X-ray Powder Diffraction and Crystal Structure Prediction for Polymorph Screening and Structure Solution in Pharmaceutical Development

Simon Bates*, R. Alex Mayo**, Zhuocen Yang**, Baimei Shi*** and Akihiro Himeda****

Abstract

Crystal structure prediction (CSP) is increasingly relevant to pharmaceutical solid-form development because it provides a computational view of the crystal-energy landscape and helps assess whether experimentally observed polymorphs are likely to represent the most stable or otherwise accessible crystal forms.[1, 2] When used together with powder X-ray diffraction (PXRD), integrated CSP workflows also offers a practical route to structure solution from powder data in cases where suitable single crystals are not available.[1, 2, 14] In this article, a CSP-assisted PXRD workflow is outlined using cimetidine as a demonstration sample. The X-ray powder data were measured on a Rigaku SmartLab system, with the CSP candidates and landscape determined from the molecular structure of cimetidine by XtalPi. Targeted CSP candidate structures were selected from the CSP landscape using two scoring parameters: comparison of simulated powder patterns with the measured powder pattern, and the similarity of the indexed unit cell from the measured data with the CSP unit cell. The candidate structures were then refined in SmartLab Studio II by whole-pattern profile fitting with the unit-cell lattice parameters and molecular Z-matrix parameters as variables. The resulting fit showed strong agreement with the measured data and the refined structure overlaid well with the published single-crystal structure. The same workflow can be extended to automated high-throughput PXRD screening by comparing measured patterns against a database of CSP-derived CIF files in real time.[9, 11, 12]

Introduction

Polymorphism remains one of the central issues in pharmaceutical development because different crystal forms of the same active ingredient can display different physical and chemical behavior.[3–6] Differences in solubility, hygroscopicity, compressibility, thermal stability, and mechanical response can influence formulation design, manufacturability, regulatory strategy, and long-term product performance.[3, 6, 13] Within a drug development process, polymorph screens are often performed at different stages of the development process with different goals in mind. For example, at an early stage the goal maybe to discover the most thermodynamically stable form. At mid stage the goal may shift towards the solid form most suited to the mode of delivery given solubility and bioavailability constraints. Later stage screens may be broader in scope and encompass additional forms that might appear during manufacturing and storage or be focused on a comprehensive Intellectual Property strategy. Accordingly, polymorph screening is not merely a discovery exercise but also a risk-mitigation activity.[3–6]

Early-stage screens trying to identify the most thermodynamically stable form often run into a persistent difficulty in solid-form screening in that metastable phases tend to crystallize readily, whereas the target lower-energy forms can be kinetically inaccessible or may appear only under limited conditions.[1, 2, 10] While it is possible to experimentally evaluate the relative stability of isolated solid forms, it remains uncertain as to whether any of the isolated solid forms are the most thermodynamically stable. CSP addresses this uncertainty by generating hypothetical crystal packings for a target molecule and ranking them according to quantities such as lattice energy and density.[1, 2] The resulting ensemble of trial structures provides an energy landscape against which experimental forms can be compared.[1, 2] In practical terms, this allows researchers to ask whether an observed form lies close to the predicted global minimum and whether low-energy unobserved forms may still exist.[1, 2, 10]

The utility of CSP is further increased when it is combined with powder diffraction measurements of the individual solid forms generated during a polymorph screen. Candidate structures produced as CIF files can be utilized to calculate theoretical PXRD patterns and compared directly with measured data. For an individual solid form, this procedure permits rapid elimination of unlikely candidate structures and provides initial structural models for potentially matching candidates.[1, 2, 14] This combined approach is particularly relevant to pharmaceutical materials and high throughput screening as the majority of isolated

* Rigaku Americas.
** XtalPi, Inc.
*** Shenzhen Jingtai Technology Co. Ltd.
**** Rigaku Corporation.

