Supplementary Information

Calibration free approaches for rapid polymorph discrimination *via* low frequency (THz) Raman spectroscopy

Magdalene W. S. Chong*[ab], Martin R. Ward[ac], Catriona McFarlan[b], Andrew J. Parrott[b], Paul Dallin[d], John Andrews[d], Iain D. H. Oswald[c] and Alison Nordon*[ab]

[a] EPSRC Future Continuous Manufacturing and Advanced Crystallisation Research Hub, University of Strathclyde, 99 George Street, Glasgow, G1 1RD, United Kingdom

[b] WestCHEM, Department of Pure and Applied Chemistry, Centre for Process Analytics and Control Technology (CPACT), University of Strathclyde, 295 Cathedral Street, Glasgow, G1 1XL, United Kingdom

[c] Strathclyde Institute of Pharmacy and Biomedical Sciences, University of Strathclyde, 161 Cathedral Street, Glasgow, G4 0RE, UK

[d] Clairet Scientific, 17/18 Scirocco Close, Moulton Park Industrial Estate, Northampton, NN3 6AP, UK

Experimental

Mefenamic acid (MFA; Sigma-Aldrich), acetone (VWR), *N*,*N*-dimethylformamide (DMF, Sigma-Aldrich), and chloroform (Alfa Aesar) were obtained from commercial sources and used as received without further purification. Deionised water was obtained from an in-house Millipore Milli-Q purification system. The cold DMF was obtained by placing a Schott bottle of a similar quantity of DMF into the freezer at the same time as the crystallisation. The general use laboratory freezer is set to −20 °C. Samples were dried in general use Memmert HCP108 laboratory ovens; the temperatures of the sample drying ovens were determined by the requirements of other users in the laboratory.

Sample 1 (MFA form I reference material). This was adapted from a reported procedure,[1] by heating to dissolve the MFA and not cooling by ice. A mixture of MFA (0.70 g) in acetone (25 mL) was heated to dissolve the MFA. The colourless solution was transferred into a freezer and left to cool overnight (*ca.* 20 h). The solid was isolated by vacuum filtration, washed with deionised water, and left to air dry to yield a pale-yellow solid (0.59 g).

Sample 31 (MFA form II reference material). A mixture of MFA (4.00 g) and DMF (20.0 g) was heated to 65 °C with stirring. The resulting yellow solution was transferred to a freezer for 30 min, after which the mixture was stirred to suspend the solid and immediately filtered to isolate the white solid from the peach/yellowish solution and washed with cold DMF. A portion of the solid (*ca.* 1/3) was dried overnight (*ca.* 22 h) in an oven (65 °C) to yield an off-white solid (1.02 g).

Validation sample (50:50 w/w mixture of forms I/II). Sample 29 was stored for 43 months (plastic tub, ambient conditions) and found to have converted to form I. Aged Sample 29 (0.1003 g) and Sample 31 (0.0997 g) were added to a vial and the solids mixed by rotating the vial (horizontal and vertical axes) to mix the contents. This sample was not analysed by THz Raman owing to the instrumentation no longer being available.

Full experimental details for all of the samples are outlined in Table S1. Preparation of Sample 2 was adapted from a reported procedure[1] and most subsequent samples were prepared using modifications of this protocol.

Typical procedure for Samples 14 to 31 (cooling in DMF): A mixture of mefenamic acid (4 g) and DMF (20 g) were heated to 65 °C with stirring. The resulting yellow solution was placed in the freezer and checked periodically every *ca.* 10 min. When nucleation was observed, the sample was removed from the freezer after 10 min and the solid suspended by stirring the mixture of white solid in a peach/yellowish solution. The solid was isolated by filtration, with washing details in Table S1. The solid was divided into three for drying at different temperatures (ambient, 30, 40, or 65 °C). Power cuts interfered with the drying of Samples 20 to 25.

Non-invasive wide-area illumination Raman spectra were acquired using a Kaiser Optical Systems PhAT probe fibre-coupled to a Kaiser Optical Systems RXN1 Raman spectrometer, with a lens giving a 6 mm spot size, focal distance of 25 cm and focussing tube of ca. 20 cm length. The PhAT probe was held in position using a custom containment chamber (Kaiser Optical Systems) and the solid material placed underneath the focussing tube. The 785 nm Kaiser Optical Systems Invictus diode laser was operated at 350 mW at source. THz Raman spectra were acquired using a procedure similar to that previously reported[2] with a 0.4 " Kaiser Optical Systems non-contact optic (NCO) attached to an Ondax THz-Raman probe head fibre-coupled to an RXN1 Raman spectrometer. The 785 nm Ondax CleanLine laser module was operated at 94 % power, giving *ca.* 78 mW at source. The solid material was placed on a laboratory jack underneath the NCO to obtain the optimal signal at a distance of *ca.* 1 cm. Both spectrometers were operated using HoloGRAMS (Kaiser Optical Systems, version 4.3) software. Each spectrum was acquired with an exposure time of 5 s and 3 accumulations; a dark current spectrum was obtained prior to spectral acquisition and automatically subtracted from the acquired spectra. The mid-frequency Raman spectra of Samples 20 to 25 were acquired using an exposure time of 10 s.

