University of Reading

Characterisation of sugars in chocolate and chocolate-manufacturing precursors by X-ray diffraction

Thesis as a collection of papers submitted for the degree of Doctor of Philosophy

School of Pharmacy, PO Box 226, Whiteknights, Reading, Berkshire, RG6 6AP

Daniel Nicholls

February 2021

Declaration

I confirm that this is my own work and all other materials and contributions from various sources are properly and fully acknowledged.

(Daniel A. Nicholls)

Acknowledgements

My gratitude must primarily go to Professor Kenneth Shankland, the supervisor of this PhD thesis. Throughout the course of these tumultuous years, he has been there with humour, support and incredible understanding of the difficulties I faced during this body of work. I am certain that I would not be writing this had I not fortuitously had him as my supervisor.

Thanks are also due to the other collaborators who provided greatly appreciated assistance across the publications included in this project, Dr Mark Spillman, Dr Carole Elleman and Dr Norman Shankland. I am also grateful for Mr Nicholas Spencer's affable and accommodating nature when providing technical support on the Chemical Analysis Facility's X-ray equipment.

The companionship of Dr Elena Kabova, Dr David Edgeley and Dr Mark Spillman (again!) provided stressful days with levity and laughter. The support, patience and encouragement of Dr Graeme Cottrell and Mrs Catherine Hale during the writing up phase of this work was crucial to its completion. To these I am thankful.

To my friends and extended family who nodded along to my ramblings and forgave my shortcomings – thank you for always being there.

Lastly, to my mum, dad and brothers, you give everything and expect nothing. Please know that your support got me through my darkest days.

Abstract

The presence of amorphous and crystalline forms of the various sugars present in chocolate plays a significant role in the production of, and taste/texture of the final product. A great deal of work has focussed on characterising the physical forms of fats in chocolate, and much less upon the sugars. This thesis focusses mainly upon the use of laboratory-based X-ray diffraction (in particular, powder X-ray diffraction), as a means of identifying and quantifying the various sugar (sucrose and lactose) crystalline forms ("phases") present in both chocolate, and the chocolate crumb that is a key intermediate in certain types of chocolate manufacturing processes. Key to this work was the development of reliable procedures for sample presentation in the transmission capillary mode of diffraction, and development of a method for reliable quantitative phase analysis of chocolate and chocolate crumb. It was found that by incorporating a carefully selected internal standard (diamond powder) in the samples being analysed, not only could accurate weight percentages be found for the various crystalline components of the chocolate, but also the percentage of sugars that were present in an amorphous form could be calculated. This approach is fully validated, including by the use of known quantities of amorphous lactose generated "in house" and provides a powerful, complementary approach to other techniques such as DSC, and is applicable to raw materials, process intermediates (chocolate crumb) and finished products.

Whilst studying the crystallisation of lactose in isolation, a powder X-ray diffraction pattern was obtained which could not be explained by any combination of the five known crystalline forms of lactose. This pattern, which resisted all attempts at powder indexing, was able to be fully characterised when a microcrystal (76 × 24 × 18 μm) was isolated from the recrystallised powder, subjected to single-crystal X-ray diffraction, and found to be a new crystalline form of αβ-D-lactose. The orientations of the numerous -OH groups in this new Z'=2 form were verified by using periodic dispersion-corrected DFT calculations. Armed with this novel crystal structure, the recrystallised powder was shown to be a three-phase mixture of crystalline lactose forms, with quantitative phase analysis showing that the novel αβ-D-lactose was the dominant component. This highlights the power of combining single-crystal X-ray diffraction and powder X-ray diffraction to solve otherwise refractory problems, shedding new light on a sugar that has been studied for decades.

Overall, the work shows the importance of careful experimental X-ray diffraction and its associated data analysis techniques in fully characterising complex mixtures of crystal forms. In particular, the quantitative phase analysis method developed herein has gone on to find routine use by Mondelez in the characterisation of many of the samples they generate.

Table of Contents

1 Introduction 2
1.1 Background 2
1.2 Chocolate 2
1.3 Chocolate Ingredients 4
1.3.1 Cocoa Butter 5
1.3.2 Sucrose 8
1.3.3 Lactose 9
1.4 Crystalline and Amorphous Forms 11
1.5 Characterisation Methods 13
1.5.1 Nuclear Magnetic Resonance (NMR) 14
1.5.2 Differential Scanning Calorimetry (DSC) 15
1.5.3 Powder X-Ray Diffraction (PXRD) 18
1.6 Project Rationale 22
1.7 Aims and Objectives 24
1.8 References 25
2 Materials, Instrumentation, and Methods 35
2.1 Materials 35
2.2 Instrumentation 35
2.2.1 Powder X-ray Diffraction 35
2.2.2 Single-crystal X-ray Diffraction 36
2.2.3 Nuclear Magnetic Resonance 36
2.2.4 Optical Hot-stage Microscopy 36
2.2.5 Software 36
2.3 Methods 37

2.3.1 Indexing 37
2.3.2 Pawley Refinement 38
2.3.3 Space Group Determination 38
2.3.4 Crystal Structure Determination 38
2.3.5 Rietveld Refinement 39
2.3.6 Quantitative Phase Analysis 39
3 Experimental Analysis of Powder Diffraction Data 41
4 Rietveld-Based Quantitative Phase Analysis of Sugars in Confectionary 67
5 Physical Characterisation of Dried Lactose Syrups: An Investigation 99
6 A new crystalline form of αβ-D-lactose prepared by oven drying a concentrated aqueous solution of D-lactose 129
7 Conclusions 149
7.1 Future work 152

Chapter 1

Introduction

1 Introduction

1.1 Background

The aim of this chapter is to briefly introduce chocolate, including its manufacture, ingredients and nature. Following this is an introduction of current and historical applications of powder X-ray diffraction, the key analytical technique used in this body of work, to chocolate. Particular attention is paid to the sugars in chocolate (sucrose and lactose) and the methods used for their characterisation and quantification in the chocolate industry.

Although there is a great deal of understanding surrounding the importance of fats in chocolate, particularly in tempering, there is little research conducted into the role of chocolate sugars in the manufacture and end-product quality of chocolate. Developing a method for chocolate sugar quantification and characterisation, and investigation into the crystallographic changes that sugars undergo during chocolate production is the main thrust of this thesis.

1.2 Chocolate

Chocolate is one of the most purchased confectionaries in the United Kingdom. In 2018, government statistics showed the manufacture of cocoa, chocolate and sugar confectionary to be worth £1.1 billion pounds to the UK economy, with £680 million being exported (Department for Environment and Department for International Trade, 2018). Furthermore, chocolate is used as an ingredient in a wide range of consumable products such as hot drinks, snacks, biscuits and cakes (B. Minifie, 1979).

Chocolate itself can be described as fine crystalline solids dispersed throughout a fatty matrix (Beckett, 2019). The key raw ingredients used in its manufacture are sugar (typically sucrose), cocoa butter, cocoa mass, additional fats (often hydrogenated fats are added to improve the heat resistance character of chocolates (Guice et al, (1959), emulsifiers (typically lecithin) and flavourings, additionally milk may be added in the case of milk chocolates (Beckett, 2019). In typical chocolate making, fermented cocoa bean nibs are ground into a cocoa mass, this mass is then pressed to release the cocoa butter which is added into a mixture of cocoa liquor (cocoa mass without the cocoa butter), sugar and milk powder (in the case of milk chocolate) (Beckett, 2011).

An alternative approach to chocolate manufacture is through the production of chocolate crumb. Chocolate crumb was first developed by several UK companies independently in the early to mid-1900s (Beckett, 2003). Where milk powders produced in the spring were susceptible to spoilage due to the presence of spore-forming bacteria (André et al., 2017), developing chocolate crumb increased the shelf life of the product to allow for manufacturers to satisfy the increased demand for chocolate during Christmas. Applying principles learnt in preserves manufacture, chocolatiers were able to lower the water activity of milk-based ingredients by drying liquid milk with sugar forming sweetened condensed milk (SCM). Typically, cocoa butter and cocoa solids are added to this SCM and further drying occurs to achieve a crystalline, brittle crumb with moisture content that varies between 0.8% and 1.5% w/w (Beckett, 2019). During the drying process temperatures of the pre-crumb solution can reach up to 124 °C; this high temperature allows for Maillard reactions to occur that do not typically occur in milk powder-based products (Schnermann and Schieberle, 1997). It is these reactions that give key "cooked" and caramelised flavours to British chocolate; indeed, switching between the use of milk powders and chocolate crumb often results in a negative response from consumers (Beckett, 2003).

The Maillard reaction is a complex series of reactions resulting from the interactions between proteins and reducing sugars when heated, often responsible for the browning observed in food (Feiner, 2006). Though the process begins simply through the formation of a glycosylamine, a further series of rearrangements leading to unstable intermediates until a variety of furan-derived compounds are formed. These continue to react and polymerise with other compounds to form insoluble and dark-coloured substances which provide more complex notes to flavour. The rate of formation of Maillard compounds is highest in high temperature, alkaline, and anhydrous environments.

