Electronic Supplementary Information:

Same or different – that is the question: identification of crystal forms from crystal structure data

Pietro Sacchi,[a] Matteo Lusi,[b] Aurora J. Cruz-Cabeza,[a]* Elisa Nauha[c] and Joel Bernstein[d]‡*

[a] Department of Chemical Engineering and Analytical Science, University of Manchester, Manchester, M13 9PL, United Kingdom. E-mail: aurora.cruzcabeza@manchester.ac.uk
[b] SSPC, the SFI Research Centre for Pharmaceuticals, Bernal Institute, Department of Chemical Sciences, University of Limerick, Co. Limerick, Ireland. E-mail: matteo.lusi@ul.ie
[c] School of Chemistry, University of Lincoln, Joseph Banks Laboratories, Green Lane, Lincoln, LN6 7DL, United Kingdom.
[d] NYU Abu Dhabi, Abu Dhabi, United Arab Emirates
‡ Professor Joel Bernstein passed away in January 2019

Table of contents

A. Calculation of rmsd for cell lengths (CL) and angles (CA) for structures in Group one

The similarity between unit cells for selected pairs is expressed by calculating the root mean squared deviations of cell axes lengths *a*, *b* and *c* (rmsd-CL) and cell angles *α*, *β* and *γ* (rmsd-CA). The same applies for the parameters of the reduced cells.

$$rmsd - CL = \sqrt{\frac{(a_1 - a_2)^2 + (b_1 - b_2)^2 + (c_1 - c_2)^2}{3}}$$

$$rmsd - CA = \sqrt{\frac{(\alpha_1 - \alpha_2)^2 + (\beta_1 - \beta_2)^2 + (\gamma_1 - \gamma_2)^2}{3}}$$

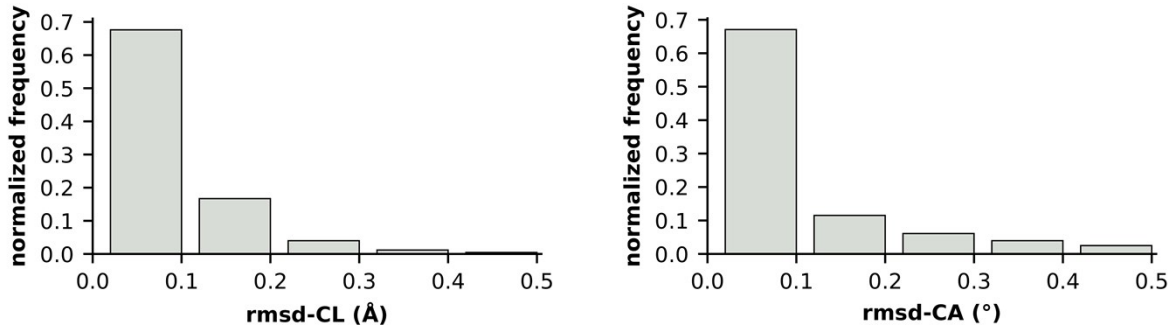


Figure S1. Histograms of rmsd-CL (bottom left) and rmsd-CA (bottom right) for a reduced 19 405 dataset of structures determined with the same space group and settings belonging to Group one (see main text).

B. Crystallographic Issues

B.1 ZEDCUG-ZEDCUG01

	S.G.	Temp. (K)	Cell. Axes (Å)	Cell. Ang. (°)	Red. Cell. Axes (Å)	Red. Cell. Ang (°)
ZEDCUG	*Cc*	RT	**a** 16.632 **b** 14.074(5) **c** 16.872(4)	**α** 90 **β** 92.98(1) **γ** 90	**a** 10.8938 **b** 10.8938 **c** 16.872	**α** 87.7256 **β** 87.7256 **γ** 80.4758
ZEDCUG01	*C*2/*c*	RT	**a** 16.632 **b** 14.074(5) **c** 16.872(4)	**α** 90 **β** 92.98(1) **γ** 90	**a** 10.8938 **b** 10.8938 **c** 16.872	**α** 87.7256 **β** 87.7256 **γ** 80.4758

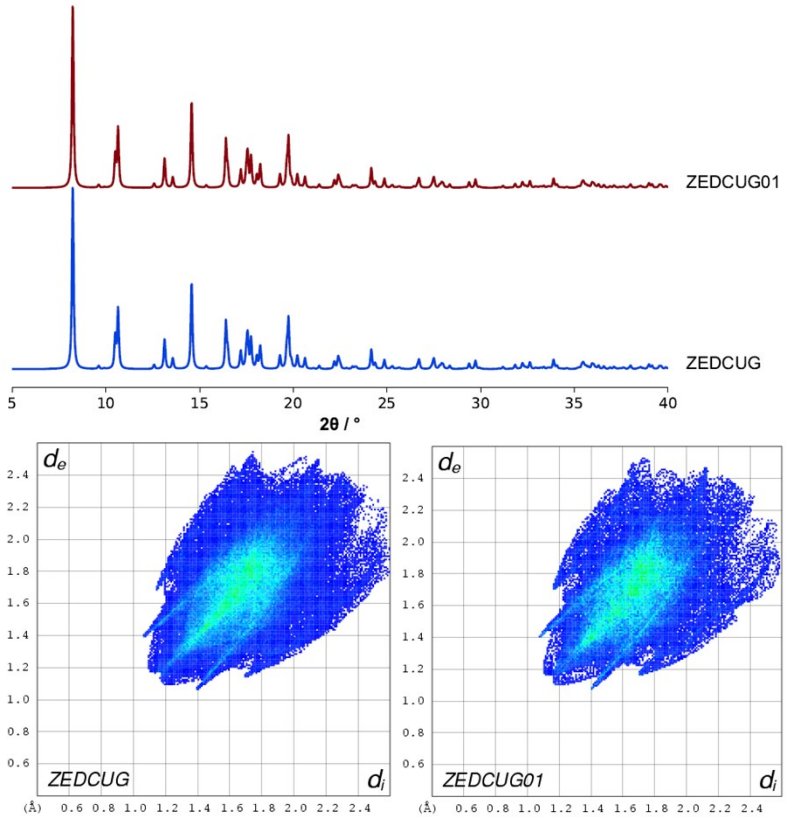
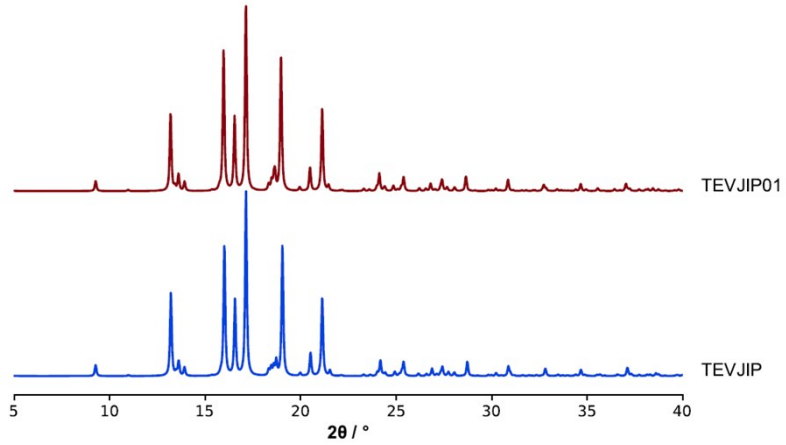


