Accurate H-atom parameters for the two polymorphs of L-histidine at 5, 105 and 295 K

Accurate Hydrogen Atom Parameters for the Two Polymorphs of *L*-Histidine at 5, 105 and 295 K

Giulia Novelli[ab]*, Charles J. McMonagle[cd], Florian Kleemiss[e], Michael Probert[d], Horst Puschmann[f], Simon Grabowsky[e], Helen E. Maynard-Casely[b], Garry J. McIntyre[b] and Simon Parsons[a]

[a]EaStCHEM School of Chemistry, University of Edinburgh, King's Buildings, West Mains Road, Edinburgh, Scotland, EH9 3FJ, United Kingdom

[b]Australian Centre for Neutron Scattering, Australian Nuclear Science and Technology Organisation, New Illawarra Road, Lucas Heights NSW 2234, Australia

[c] Swiss-Norwegian Beamlines at the European Synchrotron Radiation Facility, 71, Avenue des Martyrs, Grenoble, 38000, France

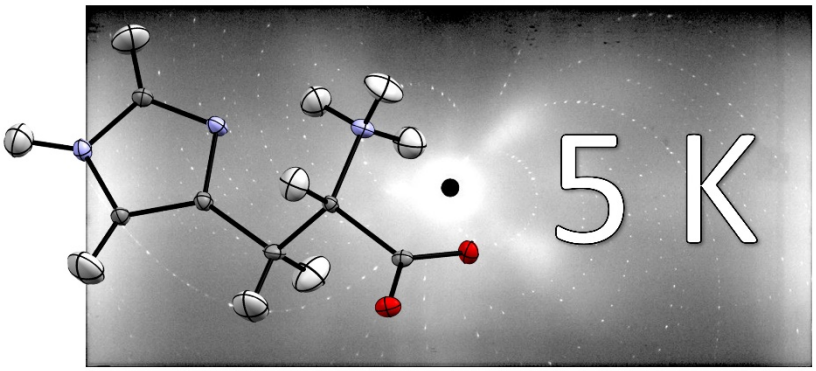
[d]School of Natural and Environmental Sciences, Newcastle University, Bedson Building, Kings Road, Newcastle upon Tyne, England, NE1 7RU, United Kingdom

[e]Department of Chemistry, Biochemistry and Pharmaceutical Sciences, University of Bern, Freiestrasse 3, 3012 Bern, Switzerland

[f]OlexSys Ltd, Chemistry Department, Durham University, South Road, Durham, England, DH1 3LE, United Kingdom

Correspondence email: g.novelli@ed.ac.uk

Synopsis Single-crystal neutron and X-ray diffraction data for the orthorhombic and monoclinic polymorphs of the amino acid *L*-histidine have been collected at 5, 105 and 295 K. The neutron structural models are compared with those refined with the new Hirshfeld atom refinement method implemented in *NoSpherA2*.



Abstract The crystal structure of the monoclinic polymorph of the primary amino acid *L*-histidine has been determined for the first time by single-crystal neutron diffraction, while that of the orthorhombic polymorph has been reinvestigated with an un-twinned crystal, improving experimental precision and accuracy. For each polymorph, neutron diffraction data were collected at 5, 105 and 295 K. Single-crystal X-ray diffraction experiments were also performed at the same temperatures. The two polymorphs, whose crystal packing is explained by intermolecular interaction energies calculated using the Pixel method, show differences in the energy and geometry of the H-bond formed along the *c*-direction. Taking advantage of the X-ray diffraction data collected at 5 K, the precision and accuracy of the new Hirshfeld atom refinement method implemented in *NoSpherA2* were probed choosing different settings of functionals and basis sets, along with the use of explicit clusters of molecules and enhanced rigid-body restraints. Equivalent atomic coordinates and anisotropic displacement parameters were compared and found to agree well with those obtained from the corresponding neutron structural models.

Keywords: amino acids; H atom parameters; neutron Laue diffraction; Hirshfeld atom refinement.

1. Introduction

Over the past 50 years, 17 of the 20 naturally-occurring amino acids, which are the building blocks of proteins, have been subjected to crystal structure determination by neutron diffraction. While the first 16 members of this collection were refined in the 1970s at the U.S.-American Brookhaven National Laboratory and the Indian Atomic Energy Laboratory, the 17th amino acid *L*-leucine, was studied only five years ago by Binns and colleagues on the Koala Laue diffractometer at ANSTO (Binns *et al.*, 2016). These neutron crystal structures continue to be topical, with glycine and *L*-alanine being the most popular with an average of ~225 citations in the past 20 years (Jönsson & Kvick, 1972, Lehmann *et al.*, 1972b). Many of the studies have been carried out to link the neutron crystal structures to the investigation of polymorphism (Görbitz, 2015), to condensed-phase NMR analysis (Gervais *et al.*, 2005, Thurber & Tycko, 2008), to periodic density functional theory calculations and as initial structures for computer modelling of proteins (Császár & Perczel, 1999, Kassab *et al.*, 2000, Berman, 2008). To date, the neutron crystal structures of *L*-isoleucine, *L*-methionine, *L*-tryptophan and the monoclinic polymorph of *L*-histidine have not been reported in the literature.

Filling the gaps in the structural database of amino acids will continue to be of high importance, especially in the context of polymorphism (Cruz-Cabeza *et al.*, 2015). In this study, we focus our attention on the two polymorphs of *L*-histidine. This amino acid contains an imidazole side chain able to coordinate Fe^{2+} efficiently in important globular proteins like haemoglobin and myoglobin. Histidine has been recently taken as a model in a combined electron diffraction, solid-state NMR, and first-principles quantum-calculations study for the understanding of H-bonding networks in nanocrystals (Guzmán-Afonso *et al.*, 2019). The amino acid is also currently being investigated as a supplement in

solid form for improving treatments based on the anti-cancer drug methotrexate or for supporting long-term neuroprotection after cerebral ischemia (Liao *et al.*, 2015, Kanarek *et al.*, 2018), and in all such applications the control and understanding of polymorphism are critical.

The X-ray crystal structures of the two polymorphs of *L*-histidine were first determined at ambient temperature by Madden and co-authors in the early 1970s (Madden, McGandy & Seeman, 1972, Madden, McGandy, Seeman, *et al.*, 1972), and are available in the Cambridge Structural Database (Groom *et al.*, 2016) as entries LHISTD01 and LHSITD10. The crystal of the monoclinic form showed lamellar twinning and the structure could not be refined below $R_1 = 0.10$ (R_1 defined as $\sum|F_o - F_c|/\sum|F_o|$). The structure was reinvestigated 20 years later by Averbuch-Pouchot with an un-twinned crystal ((Averbuch-Pouchot, 1993), LHISTD04), leading to an improvement in experimental precision and accuracy and to the location of H atoms that were missing in the previous study. Concurrently with Madden's work, Lehmann and colleagues at the Brookhaven National Laboratory refined the ambient-temperature neutron crystal structure of the orthorhombic polymorph ((Lehmann *et al.*, 1972a), LHISTD13). However, once again, crystal twinning necessitated the rejection of a quarter of the data set and the precision of the structure was lower than that obtained in the previous X-ray determination. The location of the H atoms was also less definitive than is usual in neutron diffraction analysis. More recently, the effect of high pressure on the two polymorphs of *L*-histidine has been investigated using single-crystal X-ray diffraction (Novelli *et al.*, 2020). Despite their similarity in crystal packing and intermolecular interaction energies at ambient conditions, the way in which the polymorphs respond to pressure is different. The orthorhombic form undergoes a first-order phase transition at 4.5 GPa and is significantly less compressible than the monoclinic form, which also undergoes a phase transition at 3.1 GPa.

Accurate crystal structures of amino acids derived from diffraction experiments in the ultra-low temperature regime are uncommon. A search in the Cambridge Structural Database yielded seven amino acids for which data collections have been performed below 20 K, but none of them was investigated below 10 K (see Table S1 in the Supporting Information). The oldest study includes a revision of crystal symmetry, where (Weisinger-Lewin *et al.*, 1989) investigated a solid solution of *L*-asparagine monohydrate and aspartic acid. The X-ray crystal structure of *L*-threonine, which was previously determined at ambient temperature by (Shoemaker *et al.*, 1950), was reinvestigated at 12 K by (Janczak *et al.*, 1997) for the determination of more accurate molecular geometry and of all the H atom positions. Many of these low-temperature studies aimed to determine H atom parameters for use in charge-density studies (Flaig *et al.*, 1998, Arnold *et al.*, 2000, Dittrich *et al.*, 2005, Destro *et al.*, 2008, Mondal *et al.*, 2012, Woińska *et al.*, 2016). Accurate synchrotron X-ray and neutron crystal-structure determinations were also needed to analyse the dynamics and thermodynamics of the α, β, γ - forms of glycine (Aree *et al.*, 2012, Aree *et al.*, 2013, Aree *et al.*, 2014), showing that the free-energy differences between polymorphs are determined primarily by differences in lattice and zero-point vibrational energies.

In this study, we aimed to determine the crystal structures of the two polymorphs of the amino acid *L*-histidine at 5 K. The experiments were performed using the Koala single-crystal Laue diffractometer at ANSTO. Additional data collections at 105 and 295 K were used to compare the effect of temperature on the geometry and intermolecular interaction energies in the two polymorphs. Laue methods cannot be used to determine precise absolute values of the unit-cell dimensions, therefore additional single-crystal X-ray diffraction experiments were performed at the same temperatures using the XIPHOS I diffraction facility (Newcastle University, UK, (Probert *et al.*, 2010, McMonagle & Probert, 2019)).

The X-ray crystal structures were refined with the new Hirshfeld atom refinement (HAR) method (Jayatilaka & Dittrich, 2008, Capelli, Bürgi, Dittrich*, et al.*, 2014) implemented in *NoSpherA2* (Kleemiss *et al.*, 2021) and compared with their respective neutron models. Thanks to the use of non-spherical atomic form factors, HAR is known to produce bond lengths involving H atoms with the same accuracy and precision from X-ray diffraction data as from neutron-diffraction studies for organic molecules (Woińska *et al.*, 2016). However, this does not hold for H atom anisotropic displacement parameters (ADPs) which are normally less accurate after HAR and can become skewed, oblate or non-positive definite (Dittrich *et al.*, 2017, Malaspina *et al.*, 2020). In the new *NoSpherA2* Implementation of HAR inside *Olex2* (Kleemiss *et al.*, 2021), the use of restraints for H atom ADPs within HAR has become possible. This new possibility will be tested here by applying the enhanced rigid-body restraint according to (Thorn *et al.*, 2012).

2. Experimental

2.1. Crystallisation

Crystalline *L*-histidine-h_9 ($C_6H_9N_3O_2$) and ethanol (C_2H_5OH) were purchased from Sigma-Aldrich, and sodium hexametaphospate $(NaPO_3)_6$ was purchased from Fisher Scientific and used as received.

L-histidine (0.71 g) was dissolved in deionised water (10 cm^3) at ambient temperature and the solution was split equally between two beakers. Colourless, needle-shaped crystals of the orthorhombic polymorph were obtained by slow diffusion of ethanol into one beaker. Addition of sodium hexametaphosphate (0.4g) to the other beaker, followed by slow ethanol diffusion, led to the formation of colourless and pleated crystals of the monoclinic polymorph (Averbuch-Pouchot, 1993).

Crystals of both the orthorhombic and monoclinic polymorphs were often found to be twinned or of poor quality. The monoclinic form was never found in isolation in the crystallisation vials, and screening by X-ray diffraction of suitable candidates was necessary to discriminate between the two polymorphs and identify un-twinned samples.

2.2. Data collection

2.2.1. Single-crystal neutron Laue diffraction

Several candidate crystals of each polymorph were screened by acquiring single-crystal neutron Laue patterns on the Koala Laue Diffractometer at ANSTO to select a crystal that gave small, intense Laue spots. The crystals chosen had dimensions of approximately 0.35 × 0.30 × 0.15 mm^3 and 0.45 × 0.35 × 0.30 mm^3 for the orthorhombic and monoclinic polymorphs, respectively. Each sample was attached, using a dab of fluorinated grease, to a thin aluminium plate whose normal was ~45° to the vertical instrument axis. An aluminium vacuum capsule was mounted around the sample and evacuated to isolate it from the external environment, and a helium cryorefrigerator was used then to change and maintain the temperature inside the chamber. Temperature control was achieved via a Lakeshore controller, with measurements inside the cold head being performed with a Si diode. The temperature measured by the Si diode is accurate to 0.1 K with a nominal stability of 0.05 - 0.5 K. For each polymorph, data were collected at 5, 105 and 295 K. Two orientations of each crystal, related by a roughly 180° rotation around the normal to the largest face, itself at roughly 45° to the vertical instrument axis, were needed in order to achieve the highest possible completeness. For each orientation, Laue patterns were collected by motorised step-wise rotations (denoted by ϕ) about the vertical instrument axis. The exposure times for the orthorhombic and monoclinic polymorphs were 4 and 3 hours per Laue pattern, respectively, at all temperatures. Data-collection details for all experiments are provided in Table 1. The completeness values are ~75% and ~80% for the orthorhombic and monoclinic polymorphs, respectively. Both values are close to the theoretical maximum completeness, 83.3%, of the Laue diffraction technique (Cruickshank *et al.*, 1987).