Screening powder X-ray diffraction (PXRD) patterns for Samples 1 to 4 were acquired using a Bruker D8 Advance II diffractometer equipped with a Cu Kα source (λ = 1.54056 Å). Samples were lightly ground using a pestle and mortar and deposited onto a polyimide (Kapton, 7.5 µm thickness) film mounted on a sample plate for analysis in transmission mode. Data were acquired over the 2θ range of 4 to 35 ° with a step size of 0.015 ° and 1 s per step. Samples were oscillated in the x-y plane at a speed of 0.3 mm s^{-1}. For immediate analysis of the products (remaining samples), PXRD patterns were acquired using a Bruker D2 Phaser 2nd Gen benchtop diffractometer. The samples were lightly ground using a pestle and mortar and deposited onto a silicon zero background holder. PXRD patterns were acquired over the 2θ range of 4 to 35 ° with a step size of 0.016 °, using Ni filtered Cu Kα irradiation and the X-ray source was operated at 30 kV and 10 mA. The samples were rotated at a rate of 15 rotations per minute during data acquisition.

High resolution PXRD data were collected using a Bruker D8 Advance diffractometer equipped with a Cu sealed tube source (40 kV, 50 mA) and Johanssen Kα1 monochromator. Intensities were recorded using a Lynxeye position sensitive detector. Samples were ground using a pestle and mortar before loading into a 0.7 mm diameter borosilicate glass capillary tube for data collection over the range 3-40 ° 2θ using a step size of 0.017 °.

PXRD analysis of Samples 29 and 31 (stored for 43 months) were performed using a Bruker D8 Discover diffractometer equipped with a Cu Kα source (λ = 1.54056 Å). Samples were deposited onto polymethyl methacrylate sample holders. PXRD patterns were acquired over the 2θ range of 4 to 35 ° with a step size of 0.017 °, using Ni filtered Cu Kα irradiation and the X-ray source was operated at 40 kV and 40 mA. The samples were rotated at a rate of 15 rotations per minute during data acquisition.

Rigid body Rietveld refinement of PXRD data was performed using Topas academic (Bruker, version 5) software. Reference crystal structures were obtained from the CSD database (version 5.45) and used for refinement (XYANAC08,[3] Form I; XYANAC04,[4] Form II; ZAZGAK,[4] DMF solvate). Atomic coordinates were fixed and unit cell parameters allowed to freely refine during fitting. Sample composition (component w/w%) was obtained from the output of the refinement process.

Analysis of spectral data was performed in Matlab (version R2019a, MathWorks) using the GSTools toolbox[5] to import spectral files. The spectral ranges of 6 to 200 cm^{-1} and 550 to 1700 cm^{-1} were used for analysis of the low and mid-frequency Raman spectra, respectively. Principal component analysis (PCA) was performed using PLS_Toolbox (version 8.6.2, Eigenvector Research), with standard normal variate (SNV) preprocessing and mean centring of the spectra.

Multivariate curve resolution (MCR) was performed using the graphical user interface available in PLS_Toolbox using the default options (Table S2). No preprocessing was applied to the low frequency Raman spectra. For the mid-frequency Raman spectra, an automatic weighted least squares baseline (polynomial basis, 2nd order) was applied. Where equality constraints were used, the hardness was set to “hard” and the concentration of components not present were set to zero, with the remaining values unconstrained. Where spectral estimates were used, the hardness was set to 9.5 (which is towards maximum hardness).

Methodological tips regarding the MCR modelling approaches adopted:

- In theory, MCR can work without any information. However, the model can be improved with additional information to use as constraints or initial guesses. A small number of samples were analysed by high resolution PXRD to obtain reference information to use as constraints to improve the MCR model.
- MCR can be applied to a subset of samples to obtain spectral profiles, this approach can be used to generate a spectral profile to use as an initial estimate for a phase where a pure component spectrum is not available (DMF solvate in this instance).
- To quantify components that are present in low concentrations, THz Raman would be more appropriate as it has a lower detection limit compared to mid-frequency Raman.
- For application to *in situ* samples (prior to isolation), the MCR models may need to be adapted to account for additional features arising from solute and solvent in a more complex sample matrix. THz Raman would offer significant benefits in this application as peaks would not arise from the solvent or solute (as there would not be long range order as there would be in a solid). However, signals from the liquid phase could appear as a broad background. For the mid-frequency region, the MCR model would need to discriminate between solid and solute signals and this would need evaluating during method validation.

Table S1 Details of the samples prepared: polymorphs present (from Rietveld refinement of screening PXRD, except for Samples 1, 28, and 31), preparation details (crystallisation mode, solvent used, wash solvent, and drying), and observations. Positive results (Y) are coloured green for clarity. Samples 12 and 26 were not analysed by PXRD.

