



ChemComm

**Dual Space Divergence in Small-Molecule Quasiracemates:  
Benzoyl Leucine and Phenylalanine Assemblies**

Journal:	<i>ChemComm</i>
Manuscript ID	CC-COM-12-2023-006212.R1
Article Type:	Communication

SCHOLARONE™  
Manuscripts

## COMMUNICATION

## Dual Space Divergence in Small-Molecule Quasiracemates: Benzoyl Leucine and Phenylalanine Assemblies

Received 00th January 20xx,  
Accepted 00th January 20xx

Katelyn N. Koch, Aaron J. Teo and Kraig A. Wheeler\*

DOI: 10.1039/x0xx00000x

Quasiracemic materials constructed with two points of structural difference were used to understand the role molecular shape plays in molecular assembly. Hot stage, crystallographic and occupied cavity space assessments provide insight into how imposed CH<sub>3</sub>/Cl and H/CF<sub>3</sub> structural variations placed on benzoyl leucine and phenylalanine scaffolds result in a remarkably high occurrence of cocrystal formation.

Quasiracemic materials represent a broad class of compounds created by the pairing of near enantiomers.<sup>1,2</sup> Because these materials exploit the molecular shape and chirality of the quasienantiomeric components in supramolecular assembly, it is not surprising that this approach has successfully been applied to construct material architectures that require precise control over the alignment of the building blocks.<sup>3–16</sup> Similar to the concepts of polymorphism<sup>17</sup> and crystal structure prediction<sup>18–20</sup>, the phenomenon of quasiracemate formation is straightforward and direct in principle. However, a considerable challenge exists in defining the structural boundaries and shape space of the assembled materials. The challenge arises in the design stage, where the inherent flexibility of the term quasiracemate – equimolar ratios of chemically unique compounds of opposite handedness – intentionally lacks details of the chemical frameworks, functional groups and structural difference limitations needed for successful quasiracemate formation. Literature reports suggest that the structural variation between small molecule quasienantiomers occurs at only one site with a relatively minor imposed structural difference – e.g., Cl/Br<sup>21–25</sup> and CH<sub>3</sub>/Cl<sup>21,23,26</sup> represent common pairings. Though this strategy continues to show success with quasiracemate assembly, it also presents questions relating to a broader understanding of the term quasiracemate and the degree of structural tolerance permitted during quasiracemate assembly.

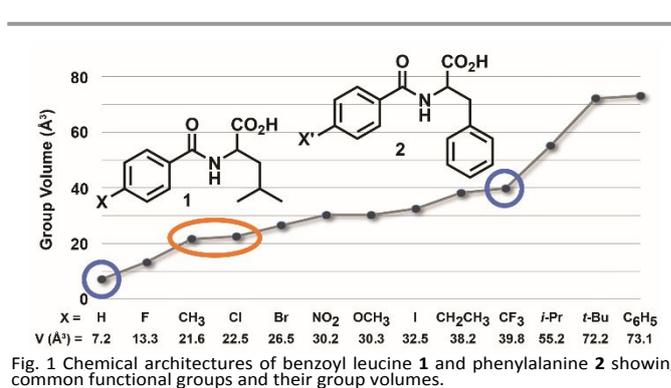


Fig. 1 Chemical architectures of benzoyl leucine **1** and phenylalanine **2** showing common functional groups and their group volumes.