Regardless of whether the milk powder or chocolate crumb route is used, further fat is added and the product is subject to further processes. Firstly, the mixture is exposed to a physically intensive grinding process in a refiner where the core ingredient particle size is reduced to less than 30 μm (Sokmen and Gunes, 2006). This particle size reduction ensures that the end product has a smooth mouthfeel without grittiness when consumed. Next, the solid particles of chocolate are thoroughly coated in fat during a process known as conching. In a conch the chocolate is consistently stirred over several hours at a temperature of approximately 30 °C. Whilst this appears

to be a simple mixing process at first, conching is essential for the flavour development in the final product and reduction of chocolate viscosity for easier downstream manufacturing by diminishing the moisture content of the mixture (Schuhmacher et al., 1996). Following conching is the key process of tempering, a process which converts the cocoa butter in the product to a more stable crystalline form which gives the end product its desired appearance, snap and melting properties, as well as giving the chocolate resistance to fat blooming and allowing for easier demoulding (Zhao, 2012). Beckett (2019) estimates that the required amount of fat crystallisation for effective tempering lies between the range of 1 and 3%. After successful tempering, chocolate can be moulded, packaged and delivered to consumers.

1.3 Chocolate Ingredients

Milk chocolate recipes vary widely between manufacturers and the whole milk content can reach up to 19% w/w of a milk chocolate bar. Table 1 shows typical percentages of components within a milk powder-based milk chocolate formulation from Fryer and Pinschower (Fryer and Pinschower, 2000). Similar proportions of each component are used in chocolate crumb-based formulations.

Table 1. Typical percentages of components in milk chocolate formulations

Component	**Milk chocolate (% w/w)**
Sucrose	48.7
Cocoa Mass	11.8
Added cocoa butter	20.0
Whole milk powder	19.1
Lecithin	0.4
Flavour compounds	< 0.1

The following section details our current understanding of the key crystallographic components of milk chocolate and where the ingredients have been less studied.

1.3.1 Cocoa Butter

In crystallography, polymorphism is the capacity for a compound to exist in different crystalline structures by changes in the arrangement of the molecules in space (Cruz-Cabeza and Bernstein, 2014). Polymorphism leads to different physicochemical properties, such as melting temperature, of the substance when solid due to the varying strength of inter- and intramolecular bonds.

The triglycerides that cocoa butter consists of exhibit complex polymorphism. Despite the numerous studies conducted on the fat, ambiguities and conflicts still exist on the true number of polymorphs that the fat exhibits (Loisel et al., 1998; Mirzaee Ghazani and Marangoni, 2021; Chapman et al., 1971; Wille and Lutton, 1966; Toro-Vazquez et al., 2004; Marangoni and McGauley, 2003). However, the majority opinion is that cocoa butter can exist in six different crystal forms, denoted by roman numerals (I, II, III, IV, V, VI) with increasing melting points (17, 22, 26, 28, 32, 36 °C). van Malssen (van Malssen et al., 1999) conducted time-resolved powder X-ray diffraction experiments to determine phase-transition behaviour in an isothermal static environment. They reported four polymorphic forms denoted as γ, α, β’ and β which each melt over a range of temperatures, a summary of these conflicting reports is shown in table 2.

Table 2. Reports of cocoa butter crystal and polymorphic forms in the literature and their melting temperatures

Wille and Lutton (1966)		van Malssen et al. (1999)	
Form nomenclature	Melting temperature (°C)	Form nomenclature	Melting temperature (°C)
I	17	γ	-5 – 5
II	22	α	17 – 22
III	26	β’	20 – 27
IV	28	β	29 – 34
V	32		
VI	36		

In Wille and Lutton’s (1966) scheme, the least stable form, I, has the crystal structure most distinct from the other polymorphs and rapidly transforms to form II. Forms II through V were generated directly from the melt or through transformation from lower stability forms, with the higher stability forms requiring longer periods of time at higher crystallisation temperatures. Form VI is very similar to form V yet exists as its own distinct crystal form obtained by transformation from form V.

In van Malssen’s work, the γ, α, and β’ forms were all found to crystallise out from the melt and then subsequently recrystallise to more stable forms (the α and β’ forms were found to transform from lower stability phases). High resolution powder X-ray diffraction data of the β’ form shows seven observed diffraction patterns. The group propose that the β’ polymorph exists as a metastable polymorph rather than the two distinct forms described by Wille and Lutton (1966) (i.e. III and IV). The β polymorph only formed from transformation of the intermediary β’

polymorph in the solid state, van Malssen et al, (1999) further propose that the V and VI forms are subphases of the β polymorph.

Consensus within the literature appears to use a combination of both classifications, denoting the forms as follows: γ(I), α(II), β'(III), β'(IV), β(V), β(VI) - though even with this classification ambiguity still exists (Gutiérrez, 2017). The stability of each polymorph is the result of the organisation of a wide range of fatty alkyl chains in space – cocoa butter itself is composed of a mixture of triacylglycerols containing palmitic acid, oleic acid, linoleic acid, stearic acid, and arachdic acid in 14 different combinations. However, the vast majority (up to 80%) of the triacylglycerols in cocoa butter are 1,2-dipalmitoyl-2-oleoyl-*m*-glycerol (POP), 1-palmitoyl-2-oleoyl-3-steroyl-*m*-glycerol (POS), and 1,3-distearoyl-2-oleoyl-*m*-glycerol (SOS) (Mirzaee Ghazani and Marangoni, 2021). The interactions between these alkyl chains at different temperatures is what causes each polymorphic arrangement to occur, Figure 1 displays an overlay of calculated PXRD patterns of three polymorphs of cocoa butter; clear differences in the patterns allows the relative intensity of each polymorph to be calculated in quantitative phase analysis.

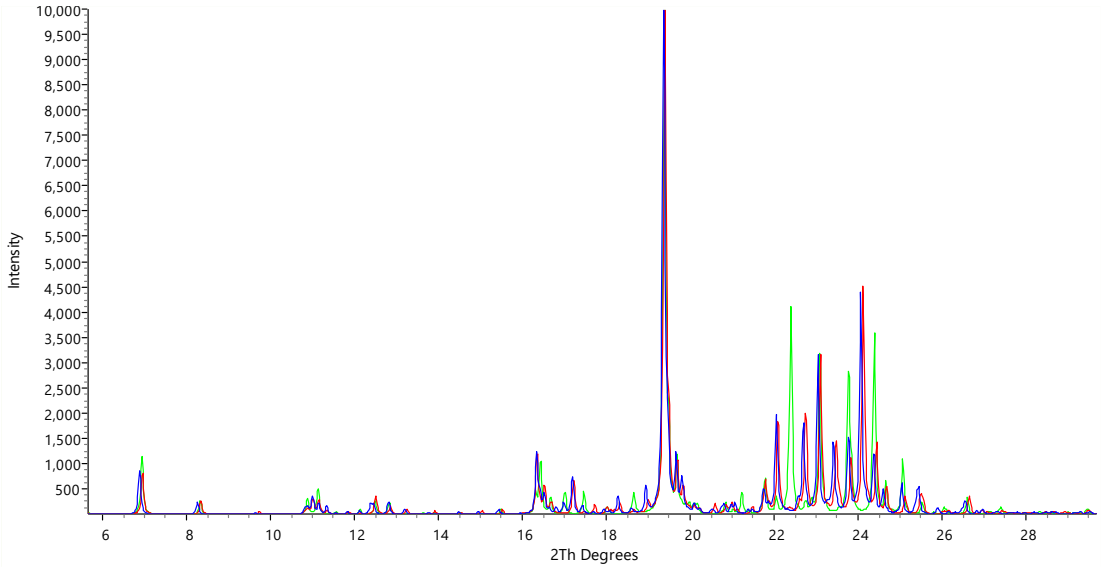


Figure 1. Calculated powder patterns of three polymorphs of cocoa butter. CSD Refcodes: β' polymorphs in red (JEMSAW) and blue (JEMSAW01), β_{VI} in green (QESJEE01) (van Mechelen et al., 2006).

The polymorphism of cocoa butter is key to the tempering process and thus key to the manufacture of quality chocolate products. Strict temperature monitoring in melting, seeding with tempered chocolate, and controlled cooling is required for the crystallisation of the desirable β(V)

polymorph which possesses the melting temperature desirable in a chocolate product. Moreover, transformation of the β(V) to β(VI) polymorph in the solid state over months the primary determining factor in advertised chocolate shelf-life labelling due to the tendency for the β_1(VI) polymorph to cause unattractive fat bloom (Alpha MOS, 2017).

Research into the role and crystallisation of cocoa butter in chocolate is still an ongoing and busy area (Yao et al., 2020; Ghazani and Marangoni, 2018; Zhao et al., 2018; Buscato et al., 2018) but for the purposes of this thesis (which is primarily concerned with chocolate sugars), the β(V) polymorph present in finished chocolate products is mostly considered. van Mechelen (van Mechelen et al., 2006) has provided crystal structures of POP, POS, and SOS, in addition to a suitable monoclinic crystal structure of Ivory Coast cocoa butter for use in powder X-ray diffraction analyses.