Sample number	Polymorphs present			Screening PXRD instrument	Crystallisation and work up				Comments
	Form I	Form II	DMF solvate		Mode	Solvent	Washing	Drying	
1	Y	N	N	D8	Cooling, overnight	acetone	water	Ambient	Pale yellow solid
2	Y	Y	Y	D8	Cooling, overnight	DMF	water	Ambient	Large pale yellow crystals
3	Y	Y	N	D8	Cooling, 6 h	DMF	None	65 °C	Off-white solid
4	Y	Y	N	D8	Cooling, 6 h	DMF	None	65 °C	Off-white solid
5	Y	Y	Y	D2	Cooling, 7 h	DMF	None	65 °C	Off-white solid
6	Y	Y	N	D2	Cooling, 7 h	DMF	Cold DMF	65 °C	Off-white solid
7	Y	N	Y	D2	Cooling, 7 h	DMF	water	65 °C	Off-white solid
8	Y	N	N	D2	Slow evaporation	$CHCl_3$	None	None	Yellowish crystals
9	Y	N	N	D2	Slow evaporation	$CHCl_3$	none	None	Off-white fine crystals
10	N	Y	Y	D2	Cooling, 4 h	DMF	Cold DMF	65 °C	White solid
11	Y	Y	Y	D2	Cooling, 4 h	DMF	Cold DMF	Ambient	Large white blocks
12					Cooling, 4 h	$CHCl_3$	None	Ambient	Off-white solid
13	Y	N	N	D2	Cooling, 2 h	$CHCl_3$	None	Ambient	Off-white solid
14	Y	Y	N	D2	Cooling, 4.5 h	DMF	None	65 °C	Off-white solid
15	Y	Y	N	D2	Cooling, 4.5 h	DMF	Cold DMF	65 °C	Off-white solid
16	N	Y	Y	D2	Cooling, 4.5 h	DMF	Cold DMF	Ambient	Large white crystals
17	Y	Y	N	D2	Cooling, 5 h	DMF	None	65 °C	Off-white solid
18	Y	Y	N	D2	Cooling, 5 h	DMF	Cold DMF	65 °C	Off-white solid
19	N	Y	Y	D2	Cooling, 5 h	DMF	Cold DMF	Ambient	Large white crystals
20	Y	Y	Y	D2	Cooling, 40 min	DMF	None	Ambient	Large white crystals
21	N	Y	Y	D2	Cooling, 40 min	DMF	None	40 °C	Off-white solid, power cut
22	Y	Y	N	D2	Cooling, 40 min	DMF	None	65 °C	Off-white solid, power cut
23	Y	Y	Y	D2	Cooling, 35 min	DMF	Cold DMF	Ambient	Off-white solid
24	Y	Y	Y	D2	Cooling, 35 min	DMF	Cold DMF	40 °C	Off-white solid, power cut
25	Y	Y	N	D2	Cooling, 35 min	DMF	Cold DMF	65 °C	Off-white solid, power cut
26					Cooling, 35 min	DMF	None	30 °C	Large white crystals
27	N	Y	Y	D2	Cooling, 35 min	DMF	None	40 °C	Off-white solid
28	Y	Y	N	D2	Cooling, 35 min	DMF	None	65 °C	Off-white solid
29	N	Y	Y	D2	Cooling, 30 min	DMF	Cold DMF	30 °C	Large white crystals
30	Y	Y	Y	D2	Cooling, 30 min	DMF	Cold DMF	40 °C	Off-white solid
31	N	Y	N	D2	Cooling, 30 min	DMF	Cold DMF	65 °C	Off-white solid

Table S2 Default settings in PLS_Toolbox for carrying out MCR.

MCR Constraints	
Contrast	Off
Closure	None
Equality	<none loaded>
Preprocessing	<none selected>
Choose Options	
Standard	
blockdetails	Standard
initmode	1
confidencelimit	0.95
alsoptions.Display	
alsoptions.waitbar	Auto
alsoptions.Non-Negativity	
alsoptions.ccon	fasternnls
alsoptions.scon	fasternnls
alsoptions.Closure	
alsoptions.closure	<none selected>
alsoptions.Initial Guess	
alsoptions.initialguessmethod	exteriorpts
alsoptions.initialguessminnorm	0.03

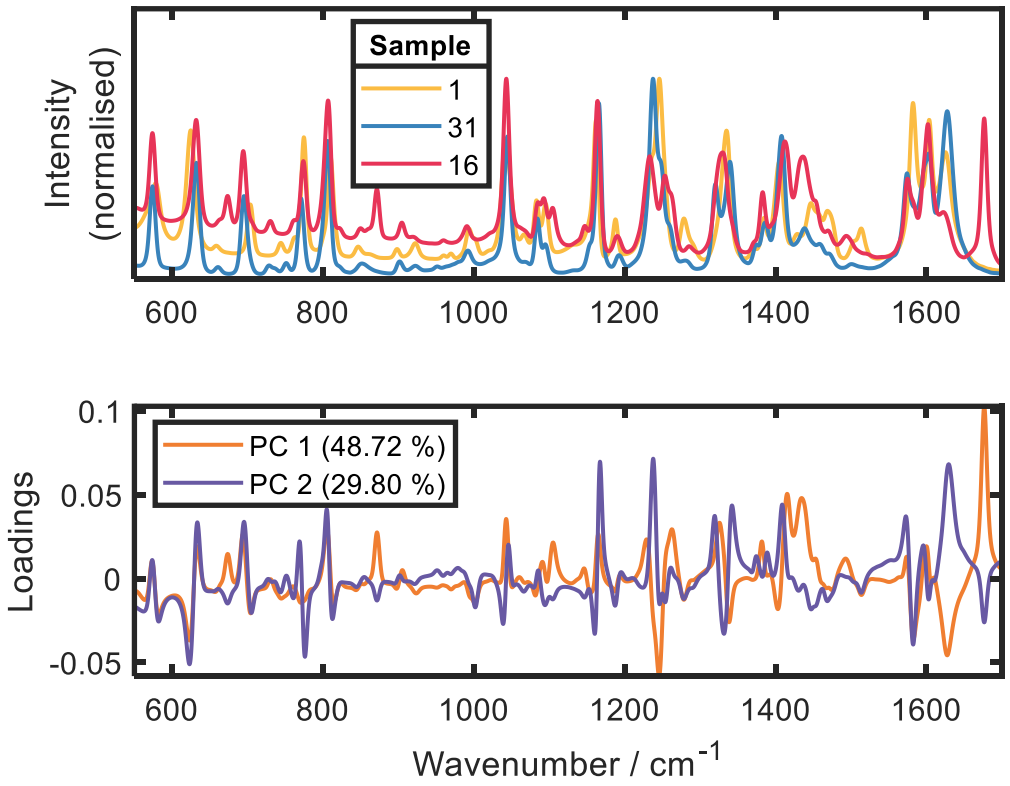

Figure S1 Overlay of (top) mid-frequency Raman spectra (normalised with respect to the maximum peak height) of phase pure Samples 1 and 31 and Sample 16 (predominantly DMF solvate, Figure S7) and (bottom) loadings in the corresponding spectral region from PCA performed on the mid-frequency Raman spectra of the 31 samples prepared.