Figure S2. Calculated powder patterns and fingerprint plots for ZEDCUG and ZEDCUG01

B.2 TEVJIP-TEVJIP01

	S.G.	Temp. (K)	Cell. Axes (Å)	Cell. Ang. (°)	Red. Cell. Axes (Å)	Red. Cell. Ang (°)
TEVJIP	$P2_1$	175	**a** 7.0752(4) **b** 8.8818(6) **c** 19.0839(14)	**α** 90 **β** 91.796(6) **γ** 90	**a** 7.0752 **b** 8.8818 **c** 19.0839	**α** 90 **β** 91.796 **γ** 90
TEVJIP01	$P2_1$	200	**a** 7.098(6) **b** 8.894(7) **c** 19.076(16)	**α** 90 **β** 91.500(14) **γ** 90	**a** 7.098(6) **b** 8.894(7) **c** 19.076(16)	**α** 90 **β** 91.500(14) **γ** 90



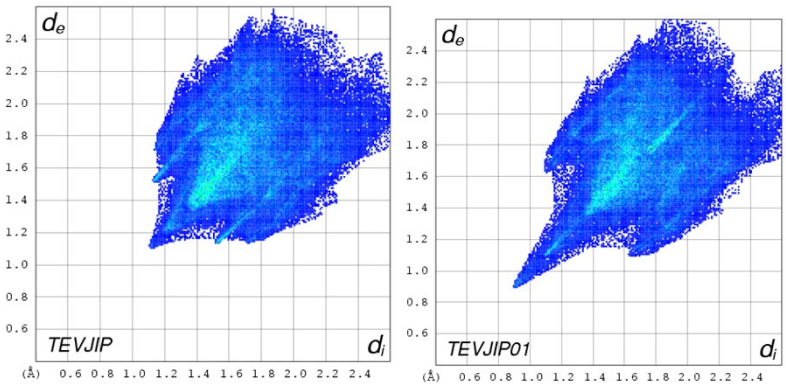


Figure S3. Calculated powder patterns and fingerprint plots for TEVJIP and TEVJIP01

B.3 TCLBEN02-TCLBEN07

	S.G.	Temp. (K)	Cell. Axes (Å)	Cell. Ang. (°)	Red. Cell. Axes (Å)	Red. Cell. Ang (°)
TCLBEN02	$P2_1/n$	175	**a** 3.7956(12) **b** 10.5175(19) **c** 9.5648(13)	**α** 90 **β** 91.28(3) **γ** 90	**a** 3.7956(12) **b** 9.5648(13) **c** 10.5175(19)	**α** 90 **β** 90 **γ** 99.723
TCLBEN07	$P2_1/n$	200	**a** 3.81117(4) **b** 10.53867(4) **c** 9.57198(3)	**α** 90 **β** 99.7088(5) **γ** 90	**a** 3.81117(4) **b** 9.57198(3) **c** 10.53867(4)	**α** 90 **β** 90 **γ** 99.7088

