energy of the mode, which is given by

$$E_j(q) = \hbar\omega_j(q)\left\{\frac{1}{2} + \frac{1}{\exp\left[\hbar\frac{\omega_j(q)}{k_BT}\right]-1}\right\}, \qquad [2]$$

where k_B is the Boltzmann constant.

From these atomic mean square displacement matrices (B_{atom}) one can compute the atomic displacement parameter (ADP) matrices by a transformation of the coordinate basis.

As it was already mentioned, equation 1 represents a relation between frequencies and MSDs (so also ADPs).The mean square displacements, and thus the ADPs are mean values of vibration over the Brillouin zone (BZ). Since the dispersion – the ***q*** dependence of the frequencies – is not accessible from the elastic scattering experiments, we use the approximation that the MSD matrices are the summation of all phonon branches, but without ***q*** dependence:

$$B_{atom}(k) = \frac{1}{Nm_k}\sum_{j}\frac{E_j}{\omega_j^2}e(k\,|\,j)e^*(k|j))^T. \qquad [3]$$

$$E_j = \hbar\omega_j\left\{\frac{1}{2} + \frac{1}{\exp\left[\hbar\frac{\omega_j}{k_BT}\right]-1}\right\}. \qquad [4]$$

The parameters refined against the observed structure factors are those usually refined in a standard structure refinement (the scaling parameters, atomic coordinates) but the refinement of the ADPs is replaced by the refinement of the frequencies ω_j of the normal mode vibrations. We include all

normal modes in the calculations of the DW factors, but we refine only a small subset of the modes with the lowest frequencies. During the refinement the normal mode vectors [$e(k|j)$ or $e(k|jq)$] are not altered and remain as computed.

Refinement details

All the initial normal-mode vectors and their frequencies at the Gamma point of BZ are obtained from the CRYSTAL17 calculations. A scaling factor of 1.0 is assigned to each frequency. Next the ADPs for all atoms (including hydrogen atoms) are calculated and are submitted with the coordinates to SHELXL[9], which calculates the structure factors and all the statistics and discrepancy factors (R, wR2). Next, the selected frequencies are optimised by refining the frequency scaling factors against the diffraction data to minimise wR2. Here, we use different optimization method than was the case of our initial refinement procedure[7]. Because NoMoRe is written in Python 3, we used the least-squares minimisation routine from the SciPy Python package (scipy.optimize.least_squares). We estimate the Jacobian matrix by a finite difference method. In this approach the Jacobian matrix is numerically estimated by a three point scheme. We used a trust region reflective algorithm The trust region for the frequency scaling parameters was set to the range 0–6. The refinements became considerably more stable, and the convergence was improved compared to the previous approach; moreover, we obtained the estimated standard deviations of the refined frequencies.

Note that in the current version of NoMoRe, the ADPs are not refined at any stage of the refinement procedure. The ADPs are always computed directly from the normal-mode frequencies, which are obtained by *CRYSTAL17* calculations for high-frequency modes and refined against XRD data for low-frequency modes.

NoMoRe web

The NoMoRe web server is written in Django and is a user-friendly server on which NoMoRe refinement can be conducted free of charge.

We used the **nomore.chem.uw.edu.pl** web server (a version of the *NoMoRe* Python program) to conduct refinements for urea, the glycine polymorphs, and 4-hydroxyacetophenone (HAP). For each dataset, we specified the number of frequencies to be refined and the measurement temperature (Table S13). Here we present the results of two types of refinement for each compound:

1. one in which only the acoustic modes are refined (the results are presented in the main manuscript),
2. one in which the acoustic mode frequencies and a few additional frequencies were refined, where we tried to find the frequency gap, at which the internal vibrations can be distinguished from external molecular vibrations, and we refined only low frequencies (FG model).

In each refinement, we started with the frequencies obtained by *CRYSTAL17* calculations. The translational acoustic modes, which are zero at the Γ point, were initialised at 50 cm^{-1}. Refinements for urea, the glycine polymorphs, and HAP were conducted against single-crystal XRD measurements from 100 K until convergence occurred.

A slightly different strategy was applied for benzoic acid (BA). This structure is disordered; therefore, we could not apply the standard version of *webNoMoRe*. Also, we could not simply compute the ADPs and refine the frequencies because the ADPs obtained from frequencies refined against XRD data for disordered structure would be biased. Therefore, we used the *mta-NoMoRe* version, in which we fit frequencies to the ADPs obtained from measurements at different temperatures. We already applied the procedure for the polymorphs of dimethyl 3,6-dichloro-2,5-dihydroxyterephthalate[10]. In that case, we could not fit many frequencies to the ADPs, as we were not refining against structure factors but rather fitting to the ADPs, and thus, we did not have as many data points as we did for the refinements against structure factors.

Table S13. Number of refined frequencies and final *wR* after NoMoRe. For all systems, we conducted two types of refinement: first by refining only the acoustic modes and second by refining additional frequencies [we found the frequency gap and refined only frequencies below the gap (FG model).]