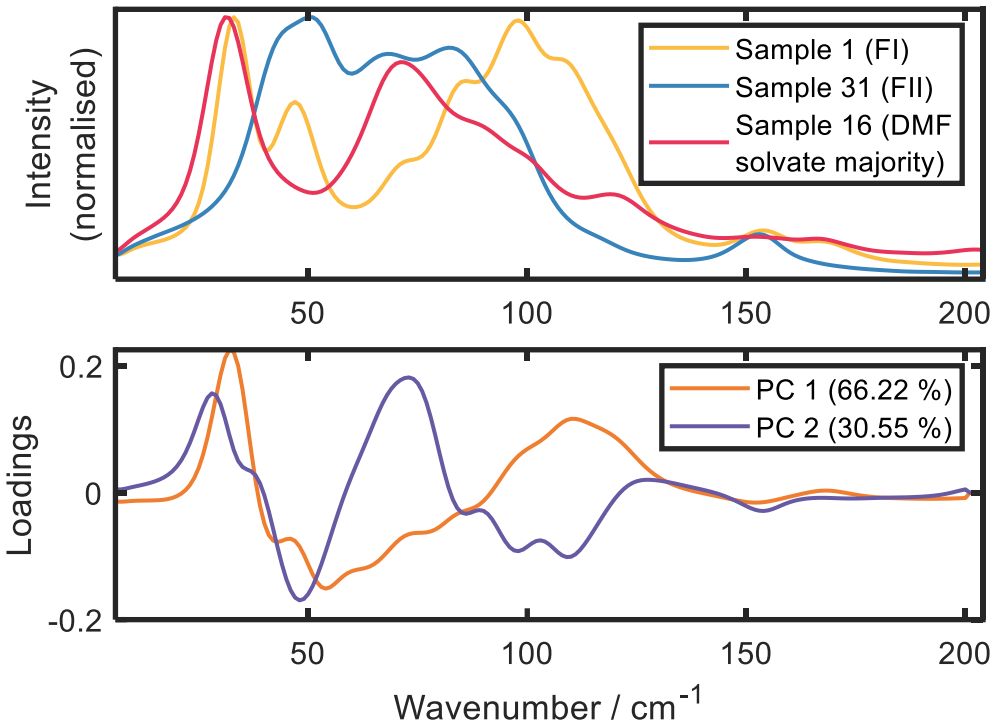

Figure S2 Overlay of (top) THz Raman spectra (normalised with respect to the maximum peak height) of phase pure Samples 1 (form I) and 31 (form II) and Sample 16 (predominantly DMF solvate, Figure S4) and (bottom) loadings in the corresponding spectral region from PCA performed on the THz Raman spectra of the 31 samples prepared.