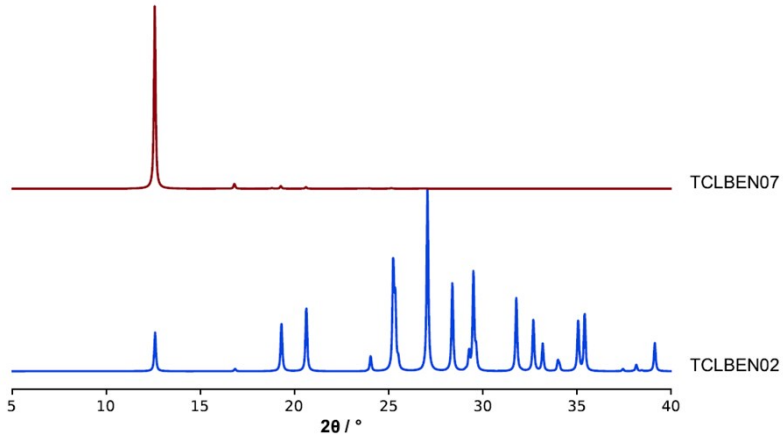


Figure S4. Calculated powder patterns for TCLBEN and TCLBEN07

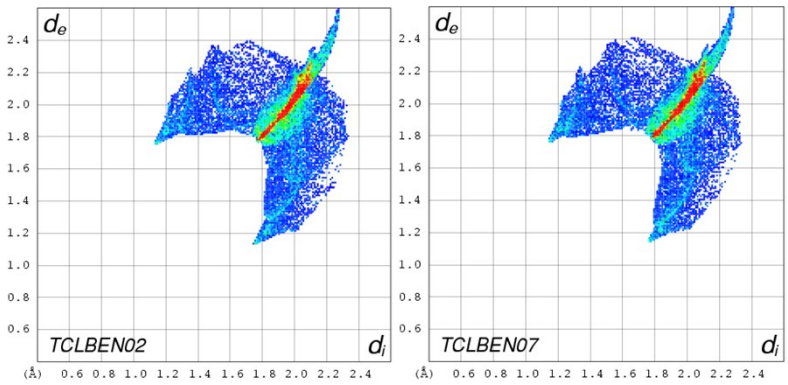


Figure S5. Fingerprint plots for TCLBEN and TCLBEN07

C. Low versus High Temperature/Pressure Structures

C.1 ACETAC-ACETAC07

	S.G.	Temp. (K)	Cell. Axes (Å)	Cell. Ang. (°)	Red. Cell. Axes (Å)	Red. Cell. Ang (°)
ACETAC*	$Pna2_1$	278	**a** 13.32(2) **b** 4.08(1) **c** 5.77(1)	**α** 90 **β** 90 **γ** 90	**a** 4.08 **b** 5.77 **c** 13.32	**α** 90 **β** 90 **γ** 90
ACETAC07	$Pna2_1$	40	**a** 13.151(3) **b** 3.923(1) **c** 5.762(1)	**α** 90 **β** 90 **γ** 90	**a** 3.923 **b** 5.762 **c** 13.151	**α** 90 **β** 90 **γ** 90

* H atoms where added manually. The automating addition with Mercury did not place the atoms correctly.

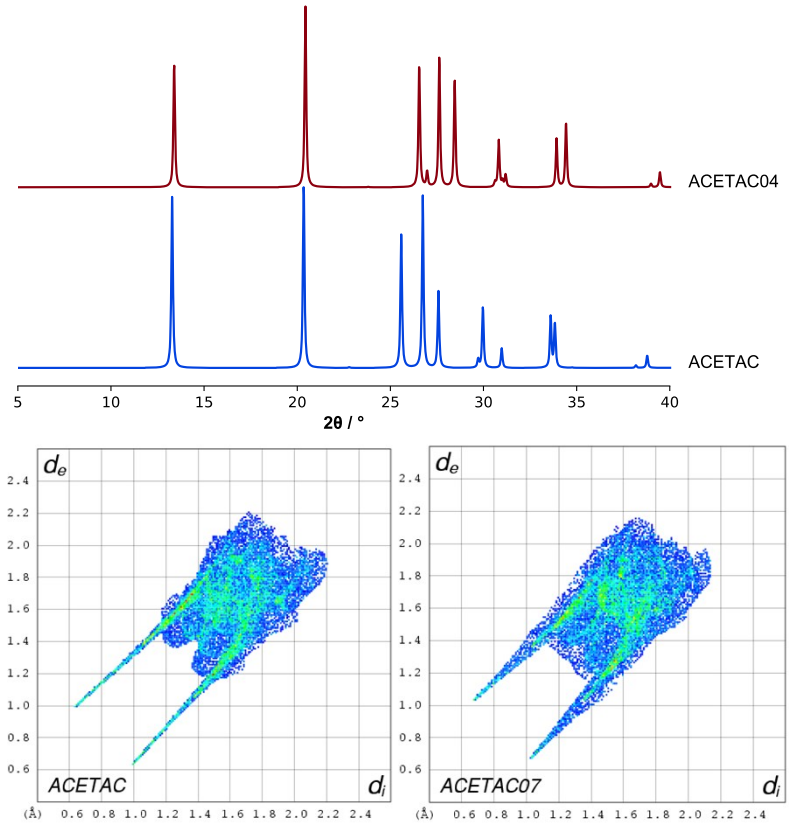


Figure S6. Calculated powder patterns and fingerprint plots for ACETAC and ACETAC04

C.2 DIJHUA-DIJHUA01

	S.G.	Temp. (K)	Cell. Axes (Å)	Cell. Ang. (°)	Red. Cell. Axes (Å)	Red. Cell. Ang (°)
DIJHUA	$P2_1/n$	123	**a** 18.767(4) **b** 10.205(2) **c** 5.6858(12)	**α** 90 **β** 93.557(4) **γ** 90	**α** 5.6858 **b** 10.205 **c** 18.767	**α** 90 **β** 93.557 **γ** 90
DIJHUA01	$P2_1/n$	293	**a** 20.498(5) **b** 9.317(3) **c** 5.9004(15)	**α** 90 **β** 93.759(5) **γ** 90	**α** 5.9004 **b** 9.317 **c** 20.498	**α** 90 **β** 93.759 **γ** 90

