ChemComm

COMMUNICATION

View Article Online
View Journal | View Issue

Cite this: *Chem. Commun.*, 2025, **61**, 925

Received 20th October 2024,
Accepted 6th December 2024

DOI: 10.1039/d4cc05591h

rsc.li/chemcomm

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]

Application of multivariate curve resolution to non-invasive Raman spectra has been investigated for rapid on-line analysis of crystallisation processes and high-throughput screening. Exploring quantification of mefenamic acid solid forms (form I, form II, and dimethylformamide solvate) from the Raman spectra indicated excellent agreement with off-line X-ray analysis.

The determination of polymorph purity is critical to manufacturing in many industries (*e.g.*, pharmaceutical, fine chemical, food, and functional materials), with the benchmark technique being powder X-ray diffraction (PXRD).[1,2] However, a key issue is that PXRD has limited *in situ* analytical capability, with delays occurring between the process and analysis[3] leading to potential phase transformations. Raman spectroscopy is an alternative technique to PXRD that can address these challenges. The ability to analyse samples directly in the reaction medium without sample preparation (*e.g.*, isolation and/or grinding)[4] and using comparatively shorter timeframes for data acquisition[5] make Raman measurements applicable to high-throughput screening.[2,6] Crucially, *in situ* Raman measurements, especially non-invasive modes,[4,7] can minimise the potential for solid form changes during isolation and analysis[8] (*e.g.*, conversion of metastable phases,[9] loss of solvent from,[10] or changes in polymorph during grinding).[11] In addition, rapid analysis can identify the presence of transient metastable phases that may not be observed during off-line analysis. The enhanced specificity and sensitivity towards polymorph discrimination (peaks arising from the crystalline solid state) offered by low frequency (THz) Raman spectroscopy has been recently demonstrated for pharmaceuticals and materials chemistry.[4,12]

Traditional methods to spectroscopically quantify different solid forms often require reference spectra for the individual phases[13] or a reference (generally off-line) technique. For example, a crystallinity index approach has been applied to hydroxyapatites,[14] conversion of amorphous indomethacin to the γ-form has been monitored using a semi-quantitative methodology,[15] and calibrations have been used to quantify the α and β polymorphs of L-glutamic acid in binary solid mixtures.[16] Calibrations to quantify transient metastable polymorphs that interconvert during analysis are particularly challenging.[17,18] Multivariate curve resolution (MCR), a chemometric method that decomposes a matrix of mixture spectra into concentration and spectral profiles,[19] provides a strategy to overcome such challenges. Application of real-world constraints (*e.g.*, non-negativity) to the MCR algorithm can lead to more chemically interpretable and meaningful results, including resolving *in situ* THz Raman spectra of solid state phase transformations.[20,21] MCR is applied to quantify polymorph contributions in Raman spectra[22] in this work to indicate the crystallisation outcomes for mefenamic acid (MFA), a non-steroidal anti-inflammatory drug that displays concomitant polymorphism where multiple phases can be isolated from the same crystallisation vessel. MCR is applied to resolve spectra where reference data are not available for all components to mimic a future scenario where unknown forms are present and demonstrate applicability to high-throughput experimentation and manufacturing.

The three known polymorphs of MFA[23] are the stable desired form I and metastable form II and form III.[24] MFA form I and form II can be prepared by recrystallisation from acetone and *N*,*N*-dimethylformamide (DMF), respectively.[25] Reference materials for MFA forms I and II were required to assess non-invasive Raman spectroscopy as a tool for screening.

[a] *EPSRC Future Continuous Manufacturing and Advanced Crystallisation Research Hub, University of Strathclyde, 99 George Street, Glasgow, G1 1RD, UK. E-mail: magdalene.chong@strath.ac.uk, alison.nordon@strath.ac.uk*
[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, UK*
[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*

† Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d4cc05591h

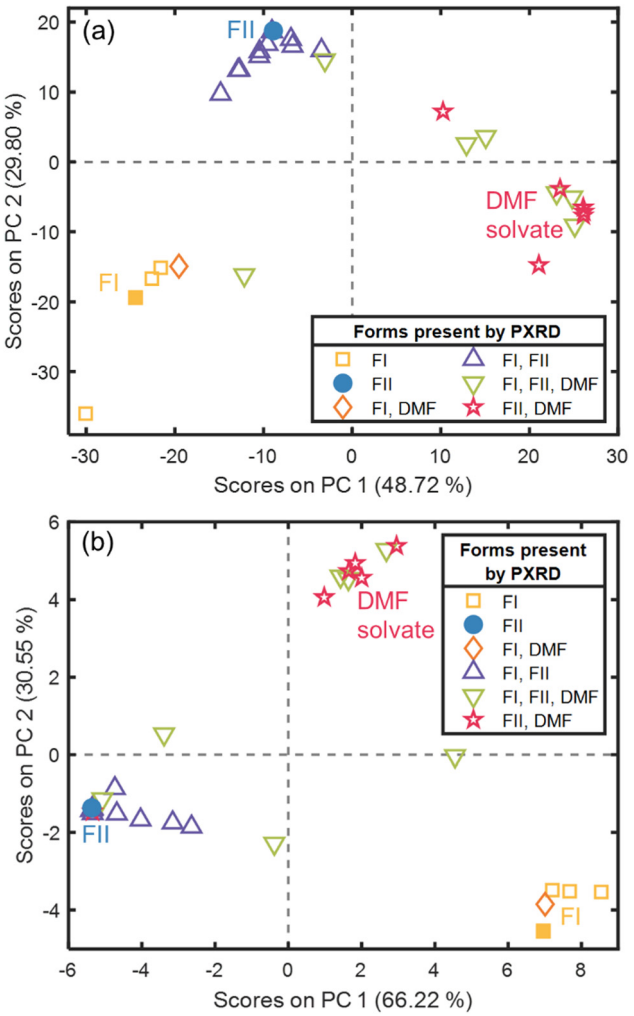

Fig. 1 Scores plot of PC 2 *vs.* PC 1 (variance in brackets) from PCA performed on the (a) mid-frequency and (b) THz Raman data for the 31 samples. The presence of form I (FI), form II (FII), and DMF solvate (denoted DMF) have been assigned from the screening PXRD, except where high resolution PXRD data was obtained (filled symbols).

Metastable form III, which was reported to form in the presence of adenine (during attempted co-crystallisation),[24] was not anticipated to be produced here but should be resolved by MCR if it was present. There is a known DMF solvate, a spectrum of which was deliberately not acquired to test the MCR method for the resolution of unknown phases. For this study, 31 samples were prepared (Table S1, ESI†) and analysed immediately after drying using low resolution screening PXRD and non-invasive Raman and THz Raman spectroscopies to accurately capture the presence of the reported metastable form II and DMF solvate.[24]

Principal component analysis (PCA) was initially employed to establish whether Raman spectroscopy (low and mid-frequency) could identify the solvate species. For both spectral regions, there are three distinct clusters in the scores for the first two principal components (PCs, Fig. 1), matching the number of phases expected in the samples. For the mid-frequency region, comparison of the loadings with reported spectra[26] show PC 2 correlated with form II peaks (632, 695, and 1574 cm^{-1}) and anticorrelated with form I peaks (624, 704, and 1582 cm^{-1}, Fig. S1, ESI†). The third cluster suggests a set of samples comprising predominantly DMF solvate with the PC 1 loadings suggesting unique peaks at 674, 872, and 1677 cm^{-1} (Fig. S1, ESI†). PC 3 is related to variations in the background fluorescence. For the THz Raman spectra, the loadings suggest PC 1 is correlated with form I peaks (33, 98, and 109 cm^{-1}) and anticorrelated with form II peaks (43, 68, and 82 cm^{-1}, Fig. S2, ESI†).[27] If the third cluster is dominated by DMF solvate, then the PC 2 loadings suggest contributions at 31 and 71 cm^{-1} (Fig. S2, ESI†). There is a rotation in the scores for the THz Raman region compared to the mid-frequency region; forms I and II separating along PC 1 instead of PC 2. The scores reflect the comparatively higher sensitivity of THz Raman to polymorphism; 96.77% and 78.58% variance is described by the first two PCs for the low and mid-frequency regions, respectively (Fig. 1). With minimal preprocessing and without using reference spectra, PCA has confirmed that the Raman spectra contain polymorph information and identified specific spectral features for the DMF solvate (for which reference data do not exist).