1.3.2 Sucrose

Chocolate confectionary mostly consists of sugars. Typically, snacking chocolate consists of approximately 50% w/w sugar primarily as sucrose, though lactose is present in milk chocolate. Sucrose is a white, solid carbohydrate with the chemical formula $C_{12}H_{22}O_{11}$ and a molecular weight of 342.30 g/mole. It is a disaccharide consisting of glycosidic-linked glucose and fructose sugars (Figure 2).

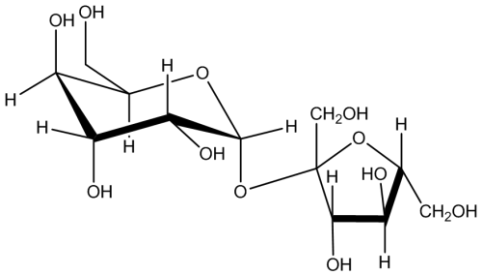


Figure 2. Sucrose molecule

The impact of sucrose on chocolate taste and sweetness is not to be understated, as a 5% w/w change in the concentration of sucrose in chocolate causes large flavour changes (Beckett, 2011). Indeed, the primary research drive concerning sucrose in chocolate is quantifying it mass in-line during chocolate manufacture (da Costa Filho, 2009). However, sucrose is not a reducing sugar and therefore does not contribute to the development of chocolate flavour in crumb by

Maillard reactions. Sucrose only exists in the crystalline state as one form, whose crystal structure was determined in 1952 (Beevers et al., 1952). For the purposes of this thesis the more recent redetermination of the structure by Hynes and Le Page will be used (Hynes and le Page, 1991).

1.3.3 Lactose

Lactose is a disaccharide present in milk chocolate from milk in the form of SCM in chocolate crumb and milk powder in traditional chocolate making. It consists of galactose and glucose linked by a glycosidic bond and appears as a white solid with the chemical formula $C_{12}H_{22}O_{11}$ and a molecular weight of 342.30 g/mole.

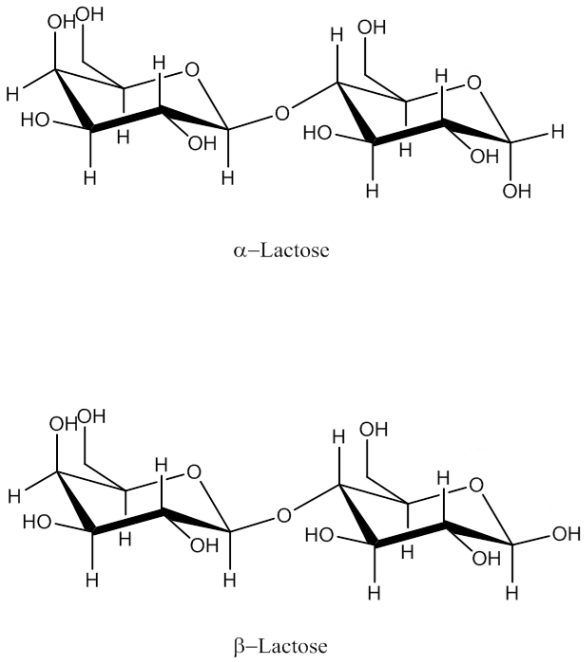


Figure 3. Lactose molecules

Lactose possesses two anomeric forms, denoted as α and β, differentiated by the positions of the hydroxyl and hydrogen groups on the free anomeric carbon (Figure 3). In solution, this anomeric carbon can switch positions causing the molecule to change from one anomer to another, a phenomenon known as mutarotation. This process happens naturally in aqueous solutions until an equilibrium is reached of 37% α-lactose and 63% β-lactose at room temperature (Fox, 1997). The α and β anomers have distinct solubility values (7 g/100 mL and 50 g/100 mL respectively) and thus produce different sweetness sensations when tasted, with β-lactose being sweeter.

However, due to the mutarotation between anomeric forms, the final solubility of lactose in solution is reported as 17 g/100 mL (Fox, 1997).

Lactose is a reducing sugar and can therefore react with the amino acids present in milk proteins at high temperatures via the Maillard reaction, leading to the development of cooked and caramelised flavours in crumb-based chocolates. These reactions are complex, with many of the flavour-generating compounds being formed in trace amounts (Davies and Labuza, 2000). Hence, the study of Maillard reactions lies outside of the scope of this thesis.

Lactose is a heavily researched sugar due to its usage as a bulking excipient in the pharmaceutical industry (Chaudhari and Patil, 2012). Several crystal forms of the lactose anomers exist, many of which have detailed methods of preparation and have been fully characterised crystallographically. The details of currently known crystal forms of lactose and associated literature are shown in Table 3.

Table 3. Reported crystalline forms of D-lactose for which crystal structures are available.

Form	Abbreviation	SpGrp	Z	Z'	T (K)	Refcode	Type	Ref.
α-lactose monohydrate	α-L.H_2O	$P2_1$	2	1	150	LACTOS11	SX	(Smith et al., 2005)
β-lactose	β-L	$P2_1$	2	1	293	BLACTO	SX	(Hirotsu and Shimada, 1974)
α-lactose (hygroscopic)	α-LH	$P2_1$	2	1	293	EYOCUQ	PXRD	(Platteau et al., 2004)
α-lactose (stable anhydrous)	α-LS	$P1$	2	2	293	EYOCUQ01	PXRD	(Platteau et al., 2005)
αβ-lactose	αβ-L_T	$P1$	2	2	120	LAKKEO01	SX	(Guiry et al., 2008)

SpGrp = Space group; Refcode = Cambridge Structural Database refcode; SX = structure obtained from single-crystal X-ray diffraction; PXRD = structure obtained from powder X-ray diffraction

1.4 Crystalline and Amorphous Forms

Crystallisation is the process where solid crystals of a substance are formed from a melt or solution and it consists of two keys steps – nucleation and crystal growth (Markov, 2016; Kashchiev, 2000). Nucleation occurs when solute molecules in solution cluster together to form nuclei (Liang and Hartel, 1991); upon these nuclei additional molecules deposit on the surface and grow the crystal further.

Crystalline substances a have defined long-range order of molecular packing that is characteristic of the polymorphic form. Indeed, different polymorphic forms exhibit different physicochemical characteristics such as solubility and melting point due to differences in long-range order. Furthermore, crystalline forms generated from solution can form intermolecular hydrogen bonds with water to produce stable hydrates, one such example being α-lactose monohydrate.

In general, amorphous materials possess no long-range order, though some materials classified as amorphous can exhibit very short-range order i.e. of the order of a few molecular dimensions. Amorphous materials are typically more soluble, hygroscopic and viscous than crystalline forms of the same substance (Yu, 2001; Willart and Descamps, 2008; Chiou and Langrish, 2007). This increase in dissolution rate of amorphous solids has led to a significant amount of interest in the formulation of solid dosage forms for pharmaceuticals. Amorphous solids are, however, metastable and hence may recrystallise when water content or temperature exceeds values dependent on the amorphous material (Jouppila et al., 1998).

Two transformation mechanisms underly the techniques behind generation of the amorphous form. Firstly, the crystalline material can be temporarily transformed into a thermodynamically unstable form through melting or dissolution, quench cooling of a melt, or rapid evaporation of a solvent. A metastable amorphous solid is formed when the rate of solidification outpaces the rate of crystallisation of the material. Secondly, the lattice of the crystalline material can be physically disrupted through mechanical methods such as milling or micronisation. It should be noted that different methods of production can lead to different physicochemical properties of the resultant amorphous solid, including the duration until

recrystallisation onset (Graeser et al., 2008). This is likely due to retention of some of the aforementioned very short-range order.

Lactose and sucrose produce relatively unstable amorphous forms; consequently they tend towards recrystallisation at low water content (5% w/w for lactose and 1% w/w for sucrose) (Kedward et al., 2000). Furthermore, it has been demonstrated (Ziegler and Hogg, 1999) by the use of differential scanning calorimetry (DSC) that amorphous lactose recrystallises during the conching phase of chocolate production by the absence of a glass transition thermal event post-conching.

The crystalline or amorphous state of lactose in milk powder chocolates can have a significant impact on the processability, stability and quality of the product. The recrystallisation of amorphous lactose in stored bulk chocolate has been suggested to be the cause of thickening of chocolate held in storage tanks. Ziegleder et al. (Ziegleder et al., 2004) propose that recrystallisation of the amorphous lactose releases moisture which then bridges milk powder particles, subsequently increasing the viscosity of the bulk chocolate. Further research has shown that usage of amorphous lactose resulted in increased viscosity when ball mill processing milk chocolate (Böhme et al., 2020). This article also showed that the moisture produced during recrystallisation of amorphous lactose caused milk fats to be released from milk powder particles, which may affect the additional quantity of fat required in downstream chocolate processes. Conversely, amorphous lactose has been shown to form a matrix dispersion of proteins, fat globules, and air which traps milk fats, thereby reducing the overall creaminess of the end product (Gaonkar and McPherson, 2006). Aguilar and Zieglar (Aguilar and Ziegler, 1995) demonstrated that the presence of amorphous lactose from spray-dried powders increased the particle size and decreased viscosity of post-refiner chocolate; they suggest that variations in chocolate processability due to flowability could be due to differences in storage and handling of milk powders resulting in variable quantities of amorphous lactose in the recipe. Recrystallisation of amorphous lactose through treatment of milk powders prior to their application in processing has been shown to desirably reduce the viscosity of chocolate, treatment of the powders through usage of a heated twin-screw extruder has been demonstrated as one potential solution (Franke and Heinzelmann, 2008).