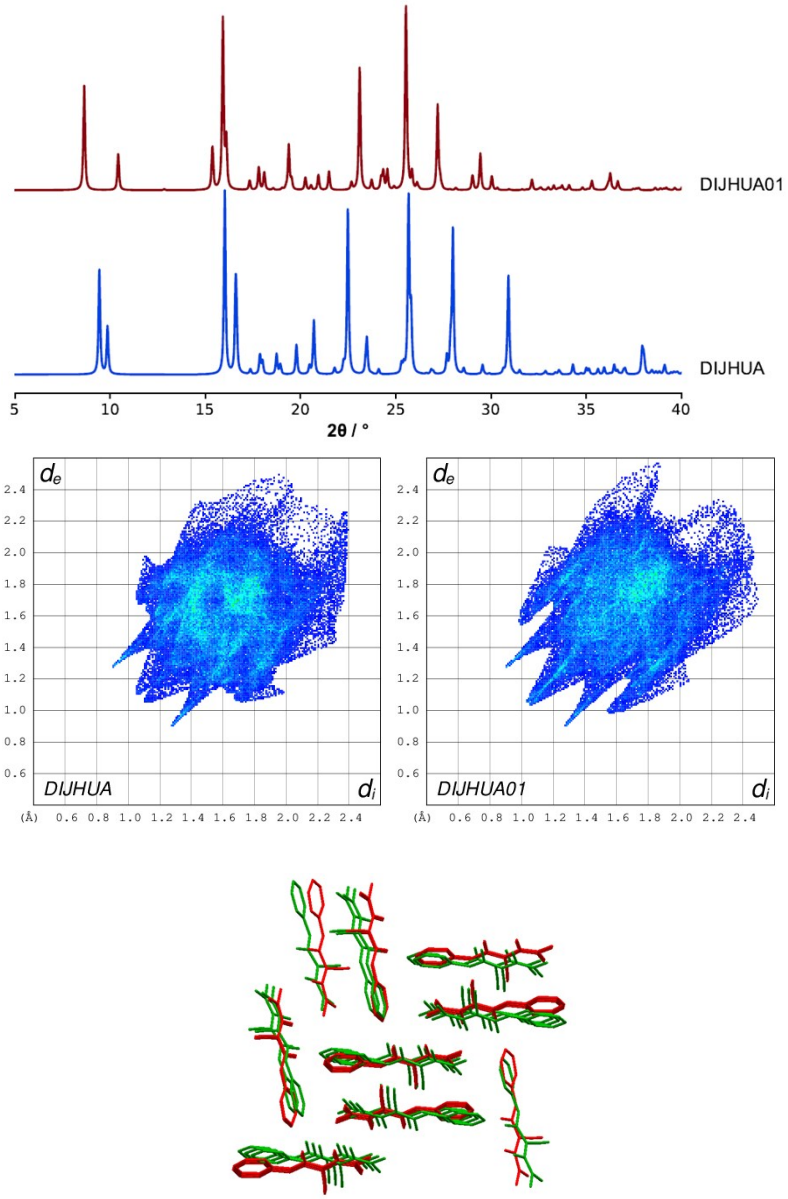


Figure S7. Calculated powder patterns, fingerprint plots and molecular overlay for DIJHUA and DIJHUA01

C.3 JAYDUI-JAYDUI06

	S.G.	Temp. (K) Press. GPa	Cell. Axes (Å)	Cell. Ang. (°)	Red. Cell. Axes (Å)	Red. Cell. Ang (°)
JAYDUI	$P2_1/n$	30 0	**a** 4.1480(14) **b** 12.612(5) **c** 6.977(3)	**α** 90 **β** 91.28(3) **γ** 90	**a** 4.1480(14) **b** 6.977(3) **c** 12.612(5)	**α** 90 **β** 90 **γ** 91.28
JAYDUI06	$P2_1/n$	RT 6	**a** 3.689(3) **b** 11.651(9) **a** 6.151(5)	**α** 90 **β** 90.48(7) **γ** 90	**a** 3.689(3) **b** 6.151(5) **c** 11.651(9)	**α** 90 **β** 90 **γ** 90.48

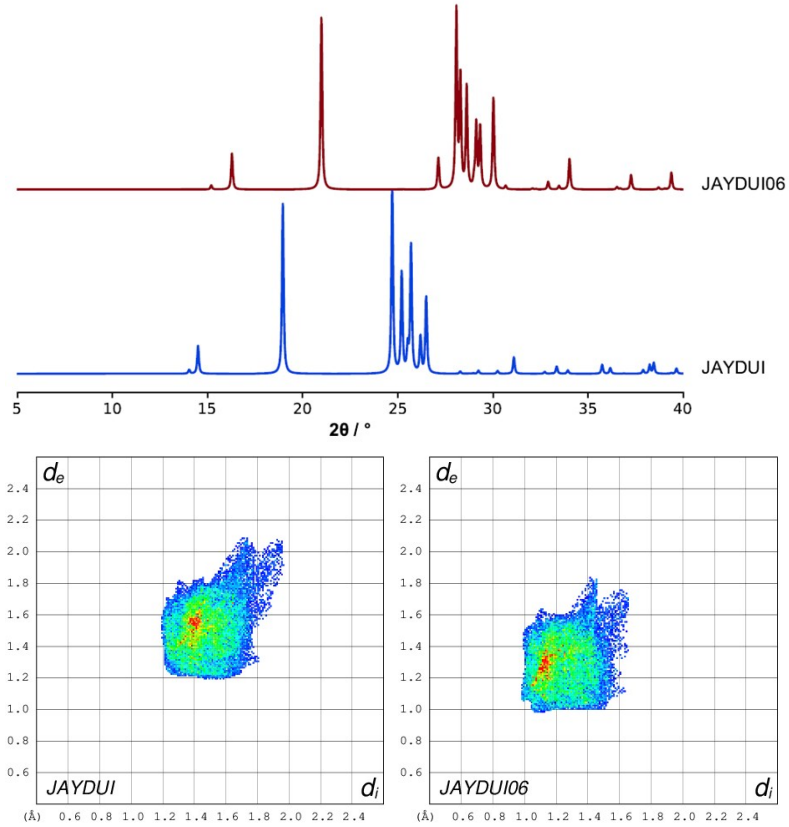


Figure S8. Calculated powder patterns and fingerprint plots for JAYDUI and JAYDUI06

D. Low versus High Temperature Phases

D.1 MOSTIX-MOSTIX01

	S.G.	Temp. (K)	Cell. Axes (Å)	Cell. Ang. (°)	Red. Cell. Axes (Å)	Red. Cell. Ang (°)
MOSTIX	$P2_12_12_1$	25	**a** 4.5386(4) **b** 14.2658(10) **c** 14.3779(11)	**α** 90 **β** 90 **γ** 90	**a** 4.5386 **b** 14.2658 **c** 14.3779	**α** 90 **β** 90 **γ** 90
MOSTIX01	$P2_12_12_1$	293	**a** 4.2464(12) **b** 15.194(5) **c** 15.069(4)	**α** 90 **β** 90 **γ** 90	**a** 4.2464 **b** 15.069 **c** 15.194	**α** 90 **β** 90 **γ** 90