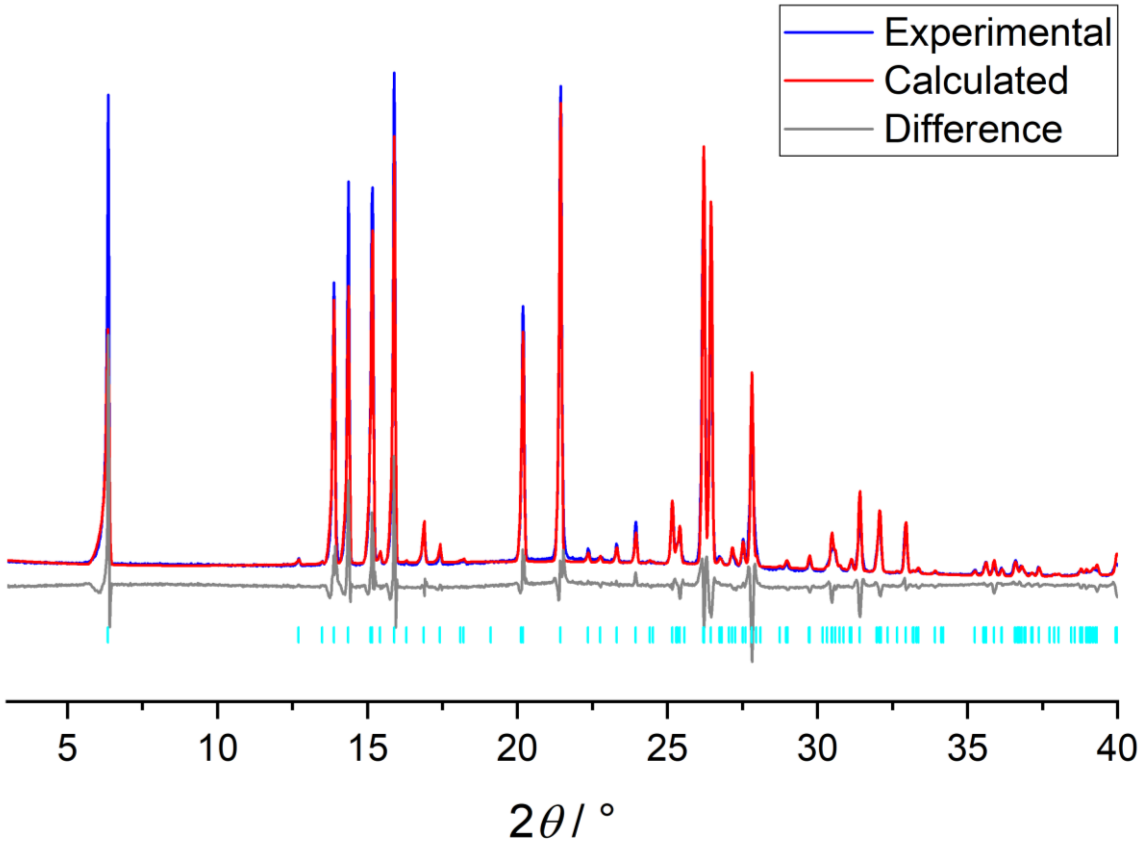

Figure S3 Rietveld refinement of mefenamic acid form I (calculated from CSD entry XYANAC08)[3] against the high resolution PXRD data for Sample 1.

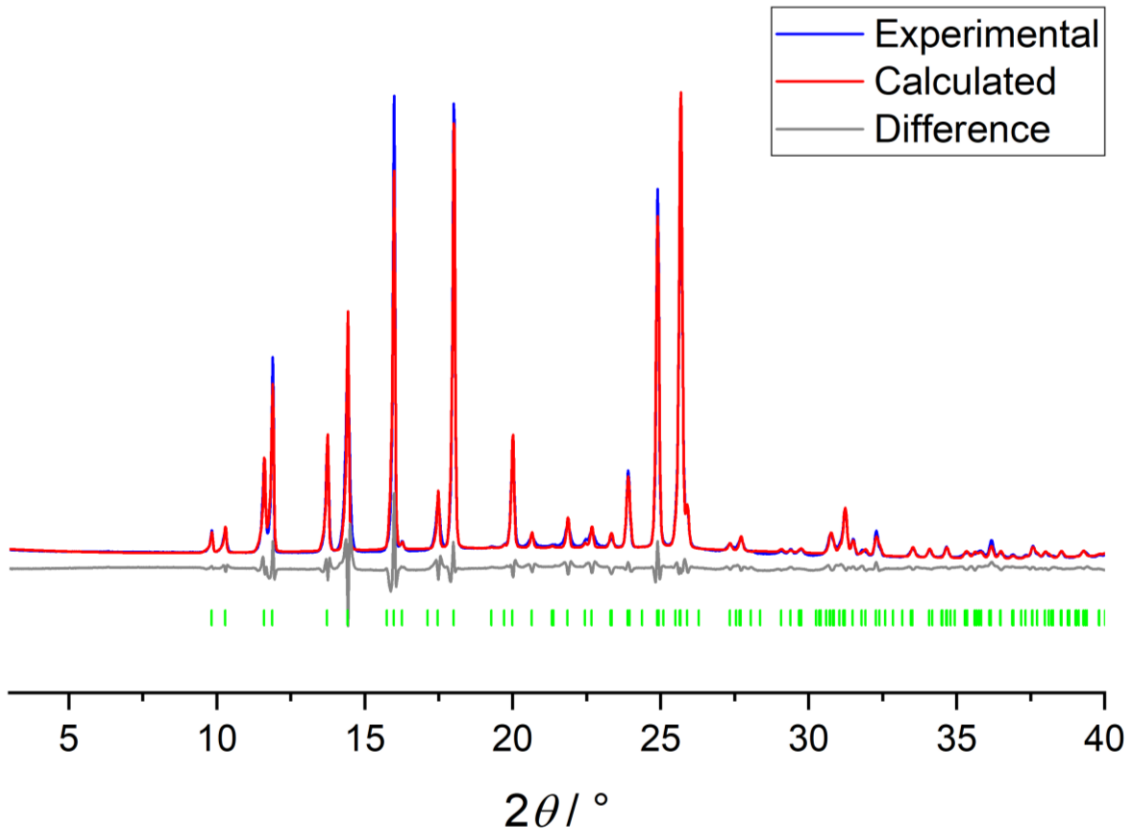

Figure S4 Rietveld refinement of mefenamic acid form II (calculated from CSD entry XYANAC04)[3] against the high resolution PXRD data for Sample 31.