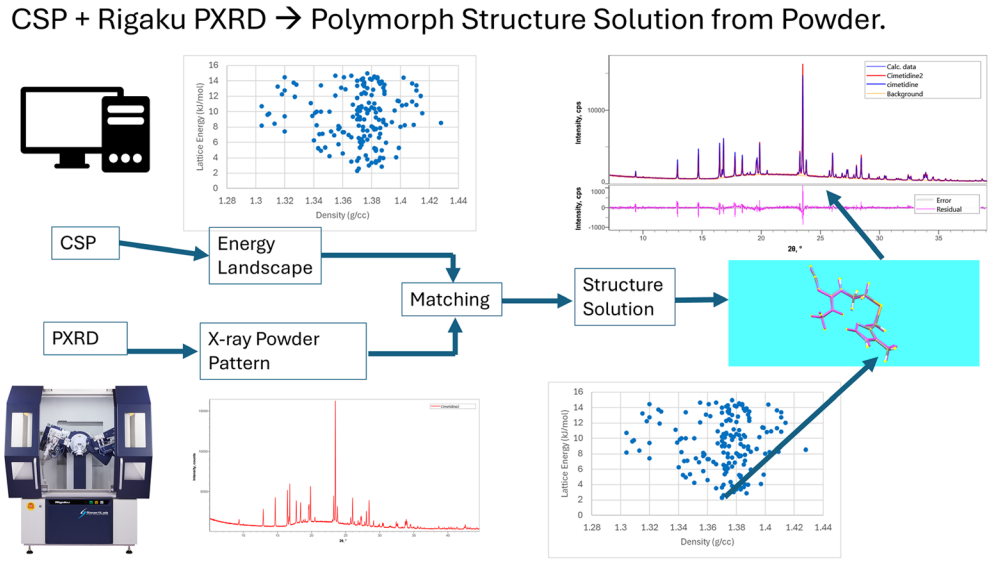

Fig. 1. Schematic representation of the CSP-assisted PXRD workflow. CSP candidate structures define an energy landscape, measured PXRD data are matched against calculated patterns, and the closest candidate is refined to obtain a powder-derived structure solution.

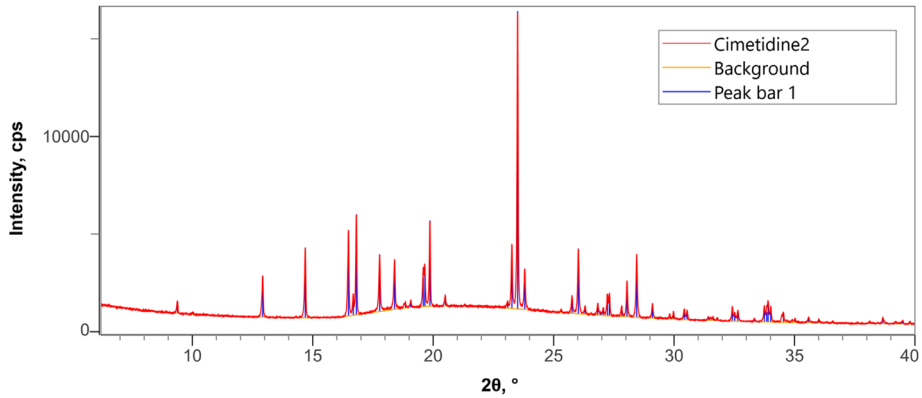

Fig. 2. Measured cimetidine powder diffraction pattern collected on a Rigaku SmartLab system in Bragg–Brentano geometry, shown with peak-search results and estimated non-crystalline background.

solids will comprise of a small amount of powder with few if any well-defined single crystals suitable for structure determination. Furthermore, routine PXRD measurements are routinely incorporated directly into solid-form screening workflows.[6], [13]

One challenge in comparing powder patterns calculated from CSP structures with measured PXRD data is that the calculated structures correspond to static or low-temperature lattices, whereas the experimental data are usually collected under ambient conditions.[15], [17] As a result, differences in lattice parameters can lead to substantial shifts in peak positions.[15], [17] Variable-cell powder-difference methods have been developed specifically to accommodate such unit-cell changes when matching experimental PXRD data to known or in silico crystal structures.[15], [16]

If the measured powder pattern can be indexed[7], [8], the resulting Bravais lattice and reduced-cell parameter ratios can be compared directly to the CSP cell, giving a more robust solution matching than simply comparing powder patterns.

The general workflow and integration of CSP into polymorph screening using powder X-ray diffraction is shown in Fig. 1.

CSP-Assisted Powder Structure Solution

A demonstration study was carried out using a powder-sample of cimetidine. The PXRD data were collected on a Rigaku SmartLab X-ray diffractometer in Bragg-Brentano geometry. The measured powder pattern together with the peak-search results and an estimate of the non-crystalline plus instrument background, are shown in Fig. 2. Using SLSII, the identified peaks were used to derive the most likely crystal unit cell utilizing established powder-pattern indexing methods.[7], [8] The measured powder pattern and indexed unit-cell parameters were then used to select candidate crystalline polymorphs from a CSP campaign.

A targeted light-weight crystal structure prediction (CSP) study was performed for cimetidine utilizing

Table 1. Indexed crystal unit cell derived from the measured data.