The phase purity of samples 1 (form I) and 31 (form II) were confirmed by Rietveld refinement[28] of the phases against the PXRD data (Fig. S3 and S4, ESI†), corroborating the PCA results. The THz Raman spectra for the form I and form II polymorphs of MFA obtained (Fig. S5, ESI†) are consistent with other studies.[27] For form I, strong peaks are observed at 33 and 47 cm^{-1} and there is a broad feature with peaks at 73, 86, 98, and 108 cm^{-1}. The form II THz Raman spectrum features a broad scattering profile[27] with peaks at 44, 50, 68, and 82 cm^{-1}. The phase pure form II polymorph was more challenging to prepare; recrystallisation *via* cooling crystallisation from DMF[25] yielded form II with the presence of form I and the DMF solvate (sample 2). A number of factors were investigated to isolate the pure form II: recrystallisation solvent,[24] cooling duration, wash solvent, and drying conditions. To successfully obtain the pure phase (sample 31), the sample was cooled for 30 min, washed with cold DMF, and dried at 65 °C. Sample 28 (Table S1, ESI†) was also analysed by high resolution PXRD, confirming the presence of a small amount of form I amongst predominantly form II (Fig. S6, ESI†). The relative mass ratios of the forms present in the remaining samples (Table S1, ESI†) were provided by Rietveld refinement[28] of the pattern obtained from screening, confirming the composition of the samples determined from PCA of the Raman spectra (Fig. 1). Interrogating the data further, form I was generally present in the mixed phase samples dried at elevated temperatures, likely owing to desolvation of the DMF solvate yielding form I.[24] Sample 16 (PC 1 score 21.05 and PC 2 score −14.74, mid-frequency; PC 1 score 2.95 and PC 2 score 5.38, low frequency, Fig. 1) was identified as comprising predominantly DMF solvate with a small amount of form II (Fig. S7, ESI†).

The mid-frequency Raman data were modelled by MCR to establish the feasibility of this approach to obtain a pure component spectrum for the DMF solvate and quantify the

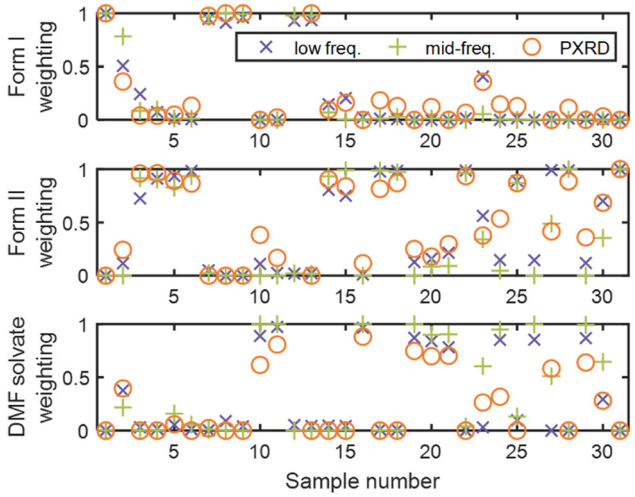

Fig. 2 Sample number and weighting of (top) form I, (middle) form II, and (bottom) the DMF solvate polymorphs of MFA present predicted by MCR modelling of the low and mid-frequency (freq.) Raman spectra and Rietveld refinement of the screening PXRD pattern. For MCR, the dataset was modelled as three components with non-negativity constraints and hard concentration constraints for samples 1, 28, and 31. Spectral estimates of the form I, form II, and DMF solvate were provided for the three components in modelling of the THz Raman data.

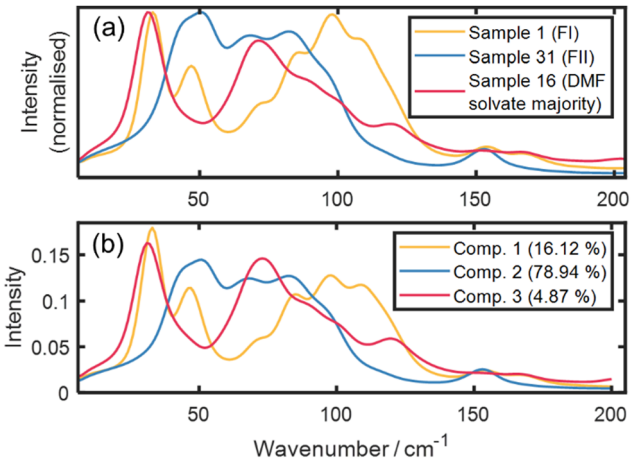

Fig. 3 Overlay of (a) THz Raman spectra (normalised with respect to the maximum peak height) of phase pure samples 1 and 31 and sample 16 (predominantly DMF solvate, Fig. S7, ESI†) and (b) spectral profiles obtained by MCR, modelling the dataset as three components (comp.) with non-negativity and equality constraints. Hard concentration constraints were applied to samples 1, 28, and 31 (from the high resolution PXRD data) and spectral estimates of the form I, form II, and DMF solvate were provided for the three components.