The physical state of sucrose can also impact the flavour and properties of chocolate products; for example, the reactive surface of amorphous sucrose absorbs flavour compounds of surrounding ingredients during chocolate refining. When chocolates are made with pre-ground sucrose, the final product is reported to have a less satisfactory flavour than when particle size reduction is applied to the chocolate mixture as a whole (Beckett, 2011). Furthermore, recrystallisation of sucrose under conditions of high humidity and/or temperature releases these compounds, reducing the flavour quality of the final product. Amorphous sucrose may also potentially improve the heat stability of chocolate – addition of ground amorphous sucrose at a level as low as 1% w/w to the post-conch chocolate stabilises the mixture, allowing the product to be enjoyed in higher temperature climates (Beckett, 2011). When considering crumb chocolates in particular, the partial amorphization of sucrose during crumb manufacture could also aid the heat stability of the final product.

Estimates of the quantities of amorphous sugars in chocolate vary wildly. Beckett (Beckett, 2011) estimates that between 30 and 90% of sugar gets converted to the amorphous form during roller refining, whilst conflicting research shows through visual X-ray diffraction analysis that any amorphous sucrose present during production of chocolate has entirely recrystallised in the end product (Gloria and Sievert, 2001). Clearly an accurate and reliable method of characterisation and quantification is essential for a consensus on the importance and presence of amorphous and crystalline sugars in chocolate.

1.5 Characterisation Methods

This section summarises the key analytical methods used in the characterisation of amorphous materials and their current application to chocolate products and precursors. Typically, analyses of amorphous materials measure changes in the physical properties such as substance density and viscosity, assessment of heat changes during thermal events, or direct assessment of the crystallinity of the substance.

Selection of appropriate analytical techniques to assess and quantify amorphous and crystalline mixtures is reliant on the sensitivity and selectivity of the method used. Choosing a technique appropriate to the material being analysed is crucial as chocolate is a complex mixture of a multitude of crystalline and amorphous phases. Analysis of the degree of crystallinity of

substances has been heavily covered in the context of pharmaceuticals, due to the importance of physical form on drug stability and the stringent requirements set out by regulatory authorities (Lehto et al., 2006; Shah et al., 2006). Food product regulation is much less concerned with the crystallinity of ingredients in a foodstuff. Furthermore, food products vary significantly more than pharmaceuticals in formulation methodology, diversity of ingredients (or in the case of pharmaceuticals, excipients), and availability of state-of-the-art analytical techniques. Consequently, a consensus on appropriate analytical technique usage for qualification and quantification of chocolate components is far from being reached.

Thermal analytical techniques including thermogravimetric analysis (TGA), differential scanning calorimetry (DSC), solution calorimetry (SolCal), and hot-stage microscopy (HSM) are thoroughly utilised in the materials science, food science and pharmaceutical industries (Giron, 2002; Craig and Reading, 2006; Plante et al., 2009; Müllertz et al., 2016; Ibañez and Cifuentes, 2001; Ramachandran et al., 2002; Roos, 2003). Spectroscopic techniques such as hydrogen and carbon nuclear magnetic resonance (NMR), near infrared (NIR) and FT-Raman spectroscopy are also extensively applied (Fossen and Andersen, 2006; Scotter, 1997; Lohumi et al., 2015; Aaltonen et al., 2008; Bugay, 2001). Powder X-ray diffraction (PXRD) is heavily used for fingerprinting and in-situ monitoring of small molecule pharmaceuticals (Fawcett et al., 2019; Thakral et al., 2018); however its applications in food science appear to fill a much smaller niche (Sanchez et al., 2018; Lamberti et al., 2004).

1.5.1 Nuclear Magnetic Resonance (NMR)

NMR is a spectroscopic technique that utilises magnetic properties of nuclei in a substance to garner information on the its structure. In the case of ^{1}H (or proton) NMR, several interactions between the nuclei and the magnetic field they are exposed to produce a spectrum that allows for assignments of hydrogen positions within a molecule. Deciphering the integrated intensity, magnetic coupling of nuclei in relation to their distance from one another in space, and chemical shift enables accurate assignment of hydrogen positions within a molecule and thus the structure of the molecule itself (Bugay, 1993).

Solid-state NMR has been used to consistently and accurately detect very low quantities (0.5%) of amorphous lactose in mixtures of crystalline and amorphous lactose without damaging

the sample assessed (Gustafsson et al., 1998) - in this study, amorphous lactose produced broader signals in the spectra. Amorim (Amorim et al., 2020) showed that by using multivariate calibration of chocolate ^{1}H-NMR spectra, the quantity of restricted trans fatty acids can be rapidly obtained. ^{1}H-NMR has also been applied to the investigation of mutarotation between the α- and β- anomers of lactose, where characteristic signals at chemical shifts of 6.3 and 6.6 ppm (the anomeric hydroxl hydrogen) were quantified (Jawad et al., 2012).

Within the remit of this thesis, NMR is used for *qualitative* purposes in determining the chemical composition of dried lactose syrups, in order to assess whether any chemical transformation through hydrolysis (or otherwise) has occurred.

1.5.2 Differential Scanning Calorimetry (DSC)

Differential scanning calorimetry is heavily used in the pharmaceutical industry to thermally quantify and characterise mixtures of crystalline and amorphous pharmaceuticals. Furthermore it is the *de facto* technique for determination of glass transition temperatures (Lehto et al., 2006). The technique involves heating or cooling a sample across a typically constant ramp rate and measuring heat changes between a sample and reference.

In analysis of amorphous materials, recrystallisation of the substance to a more stable form may occur during the heating ramp; this is reflected as an exothermic signal on the scan. Gloria and Sievert (Gloria and Sievert, 2001) leveraged DSC to analyse and compare amorphous and crystalline sucrose. Where crystalline sucrose showed a single melting endotherm at 188 °C, amorphous sucrose produced a glass transition trough at 57 °C, recrystallisation exothermic peak at 105 °C, and the onset of melting at 176 °C (Figure 4). The group utilised this characterisation to quantify the presence of amorphous sucrose during dark chocolate manufacture.

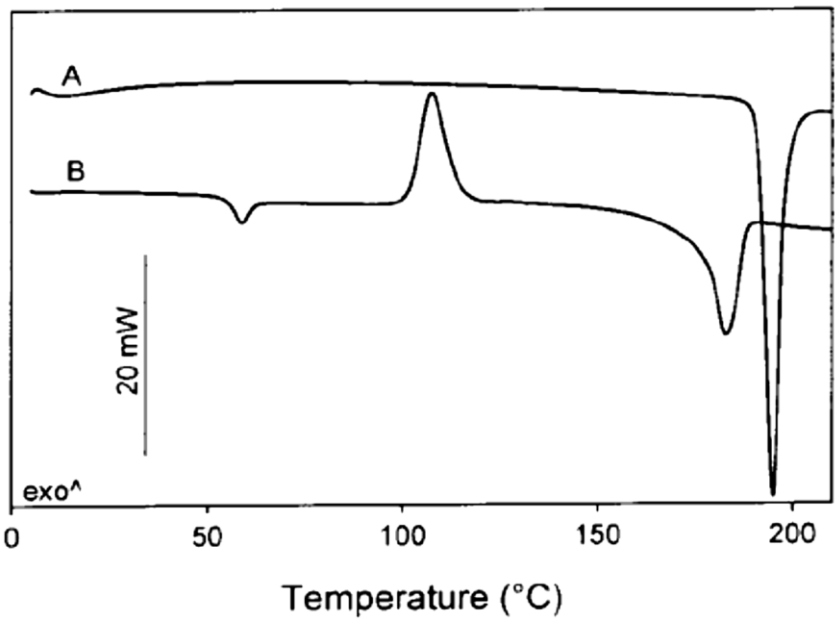


Figure 4. DSC scans of crystalline (A) and amorphous (B) sucrose. Source: (Gloria and Sievert, 2001)

The presence of amorphous sucrose and lactose is easily quantifiable when measured independently. Both amorphous sucrose and lactose have significant step signals for their glass transitions and are therefore detectable at slower ramp rates. Additionally, both sugars tend to produce recrystallisation exotherms that can be used to create a calibration curve for their quantities in the mixture (Simperler et al., 2006; Islam and Langrish, 2010). Lastly, the most stable crystalline form of lactose, α-lactose monohydrate, shows a dehydration event that can be used as a simple diagnostic assessment of its presence in a mixture (Islam and Langrish, 2010).