Unit cell	Representative unit-cell parameters	Space group
Indexing result	a = 10.389; b = 18.824; c = 6.825; beta = 106.44	$P2_1/c$

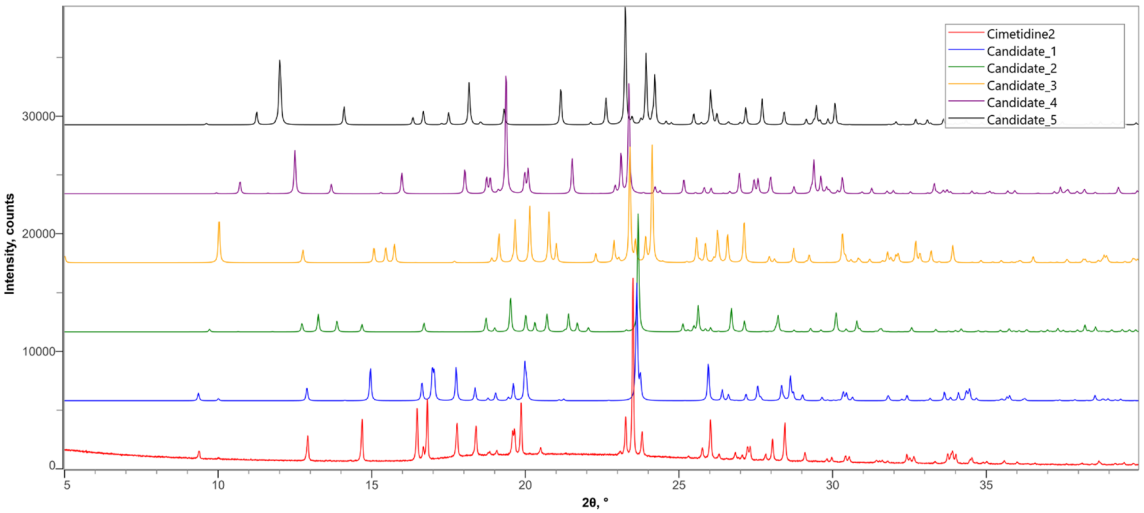

Fig. 3. Measured cimetidine powder diffraction pattern compared with calculated powder patterns derived from the five candidate CSP structures selected through initial screening.

experimental powder X-ray diffraction (PXRD) data and the two-dimensional (2D) molecular diagram as inputs. Conformational sampling and parametrization of a molecular-based force field was first performed for later crystal structure generation and initial ranking. The experimental PXRD pattern was indexed using an in-house algorithm to obtain unit-cell parameters, the space group, and Z'. To capture structures consistent with the experimental powder pattern, two parallel searching strategies were employed: fixed-lattice search using the indexed cell as input, and a global search over seven selected space groups without specific constraints. Generated structures underwent restrictive multi-stage geometry optimization and energy ranking by adopting semi-empirical and DFT-D methods, while final ranking used the optPBE-vdW level of theory. Candidate structures were then evaluated by quantitative comparison of simulated and experimental PXRD patterns. Five candidates with the best PXRD agreement were delivered; included among these is a structure consistent with the experimentally known form solved by single-crystal X-ray diffraction. From this set, five candidates whose calculated powder patterns showed promising similarity to the measured data were selected for further examination, as shown in Fig. 3. This initial rapid pre-screening of the CSP structures is important because it converts a large theoretical search space into a more manageable problem that can be addressed in standard powder-diffraction software.

Subsequently, a secondary procedure was performed to match the five candidates with the indexed Bravais lattice. Because the candidates and the indexed lattice do not necessarily correspond to the same choice of crystallographic axes, all possible axis transformations were systematically examined. The optimal axis setting for the best-matching candidate was thereby identified, resulting in the selection of candidate 1 as the final candidate. Whole-pattern profile fitting was then performed against the measured X-ray powder data using scale factor, unit-cell parameters, peak profile terms, and Z-matrix structural parameters as variables. The variables were released in incremental steps using a control macro. The refined fit reached an Rwp value of 4.59%, indicating close agreement between the measured powder pattern and the CSP-derived model; see Fig. 4. This result demonstrates that CSP candidates can serve as effective starting structures for powder-based structure solution when a direct single-crystal determination is challenging.[(1), (2), (14)]

The refined molecular geometry was also compared with the published single-crystal structure of cimetidine. The overlay in Fig. 5 shows excellent structural agreement, confirming that the powder-derived model is not merely a good pattern match but also a chemically reasonable structure model. In the context of pharmaceutical development, this capability provides a direct link between solid forms generated during a polymorph screen and a candidate crystal structure that helps position an experimentally observed form within the broader CSP energy landscape.[(1), (2), (10)]

Discussion

The cimetidine example illustrates two complementary roles for CSP in pharmaceutical PXRD workflows. First, CSP strengthens polymorph screening by making the crystal-energy landscape visible before or alongside the experimental work.[(1), (2), (10)] Second, the integrated CSP work flow provides a practical route from measured

Table 2. Candidate hypothetical cimetidine structures selected from targeted CSP analysis.