three phases present in each sample. Non-negativity constraints (applied to both spectra and concentrations, Table S2, ESI†) were used in all MCR modelling. Hard concentration (equality) constraints were applied using the information available from the high resolution PXRD data (samples 1, 28, and 31). The resulting spectral profiles of the first and third components feature peaks in the positions corresponding to Raman spectra of phase pure form I and form II, respectively (Fig. S8, ESI†). Component 2 features the peaks unique to the DMF solvate identified *via* PCA. There is reasonable agreement of the predicted relative concentrations of the three components with the screening PXRD data (Fig. 2); discrepancies potentially arise from differences in physical properties.[29] The approach was validated using a sample comprising a mixture of form I (0.50) and form II (0.50); the predicted relative concentrations (form I (0.54), form II (0.46), and DMF solvate (0.00) present in the sample) demonstrate that MCR provides an objective route to quantifying the crystalline phases present in a sample from its Raman spectrum.

The THz Raman spectra were first modelled by MCR as three components using non-negativity constraints (Table S2, ESI†) and the results evaluated by comparing the obtained spectral profiles with available reference spectra.[18,21,30] Whilst the resulting spectral profiles of the first two components exhibit characteristics of the form I and form II polymorphs (Fig. S9, ESI†), irregularities show there are limitations to adopting an unsupervised approach. To obtain an initial spectral estimate of the DMF solvate, a separate MCR model was constructed from a subset of the dataset comprising sample 31 (form II) and the mixtures of form II and the DMF solvate (samples 10, 16, 19, 21, and 29, Table S1, ESI†). As the spectral profile of component 2 exhibits characteristics of the form II polymorph (sample 31), component 1 was extracted as an estimated spectral profile for the DMF solvate (Fig. S10, ESI†). The spectra of samples 1 and 31 (Fig. S5, ESI†) were then used as spectral estimates for components 1 and 2, respectively, in the three-component model. The information from high resolution PXRD (samples 1, 28, and 31) were incorporated as hard concentration (equality) constraints. The spectral profiles thus obtained (Fig. 3) are more representative of the pure component spectra. There is generally good agreement between the concentrations predicted by the MCR model applied to the THz Raman data and Rietveld refinement of the screening PXRD pattern (Fig. 2); especially consistency between the THz Raman and X-ray analysis for whether form I is present. Whilst samples 12 and 26 were not analysed by PXRD, comparable predicted compositions to samples 13 and 29, respectively, are reasonable given the similarities in the experimental conditions (Table S1, ESI†).

Comparing the two Raman frequency regions, better agreement of the concentrations predicted from the THz Raman spectra with the X-ray data (Fig. 2) reflects the expected superior performance for the low frequency region; the associated long range vibrational modes afford higher sensitivity and selectivity towards polymorphism. In particular, for sample 28 the predicted loadings for form I of 0.01 by both THz Raman and high resolution PXRD analysis is comparable to the typical limit of detection of laboratory powder diffractometers (0.01 to 0.02 by mass).[31]

This study has shown the power of MCR applied to Raman spectra benchmarked against PXRD data to elucidate multiple crystalline forms within a sample without requiring reference data. This approach enabled extraction of information on the DMF solvate of MFA, mimicking application to transient phases that may be observed during crystallisation. The lack of sample

preparation and speed of spectral acquisition (<1 min, *vs. ca.* 40 min for PXRD) makes this an attractive method for *in situ* analysis of solid forms in high-throughput screening.

Conceptualisation: MWSC, AN; data curation, formal analysis, validation, visualisation: MWSC, MRW; funding acquisition: IDHO, AN; investigation, project administration, writing – original draft: MWSC; methodology: MWSC, CM, AJP; resources: PD, JA, AN; software: MWSC, MRW, CM, AJP; supervision: IDHO, AN; writing – review & editing: all authors.

The Engineering and Physical Sciences Research Council (EPSRC; EP/P006965/1 and EP/N015401/1), Department for Business, Energy and Industrial Strategy (BEIS), Department of Energy Security & Net Zero (DESNZ), National Nuclear Laboratory (NNL), and Centre for Process Analytics and Control Technology (CPACT) are gratefully acknowledged for funding. The experimental work was carried out in the Continuous Manufacturing and Advanced Crystallisation (CMAC) National Facility, Technology and Innovation Centre, University of Strathclyde; funded with a UK Research Partnership Institute Fund (UKRPIF) capital award, Scottish Funding Council (SFC) reference H13054, from the Higher Education Funding Council for England (HEFCE). Corin Mack and Daniel Powell are thanked for helpful discussions. Coherent are thanked for loan of an instrument.