2.2.2. Single-crystal X-ray diffraction

Data collections at 5, 105 and 295 K were performed using Mo $K\alpha$ radiation (λ = 0.7107 Å) on the two polymorphs of the amino acid *L*-histidine at the XHIPOS I facility, based at the Newcastle University, using a three-stage closed-cycle helium refrigerator. A full description of the diffractometer can be found elsewhere (Probert *et al.*, 2010, McMonagle & Probert, 2019). Crystals of sizes 0.2 × 0.1 × 0.1 mm^3 and 0.2 × 0.2 × 0.1 mm^3 were used for the orthorhombic and monoclinic forms, respectively.

Each crystal was mounted onto a sharpened graphite rod and centred optically, before collecting diffraction data at 295 K. A beryllium vacuum chamber, combined with a flexible graphite radiation shield, was then used to evacuate and isolate the sample from the external environment for the low-temperature data collections. Cooling the sample from 295 to 5 K took ~8 hours. The crystal was re-centred in the X-ray beam by diffraction and data sets were collected first at 5 K and then at 105 K. During the measurements at 2θ = 0°, the detector was placed at 65 mm from the sample in order to

prevent collisions with the beam-stop, otherwise the distance was set at 60 mm. The χ angle was kept fixed at 0°. See Table S2 in the Supporting Information for more details on the data collections.

Table 1 Collection strategy on Koala for each data set for the two polymorphs of the amino acid *L*-histidine, at 5, 105 and 295 K.

	Orientation	No. of Laue patterns	φ-range	$\Delta\varphi$	Total exposure time
Orthorhombic form					
5 K	1	14	$-90° < \phi < 90°$; $82.5° < \phi < 7.5°$	15°; 90°	56 h
	2	13	$-90° < \phi < 90°$	15°	52 h
105 K	1	13	$-90° < \phi < 90°$	15°	52 h
	2	13	$-90° < \phi < 90°$	15°	52 h
295 K	1	13	$90° < \phi < -90°$	15°	52 h
	2	16	$-90° < \phi < 90°$; $-82.5° < \phi < -52.5°$	15°; 15°	64 h
Monoclinic form					
5 K	1	10	$-90° < \phi < 90°$	20°	30 h
	2	12	$-90° < \phi < -18°$; $0° < \phi < 90°$	18°; 15°	36 h
105 K	1	14	$90° < \phi < -75°$; $-30° < \phi < -90°$	15°; 12°	42 h
	2	5	$90° < \phi < -60°$	30°	15 h
295 K	1	13	$90° < \phi < 90°$	15°	39 h
	2	13	$-90° < \phi < 90°$; $-82.5° < \phi < -97.5°$	15°; 30°	39 h

2.3. Structure analysis

2.3.1. Neutron diffraction data processing and refinement

Laue diffraction data collected on the Koala instrument were processed using the *LaueG* suite of programs (Piltz, 2018). Reflection intensities were integrated with the argonne-boxes software, a 2-D adaptation of the 3-D minimum $\sigma(I)/(I)$ algorithm described by (Wilkinson *et al.*, 1988, Prince *et al.*, 1997). Resolution limits between 0.55 and 0.65 Å were selected for the different data sets, the choice based on the shortest *d*-spacing at which 5% of the reflections in each set had $I/\sigma(I) > 3$.

Data were empirically normalised to a single common incident wavelength using the Laue normalisation software *Laue4*, where repeated observations, at different wavelengths, and equivalent reflections with wavelengths between 0.80 and 1.75 Å are compared; reflections outside this range are usually excluded from conventional analysis because of their low multiplicity or weakness. Due to the small sizes of the samples, absorption and extinction corrections were deemed unnecessary. Figure 1 shows the refined wavelength spectra collected at 295 K for each polymorph, and the nominal

instrument spectrum. The point groups used for merging were mmm and 2/m for the orthorhombic and monoclinic forms, respectively.

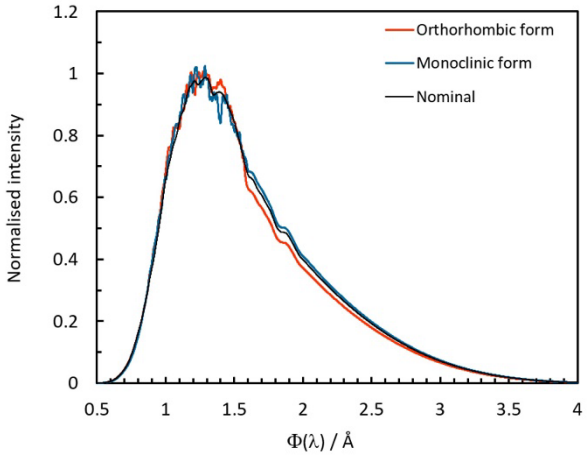


Figure 1 Refined and normalised instrument wavelength spectra for Laue diffraction data collected at 295 K for the orthorhombic and monoclinic polymorphs of the amino acid *L*-histidine. The nominal spectrum is included for comparison.

The crystal structures were refined by full-matrix least-squares on $|F|^2$ against the neutron diffraction data using *SHELXL* within the *Olex2* graphical user interface (Sheldrick, 2015, Dolomanov *et al.*, 2009). The unit-cell dimensions for each neutron diffraction data set, together with the initial atomic coordinates, were taken from the corresponding X-ray independent-atom model (see Section 2.3.2 and Table S3 in the Supporting Information). The weighting scheme used was $w = 1/[\sigma^2(|F_o|^2)+(aP)^2+bP]$, where $P = (|F_o|^2+2|F_c|^2)/3$, $a = 0.03$ and $b = 0$ (McCandlish *et al.*, 1975, Lundgren & Liminga, 1979). All atoms in the structures were refined anisotropically, which was possible without the use of any restraints. Molecular structures obtained from the neutron diffraction data for the orthorhombic and monoclinic forms, at all temperatures, are shown in Figures 2 and S1, respectively. Crystal parameters, refinement statistics and data-collection details are given in Table 2.

2.3.2. X-ray diffraction data processing and refinement

Diffraction data collected with XIPHOS I were processed using the *APEX3* suite of programs (Bruker AXS Inc., 2017). Indexing and data reduction were carried out using *SAINT*, and absorption corrections with *SADABS* (Krause *et al.*, 2015). For the monoclinic polymorph specifically, two orientations of the crystal were needed to optimise completeness (see Table S2 in the Supporting Information), and the absorption correction was performed on each data set separately. Subsequently, the un-merged intensities were scaled to each other linearly and the two data sets were merged (*XPREP*) before starting

refinement. The point groups used for merging were 222 and 2 for the orthorhombic and monoclinic forms, respectively.

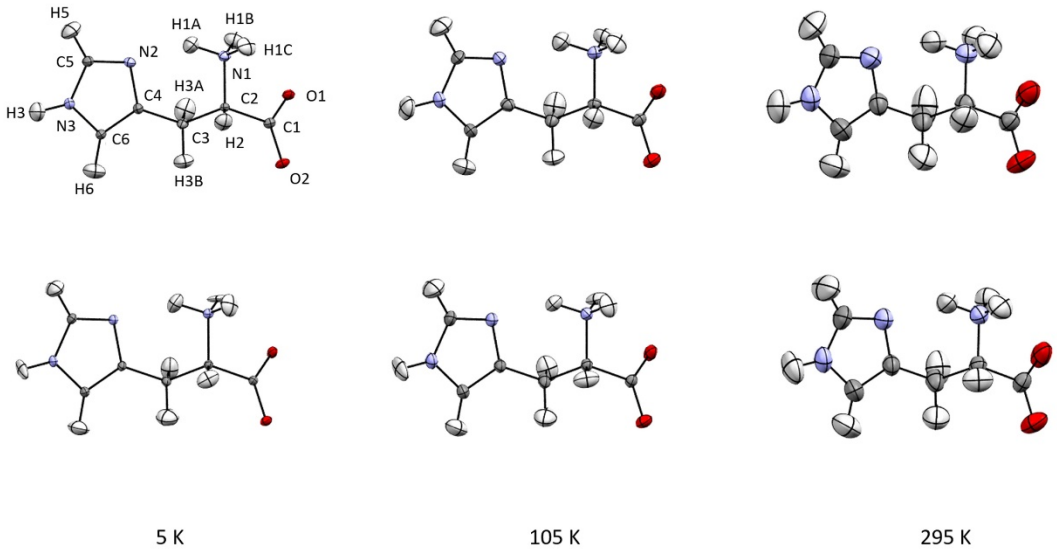


Figure 2 Molecular structures and anisotropic displacement parameters (70% probability surface) for the orthorhombic polymorph of the amino acid *L*-histidine at 5, 105 and 295 K. The top series was obtained from the neutron diffraction data, while the corresponding structures obtained from X-ray diffraction data and refined using HAR in *NoSpherA2* (M062X / x2c-TZVpp and cluster of neighbouring molecules, with a rigid-body restraint of 0.004 Å^2 applied to the 5 and 105 K structures) are shown in the bottom row. A similar representation for the monoclinic polymorph is available as Figure S1 in the Supporting Information.

Each structure was initially refined using the independent-atom model (IAM, Table S3) in *SHELXL* (Sheldrick, 2015) and then with the Hirshfeld atom model as implemented in the *NoSpherA2* module of *Olex2* (Dolomanov *et al.*, 2009, Kleemiss *et al.*, 2021) as described below. The quantum-mechanical software *ORCA* (Neese, 2012, 2017) was chosen for the wavefunction calculation, with the *Microsoft-MPI* interface used for parallelisation. The functionals PBE, B3LYP and M062X (Perdew *et al.*, 1996, Becke, 1993, Stephens *et al.*, 1994, Kim & Jordan, 1994, Zhao & Truhlar, 2008) were tested along with three different basis sets: cc-PVQZ, def2-TZVpp and x2c-TZVpp (Dunning, 1989, Weigend & Ahlrichs, 2005, Pollak & Weigend, 2017).

The refinements involving the triple-zeta basis sets were performed with and without an explicit cluster of three *L*-histidine molecules around the formula unit, corresponding to the three partners involved in H-bonding generated by the crystallographic symmetry. For all refinements, integration accuracy, self-consistent field threshold and strategy were as in (Kleemiss *et al.*, 2021); anomalous dispersion corrections and core-electron relativistic corrections were applied. Once the wavefunction outputs were obtained, *NoSpherA2* was used to perform the Hirshfeld's stockholder partitioning of the static electron density and the subsequent non-spherical atomic form-factor calculations (Hirshfeld, 1977).

Olex2.refine handled the resulting non-spherical atomic form factors to perform least-squares refinement on $|F|^2$. The weighting scheme used was $w = 1/[\sigma^2(|F_o|^2)+(aP)^2+bP]$, with $P = (|F_o|^2+2|F_c|^2)/3$, while a and b were optimised in an analysis of variance.

Table 2 Crystallographic information of the two polymorphs of the amino acid *L*-histidine from neutron diffraction data collected at 5, 105 and 295 K.