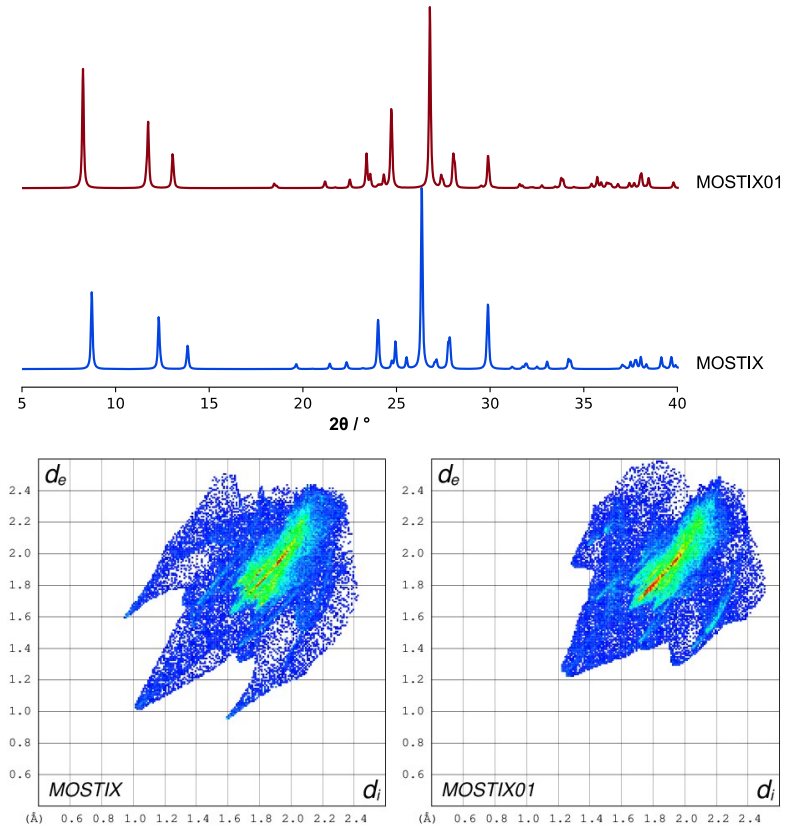


Figure S9. Calculated powder patterns and fingerprint plots for MOSTIX and MOSTIX01

D.2 ECOBUU-ECOBUU01

	S.G.	Temp. (K)	Cell. Axes (Å)	Cell. Ang. (°)	Red. Cell. Axes (Å)	Red. Cell. Ang (°)
ECOBUU	$P2_1/c$	293	**a** 3.6651(4) **b** 14.3329(13) **c** 14.0592(18)	**α** 90 **β** 94.135 **γ** 90	**a** 3.6651 **b** 14.0592 **c** 14.3329	**α** 90 **β** 90 **γ** 94.135
ECOBUU01	$P2_1/c$	180	**a** 7.1713(3) **b** 14.2753(5) **c** 14.0175(6)	**α** 90 **β** 93.449(2) **γ** 90	**a** 7.1713 **b** 14.0175 **c** 14.2753	**α** 90 **β** 90 **γ** 93.449

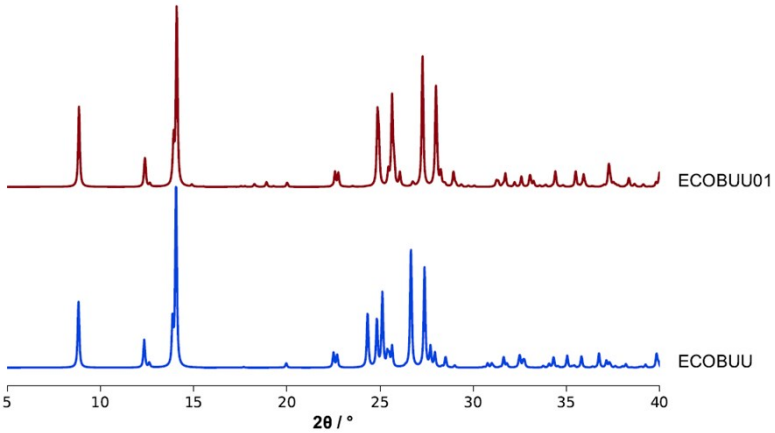


Figure S10. Calculated powder patterns for ECOBUU and ECOBUU01

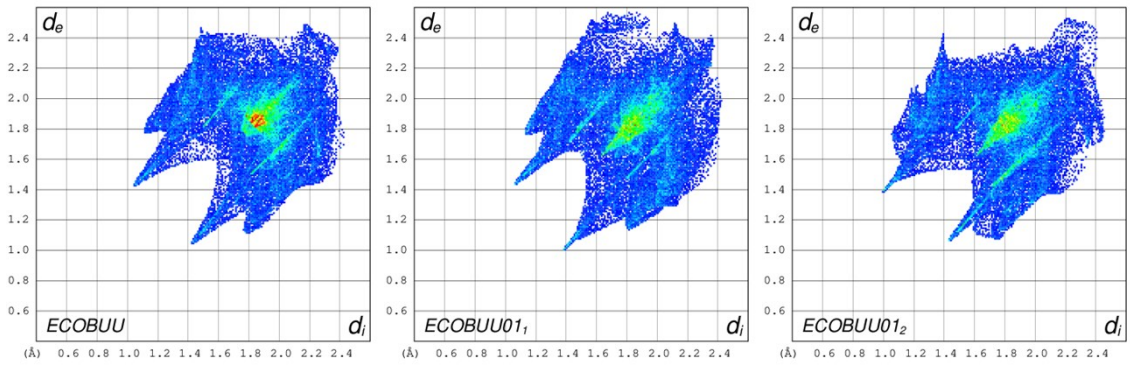


Figure S11. Fingerprint plots for ECOBUU and ECOBUU01

D.3 EJEQAL02-EJEQAL08

	S.G.	Temp. (K)	Cell. Axes (Å)	Cell. Ang. (°)	Red. Cell. Axes (Å)	Red. Cell. Ang (°)
EJEQAL02	$P2_12_12_1$	258	**a** 6.0913(4) **b** 16.727(1) **c** 23.147(1)	**α** 90 **β** 90 **γ** 90	**a** 6.0913 **b** 16.727 **c** 23.147	**α** 90 **β** 90 **γ** 90
EJEQAL02	$P2_12_12_1$	293	**a** 6.1145(2) **b** 17.2901(6) **c** 22.4334(8)	**α** 90 **β** 90 **γ** 90	**a** 6.1145 **b** 17.2901 **c** 22.4334	**α** 90 **β** 90 **γ** 90