Data availability

All data underpinning this publication are openly available from the University of Strathclyde KnowledgeBase at **https://doi.org/10.15129/f7e58e3a-af11-42e0-9925-870148b8579b**.

Conflicts of interest

There are no conflicts to declare.

Notes and references

1 A. M. Lunt, H. Fakhruldeen, G. Pizzuto, L. Longley, A. White, N. Rankin, R. Clowes, B. Alston, L. Gigli, G. M. Day, A. I. Cooper and S. Y. Chong, *Chem. Sci.*, 2024, **15**, 2456–2463.
2 R. Storey, R. Docherty, P. Higginson, C. Dallman, C. Gilmore, G. Barr and W. Dong, *Crystallogr. Rev.*, 2004, **10**, 45–56.
3 C. A. Murray, P. M. Scientists, L. Holland, R. O'Brien, A. Richards, A. R. Baker, M. Basham, D. Bond, L. D. Connor, S. J. Day, J. Filik, S. Fisher, P. Holloway, K. Levik, R. Mercado, J. Potter, C. C. Tang, S. P. Thompson and J. E. Parker, *CrystEngComm*, 2024, **26**, 753–763.
4 M. W. S. Chong, A. J. Parrott, D. J. Ashworth, A. J. Fletcher and A. Nordon, *Phys. Chem. Chem. Phys.*, 2023, **25**, 14869–14878.
5 M. K. Hatipoglu, Y. Zaker, D. R. Willett, N. Gupta, J. D. Rodriguez, S. Patankar, P. Capella and H. Yilmaz, *Anal. Chem.*, 2023, **95**, 15325–15332.
6 (*a*) L. Y. Pfund and A. J. Matzger, *ACS Comb. Sci.*, 2014, **16**, 309–313; (*b*) M. L. Peterson, S. L. Morissette, C. McNulty, A. Goldsweig, P. Shaw, M. LeQuesne, J. Monagle, N. Encina, J. Marchionna, A. Johnson, J. Gonzalez-Zugasti, A. V. Lemmo, S. J. Ellis, M. J. Cima and Ö. Almarsson, *J. Am. Chem. Soc.*, 2002, **124**, 10958–10959.
7 K. A. Esmonde-White, M. Cuellar, C. Uerpmann, B. Lenain and I. R. Lewis, *Anal. Bioanal. Chem.*, 2017, **409**, 637–649.
8 M. L. Clapham, R. E. Leighton, C. J. Douglas and R. R. Frontiera, *J. Chem. Phys.*, 2021, **155**, 234703.
9 (*a*) Y. Hu, A. Erxleben, B. K. Hodnett, B. Li, P. McArdle, Å. C. Rasmuson and A. G. Ryder, *Cryst. Growth Des.*, 2013, **13**, 3404–3413; (*b*) M. R. Chierotti, L. Ferrero, N. Garino, R. Gobetto, L. Pellegrino, D. Braga, F. Grepioni and L. Maini, *Chem. – Eur. J.*, 2010, **16**, 4347–4358.
10 R. J. Davey, M. Brychczynska, G. Sadiq, G. Dent and R. G. Pritchard, *CrystEngComm*, 2013, **15**, 856–859.
11 (*a*) M. R. Ward, S. Younis, A. J. Cruz-Cabeza, C. L. Bull, N. P. Funnell and I. D. H. Oswald, *CrystEngComm*, 2019, **21**, 2058–2066; (*b*) A. V. Trask, N. Shan, W. D. S. Motherwell, W. Jones, S. Feng, R. B. H. Tan and K. J. Carpenter, *Chem. Commun.*, 2005, 880–882.
12 (*a*) T. Salzillo, R. G. Della Valle, E. Venuti, A. Brillante, T. Siegrist, M. Masino, F. Mezzadri and A. Girlando, *J. Phys. Chem. C*, 2016, **120**, 1831–1840; (*b*) E. M. Kleist and M. T. Ruggiero, *Cryst. Growth Des.*, 2022, **22**, 939–953; (*c*) K. Bērziņš, S. J. Fraser-Miller and K. C. Gordon, *Int. J. Pharm.*, 2021, **592**, 120034.