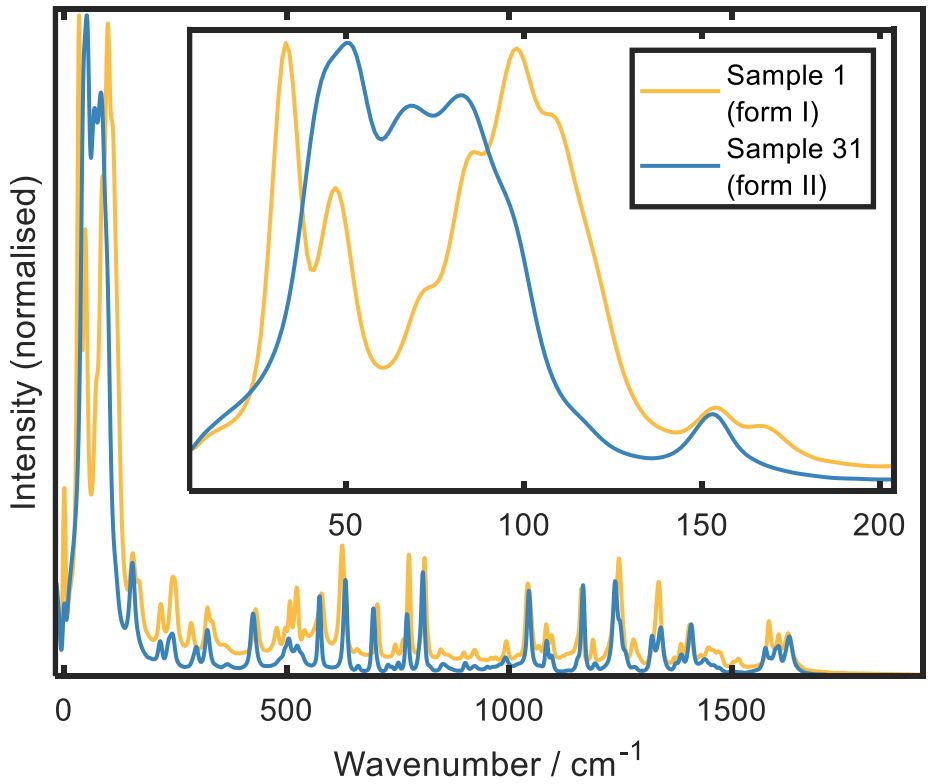

Figure S5 THz Raman spectra (normalised with respect to the maximum peak height) obtained for two polymorphs of mefenamic acid, form I and form II (Samples 1 and 31, respectively, Table S1). The low frequency region is expanded in the inset.

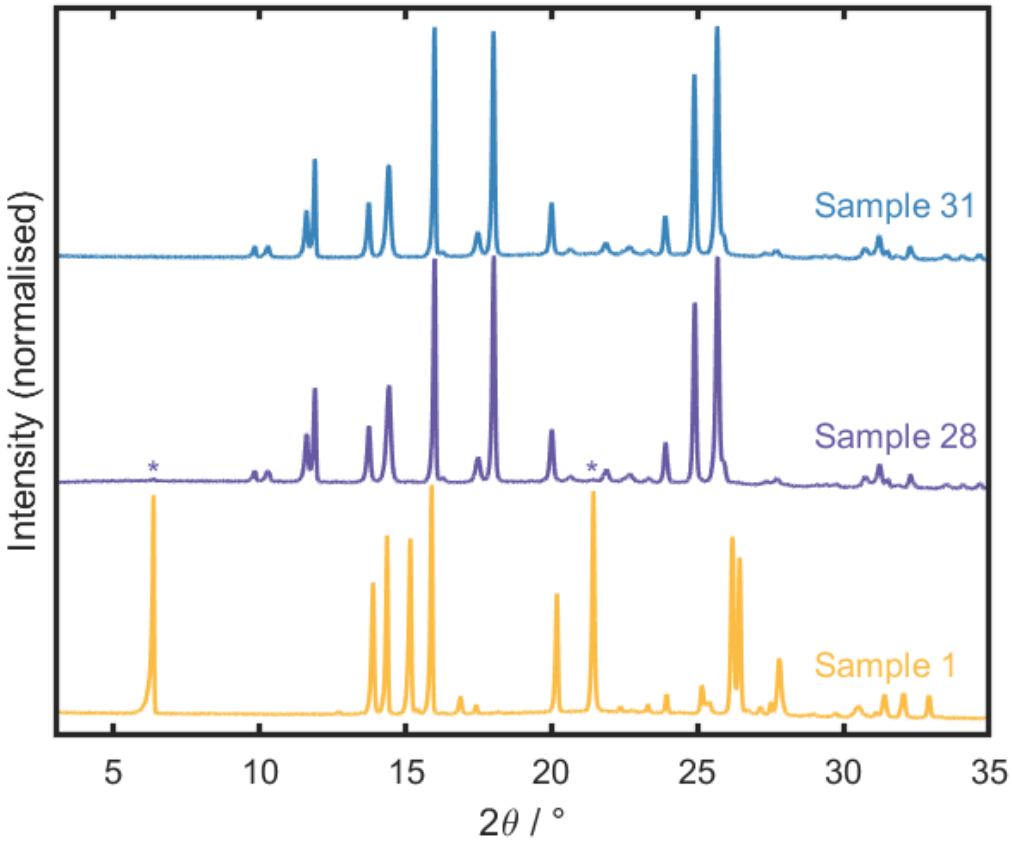

Figure S6 High resolution PXRD patterns obtained for Samples 1 (form I), 31 (form II), and 28 (form II with a small amount of form I present, form I peaks asterisked).