However, there are limitations to the usage of DSC in quantification of chocolate mixtures. Both sugars degrade shortly after melting and the resultant exotherm can make accurate enthalpy determinations difficult (Lee et al., 2011). Additionally, interactions between melting sucrose and lactose convolutes the trace produced, Figures 5 and 6 show DSC traces of crystalline sucrose and α-lactose monohydrate both independently and as a 50:50 mixture. NB: this finding is unpublished and was conducted during the iCASE ‘’placement element of the thesis.

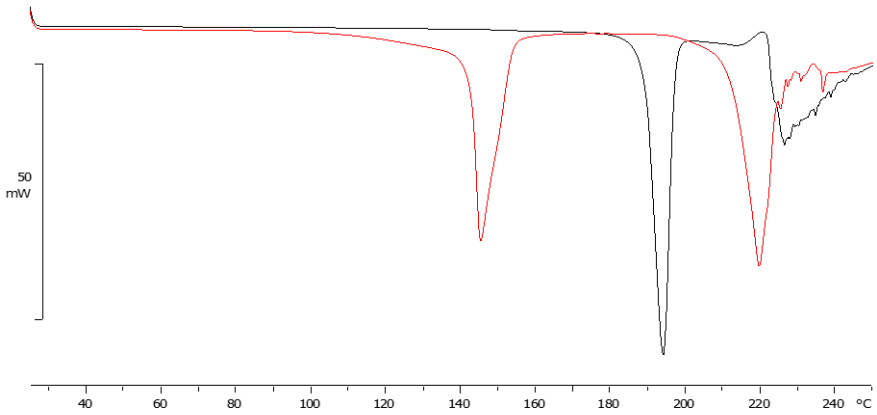


Figure 5. Overlay of the individual DSC traces of sucrose (black) and α-lactose monohydrate (red).

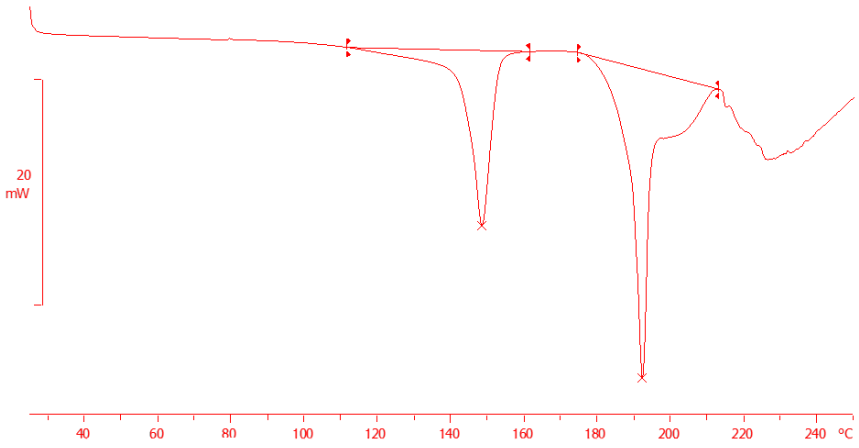


Figure 6. DSC trace of a 50:50 mixture of sucrose and α-lactose monohydrate.

When measured independently, α-lactose monohydrate shows a dehydration at 148 °C and subsequent melt peak at 220 °C, whilst sucrose shows a melt peak at 192 °C – both sugars begin to degrade at approximately 225 °C, represented by the broad and noisy peak on the DSC traces. When the sugars are measured as a 50:50 mixture the dehydration of α-lactose monohydrate is still clearly visible on the trace. However, the melt peak of lactose has shifted to a lower temperature and significant overlap between the sucrose and lactose peaks makes determination of individual heat changes impossible. Multiple heating ramp rates between 10 °C/min and 50 °C/min were tested to try to resolve the peaks but this proved unsuccessful.

When DSC was used to assess chocolate crumb a different trace was obtained, (Figure 7). Clear melt peaks are observable for both sugars, but both occur at lower temperatures than when observed in isolation and the DSC trace is complicated by significant peak broadening and overlap. Hence, generation of a calibration curve and therefore accurate quantification of their amorphous or crystalline components was not possible.

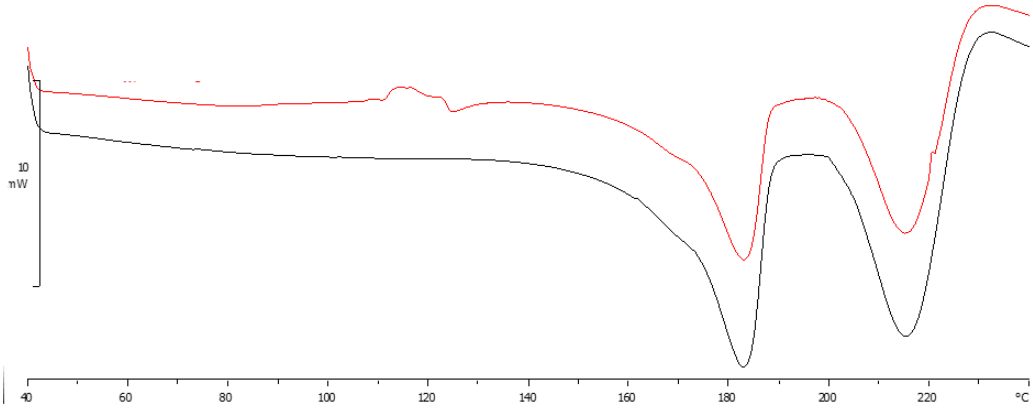


Figure 7. DSC trace of chocolate crumb (n=2).

Due to the degree of peak broadening and overlap observed in the DSC trace, it is not a suitable method for quantification of either crystalline or amorphous forms of the sugars in chocolate. Within this thesis, the use DSC is restricted to physical characterisation of dried lactose syrups.

1.5.3 Powder X-Ray Diffraction (PXRD)

Powder X-ray diffraction (PXRD) is a technique that is used primarily in the fingerprinting of crystalline materials but which has multiple applications across many industries. For a detailed introduction to the applications of PXRD, see the included published book chapter written in collaboration with Spillman and Shankland. Only a brief description of the underlying principles of PXRD is given here.

In an X-ray diffraction experiment, a sample is exposed to a beam of X-ray radiation, which interacts with the electrons within each atom and consequently diffracts. For crystalline materials, the long-range order of the atoms results in a diffraction pattern which can then be interpreted and analysed to obtain details about the crystal structure. Diffraction occurs when incident X-rays

diffract in phase, producing constructive interference when the path difference is equal to the number of wavelengths (as an integer). This is described simply by the well-known Bragg's law,

$$n\lambda=2d\cdot sin\vartheta$$

Equation 1

where n is an integer, λ is the wavelength of the incident X-rays, d is the spacing between the planes in the atomic lattice, and θ is the angle between the incident X-rays and the lattice plane.

The most widely used X-ray technique for the full 3D-characterisation of crystalline materials is single crystal X-ray diffraction. With modern X-ray diffractometers, given a single-crystal ca. 0.1 mm in all dimensions, a high resolution crystal structure can normally be obtained straightforwardly. Not only can the positions and types of all atoms in the structure be determined in the absence of any prior knowledge of the composition of the crystal, but so too can atomic displacements due to thermal vibrations. The main limitation of the technique is the need for a single crystal, growth of which is not always straightforward or indeed possible. In these cases, the power of PXRD is leveraged.

The powder diffraction method using conventional X-ray sources was developed by Debye and Scherrer in 1916. Further advances in the field, such as the development of the Rietveld method (Rietveld, 1969), availability of high-resolution laboratory equipment, and improvements in computational power have improved the number of applications and effectiveness of PXRD experiments (Spillman et al., 2020).

PXRD is the only diffraction technique suitable for the analysis of multi-phase samples (i.e. samples containing more than one crystalline form of the same, or different, materials) where the aim is to quantify each phase. Indeed, while a multitude of methods of obtaining phase-related information are available, diffraction methods are the most direct. This is because diffraction data is produced directly from the crystalline (or amorphous) structure of each phase, rather than being derived from secondary information (such as by chemical methods).

Quantification of powder diffraction data is reliant on the evaluation of each phases' contribution to the final pattern. A variety of techniques are employable and can be classified into two distinct groups:

1. Single peak methods: use a peak, or cluster of peaks, for each phase of interest to calculate the abundance. These techniques make the assumption that the peak(s) selected is (are) representative of the amount of each phase. However, this is often not the case due to phase-dependent factors (such as the tendency for certain crystal habits to organise in a non-random way, known as preferred orientation) and possible diffraction peak overlap between distinct phases in a mixture leading to inaccurate extraction of peak intensity.

2. Whole pattern methods: compare a wide range of diffraction data with a calculated pattern formed from the addition of each phase component that have been measured from pure phase samples or calculated from crystal structure information.

Though a wide range of whole pattern methods exist, a recent round robin on quantitative phase analysis (QPA) by the International Union of Crystallography determined the Rietveld method to be the most commonly used (Scarlett et al., 2002; Madsen et al., 2001). Note that application of the Rietveld method to QPA does require the knowledge of the crystal structures involved.