Unit Cell	Representative unit-cell parameters	Space group
Candidate 1	a = 6.58; b = 18.8886; c = 10.490; beta = 107.78	$P2_1/c$
Candidate 2	a = 7.66; b = 10.12; c = 9.58; alpha = 114.3; beta = 113.1; gamma = 72.4	P-1
Candidate 3	a = 17.73; b = 7.54; c = 9.12; beta = 96.42	$P2_1/c$
Candidate 4	a = 9.63; b = 8.08; c = 9.76; alpha = 114.5; beta = 112.6; gamma = 82.9	P-1
Candidate 5	a = 8.51; b = 8.11; c = 10.45; alpha = 70.43; beta = 73.4; gamma = 98.4	P-1

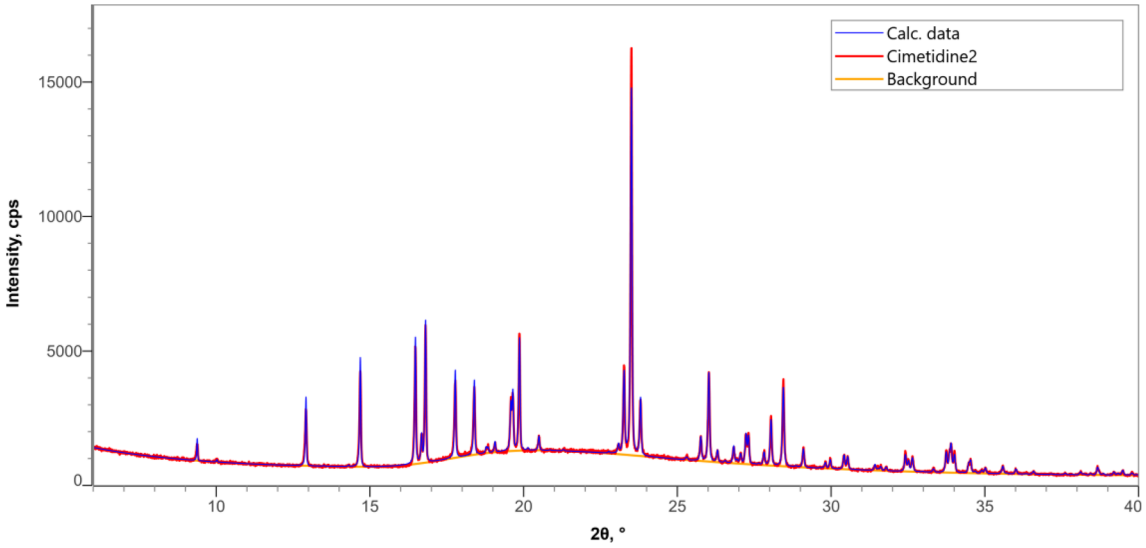

Fig. 4. Whole-pattern profile fit of the closest matching CSP candidate in SmartLab Studio II. Refinement using scale, lattice, profile, and Z-matrix parameters produced Rwp = 4.59%.

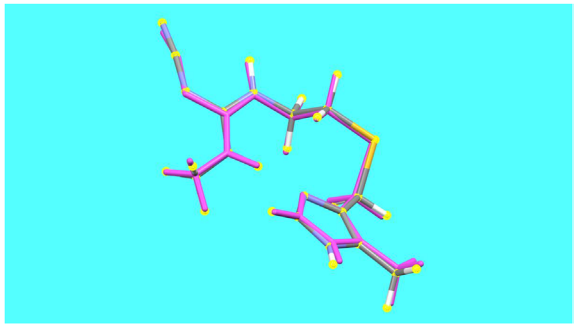

Fig. 5. Overlay of the Z-matrix refined CSP structure with the published single-crystal structure of cimetidine, showing excellent structural agreement between the powder-derived and single-crystal models.

powder data to a refined structural model.[1, 2, 14] The latter point is particularly important when crystallization yields powders only, when single crystals are too small or defective for routine work, or when rapid turnaround is required during development.[6, 13, 14]