13 Y. Sa, Y. Guo, X. Feng, M. Wang, P. Li, Y. Gao, X. Yang and T. Jiang, *New J. Chem.*, 2017, **41**, 5723–5731.
14 J. Kirkham, T. M. Korter, K. Bērziņš, C. M. McGoverin, K. C. Gordon and S. J. Fraser-Miller, *Cryst. Growth Des.*, 2023, **23**, 5748–5761.
15 P. J. Larkin, J. Wasylyk and M. Raglione, *Appl. Spectrosc.*, 2015, **69**, 1217–1228.
16 L. Helmdach, M. P. Feth and J. Ulrich, *Org. Process Res. Dev.*, 2013, **17**, 585–598.
17 M. Inoue, T. Osada, H. Hisada, T. Koide, T. Fukami, A. Roy and J. Carriere, *J. Pharm. Sci.*, 2023, **112**, 225–229.
18 K. Bērziņš, S. J. Fraser-Miller, R. Di, J. Liu, L. Peltonen, C. J. Strachan, T. Rades and K. C. Gordon, *Mol. Pharm.*, 2021, **18**, 1408–1418.
19 (*a*) A. De Juan, J. Jaumot and R. Tauler, *Anal. Methods*, 2014, **6**, 4964–4976; (*b*) A. de Juan and R. Tauler, *Anal. Chim. Acta*, 2021, **1145**, 59–78.
20 B. P. Bezerra, J. C. Fonseca, Y. S. de Oliveira, M. S. A. de Santana, K. F. Silva, B. S. Araújo and A. P. Ayala, *Vib. Spectrosc.*, 2016, **86**, 90–96.
21 (*a*) P. I. J. G. Remoto, K. Bērziņš, S. J. Fraser-Miller, T. M. Korter, T. Rades, J. Rantanen and K. C. Gordon, *Pharmaceutics*, 2023, **15**, 1526; (*b*) K. Bērziņš, P. I. J. G. Remoto, S. J. Fraser-Miller and K. C. Gordon, *Cryst. Growth Des.*, 2022, **22**, 2946–2953; (*c*) P. I. J. G. Remoto, K. Bērziņš, S. J. Fraser-Miller, T. M. Korter, T. Rades, J. Rantanen and K. C. Gordon, *Cryst. Growth Des.*, 2022, **22**, 2733–2741; (*d*) K. Bērziņš, S. J. Fraser-Miller, T. Rades and K. C. Gordon, *Mol. Pharm.*, 2019, **16**, 3678–3686.
22 (*a*) S. Piqueras, L. Duponchel, R. Tauler and A. de Juan, *Anal. Chim. Acta*, 2014, **819**, 15–25; (*b*) Y. Xie, W. Cao, S. Krishnan, H. Lin and N. Cauchon, *Pharm. Res.*, 2008, **25**, 2292–2301.
23 M. W. S. Chong, S. Ottoboni, A. R. G. Martin, D. Bowering, C. J. Price, A. Nordon, I. D. H. Oswald and M. R. Ward, *Acta Crystallogr., Sect. E: Crystallogr. Commun.*, 2022, **78**, 1209–1212.
24 S. SeethaLekshmi and T. N. Guru Row, *Cryst. Growth Des.*, 2012, **12**, 4283–4289.
25 A. J. Aguiar and J. E. Zelmer, *J. Pharm. Sci.*, 1969, **58**, 983–987.
26 (*a*) A. J. Alvarez, A. Singh and A. S. Myerson, *Cryst. Growth Des.*, 2009, **9**, 4181–4188; (*b*) I. S. Lee, A. Y. Lee and A. S. Myerson, *Pharm. Res.*, 2008, **25**, 960–968.
27 S. Roy, B. Chamberlin and A. J. Matzger, *Org. Process Res. Dev.*, 2013, **17**, 976–980.
28 H. M. Rietveld, *Phys. Scr.*, 2014, **89**, 098002.
29 Z.-P. Chen, L.-M. Li, J.-W. Jin, A. Nordon, D. Littlejohn, J. Yang, J. Zhang and R.-Q. Yu, *Anal. Chem.*, 2012, **84**, 4088–4094.
30 E. M. Paiva, R. L. Ribessi, C. F. Pereira and J. J. R. Rohwedder, *Spectrochim. Acta, Part A*, 2020, **228**, 117798.
31 (*a*) K. J. Crowley and G. Zografi, *J. Pharm. Sci.*, 2002, **91**, 492–507; (*b*) R. Surana and R. Suryanarayanan, *Powder Diffr.*, 2000, **15**, 2–6.