To investigate the structural boundaries of molecular shape-directed molecular recognition, the present work aims to develop a new approach that targets quasiracemic scaffolds designed with two points of chemical difference. We chose to combine benzoyl leucine (**1**) and phenylalanine (**2**) quasienantiomers, where the imposed topological variation of these components can be estimated by the percent difference in the *R* group volumes (% $\Delta V = 18.0$ ). Drawing from previous quasiracemate successes and a list of common functional groups (Fig. 1), the leucine and phenylalanine systems include additional functionalization using the CH<sub>3</sub>/Cl (% $\Delta V = 4.0$ ) and H/CF<sub>3</sub> (% $\Delta V = 138.7$ ) substituent pairs. While CH<sub>3</sub>/Cl<sup>21,23,26</sup> are familiar group selections with only modest size and shape space differences, the H/CF<sub>3</sub> combination is *absent* in small molecule quasiracemates, likely owing to their considerable topological differences. The design element using two sites of chemical difference provides an opportunity to examine sizable families of ten cocrystalline materials that differ iteratively in their shape space and pendant functional groups X and X', where X/X' represents the CH<sub>3</sub>/Cl and H/CF<sub>3</sub> pairings. The ten unique CH<sub>3</sub>/Cl and H/CF<sub>3</sub> compounds consist of racemates [(±)-**1**-X, (±)-**1**-X', (±)-**2**-X, (±)-**2**-X'], singly different quasiracemates [**1**-X/**1**-X', **2**-X/**2**-X', **1**-X/**2**-X, **1**-X'/**2**-X'] and doubly distinct quasiracemates [**1**-X/**2**-X', **1**-X'/**2**-X]. By examining the cocrystallization behavior using melt and solvent-assisted methods, this work seeks a deeper understanding of the role of molecular topology in molecular assembly.

Department of Chemistry, Whitworth University, 300 West Hawthorne Road, Spokane, Washington, 99251, USA. Email: kraigwheeler@whitworth.edu.

† Footnotes relating to the title and/or authors should appear here.

Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

## Hot Stage Thermomicroscopy

Video assisted hot stage microscopy offers a useful diagnostic tool for the present study by probing crystal nucleation from the melt.<sup>27–30</sup> For racemates and quasiracemates, the thermal signature from heating the two components produces a virtual melting point phase diagram showing either conglomerate formation or the growth of a new crystalline phase at the component interface. Fig. 2 depicts this process with the (±)-1-CF<sub>3</sub> and L-1-CF<sub>3</sub>/D-2-H systems, where the thermal micrographs show the emergence of two eutectic regions and a racemic or quasiracemic phase with increasing temperature.

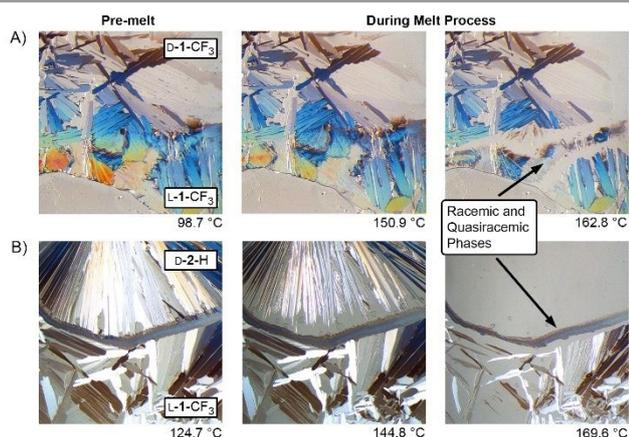


Fig. 2 Hot stage thermomicroscopy micrographs of A) L-1-CF<sub>3</sub>/D-1-CF<sub>3</sub> and B) L-1-CF<sub>3</sub>/D-2-H.

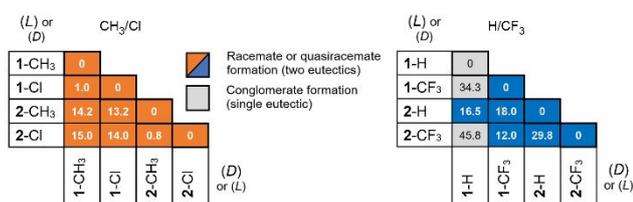


Fig. 3 Results from hot stage thermomicroscopy for the racemates and quasiracemates of **1** and **2** with the CH<sub>3</sub>/Cl and H/CF<sub>3</sub> substituent pairs. The percent difference in volumes (%ΔV) is also provided for each system.

Processing all possible CH<sub>3</sub>/Cl and H/CF<sub>3</sub> enantiomeric and quasienantiomeric sets of the components *via* the hot stage method offers insight into how systematic changes in molecular topology affect molecular recognition. Fig. 3 shows these results and the %ΔV values (±) for each system. When considering the leucine **1** and phenylalanine **2** enantiomers combined with the H, CH<sub>3</sub>, Cl and CF<sub>3</sub> substituents, all but (±)-1-H showed racemate formation. In this case, the lack of success in bringing together the L and D components highlights the inherent limitation of the hot stage method for materials that decompose at elevated temperatures.