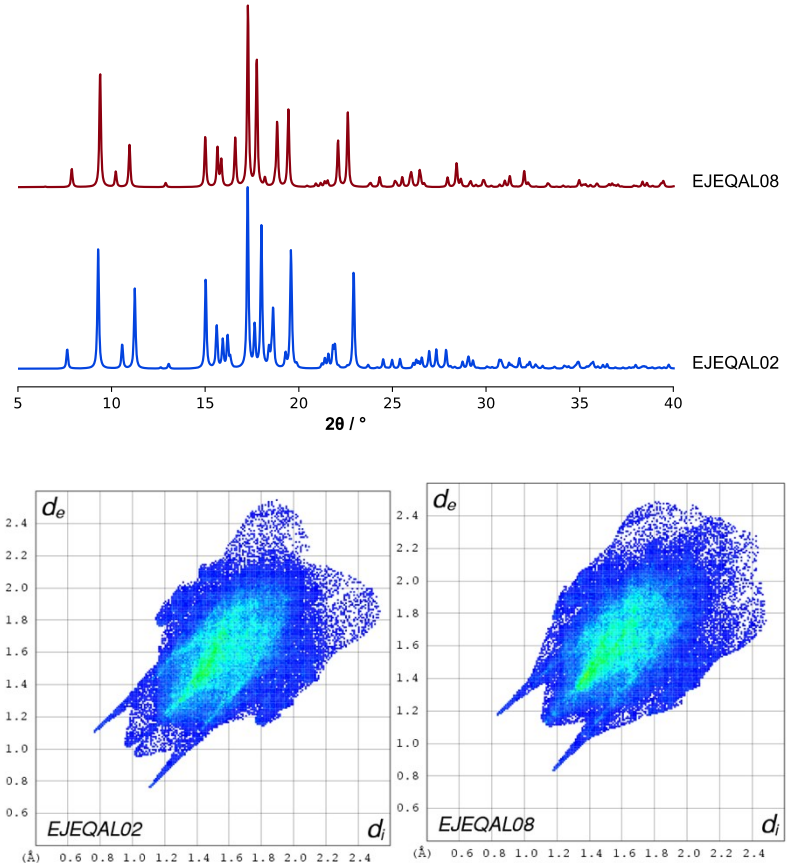
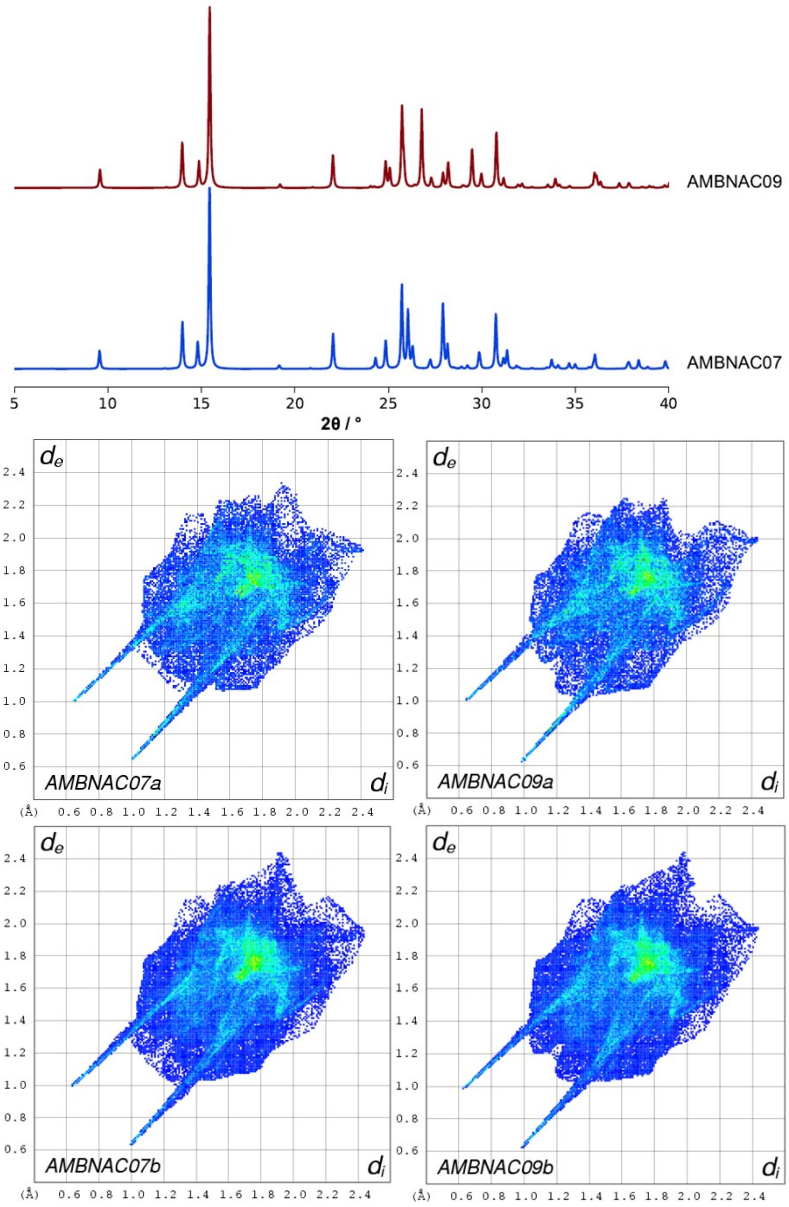


Figure S12. Calculated powder patterns and fingerprint plots for EJEQAL02 and EJEQAL08

E. Different polymorphs with very similar PXRDs

E.1 AMBNAC07-AMBNAC09

	S.G.	Temp. (K)	Cell. Axes (Å)	Cell. Ang. (°)	Red. Cell. Axes (Å)	Red. Cell. Ang (°)
AMBNAC07	$P2_1/n$	101	**a** 18.5620(11) **b** 3.7320(2) **c** 18.5680(12)	**α** 90 **β** 93.993(6) **γ** 90	**a** 3.732 **b** 18.562 **c** 18.568	**α** 93.993 **β** 90 **γ** 90
AMBNAC09	$Pna2_1$	100	**a** 26.9945(8) **b** 3.7322(3) **c** 12.6731(8)	**α** 90 **β** 90 **γ** 90	**a** 3.7322 **b** 12.6731 **c** 26.9945	**α** 90 **β** 90 **γ** 90