Temperature	**5 K**	**105 K**	**295 K**
Orthorhombic form			
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
a, b, c (Å) [a]	5.1498(2), 7.1902(2), 18.8503(6)	5.1521(10), 7.2228(2), 18.8440(6)	5.1690(3), 7.3430(5), 18.8317(13)
V (Å^3)	697.99(4)	701.23(3)	714.78(8)
Radiation type	Neutron, λ = 0.80 – 1.7 Å	Neutron, λ = 0.80 – 1.7 Å	Neutron, λ = 0.80 – 1.7 Å
Diffractometer	Koala	Koala	Koala
No. of measured, unique and observed [$I > 2\sigma(I)$] reflections	33 762, 1 375, 1 133	32 598, 1 361, 1 061	32 190, 1 104, 781
R_{int} / Completeness	0.078 / 0.758	0.079 / 0.763	0.079 / 0.750
d_{min} (Å)	0.60	0.60	0.65
R_1 [$F_o > 4\sigma(F_o)$], wR_2 (F^2), S	0.0475, 0.0692, 1.153	0.0522, 0.0739, 1.145	0.0543, 0.0779, 1 190
Weighting scheme param.[b]	a = 0.03, b = 0	a = 0.03, b = 0	a = 0.03, b = 0
No. of parameters [c]	181	181	181
No. of restraints	0	0	0
$\Delta\rho_{max}$, $\Delta\rho_{min}$ (fm Å^{-3})	1.49, –1.34	1.27, –1.10	0.67, –0.65
NPP *m, c,* r^{2d}	1.076, 0.059, 0.999	1.081, 0.010, 0.998	1.098, 0.047, 0.999
Monoclinic form			
Space group	$P2_1$	$P2_1$	$P2_1$
a, b, c (Å)	5.1651(2), 7.2324(2), 9.4957(3)	5.1656(2), 7.2761(3), 9.4978(3)	5.1854(8), 7.3998(10), 9.4976(12)
β (°)	97.065(2)	97.316(3)	98.182(10)
V (Å^3) [a]	352.03(2)	354.07(2)	360.72(9)
Radiation type	Neutron, λ = 0.80 – 1.7 Å	Neutron, λ = 0.80 – 1.7 Å	Neutron, λ = 0.80 – 1.7 Å
Diffractometer	Koala	Koala	Koala
No. of measured, unique and observed [$I > 2\sigma(I)$] reflections	16 604, 1 629, 1 267	12 588, 1 292, 939	20 381, 1 354, 914
R_{int} / Completeness	0.075 / 0.793	0.062 / 0.798	0.057 / 0.803
d_{min} (Å)	0.55	0.60	0.60
R_1 [$F_o > 4\sigma(F_o)$], wR_2 (F^2), S	0.0601, 0.0816, 1.212	0.0501, 0.0714, 1.003	0.0533, 0.0746, 1.105
Weighting scheme param. [b]	a = 0.03, b = 0	a = 0.03, b = 0	a = 0.03, b = 0
No. of parameters [c]	181	181	181
No. of restraints	1 (origin fixing)	1 (origin fixing)	1 (origin fixing)
$\Delta\rho_{max}$, $\Delta\rho_{min}$ (fm Å^{-3})	1.62, –2.00	1.12, –1.18	0.77, –1.00
NPP *m, c,* r^{2d}	1.154, 0.037, 0.999	0.932, 0.039, 0.999	1.061, 0.014, 0.999

[a] Obtained from the X-ray data collections (Table 3).

[b] The weighting scheme applied was in the form of $w = 1/[\sigma^2(|F_o|^2)+(aP)^2+bP]$, with $P = (|F_o|^2+2|F_c|^2)/3$

[c] All atoms, including H atoms, were refined anisotropically.

[d] Normal-probability-plot (NPP) parameters corresponding to gradient, m, intercept, c, and correlation coefficient, r^2

All atoms, including H atoms, were refined anisotropically. Cycles of refinement and form factor determinations were iterated until convergence was achieved. The refinements took between several minutes to hours, depending on the presence or not of molecular clusters, on an ordinary laboratory desktop computer system with an Intel i5-8350U and 16GB memory. For one of the best-performing settings of the HAR (M062X/x2c-TZVpp/cluster), RIGU restraints (Thorn *et al.*, 2012) with weights on the 1,2 and 1,3 atomic pairs respectively equal to 0.002, 0.003, 0.004 and 0.005 Å^2 were applied successively to the H atoms in the ammonium group, the imidazole ring and the carbon skeleton for testing. The results of the best-performing HAR setting (see above) are shown in Table 3 for all temperatures.

Table 3 Crystallographic information for the M062X/x2c-TZVpp/cluster HAR setting of the two polymorphs of the amino acid *L*-histidine from X-ray data collections performed at 5, 105 and 295 K.

Temperature	**5 K**	**105 K**	**295 K**
Orthorhombic form			
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
a, b, c (Å)	5.1498(2), 7.1902(2), 18.8503(6)	5.1521(10), 7.2228(2), 18.8440(6)	5.1690(3), 7.3430(5), 18.8317(13)
V (Å^3)	697.99(4)	701.23(3)	714.78(8)
Radiation type	Mo *K*α, λ = 0.7107 Å	Mo *K*α, λ = 0.7107 Å	Mo *K*α, λ = 0.7107 Å
Diffractometer	XIPHOS I	XIPHOS I	XIPHOS I
No. of measured, unique and observed [$I > 2\sigma(I)$] reflections	13 716, 2 928, 2 768	13 793, 2 942, 2 706	14 091, 3 008, 2 569
R_{int} / Completeness	0.0274 / 0.962	0.0303 / 0.964	0.0327/ 0.968
d_{min} (Å)	0.60	0.60	0.60
R_1 [$F_o > 4\sigma(F_o)$], wR_2 (F^2), *S*	0.0191, 0.0322, 1.111	0.0226, 0.0375, 1.103	0.0287, 0.0539, 1.114
Weighting scheme parameters[a]	a = 0.0100, b = 0	a = 0.0119, b = 0	a = 0.0206, b = 0.0062
No. of parameters [b]	181	181	181
No. of restraints	99	99	0
$\Delta\rho_{max}$, $\Delta\rho_{min}$ (e Å^{-3})	0.170, –0.130	0.167, –0.151	0.164, –0.170
Monoclinic form			
Space group	$P2_1$	$P2_1$	$P2_1$
a, b, c (Å)	5.1651(2), 7.2324(2), 9.4957(3)	5.1656(2), 7.2761(3), 9.4978(3)	5.1854(8), 7.3998(10), 9.4976(12)
β (°)	97.065(2)	97.316(3)	98.182(10)
V (Å^3)	352.03(2)	354.07(2)	360.72(9)
Radiation type	Mo *K*α, λ = 0.7107 Å	Mo *K*α, λ = 0.7107 Å	Mo *K*α, λ = 0.7107 Å
Diffractometer	XIPHOS I	XIPHOS I	XIPHOS I
No. of measured, unique and observed [$I > 2\sigma(I)$] reflections	11 118, 2 138, 2 087	11 193, 2 148, 2 032	11 356, 1 806, 1 683
R_{int} / Completeness	0.0254 / 0.997	0.0329 / 0.996	0.0240 / 0.869
d_{min} (Å)	0.70	0.70	0.70
R_1 [$F_o > 4\sigma(F_o)$], wR_2 (F^2), *S*	0.0167, 0.0286, 1.245	0.0218, 0.0315, 1.155	0.0247, 0.0486, 1.349
Weighting scheme parameters[a]	a = 0.0081, b = 0	a = 0.0076, b = 0	a = 0.0158, b = 0.0067
No. of parameters[b]	181	181	181
No. of restraints	100	100	1
$\Delta\rho_{max}$, $\Delta\rho_{min}$ (e Å^{-3})	0.120, –0.099	0.148, –0.119	0.104, –0.098

[a] The weighting scheme applied was in the form of $w = 1/[\sigma^2(|F_o|^2)+(aP)^2+bP]$, with $P = (|F_o|^2+2|F_c|^2)/3$
[b] All atoms, including H atoms, were refined anisotropically

Table 4 shows different settings of the HARs of the polymorphs at 5 K and Table S4 in the Supporting Information provides the performance of the M062X/x2c-TZVpp/cluster HAR setting for both polymorphs at the other temperatures. Molecular structures obtained from the X-ray data are shown in Figure 2 and S1 for the orthorhombic and monoclinic polymorphs, respectively.

Table 4 Results of HARs of the two polymorphs of *L*-histidine at 5 K.

Method - Cluster	R_1	$wR_2(F^2)$	S	$\Delta\rho_{min/max}$ (e Å^{-3})	n. p. d. ADPs	Time (min)
Orthorhombic form						
PBE / cc-PVQZ-✗	0.0198	0.0391	1.033	–0.138 / 0.209		9
PBE / def2-TZVpp-✗	0.0197	0.0386	1.031	–0.135 / 0.185		3
PBE / def2-TZVpp-✓	0.0196	0.0379	1.033	–0.130 / 0.184	H1B(f)	31
PBE / x2c-TZVpp-✓	0.0196	0.0379	1.033	–0.130 / 0.184	H1B(f)	31
B3LYP/ cc-PVQZ-✗	0.0195	0.0380	1.029	–0.141 / 0.203		40
B3LYP/ def2-TZVpp-✗	0.0194	0.0374	1.028	–0.137 / 0.176		20
B3LYP/ def2-TZVpp-✓	0.0193	0.0368	1.028	–0.131 / 0.178	H1B	82
B3LYP/ x2c-TZVpp-✓	0.0193	0.0368	1.028	–0.158 / 0.181	H1B	106
M062X / cc-PVQZ-✗	0.0193	0.0364	1.045	–0.136 / 0.198		32
M062X / def2-TZVpp-✗	0.0191	0.0326	1.129	–0.133 / 0.170		11
M062X / def2-TZVpp-✓	0.0190	0.0320	1.125	–0.128 / 0.170	H1B	122
M062X / x2c-TZVpp-✓	0.0190	0.0320	1.126	–0.128 / 0.170	H1B	119
***RIGU** 0.005*	*0.0191*	*0.0322*	*1.109*	*–0.129 / 0.169*		
***RIGU** 0.004*	*0.0191*	*0.0322*	*1.111*	*–0.130 / 0.170*		
***RIGU** 0.003*	*0.0191*	*0.0324*	*1.115*	*–0.130 / 0.169*		
***RIGU** 0.002*	*0.0193*	*0.0327*	*1.124*	*–0.131 / 0.169*		
Monoclinic form						
PBE / cc-PVQZ-✗	0.0176	0.0309	1.349	–0.106 / 0.129	H1A(f), H1B(f), H2	9
PBE / def2-TZVpp-✗	0.0176	0.0312	1.334	–0.108 / 0.129	H1A, H1B(f), H2	5
PBE / def2-TZVpp-✓	0.0176	0.0302	1.318	–0.106 / 0.133	H1B, H2	33
PBE / x2c-TZVpp-✓	0.0176	0.0302	1.318	–0.106 / 0.133	H1B, H2	36
B3LYP/ cc-PVQZ-✗	0.0173	0.004	1.307	–0.104 / 0.128	H1A, H1B(f), H2	50
B3LYP/ def2-TZVpp-✗	0.0173	0.0307	1.292	–0.106 / 0.127	H1A, H1B(f), H2	10
B3LYP/ def2-TZVpp-✓	0.0171	0.0299	1.284	–0.103 / 0.132	H1B, H2	79
B3LYP/ x2c-TZVpp-✓	0.0171	0.0299	1.284	–0.103 / 0.132	H1B, H2	85
M062X / cc-PVQZ-✗	0.0169	0.0288	1.293	–0.100 / 0.117	H1A, H1B(f), H2	38
M062X / def2-TZVpp-✗	0.0167	0.0288	1.268	–0.102 / 0.115	H1A, H1B(f), H2	9
M062X / def2-TZVpp-✓	0.0165	0.0279	1.256	–0.102 / 0.119	H1B, H2	101
M062X / x2c-TZVpp-✓	0.0165	0.0279	1.256	–0.102 / 0.119	H1B, H2	97
***RIGU** 0.005*	*0.0165*	*0.0282*	*1.233*	*–0.100 / 0.120*	*H2*	
***RIGU** 0.004*	*0.0166*	*0.0283*	*1.237*	*–0.100 / 0.120*	*H2*	
***RIGU** 0.003*	*0. 0167*	*0.0286*	*1.245*	*–0.099 / 0.120*		
***RIGU** 0.002*	*0.0168*	*0.0288*	*1.254*	*–0.099 / 0.120*		

The "f" symbols stands for atoms that are very flat, despite not being non-positive definite

2.3.3. Intermolecular interaction energy calculations

Pixel calculations were carried out with the *MrPixel* interface using the *CLP-Pixel* suite (Gavezzotti, 2005, 2007, 2011, Reeves *et al.*, 2020). The electron densities required for the calculations were obtained from *Gaussian-09* (Frisch, 2009), on a grid of dimensions 0.06 × 0.06 × 0.06 Å^3 and with the 6-31G** basis set at the MP2 level of theory. A condensation level of 3 was used. Interactions within the cluster of radius 25 Å were considered in the calculations. The results of the intermolecular

interaction energy calculations are provided in Table 5 for the neutron structures at 5 K, while the results for the ambient-temperature neutron structures are provided in Table S5 of the Supporting Information.

2.3.4. Other programs used

Mercury was used for data visualization and for the generation of Figures 2-4, 7, as well as Figures S1 and S2 in the Supporting Information (Macrae *et al.*, 2020). Intramolecular and intermolecular geometries, along with rigid-body-model libration corrections, were analysed using *PLATON* (Spek, 2009).