Compound	**Number of refined frequencies**	**Final *wR* after NoMoRe**
Urea	3 *8*	0.30 *0.17*
α-Glycine	3 *8*	0.10 *0.098*
β-Glycine	3 *7*	0.08 *0.07*
HAP	3 *32*	0.25 *0.15*
Compound	**Number of fitted frequencies**	**Final *wR* after NoMoRe**
BA	3 *6*	0.57 *0.26*

Table S14. Acoustic mode frequencies obtained from NoMoRe values
(initial value of these frequencies is 50 cm^{-1}, s.u. were obtained from covariance matrix after least-squares)

Compound	**Initial/cm^{-1}**	**Only acoustic modes refined/cm^{-1}**	**FG model/ cm^{-1}**
Urea	50	75	83
	50	44	48
	50	44	48
α-Glycine	50	61(0.93)	59(2.7)
	50	**43(0.21)**	**45(0.26)**
	50	69(1.13)	75(3.95)
β-Glycine	50	60(0.19)	53(0.32)
	50	**45(0.12)**	**46(0.19)**
	50	67(0.23)	81(0.35)
HAP	50	22	36
	50	29	50
	50	20	27
BA	50	38	24
	50	25	46
	50	26	21

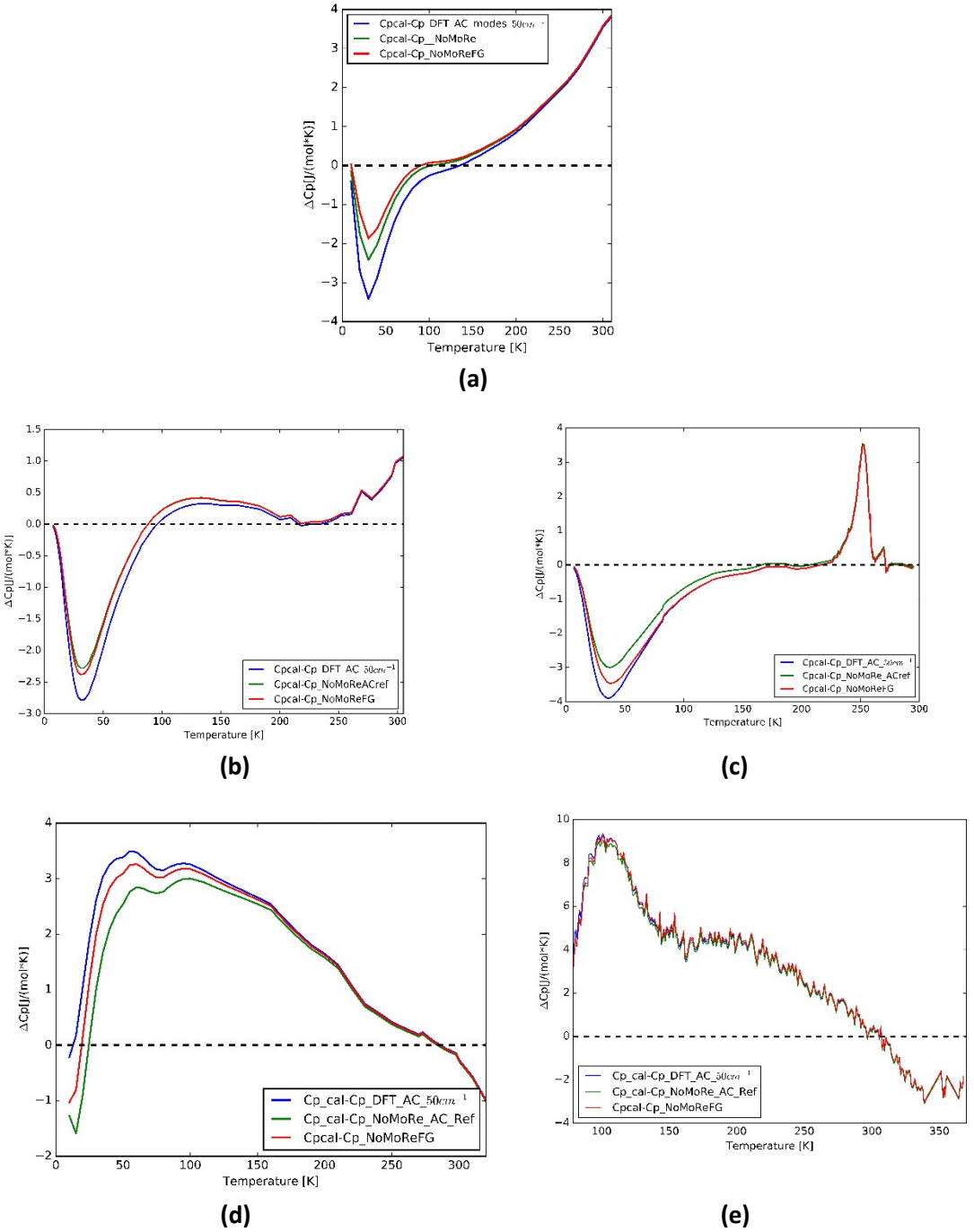

Figure S2 Comparison of two different NoMoRe models – AC and FG. Difference between heat capacity from calorimetry and: DFT Γ-point calculations with acoustic mode frequencies of 50 cm-1 (blue line), NoMoRe when only acoustic modes are refined (green line) and NoMoRe FG model (red line) (a) urea, (b) α glicyne, (c) β glicyne, (d) benzoic acid, (e) 4'-hydroxyacetophenone