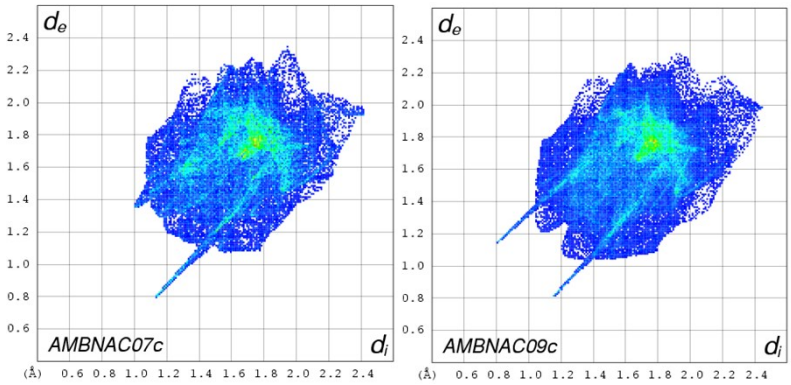


Figure S13. Calculated powder patterns and fingerprint plots for AMBNAC07 and AMBNAC09

E.2 DIWKOK-DIWKOK01

	S.G.	Temp. (K)	Cell. Axes (Å)	Cell. Ang. (°)	Red. Cell. Axes (Å)	Red. Cell. Ang (°)
DIWKOK	$P2_1$	173	**a** 6.3201(7) **b** 5.2972(8) **c** 19.429(2)	**α** 90 **β** 95.084(9) **γ** 90	**a** 5.297 **b** 6.32 **c** 19.429	**α** 95.08 **β** 90 **γ** 90
DIWKOK01	$P2_12_12_1$	173	**a** 5.2833(7) **b** 6.3043(8) **c** 38.753(6)	**α** 90 **β** 90 **γ** 90	**a** 5.283 **b** 6.304 **c** 38.753	**α** 90 **β** 90 **γ** 90

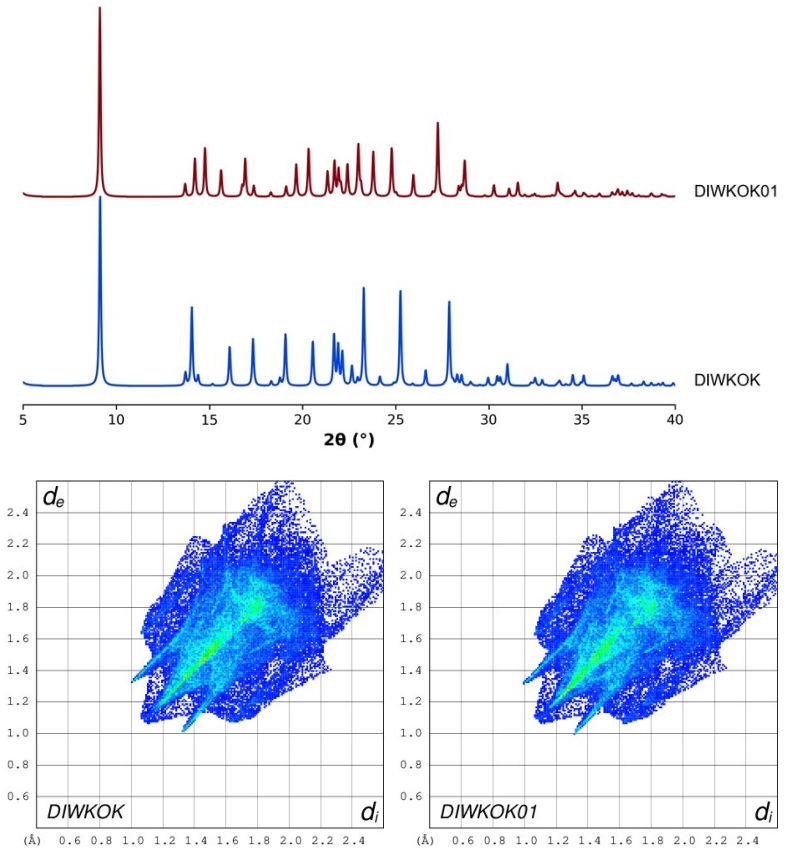


Figure S14. Calculated powder patterns and fingerprint plots for DIWKOK and DIWKOK01

E.3 DAWFUE-DAWFUE01

	S.G.	Temp. (K)	Cell. Axes (Å)	Cell. Ang. (°)	Red. Cell. Axes (Å)	Red. Cell. Ang (°)
DAWFUE	$P2_1/c$	300	**a** 68.592(4) **b** 4.8795(3) **c** 8.7780(6)	**α** 90 **β** 92.2620(10) **γ** 90	**a** 4.88 **b** 8.778 **c** 68.592	**α** 92.26 **β** 90 **γ** 90
DAWFUE01	$Pca2_1$	300	**a** 9.046(2) **b** 4.8948(12) **c** 69.909(17)	**α** 90 **β** 90 **γ** 90	**a** 4.895 **b** 9.046 **c** 69.909	**α** 90 **β** 90 **γ** 90