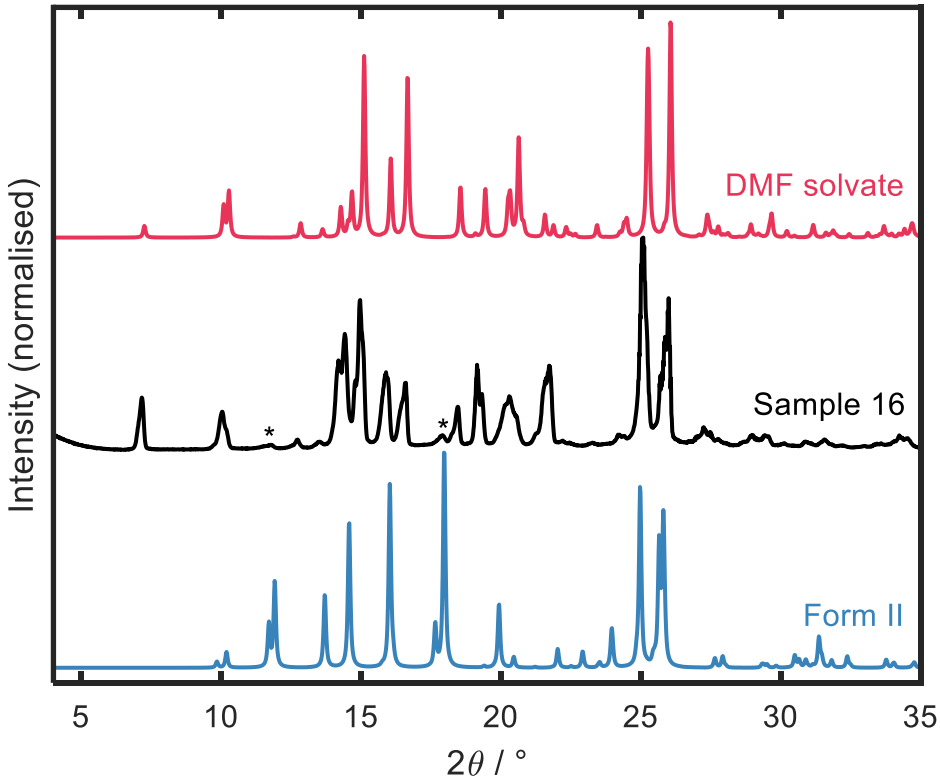

Figure S7 Screening (low resolution) PXRD pattern obtained for Sample 16 (predominantly the DMF solvate, with a small amount of form II present – asterisked peaks), overlaid with reference patterns for the two solid forms calculated from CSD entries XYANAC04 and ZAZGAK.[3]

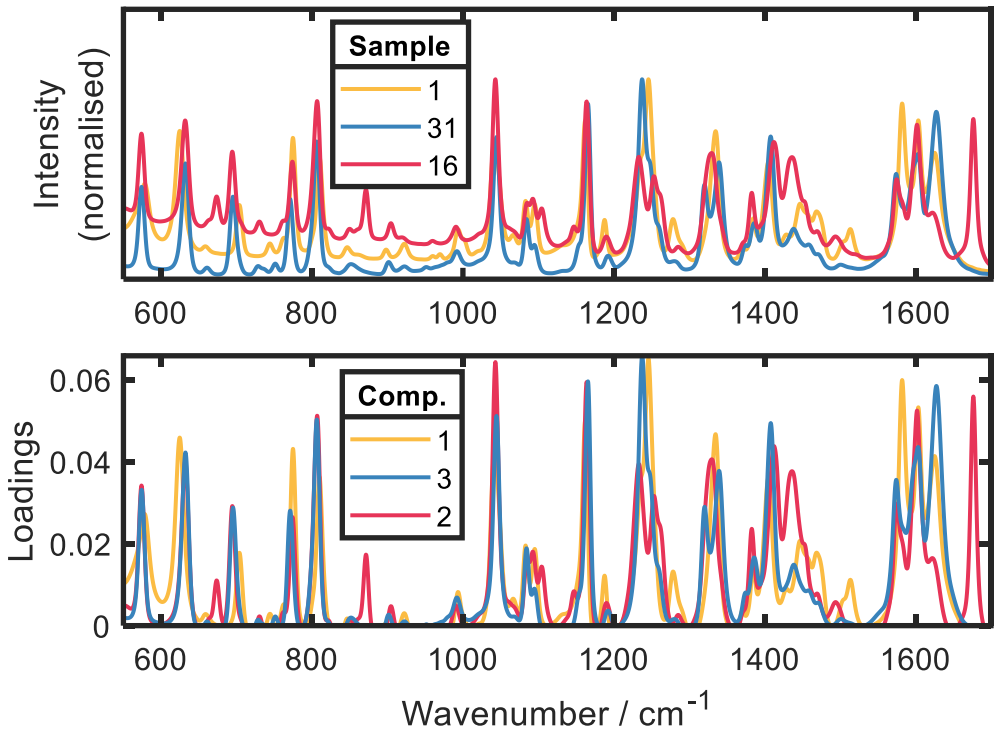

Figure S8 Overlay of (top) mid-frequency Raman spectra (normalised with respect to the maximum peak height) of phase pure Samples 1 (form I) and 31 (form II) and Sample 16 (predominantly DMF solvate, Figure S7) and (bottom) spectral profiles obtained by MCR, where components (comp.) 1 (16.39 %), 3 (60.09 %), and 2 (23.19 %) correspond to form I, form II, and the DMF solvate, respectively.