Single peak methods have been used to quantify the crystalline proportions of α-lactose monohydrate and β lactose in chocolate (Thomas et al., 2009), though these methods are purely limited to crystalline phases within the mixture and do not account for the amorphous component.

Amorphous materials do not exhibit the long-range order associated with crystalline phases. Their diffraction patterns are often said to consist of "amorphous humps", with no evidence of the peaks associated with Bragg scattering. Consequently, these substances are unable to be fully characterised crystallographically, and thus unable to be modelled using Rietveld refinement. Figure 8 shows a multiphase mixture consisting of both amorphous and crystalline components.

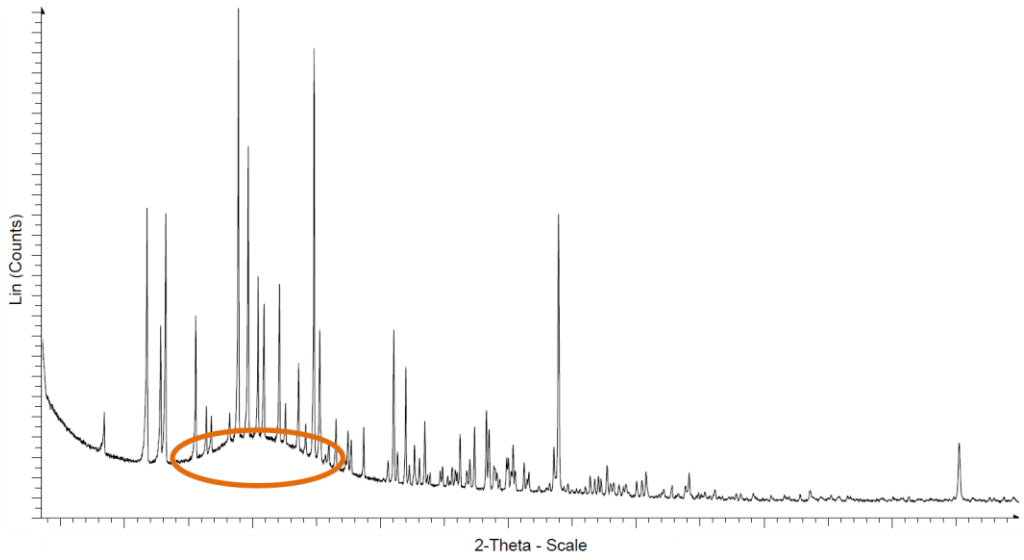


Figure 8: a multiphase mixture with both amorphous and crystalline components. The broad amorphous "hump" is highlighted in orange.

In order to calculate the absolute percentages of each phase in a multiphase mixture, the amorphous material must be accounted for. Methods used in quantification of amorphous phases can be divided into two categories:

1. *Direct methods* – infer the quantity of amorphous material directly from the pattern. These methods are not always possible, as software used to model amorphous phase contributions find challenge in distinguishing between phases (due to the nature of broad, weak amorphous humps).

2. *Indirect methods* – use a standard material of known 100% crystallinity. Relative percentages of each phase determined by Rietveld refinement can then be placed on an absolute scale by reference to the percentage of standard introduced.

The internal standard method is an indirect method that only analyses the crystalline components of the mixture. A sample is spiked with a standard assumed to be of 100% crystallinity, by use of the following equation the absolute weight fraction of each crystalline phase can be counted (i.e. the sample excluding the standard):

$$Wx(abs) = Wx(calc) \cdot Ws(known) Ws(calc)$$

Equation 2

where *Wx(abs)* is the absolute weight fraction of phase *x*, *Wx(calc)* is the calculated relative weight fraction of phase *x*, *Ws(known)* is the known weight fraction of the standard and *Ws(calc)* is the calculated relative weight fraction of the standard.

Thus, by obtaining the absolute weight fraction of each crystalline phase within the sample, the amorphous concentration can be calculated as the weight fraction not accounted for by crystalline phases in the mixture. If the original composition of the mixture assessed is known (or if this is analysed by use of complementary techniques such as HPLC) the contribution of each component to total amorphous content can then be simply elucidated.

Selection of an appropriate internal standard is paramount to obtaining accurate results in Rietveld-based QPA, primarily because of microabsorption effects. These microabsorption effects result when a multiphase sample contains elements with highly variable absorption coefficients. Weakly absorbing phases cause a greater degree of X-ray scattering and hence are over-estimated when their weight fraction is calculated (the converse is also true). Hence it is desirable to assess a mixture with similar absorption coefficients as the effects become negligible. Thus we have chosen diamond powder to utilise as a standard when conducting QPA on chocolate-based mixtures because chocolate (as a mixture of organic phases) possess primarily carbon, hydrogen, nitrogen and oxygen atoms; diamond is of course pure carbon. An associated benefit of utilising diamond as an internal standard is that it produces few diffraction peaks at high 2θ angles, and therefore the likelihood of it interfering with analysis of diffraction peaks attributable to the chocolate's crystalline components is small.

1.6 Project Rationale

Accessibility to high fat and high sugar products, combined with a more sedentary lifestyle, has led to a worldwide obesity pandemic. Obesity rates have risen significantly over the past 35 years – in 1986, 1 in 200 American adults were obese; as of 2017 that number has risen to 1 in 5 (Agha and Agha, 2017). Similar trends are observed in the UK (Department of Health & Social Care, 2020), with two thirds of UK adults weighing over their healthy weight recommendation. Moreover, people who are overweight or obese are more likely to be admitted to hospital, receive intensive care treatment, and die if they contract COVID-19 (Public Health England, 2020a).

The UK government is responding to this crisis through several measures. Most relevant is the banning of advertising on television and online of products that are high in fat, sugar or salt, as well as the introduction of legislation to restrict promotional offers on these products and their placement in prominent locations in stores (Government Office for Science, 2009). Recently, the government encouraged manufacturers to reduce the quantity of sugar in their products, including chocolate manufacturers (Public Health England, 2020b). Chocolate manufacturers are responding to this voluntary program by reduction of the sugar content in chocolate by several means.

Replacement of sugar with artificial sweeteners or other natural alternatives is one approach confectionary companies are taking (Saputro et al., 2017; Cikrikci et al., 2017; di Monaco et al., 2018). However, replacement of sucrose (of which chocolate is typically comprised of 50% w/w) with significantly sweeter (by weight) artificial sweeteners leads to significant reductions in the bulk weight of the end product. Additionally, some sweeteners, such as xylitol, provide an undesirable cooling sensation when eaten. Furthermore, the government places restrictions on the daily quantity of artificial sweetener appropriate for ingestion due to the laxative properties of many of the compounds (Beckett, 2011).

Some chocolate companies have attempted alternative approaches to sugar reduction. Nestle produced a chocolate containing porous sucrose which cut the sugar content of these products by nearly a third. However, poor sales have led to the product being pulled from stores (BBC News, 2020). The usage of the amorphous form of sucrose may be an alternative approach to sugar reduction. The reactive surface of amorphous sucrose can absorb nearby chocolate flavours producing a more intense flavour. Moreover, amorphous sugars dissolve more quickly on the tongue, providing a sweeter sensation when eaten. Finally, amorphous sugars can provide a more cohesive mouthfeel due to the faster absorption of water. Amorphous sugars are thought to be a component of the differences in flavour observed in chocolate crumb chocolates (Beckett, 2019).

Amorphous sugars are not a perfect solution to sugar reduction in chocolate. Their utilisation can affect the rheological properties of the chocolate intermediaries, causing difficulties in downstream processes due to increased viscosity. Furthermore, the increased hygroscopic properties of amorphous sugars can cause difficulties in storage and flavour – high water content

chocolates often give rise to a sticky mouthfeel (B. W. Minifie, 1979). Literature surrounding the impact of sugar crystallinity on the processing and end product is often conflicting; this in part is due to the lack of an accurate method of characterisation and quantification of sugar crystallinity. The development of a reproducible, accurate method to quantify amorphous and crystalline sugars in chocolate and chocolate precursors is key to facilitating further investigations into their role in processing and end-product quality.

Additionally, the crystallisation of chocolate sugars in chocolate crumb is not well understood. The crumb drying process subjects the sugars in milk chocolate crumb to a variety of high temperatures across a range of drying durations. As a result, understanding the crystallographic changes that occur during this intensive drying through observation and qualification of simpler systems is essential to gain insight into the changes that sugars in chocolate undergo during crumb manufacture.

To establish an approach that will fulfil the need for sucrose and lactose amorphous and crystalline quantification, and to further investigate the physical and chemical changes chocolate crumb sugars sustain during manufacture, the following aims and objectives are proposed.

1.7 Aims and Objectives

The original aim of this work was to try to obtain deeper insights into the nature of sucrose and lactose throughout some of the processing steps in the chocolate production process and also to potentially explore the use of alternative sweeteners. As the work developed, and following an extensive literature survey, it became clear that the crystallography of the sugars (and of lactose in particular) was not well understood or utilised in much of the research into chocolate. The aim of the work then became more focussed on the use of X-ray diffraction to better establish and demonstrate its power in this particular field.