Computational cost of NoMoRe

The most expensive (in terms of computational time) part of NoMoRe is the periodic ab-initio DFT frequency calculations, which are the starting point for the refinement. For such frequency calculations the computational cost will strongly depend on the number of atoms. For example, we can compare the computational costs of our approach with the full ab-initio lattice-dynamical approach in the quasi-harmonic approximation (QHA): In the NoMoRe case we need to optimize the structure and calculate the frequencies for one volume only (corresponding to the experimental cell parameters). In the case of the QHA, the geometry optimization and frequency calculations should be done for a few different cell volumes (contracted and expanded, according to CRYSTAL manual 4 volumes are the minimum). Moreover, we are utilizing only calculations from Gamma point, which is much faster than conducting supercell calculations (as it was done for urea by Erba[11]).

Heat capacity estimation:

To estimate the heat capacity we applied the procedure proposed by Aree and Bürgi[12]. First we calculated $\bar{C}_v$ as:

$$\bar{C}_v = \sum_{i=1}^{3} 3R\left(\frac{T}{1.437\bar{\nu}_{D,i}}\right)^3 \int_0^{1.437\bar{\nu}_{D,i}/T} \frac{x^4 e^x}{(e^x-1)^2}dx + \sum_j mR\left(\frac{\theta_{E_j}}{T}\right)^2 \frac{exp\left(\frac{\theta_{E_j}}{T}\right)}{\left[exp\left(\frac{\theta_{E_j}}{T}\right)-1\right]^2},$$

where R stands for the gas constant and T is the absolute temperature. The first part of this equation refers only to acoustic frequencies, and $\nu_{D,i}$ is the Debye frequency, defined as $\nu_{D,i}$ =1.616$\nu_{E,i}$, whereas the second part the sum runs over the unique normal modes j, m is their degeneracy, θ_{E_j}=1.437 ν_{eff}(j), $\theta_{E_j}$is the Einstein temperature and ν is the wavenumber in cm^{-1}. The frequencies of intramolecular vibrations are overestimated before we calculated C_v we applied a scale factor correction for high vibrational frequencies: the frequencies higher than 500 cm^{-1} were multiplied by 0.956. The difference between C_v and C_p has been approximated with the Nernst–Lindemann relation which is based on two quantities: the melting point, and a universal constant (1.63 × 10^{-2} K mol cal^{-1}).

References:

1. G. Sheldrick, *Acta Crystallographica Section A*, 2015, **71**, 3-8.
2. C. R. Lee, W. Yang and R. G. Parr, *Phys. Rev. B*, 1988, **37**, 785.
3. R. Dovesi, V. R. Saunders, C. Roetti, R. Orlando, C. M. Zicovich-Wilson, F. Pascale, B. Civalleri, K. Doll, N. M. Harrison, I. J. Bush, P. D'Arco, M. Llunell, M. Causà, Y. Noël, L. Maschio, A. Erba, M. Rerat and S. Casassa, *Journal*, 2017.
4. N. Wahlberg, P. Ciochoń, V. Petriĉek and A. Ø. Madsen, *Crystal Growth & Design*, 2014, **14**, 381-388.
5. C. Quarti, A. Milani, B. Civalleri, R. Orlando and C. Castiglioni, *The Journal of Physical Chemistry B*, 2012, **116**, 8299-8311.
6. A. O. Madsen and A. A. Hoser, *Journal of Applied Crystallography*, 2014, **47**, 2100-2104.
7. A. A. Hoser and A. O. Madsen, *Acta Crystallographica Section A*, 2016, **72**, 206-214.
8. A. Ø. Madsen, in *Modern Charge-Density Analysis*, Dordrecht: Springer Netherlands., 2011, DOI: doi:10.1007/978-90-481-3836-4_3, pp. pp. 133–163.
9. G. M. Sheldrick, *Acta Cryst.*, 2008, **A64**, 112-122.
10. P. M. Kofoed, A. A. Hoser, F. Diness, S. C. Capelli and A. O. Madsen, *IUCrJ*, 2019, **6**.
11. A. Erba, J. Maul and B. Civalleri, *Chemical Communications*, 2016, **52**, 1820-1823.
12. T. Aree and H.-B. Buergi, *Journal of Physical Chemistry B*, 2006, **110**, 26129-26134.