For the quasiracemates, the hot stage successes correlate well with the %ΔV values. The six CH<sub>3</sub>/Cl systems vary from %ΔV = 0.8 to 15.0% with each forming cocrystalline materials. Combining the H/CF<sub>3</sub> quasienantiomers (%ΔV = 12.0–45.8) tells a different story, where four of these systems result in quasiracemates (*i.e.*, 1-H/2-H, 2-H/1-CF<sub>3</sub>, 2-H/2-CF<sub>3</sub>, 1-CF<sub>3</sub>/2-CF<sub>3</sub>). The remaining 1-H/1-CF<sub>3</sub> and 1-H/2-CF<sub>3</sub> compounds

exhibit the most significant %ΔV range (34.3 and 45.8%, respectively) and do not achieve quasiracemic crystals from the melt.

## Crystallography

Given the range of chemical variation for the CH<sub>3</sub>/Cl system, it is somewhat surprising that each of the ten targeted compounds gave suitable samples for crystallographic assessment (Table S1). These structures lack disorder and align components in one of two crystalline phases. The majority crystallize into Form I with either space group *P*2<sub>1</sub>/*c* (racemates) or *P*2<sub>1</sub> (quasiracemates) and display similar unit cell parameters, crystal packing and *R*<sub>2</sub><sup>2</sup>(10) N-H...O=C<sub>carboxyl</sub> and *C*(7) (O-H...O=C<sub>amide</sub>) hydrogen-bond motifs<sup>31,32</sup> (Fig. 4a, Table S2). Many of these structures that take on Form I also exhibit close π-π stacking of the benzoyl groups (Table S3). Interestingly, those structures constructed from only leucine **1** components [*i.e.*, (±)-1-CH<sub>3</sub>, (±)-1-Cl and D-1-CH<sub>3</sub>/L-1-Cl] assemble differently *via* Form II (space groups *P*2<sub>1</sub>/*c* and *P*2<sub>1</sub>) with *R*<sub>2</sub><sup>2</sup>(8) carboxylic acid O-H...O=C hydrogen-bonded dimers and N-H...O=C<sub>amide</sub> *C*(4) interactions (Fig. 4b). The *R*<sub>2</sub><sup>2</sup>(10) (Form I) and *R*<sub>2</sub><sup>2</sup>(8) (Form II) motifs are positioned on inversion symmetry elements for the racemates and approximate inversion relationships for the quasiracemic systems.

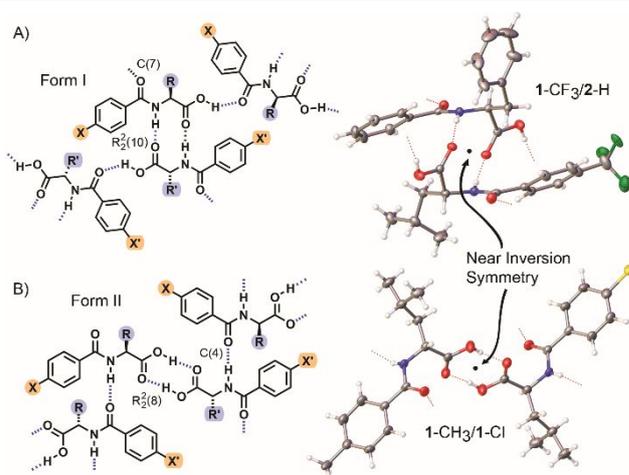


Fig. 4 Depictions of the hydrogen-bond motifs (Form I and II) and the crystal structures of quasiracemates D-1-CF<sub>3</sub>/L-2-H and D-1-CH<sub>3</sub>/L-1-Cl.