3. Results and discussion

3.1. Neutron structure of the polymorphs at 5 K

At 5 K, the orthorhombic form of *L*-histidine belongs to the space group $P2_12_12_1$, with four molecules in the unit cell ($Z' = 1$) and unit-cell parameters a = 5.1498(2) Å, b = 7.1902(2) Å, c = 18.8503(6) Å and V = 697.99(4) Å^3. The monoclinic form belongs to the space group $P2_1$, with two molecules in the unit cell ($Z' = 1$) and unit-cell parameters: a = 5.1651(2) Å, b = 7.2324(2) Å, c = 9.4957(3) Å, β = 97.065(2) ° and V = 352.02(2) Å^3. The two unit cells are related by the relationship: $a_{ortho} \approx a_{mono}$, $b_{ortho} \approx b_{mono}$, $c_{ortho} \approx 2c_{mono}\cos(\beta\text{-}90°)$ (Lehmann *et al.*, 1972a). The orthorhombic form, which corresponds to the thermodynamic phase, has higher density and a more negative lattice energy (1.476 g cm^{-3} and −305.8 kJ mol^{-1}, the energy being calculated with the Pixel method), while the monoclinic form is the kinetic phase (1.464 g cm^{-3} and −302.6 kJ mol^{-1}).

The molecules in both phases are in the zwitterionic form with their conformation stabilised by an intramolecular H-bond between the ammonium-group hydrogen H1A and the nitrogen N2 belonging to the imidazole ring (Figure 2 and S1). Most of the intramolecular bond lengths and angles determined in the neutron diffraction study differ no more than 3σ between the two polymorphs. However, there is a small difference in the <C5–N3–H3 and <C6–N3–H3 angles, which are more similar in the orthorhombic form (126.3(2)° and 126.6(2)°) than in the monoclinic form (124.0(2)° and 128.5(2)°). There are other significant differences in the dihedral angles that define the torsion of the imidazole ring about the backbone (such as C3–C4–C6–N3, C2–C3–C4–N2 and C2–C3–C4–C6, which differ by 1.2(2)°, 2.2(3)° and 3.7(2)°, respectively). When the chemically equivalent atoms of the two molecular structures are superimposed, the root-mean-square fit is 0.041 Å.

The intermolecular interaction energies within the first coordination sphere of the two polymorphs, which contains 12 molecules in six symmetry-related pairs, were calculated with the Pixel method (Table 5). The molecule-molecule energies are grouped into six symmetry-equivalent pairs labelled A/A′, B/B′ *etc,* and they are broken down into their Coulombic, polarisation, dispersion and repulsion terms. Three conventional H-bonds (described as contacts A, B and C) are formed by each molecule in

both polymorphs, two originating from the ammonium group (N1), another formed by the nitrogen atom (N3) belonging to the imidazole ring, and all having a carboxylate atom as acceptor.

The comparison between the H-bonding and packing in the two phases, presented below, is informed by comments made by a referee to our earlier paper (Novelli *et al.*, 2020). The crystal packing in the *ab*-layers by the N1–H1B···O2^{ii} and N1–H1C···O2^{iii} H-bonds (contacts B and C, respectively) is identical in the orthorhombic (Figure 3a, B [(ii) $-1+x, y, z$] and C [(iii) $1-x, -1/2+y, 1/2-z$]) and monoclinic (Figure 3b, B [(ii) $-1+x, y, z$] and C [(iii) $2-x, 1/2+y, 2-z$]) forms.

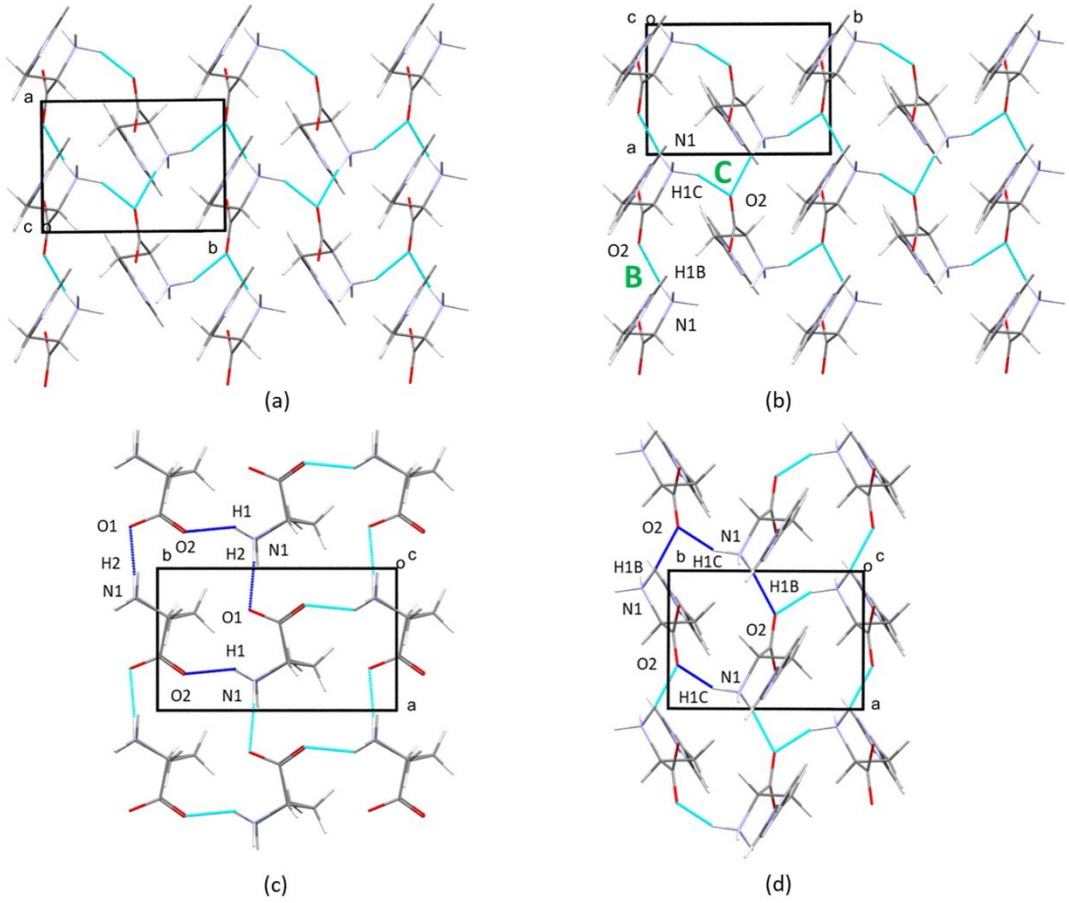


Figure 3 H-bonding in *L*-histidine at 5 K, as viewed along **c**, for the orthorhombic (a) and monoclinic (b) polymorphs. (c) and (d) show the *ab*-layers in *L*-proline and *L*-histidine, respectively. (d) Corresponds to the monoclinic form of *L*-histidine rotated 180° about the *a*-axis. The crystal structure of *L*-proline corresponds to the entry PROLIN05 in the Cambridge Structural Database. The contacts B and C referred to in the text and in Table 5 are labelled in green in (b).

Table 5 Interactions in the first molecular coordination sphere of the polymorphs of the amino acid *L*-histidine, at 5 K. All energies are in kJ mol^{-1}. Calculations were performed on the experimental neutron structures. The symmetry operators relate the central molecule to the other molecules in the first coordination sphere, for a total of 12 intermolecular contacts. H-atom-to-acceptor distances are reported, for each contact, in Ångstrom.

Label	Symmetry Operator	T	Centroid Distance	Pixel					Intermolecular Contact, direction
		(K)	(Å)	Coulombic	Polarisation	Dispersion	Repulsion	Total	(Å)
Orthorhombic form									
H-bonds									
A / A′	3/2 – *x*, 1 – *y*, – 1/2 + *z* 3/2 – *x*, 1 – *y*, 1/2 + *z*	5	9.448	–110.6	–36.3	–12.2	59.7	–99.4	N3–H3···O1^{i} = 1.719(4), along **c**
B / B′	– 1 + *x*, *y*, *z* 1 + *x*, *y*, *z*	5	5.150	–21.7	–34.6	–25.7	49.5	–32.5	N1–H1B···O2^{ii} = 1.817(4), along **a**
C / C′	1 – *x*, – 1/2 + *y*, 1/2 – *z* 1 – *x*, 1/2 + *y*, 1/2 – *z*	5	7.046	–116.4	–42.3	–21.4	50.7	–129.3	N1–H1C···O2^{iii} = 1.805(4), along **b**
Electrostatic Interactions									
D / D′	2 – *x*, – 1/2 + *y*, 1/2 – *z* 2 – *x*, 1/2 + *y*, 1/2 – *z*	5	6.742	–50	–15.1	–12.7	11.7	–65.9	N1H1X···O1 (X = A, B, C) = 2.405 - 3.054
E / E′	1/2 + *x*, 3/2 – *y*, 1– *z* – 1/2 + *x*, 3/2 – *y*, 1– *z*	5	6.383	–16.4	–7.2	–19.6	17.9	–25.2	C6H6···ring = 2.607
F / F′	– 1/2 + *x*, 1/2 – *y*, 1– *z* 1/2 + *x*, 1/2 – *y*, 1– *z*	5	5.774	21.4	–7	–23.2	13.7	4.9	NH_3^+···NH_3^+ = 7.590, NH_3^+···ring = 4.954
Monoclinic form									
H-bonds									
A / A′	*x*, *y*, – 1 + *z* *x*, *y*, 1 + *z*	5	9.496	–101.1	–34.3	–12.2	56	–91.6	N3–H3···O1^{i} = 1.745(5), along **c**
B / B′	– 1 + *x*, *y*, *z* 1 + *x*, *y*, *z*	5	5.165	–23.5	–35.1	–25.1	50.4	–33.2	N1–H1B···O2^{ii} = 1.806(4), along **a**
C / C′	2 – *x*, 1/2 + *y*, 2 – *z* 2 – *x*, – 1/2 + *y*, 2 – *z*	5	7.068	–121.2	–45.3	–21.6	58.9	–129.2	N1–H1C···O2^{iii} = 1.762(4), along **b**
Electrostatic interactions									
D / D′	1 – *x*, – 1/2 + *y*, 2 – *z* 1 – *x*, 1/2 + *y*, 2 – *z*	5	6.870	–46.2	–12.9	–10.7	8.5	–61.3	N1H1X···O1 (X = A, B, C) = 2.480 - 3.075
E / E′	1 – *x*, 1/2 + *y*, 1 – *z* 1 – *x*, – 1/2 + *y*, 1 – *z*	5	5.692	–6.6	–8	–24.2	15.5	–23.3	C6H6···ring = 2.612
F / F′	2 – *x*, – 1/2 + *y*, 1 – *z* 2 – *x*, 1/2 + *y*, 1 – *z*	5	6.458	9.5	–6	–18.4	14.9	0	NH_3^+···NH_3^+ = 8.190, NH_3^+···ring = 4.856

Of the other amino acids, only the *ab*-layer in *L*-proline (PROLIN05:(Koenig *et al.*, 2018)) has a similar arrangement to that in histidine. *L*-histidine and *L*-proline both have only two H atoms at the ammonium nitrogen atom N1 that are available for intermolecular interactions. As shown in Figure 3c for *L*-proline, these H atoms (H1 and H2) interact with different O atoms of different carboxylic groups (O2 and O1, respectively). A similar comment applies to *L*-histidine, but in its case, the two H atoms (H1B and H1C) interact with the same O atom ($O2^{ii}$ and $O2^{iii}$) of different carboxylic groups (Figure 3d). The total molecule-molecule energies of interactions B and C are very similar in the two polymorphs of *L*-histidine, being –32.5 kJ mol^{-1} and –129.3 kJ mol^{-1} for the orthorhombic form, and –33.2 kJ mol^{-1} and –129.2 kJ mol^{-1} for the monoclinic form, respectively. A similar comment applies to the H-atom-to-acceptor geometries.

The difference in crystal packing between the polymorphs is evident in the way the *ab*-layers are stacked along the *c*-axis via the N3–H3···$O1^{i}$ H-bond, which is shown in Figure 4 as contact A ([(i) 3/2 – *x*, 1 – *y*, – 1/2 + *z*] and [(i) *x*, *y*, – 1 + *z*] for the orthorhombic and monoclinic forms, respectively). In the orthorhombic form, the 2_1-screw axis rotates every other molecule, while in the monoclinic form all the molecules are in the same orientation and related by lattice translations. This particular relationship between two polymorphs is unique to *L*-histidine among the amino acids, the closest analogy being that of the monoclinic and triclinic forms of *DL*-valine (Mallikarjunan & Rao, 1969, Dalhus & Görbitz, 1996). Contact A is slightly different in the two polymorphs, being in the orthorhombic form a more stable and closer contact (–99.4 kJ mol^{-1}, 1.719(4) Å and < N3–H3···$O1^{i}$ angle of 172.1(3)°) than that in the monoclinic form (–91.6 kJ mol^{-1}, 1.745(5) Å and 161.8(3)°).