Whilst studying the crystallisation of sugars in isolation (on the basis that it was likely to be easier to understand systems in isolation, rather than as parts of a complex mixture) a powder X-ray diffraction pattern for lactose was obtained which could not be explained by any combination of known crystalline forms. This pattern, which resisted all attempts at indexing, was

finally fully characterised when a microcrystal (76 × 24 × 18 μm) was isolated from the recrystallised powder and found to be a new crystalline form of αβ-D-lactose. Armed with this knowledge, the recrystallized powder was shown to be a three-phase mixture of lactose forms, with the novel form as the dominant component.

The main objectives then became:

Objective one: To apply PXRD to sugar raw materials, chocolate intermediates and chocolate samples. In particular, to establish whether or not conventional laboratory-based powder X-ray diffractometers (of a type that could be easily used by industry) are sufficiently powerful to enable reliable and informative characterisation of these materials and samples.

Objective two: To develop a reliable methodology for the accurate quantification of crystalline sugars in chocolate intermediates and chocolate samples.

Objective three: To use diffraction to investigate the phase behaviour of sugars upon drying of syrups under conditions much simpler than, but closely related to, those experienced during the manufacture of chocolate crumb.

1.8 References

Aaltonen, J. et al. (2008) Perspectives in the use of spectroscopy to characterise pharmaceutical solids. *International journal of pharmaceutics*. 364 (2), 159–169.

Andre, S. et al. (2017) Spore-forming bacteria responsible for food spoilage. *Research in Microbiology*. 168 (4), 379-387.

Agha, M. & Agha, R. (2017) The rising prevalence of obesity: part A: impact on public health. *International journal of surgery. Oncology*. 2 (7), 17.

Aguilar, C. A. & Ziegler, G. R. (1995) Viscosity of molten milk chocolate with lactose from spray-dried whole-milk powders. *Journal of food science*. 60 (1), 120–124.

Alpha MOS (2017) *Evaluation of chocolate bloom over shelf-life*

Amorim, T. L. et al. (2020) Screening method for determination of C18:1 trans fatty acids positional isomers in chocolate by H-1 NMR and chemometrics. *Food Science and Technology*. 131.

BBC News (2020) *Nestle axes low sugar chocolate due to weak sales* [online]. Available from: https://www.bbc.co.uk/news/business-51439407 (Accessed 20 February 2021).

Beckett, S. T. (2011) *Industrial chocolate manufacture and use*. John Wiley & Sons.

Beckett, S. T. (2003) Is the taste of British milk chocolate different? *International Journal of Dairy Technology*. 56 (3), 139–142.

Beckett, S. T. (2019) *The science of chocolate*. Royal Society of Chemistry.

Beevers, C. A. et al. (1952) The crystal structure of sucrose. *Acta Crystallographica*. 5 (5), 689–690.

Böhme, B. et al. (2020) Liberation of fat from milk powder particles during chocolate processing through moisture-induced lactose crystallisation. *LWT*.

Bugay, D. E. (2001) Characterization of the solid-state: spectroscopic techniques. *Advanced Drug Delivery Reviews*. 48 (1), 43–65.

Bugay, D. E. (1993) Solid-State Nuclear-Magnetic Resonance Spectroscopy - Theory and Pharmaceutical Applications. *Pharmaceutical Research*. 10 (3), 317–327.

Buscato, M. H. M. et al. (2018) Delaying fat bloom formation in dark chocolate by adding sorbitan monostearate or cocoa butter stearin. *Food chemistry*. 256390–396.

Chapman, G. M. et al. (1971) Cocoa butter and confectionery fats. Studies using programmed temperature X-ray diffraction and differential scanning calorimetry. *Journal of the American Oil Chemists Society*. 48 (12), 824–830.

Chaudhari, S. P. & Patil, P. S. (2012) Pharmaceutical excipients: a review. *Int J Adv Pharm Biol Chem*. 1 (1), 21–34.

Chiou, D. & Langrish, T. (2007) Crystallization of Amorphous Components in Spray-Dried Powders. *Drying Technology*. 251427–1435.

Cikrikci, S. et al. (2017) Physical characterization of low-calorie chocolate formulations. *Journal of Food Measurement and Characterization*. 11 (1), 41–49.

da Costa Filho, P. A. (2009) Rapid determination of sucrose in chocolate mass using near infrared spectroscopy. *Analytica Chimica Acta*. 631 (2), 206–211.

Craig, D. Q. M. & Reading, M. (2006) *Thermal analysis of pharmaceuticals*. CRC press.

Cruz-Cabeza, A. J. & Bernstein, J. (2014) Conformational polymorphism. *Chemical reviews*. 114 (4), 2170–2191.

Davies, C. & Labuza, T. (2000) *The Maillard Reaction Application to Confectionery Products*.

Department for Environment, F. & R. A. & Department for International Trade (2018) *UK's artisan chocolates an international favourite this Easter*

Department of Health & Social Care (2020) *Tackling obesity: empowering adults and children to live healthier lives*. [online]. Available from: https://www.gov.uk/government/publications/tackling-obesity-government-strategy/tackling-obesity-empowering-adults-and-children-to-live-healthier-lives.

Fawcett, T. G. et al. (2019) A practical guide to pharmaceutical analyses using X-ray powder diffraction. *Powder Diffraction*. 34 (2), 164–183.

Feiner, G. (2006) Meat Products Handbook. *Food Science, Technology, and Nutrition.*

Fossen, T. & Andersen, Ø. M. (2006) Spectroscopic techniques applied to flavonoids. *Flavonoids: chemistry, biochemistry and applications*. 37–142.

Fox, P. F. (1997) *Advanced dairy chemistry: lactose, water, salts and vitamins.* Volume 3. Springer.

Franke, K. & Heinzelmann, K. (2008) Structure improvement of milk powder for chocolate processing. *International Dairy Journal*. 18 (9), 928–931.

Fryer, P. & Pinschower, K. (2000) The materials science of chocolate. *Mrs Bulletin*. 25 (12), 25–29.

Gaonkar, A. G. & Mcpherson, A. (2006) *Ingredient interactions*. Second edition. CRC Press.

Ghazani, S. M. & Marangoni, A. G. (2018) New insights into the β polymorphism of 1, 3-palmitoyl-stearoyl-2-oleoyl glycerol. *Crystal Growth & Design*. 18 (9), 4811–4814.

Giron, D. (2002) Applications of thermal analysis and coupled techniques in pharmaceutical industry. *Journal of Thermal Analysis and Calorimetry*. 68 (2), 335–357.

Gloria, H. & Sievert, D. (2001) Changes in the Physical State of Sucrose during Dark Chocolate Processing. *Journal of Agricultural and Food Chemistry*. 49 (5), 2433–2436.

Government Office for Science (2009) *Foresight - tackling obesities: future choices - modelling future trends in obesity & their impact on health.* [online]. Available from: www.bis.gov.uk/assets/bispartners/foresight/docs/obesity/14.pdf.

Graeser, K. A. et al. (2008) Physicochemical Properties and Stability of Two Differently Prepared Amorphous Forms of Simvastatin. *Crystal Growth & Design*. 8 (1), 128–135.

Guice, W. A. et al. (1959) Addition of hydrogenated fats to chocolate to improve heat resistance. *Journal of the American Oil Chemists' Society*. 36, 4-8.

Guiry, K. P. et al. (2008) Effect of 1-Deoxy-d-lactose upon the Crystallization of d-Lactose. *Crystal Growth and Design*. 8 (11), 3927–3934.

Gustafsson, C. et al. (1998) Comparison of solid-state NMR and isothermal microcalorimetry in the assessment of the amorphous component of lactose. *International Journal of Pharmaceutics.* 174 (1–2), 243–252.

Gutiérrez, T. J. (2017) State-of-the-art chocolate manufacture: a review. *Comprehensive Reviews in Food Science and Food Safety*. 16 (6), 1313–1344.

Hirotsu, K. & Shimada, A. (1974) The crystal and molecular structure of β-lactose. *Bulletin of the Chemical Society of Japan*. 47 (8), 1872–1879.

Hynes, R. C. & le Page, Y. (1991) Sucrose, a convenient test crystal for absolute structures. *Journal of Applied Crystallography*. 24 (4), 352–354.

Ibañez, E. & Cifuentes, A. (2001) New analytical techniques in food science. *Critical reviews in food science and nutrition*. 41 (6), 413–450.

Islam, M. I. U. & Langrish, T. A. G. (2010) An investigation into lactose crystallization under high temperature conditions during spray drying. *Food Research International*. 43 (1), 46–56.

Jawad, R. et al. (2012) The Measurement of the beta/alpha Anomer Composition Within Amorphous Lactose Prepared by Spray and Freeze Drying Using a Simple H-1-NMR Method. *Pharmaceutical Research*. 29 (2), 511–524.

Jouppila, K. et al. (1998) Factors affecting crystallization and crystallization kinetics in amorphous corn starch. *Carbohydrate polymers*. 36 (2–3), 143–149.

Kashchiev, D. (2000) *Nucleation*. Elsevier.