Similar to the hot stage experiments, crystal growth of the H/CF<sub>3</sub> family proved challenging compared to the CH<sub>3</sub>/Cl system. Three racemates [(±)-1-H, (±)-1-CF<sub>3</sub>, (±)-2-H] and four quasiracemates [(D-2-H/L-2-CF<sub>3</sub>, L-1-H/D-2-H, L-1-CF<sub>3</sub>/D-2-CF<sub>3</sub>, D-1-CF<sub>3</sub>/L-2-H)] gave suitable crystals for structure determination with an increased occurrence of Form I [(±)-1-H, (±)-2-H, D-2-H/L-2-CF<sub>3</sub>, L-1-H/D-2-H, L-1-CF<sub>3</sub>/D-2-H] relative to Form II [(±)-1-CF<sub>3</sub>]. The structure of L-1-CF<sub>3</sub>/D-2-CF<sub>3</sub> displays a third crystalline phase that exhibits matching hydrogen-bond contacts to Form I but deviates in molecular alignment resulting from F...F or F...H interactions. Crystal growth experiments involving (±)-2-CF<sub>3</sub> and quasiracemates 1-H/1-CF<sub>3</sub> and 1-H/2-CF<sub>3</sub> did not produce suitable samples for crystallographic

assessment even after multiple attempts using various solvent systems and methods.

The extent of isomorphism observed within Form I and II was further studied and confirmed using the Crystal Structure Similarity search feature in CCDC-Mercury (v2022.3.0, ref. 33) (Table S4). This assessment method quantitatively measures the crystal packing likeness for pairs of structures that, when applied to systems **1** and **2**, showed the expected trends with distinctions related to Leu/Phe, CH<sub>3</sub>/Cl, and H/CF<sub>3</sub> structural differences. In view of the structural similarity of the components and their propensity to align with one of two primary motifs, the likelihood of polymorphism should be high. While it is possible Form I and II represent two potential polymorphic phases, this study did not encounter polymorphism.

Another important structural feature of the H/CF<sub>3</sub> system is the presence of disorder in the leucine [(±)-1-CF<sub>3</sub>, L-1-H/D-2-H and D-1-CF<sub>3</sub>/L-2-H] and CF<sub>3</sub> [L-1-CF<sub>3</sub>/D-2-CF<sub>3</sub> and D-1-CF<sub>3</sub>/L-2-H] groups. While CF<sub>3</sub> disorder is pervasive in the crystal structures of aryl CF<sub>3</sub> groups (Cambridge Crystallographic Database, v5.43, ref. 34), where 42% of the deposited 5804 structures were modeled with disorder, only 10% of the leucine entries contain disorder. The observation of disorder in several of our structures is potentially significant when considering how the imposed topological difference of the leucine **1** and phenylalanine **2** quasisenantiomeric components (%ΔV<sub>Leu-Phe</sub> = 18.0%) are accommodated during the molecular assembly process.

Lattice energy calculations using Crystal Explorer<sup>35</sup> (Gaussian<sup>36</sup>, B3LYP/6-31g(d,p)) and differential scanning calorimetry (DSC) give details on the crystal energetics of these systems. Collectively, the structures constructed solely from phenylalanine **2** achieve ~20 kJ/mol greater stability than those restricted to the leucine **1** framework (Table S5). Even so, the lattice energy ( $E_{Latt}$ ) calculations for the quasiracemates are, in many cases, comparable regardless of component selection. An  $E_{Latt}$  of -177 kJ/mol ( $E_{Latt}$ (DSC) = -68 kJ/mol) for quasiracemate 1-Cl/2-CH<sub>3</sub> and that for (±)-1-Cl and (±)-2-CH<sub>3</sub> are -164 kJ/mol ( $E_{Latt}$ (DSC) = -65 kJ/mol) and -190 kJ/mol ( $E_{Latt}$ (DSC) = -91 kJ/mol), respectively. This trend in  $E_{Latt}$  values suggests that the observed structural mimicry of inversion symmetry in the quasiracemates results from the favorable close alignment achieved in the racemate examples.

## Occupied Cavity Spaces

Void space determinations reveal the crystal packing efficiencies of the leucine (Leu) and phenylalanine (Phe) side chains. Given the isomorphous nature of many of these structures and that the Leu and Phe *R* groups differ significantly in size and shape (%ΔV of 18%), we questioned how the crystal assemblies of these systems can accommodate quasisenantiomeric components with drastically different shape spaces. Fig. 5a shows the result obtained using CCDC-Mercury to determine the cavity space occupied by the Leu (261.6 Å<sup>3</sup>) and Phe (308.9 Å<sup>3</sup>) *R* groups of quasiracemate 1-CH<sub>3</sub>/2-Cl. Applying Kitaigorodskii's packing coefficient ( $C_p$ ) approach<sup>37,38</sup> to cavity