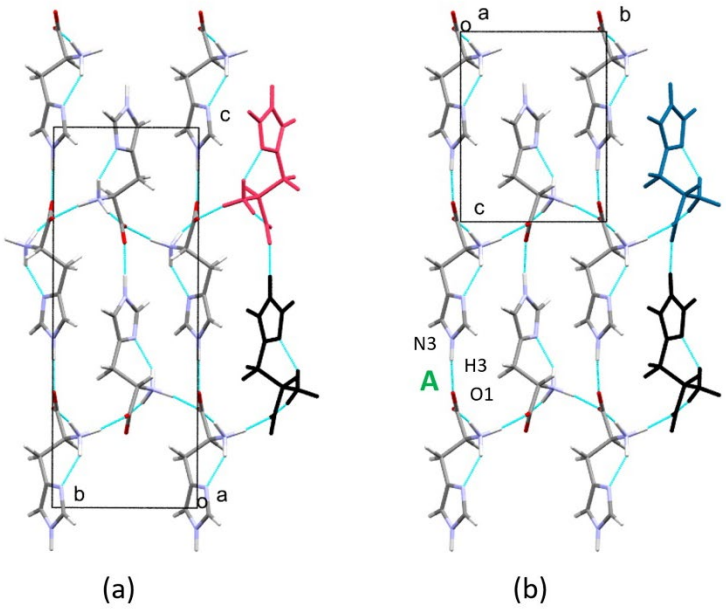


Figure 4 H-bonding in *L*-histidine at 5 K, as viewed along **a**, for the orthorhombic (a) and monoclinic (b) polymorphs. Contact A referred to in the text and in Table 5 is labelled in green in (b).

3.2. Neutron structures of the polymorphs at 295 K and the effect of temperature

The intramolecular geometries of the two polymorphs at ambient temperature are essentially unchanged from those at 5 K. When the two neutron molecular structures at 5 and 295 K are overlapped for each polymorph separately, the root-mean-square fit of all chemically equivalent atoms is 0.018 Å. Higher thermal motion at 295 K decreases the apparent absolute value of the bond lengths slightly (see Tables S6 and S7), with the largest variations being in the O1–C1, –0.016 and –0.011 Å, and N1–H1B, –0.037 and –0.017 Å bond lengths for the orthorhombic and monoclinic forms, respectively. A rigid-body-model libration correction is able to account for some of the differences (Busing & Levy, 1964, Schomaker & Trueblood, 1968, Hirshfeld, 1976), decreasing the largest variations in the O1–C1 bond to –0.012 and –0.007 Å, and in the N1–H1B bond to –0.034 and –0.013 Å, for the orthorhombic and monoclinic forms, respectively.

The ambient-temperature neutron crystal structure of the orthorhombic form was first determined by (Lehmann *et al.*, 1972a) and is available in the Cambridge Structural Database as entry LHISTD13. The two structures are in good agreement, with all the intramolecular parameters varying no more than 3σ, though the present structure has standard uncertainties improved (decreased) by ~50%. The size of the crystal in (Lehmann *et al.*, 1972a) work was ~120 times bigger in volume (1.9 mm^3) than that used in this study.

The two ambient-temperature molecular structures show the same differences in angles and torsions between the two polymorphs as encountered for the 5 K neutron structures, and when they are overlapped the root-mean-square fit of all chemically equivalent atoms is 0.052 Å. The crystal packing of both polymorphs remains unchanged in terms of symmetry, molecular orientations and positions to the 5 K structures. The H-bonding patterns are also similar, with generally small increases in the molecule-molecule total energy and in the H-atom-to-acceptor distance. A detailed description of each contact within the first coordination sphere of both polymorphs at 295 K is provided in Table S5. The differences seen at 5 K in the energy and geometry of contact A are still present at ambient temperature (–95.7 kJ mol^{-1}, 1.745(8) Å and <N3–H3···O1^{i} angle of 172.8(5)° for the orthorhombic form; –90.1 kJ mol^{-1}, 1.756(5) Å and 163.3(4)° for the monoclinic form). The total lattice energies of the two polymorphs become more positive with temperature, being –303.3 kJ mol^{-1} and –298.1 kJ mol^{-1} for the orthorhombic and monoclinic forms, respectively. The lattice-energy difference of 5.2 kJ mol^{-1} is in good agreement with that calculated by (Novelli *et al.*, 2020) for the two polymorphs at a similar temperature (3.6 kJ mol^{-1} at 298 K).

Neutron Laue diffraction patterns collected during heating and cooling of both polymorphs showed no additional spots, which might have been associated with discontinuous changes in unit-cell parameters and volume in the crystal structure. The X-ray-determined volume and unit-cell dimensions of both polymorphs respond in a similar way to temperature (Figure 5). In the orthorhombic form, the volume increases by 2.41% from 5 K to 295 K, the *c*-axis remains essentially unchanged (–0.10%), while the other axes increase with *a* (+0.37%) and *b* (+2.13%) showing the smallest and biggest thermal expansion, respectively. Similar comments apply to the monoclinic form, the values being +2.41%, +0.39%, +2.31%, +0.02% and +1.15% for the volume, *a*, *b*, *c* - axes and β angle, respectively.

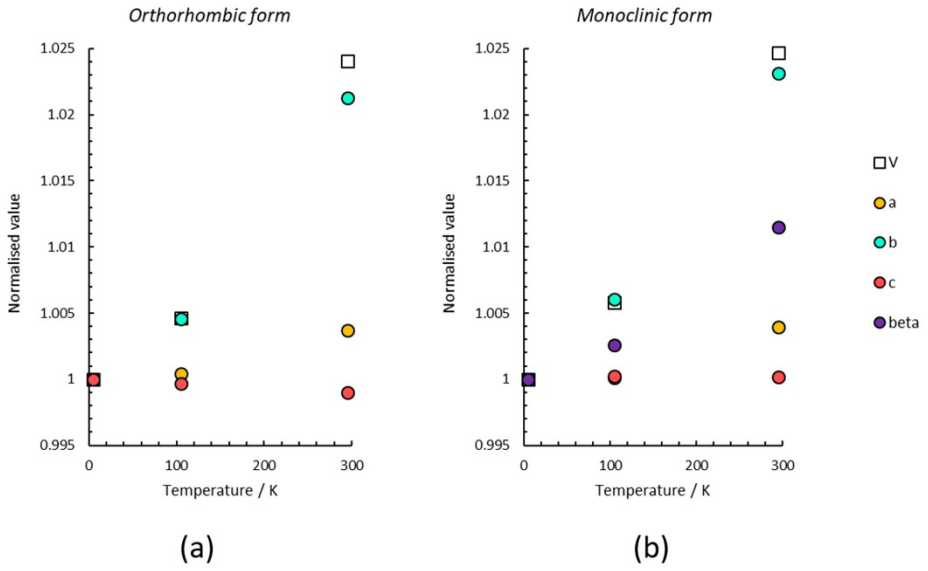


Figure 5 Variation of the normalised (X-ray) lattice parameters and unit-cell volume (*V*) of the orthorhombic (a) and monoclinic (b) polymorphs of *L*-histidine as a function of temperature.

Prior differential scanning calorimetry (DSC) analysis of powder samples performed by (Supriya *et al.*, 2018) show the absence of phase transition in the two polymorphs between ~313 and 573 K. Other studies have revealed discontinuities in the Raman spectra of the orthorhombic form at 165 K and an enthalpic anomaly at around 160 K also by DSC (De Sousa *et al.*, 2013), suggesting a conformational phase transition. Further and more detailed temperature-dependent diffraction work around 165 K is needed to understand fully the phase behaviour of the orthorhombic form of *L*-histidine. However, if there is a conformational change then it must be quite subtle

3.3. Precision and accuracy of Hirshfeld atom refinements with *NoSpherA2*

Single-crystal X-ray diffraction data were used to test the precision and accuracy of the new Hirshfeld atom refinement (HAR) method for X-ray crystal structure determinations, implemented in *NoSpherA2* (Kleemiss *et al.*, 2021), against the results from the neutron diffraction study. Different settings of

functionals and basis sets, along with the inclusion of explicit clusters of interacting molecules and the enhanced rigid-bond restraints (RIGU, (Thorn *et al.*, 2012)), were used for the refinement of all structural parameters, including H atom positions and anisotropic displacement parameters (see Section 2.3.2 for more details).

The first comparison considers a series of agreement factors, such as R_1, wR_2 (F^2), goodness-of-fit (S) and residual electron density ($\Delta\rho$). These indicators, which are shown in Table 4 for the polymorphs at 5 K, provide an overall indication of the accuracy of the HAR models. All the HAR settings yielded comparable results and significant improvements over the IAM refinements. The HAR models have R_1 factors between 0.0190 and 0.0198 for the orthorhombic form, and between 0.0165 and 0.0176 for the monoclinic form. These values are at least 0.0111 smaller than those obtained in the corresponding IAM refinements (R_1 = 0.0309 and 0.0288, respectively). A similar comment applies to the $wR_2(F^2)$ factors, where the improvement extends to 50%. Accordingly, lower values of residual electron density observed for the HARs indicate their ability to account for aspherical features, bond and lone pairs, which are neglected when spherical atom scattering factors are used. For the orthorhombic form, the HARs lead to average max/min residual densities of $\overline{\Delta\rho_{max}}$= 0.184 e Å^{-3} and $\overline{\Delta\rho_{min}}$ = –0.135 e Å^{-3}, while the corresponding values for the IAM refinement are 0.451 e Å^{-3} and –0.223 e Å^{-3}. For the monoclinic form, the HARs lead to average max/min residual densities of $\overline{\Delta\rho_{max}}$= 0.126 e Å^{-3} and $\overline{\Delta\rho_{min}}$ = –0.104 e Å^{-3}, while the corresponding values for the IAM refinement are 0.400 e Å^{-3} and –0.180 e Å^{-3}.

Final goodness-of-fit values are always close to unity and remain similar to those obtained in the IAM refinements. The *SHELX* weighting scheme was used for all refinements but the optimised a and b parameters were closer to zero for the HAR refinements, making the weights more similar to $1/[\sigma^2(|F_o|^2)]$. Specifically, the a and b parameters in the IAM refinements were a = 0.0481/b = 0.0464 and 0.0413/0.0573 for the orthorhombic and monoclinic forms, respectively. These values are considerably larger than those of the HAR refinements (0.0100/0 and 0.0081/0, respectively).

All basis sets used in this study performed similarly in terms of agreement factors. Among the different HAR settings, those carried out with the 'Minnesota' functional M062X were found to have consistently lower values of R factors and residual electron density. The refinements with the PBE functional were the fastest.

For some of the HAR settings, the wavefunction calculations were performed on a cluster of *L*-histidine molecules formed by contacts A, B and C to account for any electron density distortions due to H-bonding. The presence of the molecular cluster led to longer computational times but also to slightly lower R factors and to higher accuracy in the element–H bond lengths. For example, in the case of the orthorhombic form, the H1X–N1 (X = A, B, C) and N3–H3 bond lengths are found to be 1.018(9) Å,

1.012(8) Å, 1.009(9) Å and 1.018(8) Å for the M062X/def2-TZVpp/cluster HAR setting. However, for the same HAR performed on a single molecule the bond lengths are 1.020(9) Å, 0.998(9) Å, 1.000(9) Å and 1.008(9) Å, respectively. The differences are not statistically significant, but the absolute values calculated using a cluster model are consistently closer to the values derived from neutron diffraction (1.042(4) Å, 1.042(4) Å, 1.039(4) Å and 1.054(4) Å), which are systematically longer than the HAR-derived bond lengths. These comments apply also to the monoclinic form and are in agreement with the results observed by (Fugel *et al.*, 2018).

For the sake of brevity, a comparison between the atomic coordinates, intramolecular bond distances and anisotropic displacement parameters of the neutron and X-ray models will be discussed in the next section for the 5 K crystal structures using the unrestrained HAR M062X/x2c-TZVpp/cluster as the X-ray model exclusively.

3.3.1. Comparison of atomic coordinates and intramolecular bond distances

Figure 6 shows two different ways to assess the agreement between the atomic positions of the X-ray and neutron models. In the top graphs, the direct comparison between fractional atomic coordinates of all atoms closely follow the $y = x$ trend line. Regression yields a correlation coefficient (r^2) equal to 0.9999 for both polymorphs with intercepts (c) of 0.0021 and 0.0017 and gradients (m) of 0.9957 and 0.9968 for the orthorhombic and monoclinic forms, respectively. Half-normal probability plots have correlation coefficients of 0.9844 and 0.9701, while the gradients (2.0146 and 2.4976) indicate that the standard uncertainties of the differences may be underestimated.