From an instrumentation perspective, the workflow is well matched to Rigaku's SmartLab platform because measurement, pattern analysis, and candidate refinement can be integrated within the same automated analysis environment. From a development perspective, the ability to combine CSP outputs with automated PXRD screening suggests a scalable path for solid-form programs in which many samples must be assessed quickly and consistently.[9, 11, 12] Rather than treating CSP as a purely theoretical exercise, the approach uses prediction as a practical filter and refinement starting point for real diffraction data.[1, 2, 9]

More broadly, pharmaceutical solid-form programs benefit from early and systematic polymorph screening because late-emerging forms can create technical, regulatory, and commercial risk.[3–6, 10] The availability of robust PXRD methods, combined with automated or high-throughput data acquisition and analysis, makes it increasingly realistic to evaluate large numbers of candidate forms and formulations within development timescales.[11–13]

Conclusion

Crystal structure prediction can make a direct and practical contribution to pharmaceutical solid-form development when combined with Rigaku powder X-ray diffraction.[1, 2, 9] In the cimetidine demonstration, CSP-derived candidates reduced the structure-solution problem to a manageable set of plausible models, and refinement in SmartLab Studio II yielded a high-quality fit to the measured powder pattern, with a powder-derived structure consistent with the published single-crystal result.[1, 2, 14] The same methodology can be extended to automated high-throughput screening by matching measured patterns against a database of CSP-generated

CIF files.[9), 11), 12)] Taken together, these results show that CSP-assisted PXRD offers a robust route to polymorph assessment, structure verification, and scalable screening in pharmaceutical development.[2), 3), 6), 13)]

References

(1) S. L. Price: *Chem. Soc. Rev.*, **43** (2014), 2098–2111.
(2) S. L. Price, D. E. Braun and S. M. Reutzel-Edens: *Chem. Commun.*, **52** (2016), 7065–7077.
(3) S. Byrn, R. Pfeiffer, M. Ganey, C. Hoiberg and G. Poochikian: *Pharm. Res.*, **12** (1995), 945–954.
(4) G. P. Stahly: *J. Pharm. Sci. Technol. Jpn.*, **66** (2006), 435–439.
(5) G. P. Stahly: *Cryst. Growth Des.*, **7** (2007), 1007–1026.
(6) L. Yu, S. M. Reutzel and G. A. Stephenson: *Pharm. Sci. Technol. Today*, **1** (1998), 118–127.
(7) A. Boultif and D. Louer: *J. Appl. Cryst.*, **37** (2004), 724–731.
(8) P.-E. Werner, L. Eriksson and M. Westdahl: *J. Appl. Cryst.*, **18** (1985), 367–370.
(9) C. L. Burcham, M. F. Doherty, B. G. Peters, S. L. Price, M. Salvalaglio, S. M. Reutzel-Edens, L. S. Price, R. K. R. Addula, N. Francia, V. Khanna and Y. Zhao: *Cryst. Growth Des.*, **24** (2024), 5417–5438.
(10) R. M. Bhardwaj, J. A. McMahon, J. Nyman, L. S. Price, S. Konar, I. D. H. Oswald, C. R. Pulham, S. L. Price and S. M. Reutzel-Edens: *J. Am. Chem. Soc.*, **141** (2019), 13887–13897.
(11) M. Reinle-Schmitt, D. Sisak Jung, M. Morin, F. N. Costa, N. Casati and F. Gozzo: *Int. J. Pharm. X*, **6** (2023), 100221.
(12) K. Egusa, F. Okazaki, J. Schiewe, U. Werthmann and M. Wolkenhauer: *Drugs R D*, **17** (2017), 413–418.
(13) N. K. Thakral, R. L. Zanon, R. C. Kelly and S. Thakral: *J. Pharm. Sci.*, **107** (2018), 2969–2982.
(14) H. Polyzois, R. Guo, V. K. Srirambhatla, M. Warzecha, E. Prasad, A. Turner, G. W. Halbert, P. Keating, S. L. Price and A. J. Florence: *Cryst. Growth Des.*, **22** (2022), 4146–4156.
(15) Arie van der Lee and Dan G. Dumitrescu: *Chem. Sci.*, **12** (2021), 8537–8547.
(16) R. A. Mayo, K. M. Marczenko and E. R. Johnson: *Chem. Sci.*, **14** (2023), 4777–4785.
(17) R. A. Mayo, A. Otero-de-la-Roza and E. R. Johnson: *CrystEngCommun*, **24** (2022), 8326–8338.