spaces allows a direct comparison of these crystallographic voids to the literature volumes of the Leu ( $V_{CH_2(CH_3)_2} = 71.6 \text{ \AA}^3$ ) and Phe ( $V_{CH_2C_6H_5} = 85.8 \text{ \AA}^3$ ) groups<sup>39</sup>, where  $C_p = V_{group}/V_{cavity}$  (Table S6). In the case of 1-CH<sub>3</sub>/2-Cl, the  $C_p$ (Leu) and  $C_p$ (Phe) values for the two components are 0.55 and 0.56 Å<sup>3</sup>, respectively, indicating comparable packing efficiencies for these different amino acid side chains. From inspection of the entire CH<sub>3</sub>/Cl system, the structures with at least one phenylalanine **2** component align the Leu side chains with  $C_p$  values ranging from 0.54 to 0.55 (Fig. 5B); whereas crystals composed of only leucine **1** (*i.e.*, (±)-1-CH<sub>3</sub>, (±)-1-Cl and 1-CH<sub>3</sub>/1-Cl) pack more efficiently ( $C_p = 0.60$ -0.66). This result seems counterintuitive since sterically encumbered groups such as the leucine side chain should translate to less effective packing. A similar trend in  $C_p$  values is observed for the H/CF<sub>3</sub> system with the exception of (±)-1-CF<sub>3</sub>, 1-H/2-H and 1-CF<sub>3</sub>/2-H structures. In those cases, the %ΔV values are less (0.49-0.56) and indicate additional space required to accommodate the disordered Leu groups.

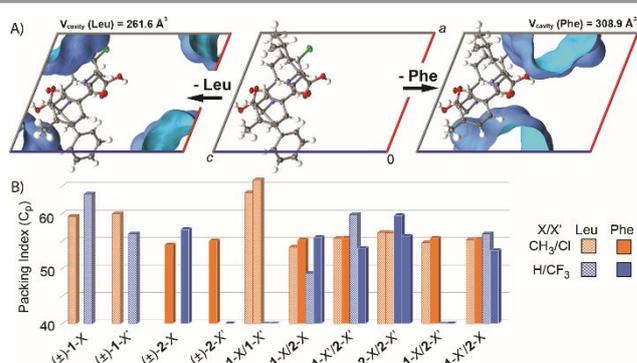


Fig. 5 Leu and Phe A) cavity spaces for quasiracemate D-1-CH<sub>3</sub>/L-2-Cl using a 1.4 Å probe radius 0.1 Å grid spacing and B) packing coefficients for the CH<sub>3</sub>/Cl and H/CF<sub>3</sub> systems.

When considering all CH<sub>3</sub>/Cl and H/CF<sub>3</sub> quasiracemic structures containing both Leu and Phe and excluding those with disorder, the packing efficiencies of the Leu ( $C_p$ (ave) = 0.56 with a range of 0.54-0.60) and Phe ( $C_p$ (ave) = 0.55, 0.54-0.56) groups are remarkably similar. Outside of the leucine dominant 1-CH<sub>3</sub>/1-Cl quasiracemate ( $C_p = 0.64$  and 0.66), the use of Form I or II does not meaningfully affect packing efficiencies. When paired with phenylalanine **2** components, these results show that the cavity occupied by the Leu group is larger, with packing efficiencies similar to Phe. This similarity effectively creates crystal architectures that more closely mimic the spatial distribution of the enantiomers found in the racemates of **1** and **2**.

We have described a fundamentally different approach to creating quasiracemic materials. Most previous approaches pair quasisenantiomeric components where the structural difference is limited to one site and sterically similar groups.

By cocrystallizing the benzoyl leucine **1** and phenylalanine **2** quasisenantiomeric components *via* robust hydrogen-bond motifs, these systems are able to accommodate additional CH<sub>3</sub>/Cl and H/CF<sub>3</sub> substitutions and large spatial differences between the quasiracemic components. Hot stage thermomicroscopy and crystallographic approaches reveal the structural boundaries for these systems and the structural

preference for near-inversion symmetry. Void space determinations of the Leu and Phe side groups for the quasiracemates indicate that smaller leucine component **1** takes up a larger space than anticipated to accommodate quasiracemate formation.