Superposition of the neutron and X-ray molecular structures yields the root-mean-square fit values of 0.022 Å and 0.023 Å, in the orthorhombic and monoclinic forms, respectively. The wRMSD ratios ($\langle(\Delta P_{X\text{-}N})^2/[\text{s.u.}(P_X)^2+\text{s.u.}(P_N)^2]\rangle^{1/2}$, where P corresponds to a given structural parameter) are found to be 1.16 and 2.45 for the non-H atoms, and 2.39 and 2.25 for the H atoms, in the orthorhombic and monoclinic forms, respectively, indicating that the differences are within the statistical uncertainties. A further indication of the agreement between the X-ray and neutron models is provided by the Pixel calculations, which are very sensitive to H atom positions. The total lattice energies of the two polymorphs at 5 K, calculated using the HAR models, are found to be similar to those of the respective neutron models, being −304.7(+1.1) kJ mol^{-1} and −301.1(+1.5) kJ mol^{-1} for the orthorhombic and monoclinic forms, respectively.

Tables S6 and S7 in the Supporting Information compare the bond lengths obtained from the neutron diffraction data of the two polymorphs of *L*-histidine at 5 K, and corresponding values from the unrestrained HAR setting. Overall, the standard uncertainties (s.u.) on the element–element bond lengths are less precise in the neutron models (s.u. = 0.003 Å) than in the X-ray models (0.0007 Å), while the opposite is observed for the element–H bond lengths (0.004 Å and 0.008 Å, respectively). In

the latter case, the absolute values of the bond lengths are also consistently smaller in the X-ray models, an effect that improves slightly across the HAR settings with the use of a cluster of neighbouring molecules during refinement (see above). The agreement between the element–element bond lengths of the X-ray and neutron models is very high (all equivalent distances differing no more than 2σ), while that of the element–H bond lengths includes some maximum variations of 5σ in the C3–H3B and N3–H3 bonds in the orthorhombic and monoclinic forms, respectively. Corresponding *w*RMSD ratios were found to be 1.02 and 1.25 for the element–element bond lengths, and 3.14 and 2.94 for the element–H bond lengths, in the orthorhombic and monoclinic forms, respectively. These values are close to those obtained by (Kleemiss *et al.*, 2021) during the validation of different *NoSpherA2* HAR models against their corresponding neutron models, at the same temperatures. In the case of the amino acid *L*-alanine (23 K, (Destro *et al.*, 1988, Malaspina *et al.*, 2019)), the *w*RMSD ratios for the element–H bond lengths were between 1.823 and 1.896. Similar values for the dipeptide glycyl-*L*-alanine (12 K, (Capelli, Bürgi, Dittrich*, et al.*, 2014, Capelli, Bürgi, Mason*, et al.*, 2014)) are between 1.152 and 1.309. More details and agreement statistics are available in pages 13-42 in the Supporting Information of (Kleemiss *et al.*, 2021).

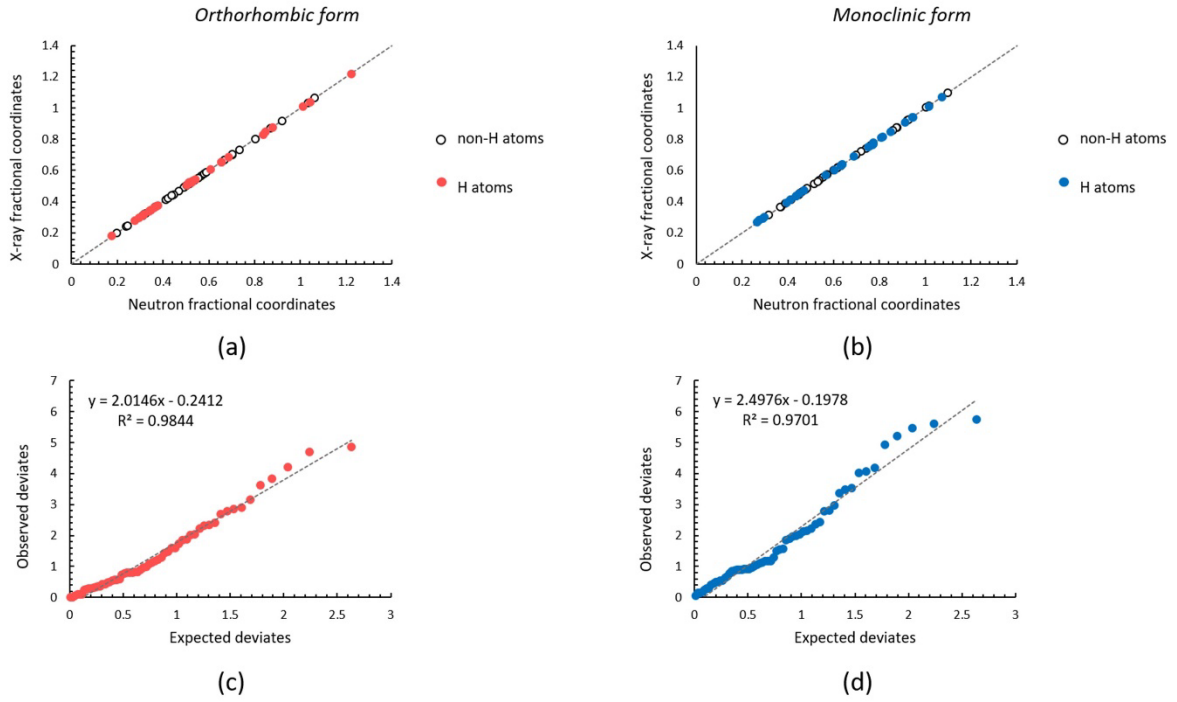


Figure 6 Direct comparison and half-normal-probability plots of the neutron and X-ray fractional atomic coordinates for the orthorhombic (a and c, respectively) and monoclinic (b, d) polymorphs of the amino acid *L*-histidine at 5 K. The X-ray model corresponds to the unrestrained M062X/x2c-TZVpp/cluster HAR setting. In the top graphs, the grey dotted line represents the line y = x, while in the bottom graphs the grey dotted line represent the fitting of the data. In the top graphs, the errors on the parameters are smaller than the dots.

3.3.2. Comparison of anisotropic displacement parameters and the use of restraints

The ADPs of the H atoms belonging to the ammonium group (H1A, H1B and H1C), the imidazole ring (H3) and the α carbon (H2) became non-positive definite at 5 and 105 K in the HAR refinements (Table 4). Statistically, it is arguable that this effect should not be regarded as problematic, as the ADPs are tiny at 5 K and within error of being positive definite for all atoms. Nevertheless, non-positive definite ADPs are physically unreasonable, and therefore attempts to improve the models were performed using enhanced rigid body restraints applied to the local environment of the H atoms (RIGU). Different values of the restraint weight were tested, ranging from 0.005 Å^2 to 0.002 Å^2. Figure 7 shows the different models obtained for the orthorhombic form at 5 K (Figure S2 in the Supporting Information shows a similar representation for the monoclinic form). In both polymorphs, as the restraint becomes tighter, the non-positive-definite ellipsoids disappear. There is hardly any effect on the agreement factors, R_1 going from 0.0190 to 0.0191-0.0193 and from 0.0165 to 0.0165-0.0168 for the orthorhombic and monoclinic forms, respectively), and the residual electron density is unchanged. The restraints moved the goodness-of-fit values closer to unity (going from 1.126 to 1.109-1.124 and from 1.256 to 1.233-1.254). The biggest effect of RIGU on the bond lengths is observed for the N–H bonds, which become closer to the neutron values as the restraints tighten. In the orthorhombic form, the N1–H1A and N1–H1C distances increase to 1.021(+0.003) Å, and 1.010(+0.001) Å with a restraint weight of 0.002 Å^2, respectively. Equivalent values for the monoclinic form are 1.017(+0.006) Å and 1.026(+0.009) Å, respectively.

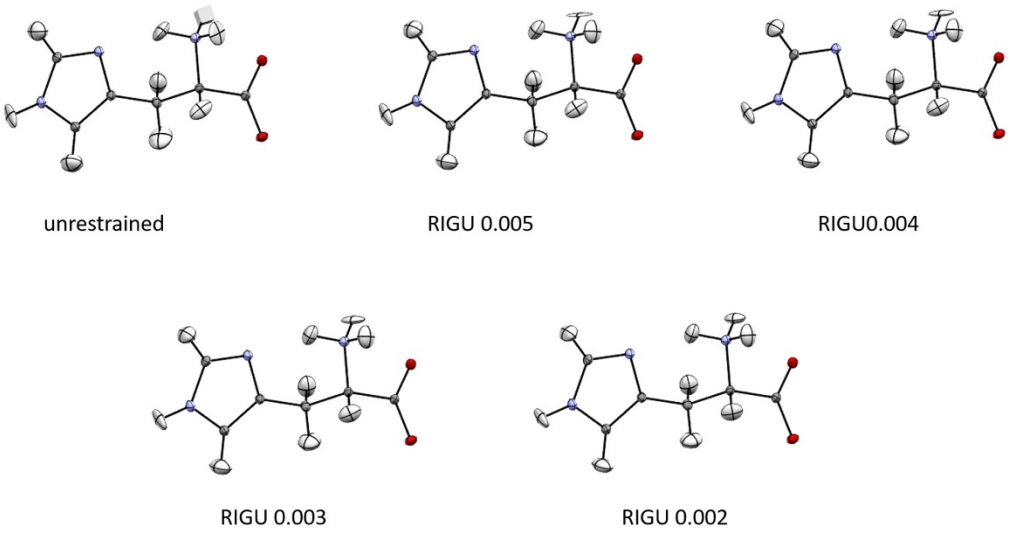


Figure 7 Molecular structures and anisotropic displacement parameters (70% probability surface) for the orthorhombic polymorph of the amino acid *L*-histidine at 5 K. The models were obtained from the X-ray data and refined using HAR in *NoSpherA2* (M062X / x2c-TZVpp and cluster of neighbouring molecules), different values of the RIGU restraint were applied (the weights in the figure are reported in Å^2). In the unrestrained model, the square representing the H1B atom indicates a non-positive-definite ellipsoid. A similar representation for the monoclinic polymorph at the same temperature is available as Figure S2 in the Supporting Information.

Figure 8 shows the agreement between the ADPs of the X-ray and neutron models, with and without the application of restraints, using direct comparisons and normal probability plots. When equivalent ADPs are plotted against each other for the unrestrained models, r^2 values of 0.7356 and 0.7489 are obtained in the orthorhombic and monoclinic forms, respectively. These values, which are determined mainly by the large deviations in the H atoms ADPs, suggest a more complicated relationship than that seen for the atomic coordinates. Additionally, the value of the gradient (1.5946 for the orthorhombic form, 1.4534 for the monoclinic form) in the normal probability plots indicate that the standard deviation of the difference between equivalent parameters may be underestimated. ADPs are known to differ when determined by X-ray and neutron diffraction (Blessing, 1995), but the linearity of the normal probability plots and their near-origin intercepts suggests that systematic effects are quite low in this comparison. The agreement between equivalent ADPs becomes better with the use of the restraint RIGU, with the correlation coefficient for direct comparison, r^2, increasing to 0.8228 and 0.8618 for the orthorhombic and monoclinic forms, respectively. For the monoclinic form this effect is also seen in the normal probability plot, $r^2 = 0.9906$, while for the other polymorph the value remains unchanged, $r^2 = 0.9807$. Overall, although differences in ADPs between the HAR and neutron models are present, their magnitudes are within experimental uncertainties. The wRMSD ratios for the ADPs in the unrestrained models were found to be 1.66 and 1.47 for the non-H atoms, and 1.54 and 1.48 for the H atoms for the orthorhombic and monoclinic forms, respectively. Corresponding values in the restrained models are identical. As was the case for the element–H bond length, the values of wRMSD ratios for the ADPs obtained by (Kleemiss et al., 2021) using different HAR models are very close to those in this study, being 2.782-2.783 and 2.579-2.582 for the non-H atoms, while 1.930-1.938 and 1.385-1.386 for the H atoms, in *L*-alanine and glycyl-*L*-alanine, respectively.