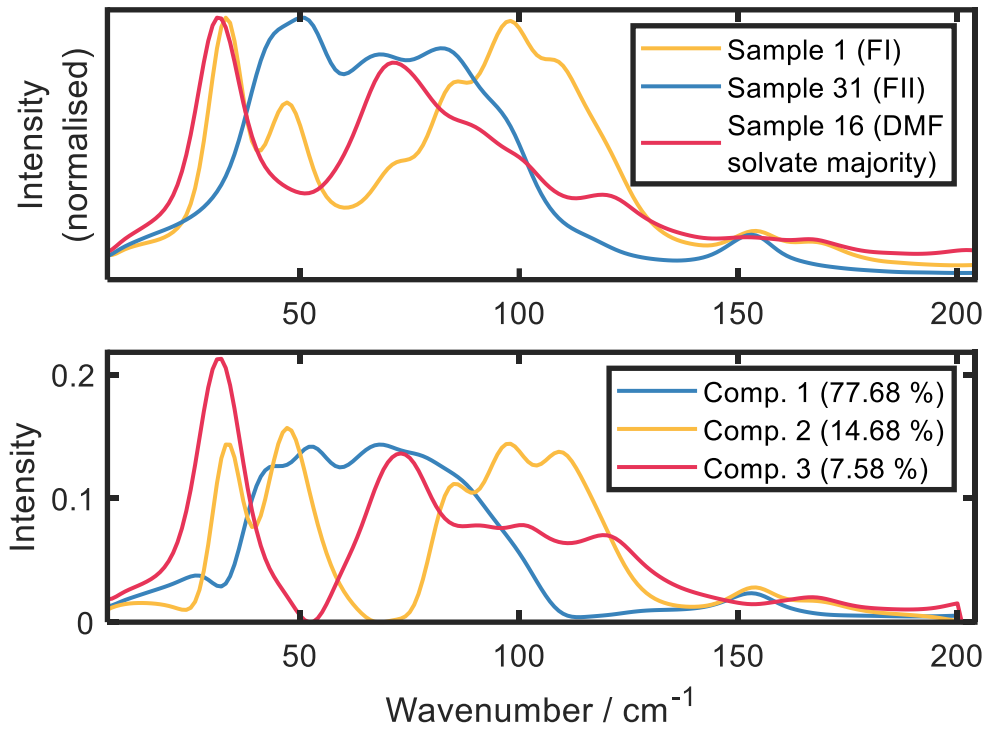

Figure S9 Overlay of (top) THz Raman spectra (normalised with respect to the maximum peak height) of phase pure Samples 1 (form I) and 31 (form II) and Sample 16 (predominantly DMF solvate, Figure S7) and (bottom) spectral profiles obtained by MCR, modelling the dataset as three components (comp.) and using non-negativity constraints in addition to the software default settings (Table S2).

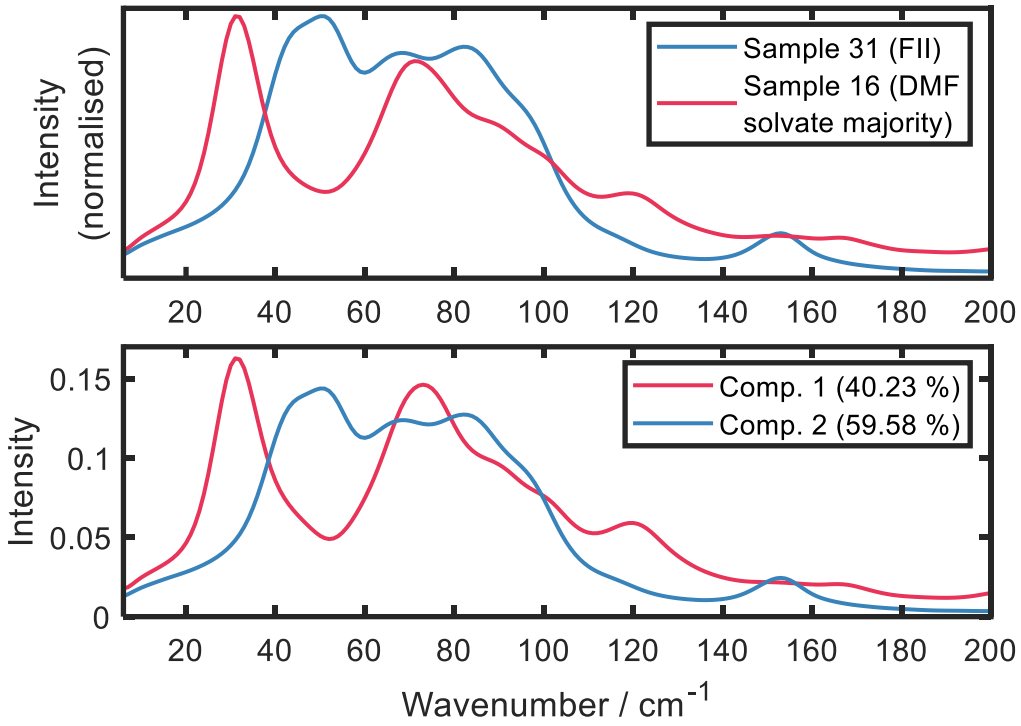

Figure S10 Overlay of (top) THz Raman spectra (normalised with respect to the maximum peak height) of Sample 31 (form II) and Sample 16 (predominantly DMF solvate, Figure S7) and (bottom) spectral profiles obtained by MCR, modelling Sample 31 (form II) and the form II DMF solvate mixture samples (Samples 10, 16, 19, 21, and 29) as two components and using non-negativity constraints in addition to the software default settings (Table S2).

References

1. A. J. Aguiar and J. E. Zelmer, *J. Pharm. Sci.*, 1969, **58**, 983–987.
2. M. W. S. Chong, A. J. Parrott, D. J. Ashworth, A. J. Fletcher and A. Nordon, *Phys. Chem. Chem. Phys.*, 2023, **25**, 14869–14878.
3. A. Perez-Benitez, I. Angel Nieto, S. Bernes, *CSD Communication*, 2021, CCDC 2059210.
4. S. SeethaLekshmi and T. N. Guru Row, *Cryst. Growth Des.*, 2012, **12**, 4283–4289.
5. K. De Gussem, J. De Gelder, P. Vandenabeele and L. Moens, *Chemom. Intell. Lab. Syst.*, 2009, **95**, 49–52.