### Author Contributions

Conceptualization, K.A.W.; Methodology, K.A.W.; Formal analysis, K.N.K., A.J.T. and K.A.W.; Funding acquisition, K.A.W.; Investigation, K.N.K., A.J.T. and K.A.W.; Supervision, K.A.W.; Project administration, K.A.W.; Validation, K.N.K., A.J.T. and K.A.W.; Visualization, K.N.K., A.J.T. and K.A.W.; Writing – original draft, K.A.W.; Writing – review and editing, K.N.K., A.J.T. and K.A.W. All authors have read and agreed to the published version of the manuscript.

### Conflicts of interest

There are no conflicts to declare.

### Acknowledgment

We gratefully acknowledge financial support from the National Science Foundation (DMR2304042 and CHE1827313).

### Notes and references

‡ The percent difference in volume (% $\Delta V$ ) was determined as  $[(|V_1 - V_2| / (V_1 + V_2) / 2) * 100]$ , where  $V_1 = V_X + V_R$  and  $V_2 = V_{X'} + V_{R'}$  are the summation of the cavity volumes for the pendant substituents ( $X$  and  $X'$ ) and amino acid side chains ( $R$  and  $R'$ ).

- 1 A. Fredga, *Bulletin de la Societe Chimique de France*, 1973, **1**, 173–182.
- 2 Q. Zhang and D. P. Curran, *Chem. Eur. J.*, 2005, **11**, 4866–4880.
- 3 M. M. H. Smets, E. Kalkman, A. Krieger, P. Tinnemans, H. Meekes, E. Vlieg and H. M. Cuppen, *IUCrJ*, 2020, **7**, 331–341.
- 4 U. K. Shrestha, A. E. Gollither, T. D. Newar, F. O. Holguin and W. A. Maio, *J. Org. Chem.*, 2021, **86**, 11086–11099.
- 5 Q. Ronzon, W. Zhang, T. Charote, N. Casaretto, G. Frison and B. Nay, *Angew Chem Int Ed*, DOI:10.1002/anie.202212855.
- 6 D. F. Bánhegyi, E. Fogassy, J. Madarász and E. Pálovics, *Molecules*, 2022, **27**, 3134.
- 7 D. F. Bánhegyi, E. Fogassy and E. Pálovics, *Symmetry*, 2021, **13**, 1516.
- 8 T. Nakakoji, H. Sato, D. Ono, H. Miyake, E. Mieda, S. Shinoda, H. Tsukube, H. Kawasaki, R. Arakawa and M. Shizuma, *RSC Adv.*, 2021, **11**, 36237–36241.
- 9 T. Nakakoji, H. Sato, D. Ono, H. Miyake, S. Shinoda, H. Tsukube, H. Kawasaki, R. Arakawa and M. Shizuma, *Chem. Commun.*, 2020, **56**, 54–57.
- 10 L. Lodhi, J. P. Yadav, T. Yamazaki, N. T. Duong, S. L. Poojary, K. K. Dey, Y. Nishiyama and M. Ghosh, *J. Phys. Chem. C*, 2022, **126**, 17291–17305.
- 11 A. Hazari, M. R. Sawaya, N. Vlahakis, T. C. Johnstone, D. Boyer, J. Rodriguez, D. Eisenberg and J. A. Raskatov, *Chem. Sci.*, 2022, **13**, 8947–8952.
- 12 Y.-H. Huang, Q. Du, Z. Jiang, G. J. King, B. M. Collins, C. K. Wang and D. J. Craik, *Molecules*, 2021, **26**, 5554.
- 13 C. Zuo, W.-W. Shi, X.-X. Chen, M. Glatz, B. Riedl, I. Flamme, E. Pook, J. Wang, G.-M. Fang, D. Bierer and L. Liu, *Sci. China Chem.*, 2019, **62**, 1371–1378.
- 14 A. Kai, B. D. Egleston, A. Tarzia, R. Clowes, M. E. Briggs, K. E. Jelfs, A. I. Cooper and R. L. Greenaway, *Adv Funct Materials*, 2021, **31**, 2106116.