Similar comments apply to the data collected at 105 K and 295 K, which were refined exclusively using the M062X/x2c-TZVpp/cluster HAR setting. Improvements in agreement factors are smaller than those observed in the 5 K structures but still present (see Table S3 in the Supporting Information). For each temperature, when the neutron and X-ray molecular structures are overlapped, the root-mean-square fit of all chemically equivalent atoms is between 0.022 Å and 0.029 Å. Non-positive definite ADPs were observed for the orthorhombic (H1C, H3) and monoclinic (H1B, H2) forms at 105 K, which could be removed by applying mild RIGU restraints (0.004 or 0.005 Å^2) to the atoms of interest. In the ambient temperature structures, this feature disappears and all the atoms have positive definite and reasonable ellipsoids without the use of any restraints. Figures S3 and S4 in the Supporting Information show similar trends as those seen at 5 K for the comparison between X-ray and neutron ADPs, at 105 K and 295 K, respectively.

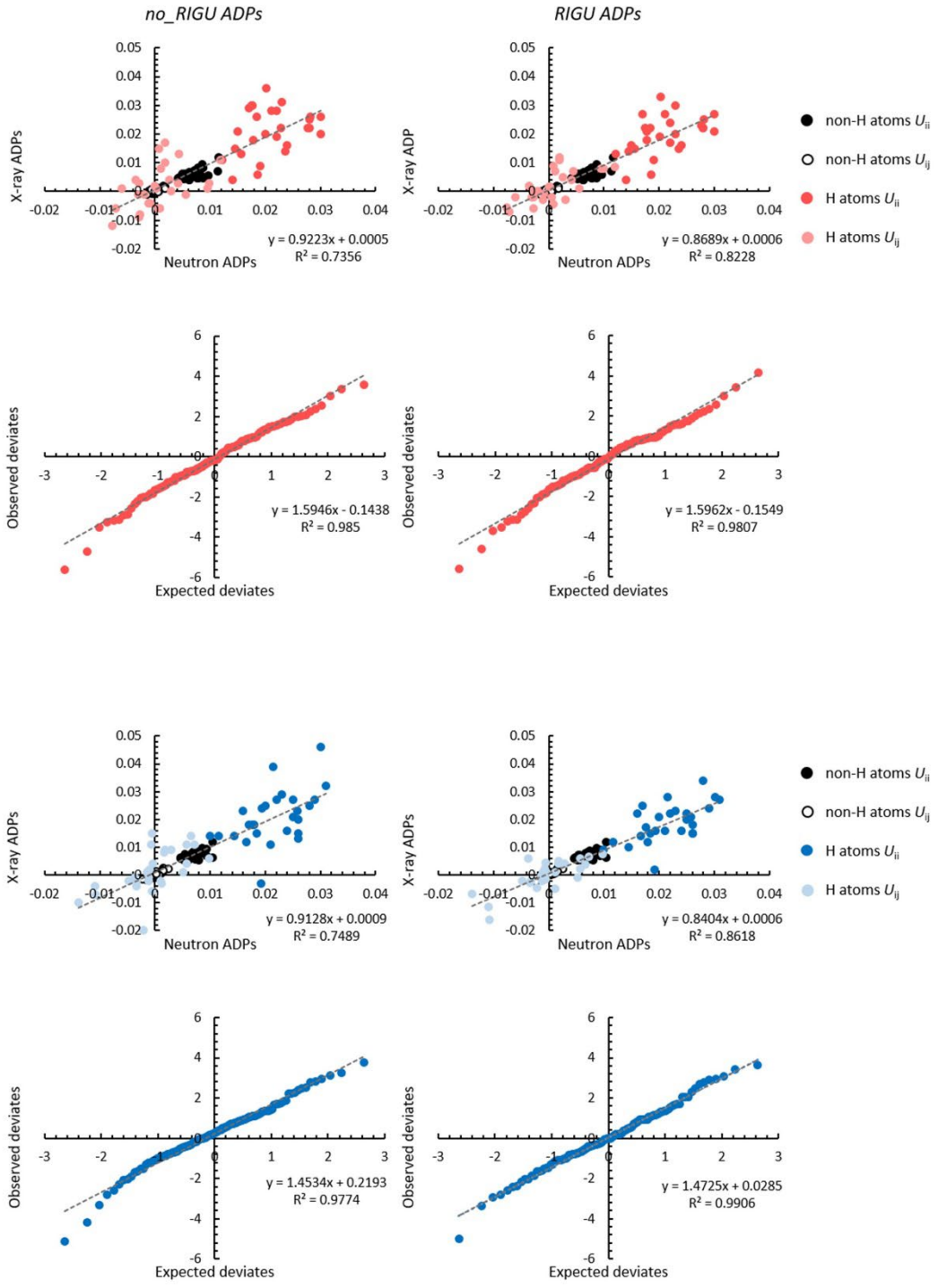


Figure 8 Direct comparison and normal probability plots of the neutron and X-ray ADPs for the orthorhombic (red) and monoclinic (blue) polymorphs of the amino acid *L*-histidine at 5 K. The X-ray model corresponds to the unrestrained M062X/x2c-TZVpp/cluster HAR setting. The left side represents unrestrained models, while the right side shows the results for the application of the RIGU restraint on the orthorhombic (0.004 Å^2) and monoclinic (0.003 Å^2) polymorphs. In the direct comparison of equivalent parameters, the three diagonal U_{ii} ADPs were grouped together as well as the three off-diagonal U_{ij} ADPs for the H and non-H atoms.

4. Conclusion

The Brookhaven and Trombay studies performed in the 1970s on the neutron crystal structures of the naturally occurring amino acids required single crystals with volumes of at least of 10 mm^3 (Binns *et al.*, 2016). Twinning or poor crystal quality led to the structures of *L*-histidine, *L*-isoleucine, *L*-methionine and *L*-leucine either suffering from low precision or remaining undetermined. As demonstrated by the investigation of *L*-leucine (Binns *et al.*, 2016), the use of modern thermal-neutron Laue diffraction, coupled with advances in image-plate technology (Cole *et al.*, 2001, McIntyre *et al.*, 2006, Aznavour *et al.*, 2008, Edwards, 2011), has extended the range of applicability of neutron crystallography to samples with volumes frequently less than 0.1 mm^3.

In this study, the crystal structures of the two polymorphs of *L*-histidine were determined at 5, 105 and 295 K. The ability to use a small sample size enabled un-twinned crystals to be used for data collection. The resulting structures yield geometric parameters with sufficient precision and accuracy for inclusion in restraints libraries for macromolecular structure modelling, with the estimated standard deviations on element–H bond lengths ranging from 0.004-0.005 Å at 5 K and 0.004-0.008 Å at 295 K. Pixel calculations at 5 and 295 K reveal significant differences between the two polymorphs in the geometry and the molecule-molecule energy of the N3–H3···$O1^{i}$ H-bond, formed along the *c*-direction.

Single-crystal X-ray diffraction experiments were also performed, at the same temperatures, on the XIPHOS I facility. Taking advantage of the X-ray data collections at 5 K, the precision and accuracy of the new HAR method implemented in *NoSpherA2* were probed. Overall, there is a good agreement between the HAR and neutron models. X-ray diffraction studies based on HAR in the new *NoSpherA2* implementation can be seen to approach the accuracy of neutron data, and may present an alternative when neutron studies are not possible or too difficult, such as for the remaining missing neutron studies of the amino acids *L*-isoleucine, *L*-methionine and *L*-tryptophan.

Between 5 and 295 K in both phases there is an apparent decrease in the bond lengths with average values being −0.005 Å and −0.007 Å for the element–element and element–H bond lengths, respectively. This well-known effect occurs as the result of higher thermal motion at 295 K, and is of clear relevance to studies that aim to 'benchmark' the performance of computational geometry optimisation methods, many of which neglect thermal effects and refer to a nominal temperature of 0 K. The application of first-principles calculations, such as PBE-TS, to study the structure and properties of crystalline organic materials has gained popularity in the past decade (Funnell *et al.*, 2010, Hunter *et al.*, 2013, Moggach *et al.*, 2015, Budd *et al.*, 2015, Price *et al.*, 2016, Funnell *et al.*, 2019, Fortes, 2020). A study performed seven years ago in our own laboratory (Binns *et al.*, 2014) assessed the performance of DFT-D in reproducing the crystal structures of 30 small organic molecules. The study used data from structures in the Cambridge Structural Database between 2 and 50 K, but there were only four entries below 10 K.

Thermal-neutron Laue diffraction techniques, pioneered by the LADI and VIVALDI instruments at the Institut Laue-Langevin (Cipriani *et al.*, 1996, Myles *et al.*, 1997, Wilkinson *et al.*, 2002, McIntyre *et al.*, 2006), and developed further on the KOALA diffractometer at ANSTO (Edwards, 2011), can now be applied to much smaller samples than was hitherto the case to yield crystal structures below 10 K. The combination of these data with X-ray diffraction data collected at the same temperatures, and with the advent of new X-ray methods for more accurate structure refinements from X-ray diffraction data such as HAR, will likely lead to a step change in the availability of high-accuracy structural data at temperatures approaching absolute zero.

Acknowledgements We thank the referee to our earlier paper (Novelli *et al.*, 2020) for insight into the crystal packing of the two polymorphs of *L*-histidine. We thank the Australian Centre for Neutron Scattering, ANSTO, for the allocation of neutron beam time under proposals 7449 and 8110 and for studentship funding to GN, who was also supported by The University of Edinburgh and the Engineering and Physical Sciences Research Council (UK).

References

Aree, T., Bürgi, H.-B. & Capelli, S. C. (2012). *The Journal of Physical Chemistry A* **116**, 10037-10037.
Aree, T., Bürgi, H.-B., Chernyshov, D. & Törnroos, K. W. (2014). *The Journal of Physical Chemistry A* **118**, 9951-9959.
Aree, T., Bürgi, H.-B., Minkov, V. S., Boldyreva, E. V., Chernyshov, D. & Törnroos, K. W. (2013). *The Journal of Physical Chemistry A* **117**, 8001-8009.
Arnold, W. D., Sanders, L. K., McMahon, M. T., Volkov, A. V., Wu, G., Coppens, P., Wilson, S. R., Godbout, N. & Oldfield, E. (2000). *Journal of the American Chemical Society* **122**, 4708-4717.
Averbuch-Pouchot, M. T. (1993). *Zeitschrift für Kristallographie - Crystalline Materials* **207**, 111-120.
Aznavour, K., Yousufuddin, M., Bau, R., McIntyre, G. J., Mason, S. A., Chen, X., Lim, S., Plečnik, C. E., Liu, S., Du, B., Meyers, E. A. & Shore, S. G. (2008). *Journal of Molecular Structure* **890**, 277-280.
Becke, A. D. (1993). *The Journal of Chemical Physics* **98**, 5648-5652.
Berman, H. M. (2008). *Acta Crystallographica Section A* **64**, 88-95.
Binns, J., Healy, M. R., Parsons, S. & Morrison, C. A. (2014). *Acta Crystallographica Section B* **70**, 259-267.
Binns, J., Parsons, S. & McIntyre, G. J. (2016). *Acta Crystallographica Section B* **72**, 885-892.
Bruker AXS Inc., M., Wisconsin, USA (2017).
Budd, L. E., Ibberson, R. M., Marshall, W. G. & Parsons, S. (2015). *Chemistry Central Journal* **9**, 18.
Busing, W. R. & Levy, H. A. (1964). *Acta Crystallographica* **17**, 142-146.
Capelli, S. C., Bürgi, H.-B., Dittrich, B., Grabowsky, S. & Jayatilaka, D. (2014). *IUCrJ* **1**, 361-379.
Capelli, S. C., Bürgi, H.-B., Mason, S. A. & Jayatilaka, D. (2014). *Acta Crystallographica Section C* **70**, 949-952.
Cipriani, F., Castagna, J. C., Wilkinson, C., Oleinek, P. & Lehmann, M. S. (1996). *Journal of Neutron Research* **4**, 79-85.
Cole, J. M., McIntyre, G. J., Lehmann, M. S., Myles, D. A. A., Wilkinson, C. & Howard, J. A. K. (2001). *Acta Crystallographica Section A* **57**, 429-434.
Cruickshank, D. W. J., Helliwell, J. R. & Moffat, K. (1987). *Acta Crystallographica Section A* **43**, 656-674.
Cruz-Cabeza, A. J., Reutzel-Edens, S. M. & Bernstein, J. (2015). *Chemical Society Reviews* **44**, 8619-8635.
Császár, A. G. & Perczel, A. (1999). *Progress in Biophysics and Molecular Biology* **71**, 243-309.
Dalhus, B. & Görbitz, C. H. (1996). *Acta Crystallographica Section C* **52**, 1759-1761.
De Sousa, G. P., Freire, P. T. C., Filho, J. M., Melo, F. E. A. & Lima, C. L. (2013). *Brazilian Journal of Physics* **43**, 137-144.
Destro, R., Marsh, R. E. & Bianchi, R. (1988). *The Journal of Physical Chemistry* **92**, 966-973.
Destro, R., Soave, R. & Barzaghi, M. (2008). *The Journal of Physical Chemistry B* **112**, 5163-5174.
Dittrich, B., Hubschle, C. B., Messerschmidt, M., Kalinowski, R., Girnt, D. & Luger, P. (2005). *Acta Crystallographica Section A* **61**, 314-320.
Dittrich, B., Lübben, J., Mebs, S., Wagner, A., Luger, P. & Flaig, R. (2017). *Chemistry – A European Journal* **23**, 4605-4614.
Dolomanov, O. V., Bourhis, L. J., Gildea, R. J., Howard, J. A. K. & Puschmann, H. (2009). *Journal of Applied Crystallography* **42**, 339-341.
Dunning, T. H. J. (1989). *The Journal of Chemical Physics* **90**, 1007-1023.
Edwards, A. J. (2011). *Australian Journal of Chemistry* **64**, 869-872.
Flaig, R., Koritsanszky, T., Zobel, D. & Luger, P. (1998). *Journal of the American Chemical Society* **120**, 2227-2238.
Fortes, A. (2020). *Acta Crystallographica Section E* **76**, 1062-1069.
Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Petersson, G. A., Nakatsuji, H., Li, X., Caricato, M., Marenich, A. V., Bloino, J., Janesko, B. G., Gomperts, R., Mennucci, B., Hratchian, H. P., Ortiz, J. V., Izmaylov, A. F., Sonnenberg, J. L., Williams, Ding, F., Lipparini, F., Egidi, F., Goings, J., Peng, B., Petrone, A., Henderson, T., Ranasinghe, D., Zakrzewski, V. G., Gao, J., Rega, N., Zheng, G., Liang, W., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Throssell, K., Montgomery, J. A. Jr, Peralta, J. E., Ogliaro, F., Bearpark, M. J., Heyd, J. J.,