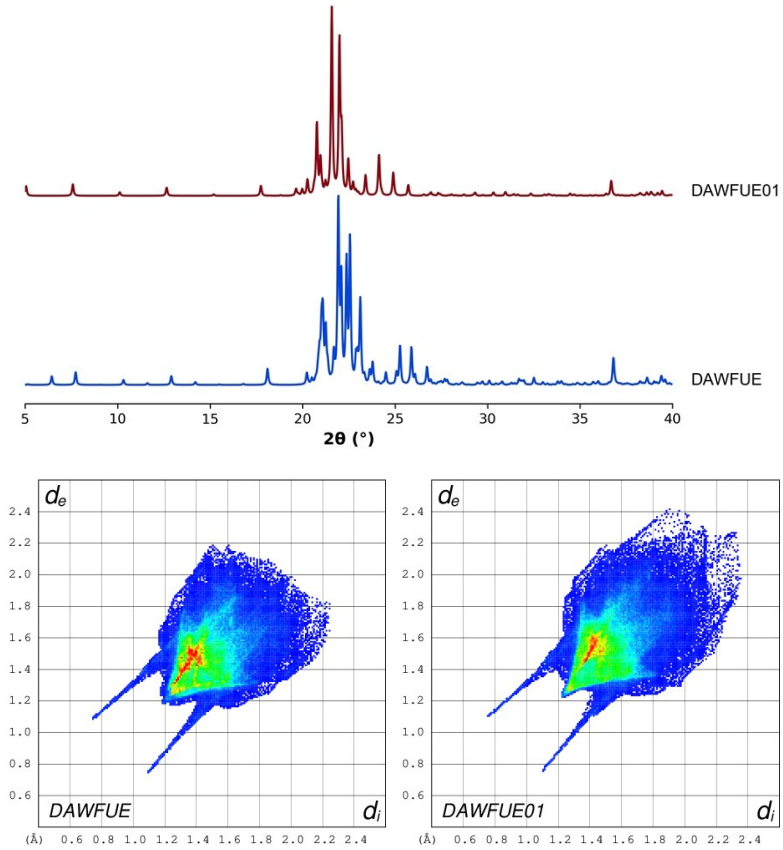


Figure S15. Calculated powder patterns and fingerprint plots for DAWFUE and DAWFUE01

E.4 XOCVAN–XOCVAN02

The structures of (S)-2-Phenyl-1-(4-nitrobenzenesulfonyl)aziridine, is collected at 122 and 100 K, XOCVAN[56] and XOCVAN02[57] respectively. As in the previous case, the low temperature structure has a lower symmetry and a double unit cell with Z´ = 2, thought the β angle close to 90° and the R-factor of 9.6 could suggest of a wrong space group choice (table 4). Differences are visible in the calculated PXRDs despite the R value of 0.98. FPs appear rather similar which might indicate a missed higher symmetry (Figure 5). On the other hand, the crystal packing similarity show a match of 9 out of 20 molecules with an rmsd-r[20] of 0.08 suggesting that the structural difference is real.

	S.G.	Temp. (K)	Cell. Axes (Å)	Cell. Ang. (°)	Red. Cell. Axes (Å)	Red. Cell. Ang (°)
XOCVAN	$P2_1$	122	**a** 6.8140(5) **b** 7.4770(7) **c** 13.4630(9)	**α** 90 **β** 95.487(5) **γ** 90	**a** 6.814 **b** 7.477 **c** 13.463	**α** 90 **β** 95.487 **γ** 90
XOCVAN02	$P2_1$	100	**a** 6.7710(3) **b** 6.8040(15) **c** 7.4480(3)	**α** 90 **β** 89.867(3) **γ** 90	**a** 6.771 **b** 7.448 **c** 26.804	**α** 90 **β** 90 **γ** 90.133

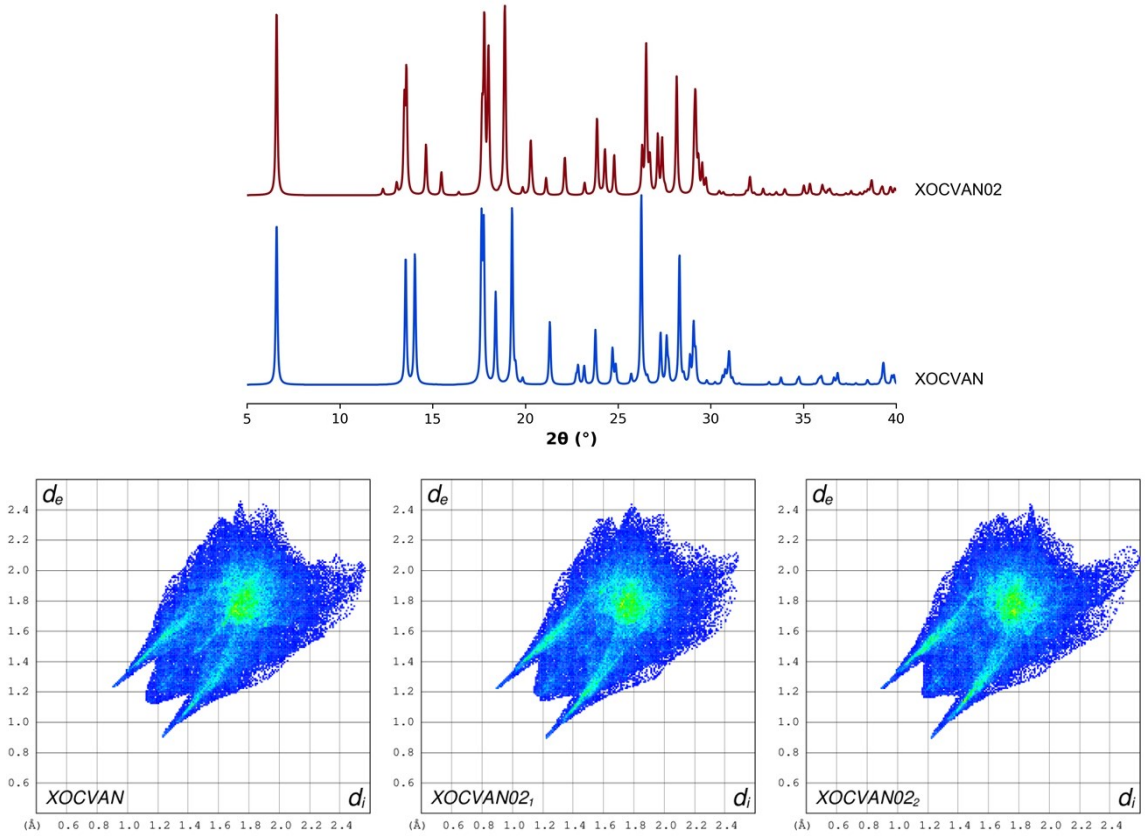


Figure S16. Calculated powder patterns and fingerprint plots for XOCVAN and XOCVAN02