- 15 D. Wang, Y. Xin, D. Yao, X. Li, H. Ning, H. Zhang, Y. Wang, X. Ju, Z. He, Z. Yang, W. Fan, P. Li and Y. Zheng, *Advanced Functional Materials*, 2022, **32**, 2104162.
- 16 C. Ke, Z. Liu, S. Ruan, X. Feng and X. Liu, *Org. Chem. Front.*, 2021, **8**, 5705–5709.
- 17 J. Bernstein, *Polymorphism in Molecular Crystals*, Oxford University Press, Oxford, 2002, vol. 14.
- 18 S. Yang and G. M. Day, *J. Chem. Theory Comput.*, 2021, **17**, 1988–1999.
- 19 S. L. Price and S. M. Reutzel-Edens, *Drug Discovery Today*, 2016, **21**, 912–923.
- 20 S. L. Price, D. E. Braun and S. M. Reutzel-Edens, *Chem. Commun.*, 2016, **52**, 7065–7077.
- 21 J. T. Cross, N. A. Rossi, M. Serafin and K. A. Wheeler, *CrystEngComm*, 2014, **16**, 7251–7258.
- 22 S. L. Fomulu, M. S. Hendi, R. E. Davis and K. A. Wheeler, *Crystal Growth & Design*, 2002, **2**, 645–651.
- 23 A. M. Lineberry, E. T. Benjamin, R. E. Davis, W. S. Kassel and K. A. Wheeler, *Crystal Growth and Design*, 2008, **8**, 612–619.
- 24 J. M. Spaniol and K. A. Wheeler, *RSC Adv.*, 2016, **6**, 64921–64929.
- 25 F. Toda, K. Tanaka, H. Miyamoto, H. Koshima, I. Miyahara and K. Hirotsu, *J. Chem. Soc., Perkin Trans. 2*, 1997, 1877–1886.
- 26 A. K. Brandt, D. J. Boyle, J. P. Butler, A. R. Gillingham, S. E. Penner, J. M. Spaniol, A. K. Stockdill, M. M. Vanderwall, A. Yeraly, D. R. Schepens and K. A. Wheeler, *Crystals*, 2021, **11**, 1596.
- 27 A. Kumar, P. Singh and A. Nanda, *Appl. Microsc.*, 2020, **50**, 12.
- 28 I. C. Tinsley, J. M. Spaniol and K. A. Wheeler, *Chem. Commun.*, 2017, **53**, 4601–4604.
- 29 L. Kofler and Kofler, A., *Thermal micromethods for the study of organic compounds and their mixtures*, Wagner, Innsbruck, 1952.
- 30 W. C. McCrone, *Fusion Methods in Chemical Microscopy*, Wiley, New York, 1957.
- 31 M. C. Etter, J. C. MacDonald and J. Bernstein, *Acta Crystallographica Section B Structural Science*, 1990, **46**, 256–262.
- 32 J. Bernstein, R. E. Davis, L. Shimoni and N.-L. Chang, *Angewandte Chemie International Edition in English*, 1995, **34**, 1555–1573.
- 33 C. F. Macrae, I. Sovago, S. J. Cottrell, P. T. A. Galek, P. McCabe, E. Pidcock, M. Platings, G. P. Shields, J. S. Stevens, M. Towler and P. A. Wood, *J Appl Cryst*, 2020, **53**, 226–235.
- 34 C. R. Groom, I. J. Bruno, M. P. Lightfoot and S. C. Ward, *Acta Cryst B*, 2016, **72**, 171–179.
- 35 P. R. Spackman, M. J. Turner, J. J. McKinnon, S. K. Wolff, D. J. Grimwood, D. Jayatilaka and M. A. Spackman, *J Appl Cryst*, 2021, **54**, 1006–1011.
- 36 M. J. Frisch, et al., *Gaussian 16 Rev. C.01* 2016.
- 37 A. I. Kitaigorodski, *Molecular Crystals and Molecules*, Academic Press, New York, 1973.
- 38 Y. L. Slovokhotov, *Struct Chem*, 2019, **30**, 551–558.
- 39 S. Sirimulla, M. Lerma and W. C. Herndon, *J. Chem. Inf. Model.*, 2010, **50**, 194–204.