Brothers, E. N., Kudin, K. N., Staroverov, V. N., Keith, T. A., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A. P., Burant, J. C., Iyengar, S. S., Tomasi, J., Cossi, M., Millam, J. M., Klene, M., Adamo, C., Cammi, R., Ochterski, J. W., Martin, R. L., Morokuma, K., Farkas, O., Foresman, J. B. & Fox, D. J. (2009). *Gaussian 09, Gaussian Inc.: Wallingford, Connecticut, USA.*
Fugel, M., Jayatilaka, D., Hupf, E., Overgaard, J., Hathwar, V. R., Macchi, P., Turner, M. J., Howard, J. A. K., Dolomanov, O. V., Puschmann, H., Iversen, B. B., Bürgi, H.-B. & Grabowsky, S. (2018). *IUCrJ* **5**, 32-44.
Funnell, N. P., Bull, C. L., Ridley, C. J. & Capelli, S. (2019). *CrystEngComm* **21**, 4473-4483.
Funnell, N. P., Dawson, A., Francis, D., Lennie, A. R., Marshall, W. G., Moggach, S. A., Warren, J. E. & Parsons, S. (2010). *CrystEngComm* **12**, 2573-2583.
Gavezzotti, A. (2005). *Zeitschrift für Kristallographie - Crystalline Materials* **220**, 499.
Gavezzotti, A. (2007). *Molecular aggregation: structure analysis and molecular simulation of crystals and liquids*. Oxford, UK: Oxford University Press.
Gavezzotti, A. (2011). *New Journal of Chemistry* **35**, 1360-1368.
Gervais, C., Dupree, R., Pike, K. J., Bonhomme, C., Profeta, M., Pickard, C. J. & Mauri, F. (2005). *The Journal of Physical Chemistry A* **109**, 6960-6969.
Görbitz, C. H. (2015). *Crystallography Reviews* **21**, 160-212.
Groom, C. R., Bruno, I. J., Lightfoot, M. P. & Ward, S. C. (2016). *Acta Crystallographica Section B* **72**, 171-179.
Guzmán-Afonso, C., Hong, Y.-l., Colaux, H., Iijima, H., Saitow, A., Fukumura, T., Aoyama, Y., Motoki, S., Oikawa, T., Yamazaki, T., Yonekura, K. & Nishiyama, Y. (2019). *Nature Communications* **10**, 3537.
Hirshfeld, F. L. (1976). *Acta Crystallographica Section A* **32**, 239-244.
Hirshfeld, F. L. (1977). *Theoretica Chimica Acta* **44**, 129-138.
Hunter, S., Sutinen, T., Parker, S. F., Morrison, C. A., Williamson, D. M., Thompson, S., Gould, P. J. & Pulham, C. R. (2013). *The Journal of Physical Chemistry C* **117**, 8062-8071.
Janczak, J., Zobel, D. & Luger, P. (1997). *Acta Crystallographica Section C* **53**, 1901-1904.
Jayatilaka, D. & Dittrich, B. (2008). *Acta Crystallographica Section A* **64**, 383-393.
Jönsson, P.-G. & Kvick, Å. (1972). *Acta Crystallographica Section B* **28**, 1827-1833.
Kanarek, N., Keys, H. R., Cantor, J. R., Lewis, C. A., Chan, S. H., Kunchok, T., Abu-Remaileh, M., Freinkman, E., Schweitzer, L. D. & Sabatini, D. M. (2018). *Nature* **559**, 632-636.
Kassab, E., Langlet, J., Evleth, E. & Akacem, Y. (2000). *Journal of Molecular Structure: THEOCHEM* **531**, 267-282.
Kim, K. & Jordan, K. D. (1994). *The Journal of Physical Chemistry* **98**, 10089-10094.
Kleemiss, F., Dolomanov, O. V., Bodensteiner, M., Peyerimhoff, N., Midgley, L., Bourhis, L. J., Genoni, A., Malaspina, L. A., Jayatilaka, D., Spencer, J. L., White, F., Grundkötter-Stock, B., Steinhauer, S., Lentz, D., Puschmann, H. & Grabowsky, S. (2021). *Chemical Science* **12**, 1675-1692.
Koenig, J. J., Neudorfl, J.-M., Hansen, A. & Breugst, M. (2018). *Acta Crystallographica Section E* **74**, 1067-1070.
Krause, L., Herbst-Irmer, R., Sheldrick, G. M. & Stalke, D. (2015). *Journal of Applied Crystallography* **48**, 3-10.
Lehmann, M. S., Koetzle, T. F. & Hamilton, W. C. (1972a). *International Journal of Peptide and Protein Research* **4**, 229-239.
Lehmann, M. S., Koetzle, T. F. & Hamilton, W. C. (1972b). *Journal of the American Chemical Society* **94**, 2657-2660.
Liao, R.-J., Jiang, L., Wang, R.-R., Zhao, H.-W., Chen, Y., Li, Y., Wang, L., Jie, L.-Y., Zhou, Y.-D., Zhang, X.-N., Chen, Z. & Hu, W.-W. (2015). *Scientific Reports* **5**, 15356.
Lundgren, J.-O. & Liminga, R. (1979). *Acta Crystallographica Section B* **35**, 1023-1027.
Macrae, C. F., Sovago, I., Cottrell, S. J., Galek, P. T. A., McCabe, P., Pidcock, E., Platings, M., Shields, G. P., Stevens, J. S., Towler, M. & Wood, P. A. (2020). *Journal of Applied Crystallography* **53**, 226-235.
Madden, J. J., McGandy, E. L. & Seeman, N. C. (1972). *Acta Crystallographica Section B* **28**, 2377-2382.
Madden, J. J., McGandy, E. L., Seeman, N. C., Harding, M. M. & Hoy, A. (1972). *Acta Crystallographica Section B* **28**, 2382-2389.
Malaspina, L. A., Hoser, A. A., Edwards, A. J., Woińska, M., Turner, M. J., Price, J. R., Sugimoto, K., Nishibori, E., Bürgi, H.-B., Jayatilaka, D. & Grabowsky, S. (2020). *CrystEngComm* **22**, 4778-4789.
Malaspina, L. A., Wieduwilt, E. K., Bergmann, J., Kleemiss, F., Meyer, B., Ruiz-López, M. F., Pal, R., Hupf, E., Beckmann, J., Piltz, R. O., Edwards, A. J., Grabowsky, S. & Genoni, A. (2019). *The Journal of Physical Chemistry Letters* **10**, 6973-6982.
Mallikarjunan, M. & Rao, S. T. (1969). *Acta Crystallographica Section B* **25**, 296-303.
McCandlish, L. E., Stout, G. H. & Andrews, L. C. (1975). *Acta Crystallographica Section A* **31**, 245-249.
McIntyre, G. J., Lemée-Cailleau, M.-H. & Wilkinson, C. (2006). *Physica B: Condensed Matter* **385-386**, 1055-1058.
McMonagle, C. J. & Probert, M. R. (2019). *Journal of Applied Crystallography* **52**, 445-450.
Moggach, S. A., Marshall, W. G., Rogers, D. M. & Parsons, S. (2015). *CrystEngComm* **17**, 5315-5328.
Mondal, S., Prathapa, S. J. & van Smaalen, S. (2012). *Acta Crystallographica Section A* **68**, 568-581.
Myles, D. A. A., Bon, C., Langan, P., Cipriani, F., Castagna, J. C., Lehmann, M. S. & Wilkinson, C. (1997). *Physica B: Condensed Matter* **241-243**, 1122-1130.
Neese, F. (2012). *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2**, 73-78.
Neese, F. (2017). *Wiley Interdisciplinary Reviews: Computational Molecular Science* **8**, e1327.
Novelli, G., Maynard-Casely, H. E., McIntyre, G. J., Warren, M. R. & Parsons, S. (2020). *Crystal Growth & Design* **20**, 7788-7804.
Perdew, J. P., Burke, K. & Ernzerhof, M. (1996). *Physical Review Letters* **77**, 3865-3868.
Piltz, R. (2018). *Journal of Applied Crystallography* **51**, 963-965.
Pollak, P. & Weigend, F. (2017). *Journal of Chemical Theory and Computation* **13**, 3696-3705.
Price, S. L., Braun, D. E. & Reutzel-Edens, S. M. (2016). *Chemical Communications* **52**, 7065-7077.
Prince, E., Wilkinson, C. & McIntyre, G. J. (1997). *Journal of Applied Crystallography* **30**, 133-137.
Probert, M. R., Robertson, C. M., Coome, J. A., Howard, J. A. K., Michell, B. C. & Goeta, A. E. (2010). *Journal of Applied Crystallography* **43**, 1415-1418.
Reeves, M. G., Wood, P. A. & Parsons, S. (2020). *Journal of Applied Crystallography* **53**, 1154-1162.
Schomaker, V. & Trueblood, K. N. (1968). *Acta Crystallographica Section B* **24**, 63-76.
Sheldrick, G. (2015). *Acta Crystallographica Section C* **71**, 3-8.
Shoemaker, D. P., Donohue, J., Schomaker, V. & Corey, R. B. (1950). *Journal of the American Chemical Society* **72**, 2328-2349.
Spek, A. (2009). *Acta Crystallographica Section D* **65**, 148-155.

Stephens, P. J., Devlin, F. J., Chabalowski, C. F. & Frisch, M. J. (1994). *The Journal of Physical Chemistry* **98**, 11623-11627.
Supriya, S., Sivan, S. & Srinivasan, K. (2018). *Crystal Research and Technology* **53**, 1700239.
Thorn, A., Dittrich, B. & Sheldrick, G. M. (2012). *Acta Crystallographica Section A* **68**, 448-451.
Thurber, K. R. & Tycko, R. (2008). *Journal of Magnetic Resonance* **195**, 179-186.
Weigend, F. & Ahlrichs, R. (2005). *Physical Chemistry Chemical Physics* **7**, 3297-3305.
Weisinger-Lewin, Y., Frolow, F., McMullan, R. K., Koetzle, T. F., Lahav, M. & Leiserowitz, L. (1989). *Journal of the American Chemical Society* **111**, 1035-1040.
Wilkinson, C., Cowan, J. A., Myles, D. A. A., Cipriani, F. & McIntyre, G. J. (2002). *Neutron News* **13**, 37-41.
Wilkinson, C., Khamis, H. W., Stansfield, R. F. D. & McIntyre, G. J. (1988). *Journal of Applied Crystallography* **21**, 471-478.
Woińska, M., Grabowsky, S., Dominiak, P. M., Woźniak, K. & Jayatilaka, D. (2016). *Science Advances* **2**, e1600192.
Zhao, Y. & Truhlar, D. G. (2008). *Theoretical Chemistry Accounts* **120**, 215-241.