Kedward, C. J. et al. (2000) Crystallization kinetics of amorphous lactose as a function of moisture content using isothermal differential scanning calorimetry. *Journal of Food Science*. 65 (2), 324–328.

Lamberti, M. et al. (2004) Starch transformation and structure development in production and reconstitution of potato flakes. *LEBENSMITTEL-WISSENSCHAFT UND-TECHNOLOGIE-FOOD SCIENCE AND TECHNOLOGY*. 37 (4), 417–427.

Lee, J. W. et al. (2011) Investigation of the Heating Rate Dependency Associated with the Loss of Crystalline Structure in Sucrose, Glucose, and Fructose Using a Thermal Analysis Approach (Part I). *Journal of Agricultural and Food Chemistry*. 59 (2), 684–701.

Lehto, V.-P. et al. (2006) The comparison of seven different methods to quantify the amorphous content of spray dried lactose. *Powder Technology*. 167 (2), 85–93.

Liang, B. & Hartel, R. W. (1991) Techniques for developing nucleation and growth kinetics from MSMPR data for sucrose crystallization in the presence of growth rate dispersion. *Journal of Crystal Growth*. 108 (1–2), 129–142.

Lohumi, S. et al. (2015) A review of vibrational spectroscopic techniques for the detection of food authenticity and adulteration. *Trends in Food Science & Technology*. 46 (1), 85–98.

Loisel, C. et al. (1998) Phase transitions and polymorphism of cocoa butter. *Journal of the American Oil Chemists' Society*. 75 (4), 425–439.

Madsen, I. C. et al. (2001) Outcomes of the International Union of Crystallography Commission on Powder Diffraction round robin on quantitative phase analysis: samples 1a to 1h. *Journal of Applied Crystallography*. 34 (4), 409–426.

van Malssen, K. et al. (1999) Phase behavior and extended phase scheme of static cocoa butter investigated with real-time X-ray powder diffraction. *Journal of the American Oil Chemists' Society*. 76 (6), 669–676.

Marangoni, A. G. & McGauley, S. E. (2003) Relationship between crystallization behavior and structure in cocoa butter. *Crystal Growth & Design*. 3 (1), 95–108.

Markov, I. V. (2016) *Crystal growth for beginners: fundamentals of nucleation, crystal growth and epitaxy*. World scientific.

van Mechelen, J. B. et al. (2006) Structures of mono-unsaturated triacylglycerols. II. The [beta]2 polymorph. *Acta Crystallographica Section B*. 62 (6), 1131–1138.

Minifie, B. (1979) A review of the processes for the manufacture of milk chocolate. *Manuf. Confect*. 59 (10), 43–48.

Minifie, B. W. (1979) *Chocolate, cocoa and confectionary: science and technology*. Second Edition.

Mirzaee Ghazani, S. & Marangoni, A. (2021) Molecular Origins of Polymorphism in Cocoa Butter. *Annual Review of Food Science and Technology*. 12.

di Monaco, R. et al. (2018) Strategies to reduce sugars in food. *Current Opinion in Food Science*. 19 (SI), 92–97.

Müllertz, A. et al. (2016) *Analytical techniques in the pharmaceutical sciences*. Springer.

Plante, A. F. et al. (2009) Application of thermal analysis techniques in soil science. *Geoderma*. 153 (1–2), 1–10.

Platteau, C. et al. (2004) Ab initio structure determination of the hygroscopic anhydrous form of α-lactose by powder X-ray diffraction. *Acta Crystallographica Section B: Structural Science*. 60 (4), 453–460.

Platteau, C. et al. (2005) Structure determination of the stable anhydrous phase of α-lactose from X-ray powder diffraction. *Acta Crystallographica Section B: Structural Science*. 61 (2), 185–191.

Public Health England (2020a) *Excess Weight and COVID-19 - Insights from new evidence*. [online]. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/907966/PHE_insight_Excess_weight_and_COVID-19__FINAL.pdf.

Public Health England (2020b) *Sugar reduction - Report on progress between 2015 and 2019*. [online]. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/925027/SugarReportY3.pdf.

Ramachandran, V. S. et al. (2002) *Handbook of thermal analysis of construction materials*. William Andrew.

Rietveld, H. (1969) A profile refinement method for nuclear and magnetic structures. *Journal of Applied Crystallography*. 2 (2), 65–71.

Roos, Y. (2003) Thermal analysis, state transitions and food quality. *Journal of Thermal Analysis and Calorimetry*. 71 (1), 197–203.

Sanchez, J. O. et al. (2018) Degradation of L-Ascorbic Acid in the Amorphous Solid State. *Journal of Food Science*. 83 (3), 670–681.

Saputro, A. D. et al. (2017) Quality attributes of dark chocolates formulated with palm sap-based sugar as nutritious and natural alternative sweetener. *European Food Research and Technology*. 243 (2), 177–191.

Scarlett, N. V. Y. et al. (2002) Outcomes of the International Union of Crystallography Commission on powder diffraction round robin on quantitative phase analysis: Samples 2, 3, 4, synthetic bauxite, natural granodiorite and pharmaceuticals. *Journal of Applied Crystallography*. 35 (4), 383–400.

Schnermann, P. & Schieberle, P. (1997) Evaluation of key odorants in milk chocolate and cocoa mass by aroma extract dilution analyses. *Journal of Agricultural and Food Chemistry*. 45 (3), 867–872.

Schuhmacher, K. et al. (1996) El gran libro del chocolate. *León: Everest*.

Scotter, C. N. G. (1997) Non-destructive spectroscopic techniques for the measurement of food quality. *Trends in food science & technology*. 8 (9), 285–292.

Shah, B. et al. (2006) Analytical techniques for quantification of amorphous/crystalline phases in pharmaceutical solids. *Journal of Pharmaceutical Sciences*. 95 (8), 1641–1665.

Simperler, A. et al. (2006) Glass transition temperature of glucose, sucrose, and trehalose: An experimental and in silico study. *Journal of Physical Chemistry B*. 110 (39), 19678–19684.

Smith, J. H. et al. (2005) α-lactose monohydrate: a redetermination at 150 K. *Acta Crystallographica Section E: Structure Reports Online*. 61 (8), 2499–2501.

Sokmen, A. & Gunes, G. (2006) Influence of some bulk sweeteners on rheological properties of chocolate. *Food Science and Technology*. 39 (10), 1053–1058.

Spillman, M., Nicholls, D. & Shankland, K. (2020) Experimental analysis of powder diffraction data. *Handbook on Big Data and Machine Learning in the Physical Sciences*. 1-23

Thakral, N. K. et al. (2018) Applications of Powder X-Ray Diffraction in Small Molecule Pharmaceuticals: Achievements and Aspirations. *Journal of Pharmaceutical Sciences.* 107 (12), 2969–2982.

Thomas, N. R. et al. (2009) Quantitative X-ray diffraction determination of alpha-lactose monohydrate and beta-lactose in chocolate. *Journal of Food Science*. 74 (7), C513-8.

Toro-Vazquez, J. F. et al. (2004) Rheometry and polymorphism of cocoa butter during crystallization under static and stirring conditions. *Journal of the American Oil Chemists' Society*. 81 (2), 195–202.

Van Mechelen, J.B. et al. (2006). Structures of mono-unsaturated triacylglycertols. I. The β_1 polymorph. *Acta Crystallographica B*. 62, 1121-1130.

Willart, J. F. & Descamps, M. (2008) Solid State Amorphization of Pharmaceuticals. *Molecular Pharmaceutics*. 5 (6), 905–920.

Wille, R. L. & Lutton, E. S. (1966) Polymorphism of cocoa butter. *Journal of the American Oil Chemists' Society*. 43 (8), 491–496.

Yao, Y. et al. (2020) Dynamic changes in the triacylglycerol composition and crystallization behaviour of cocoa butter. *LWT*.

Yu, L. (2001) Amorphous pharmaceutical solids: preparation, characterization and stabilization. *Advanced drug delivery reviews*. 48 (1), 27–42.

Zhao, H. et al. (2018) Effects of fat polymorphic transformation and nonfat particle size distribution on the surface changes of untempered model chocolate, based on solid cocoa mass. *Journal of Food Science*. 83 (4), 998–1004.

Zhao, Y. (2012) *Specialty foods: processing technology, quality, and safety*. CRC Press.

Ziegleder, G. et al. (2004) Thickening of molten white chocolates during storage. *Food Science and Technology*. 37 (6), 649–656.

Ziegler, G. & Hogg, R. (1999) Particle size reduction. *Industrial chocolate manufacture and use*. 115–136.

Chapter 2

Materials, Instrumentation, and Methods

2 Materials, Instrumentation, and Methods

This chapter covers high-level details of the materials and methods used throughout the work. Experimental details for the main experiments undertaken are described in their individual chapters. Citations to important pieces of software or techniques used in the work are given in the relevant chapters in which they are used.

2.1 Materials

All chemicals used throughout this work were purchased from Sigma-Aldrich and used "as received" without further purification. Of particular note is the synthetic monocrystalline diamond powder (1 micron), Sigma-Aldrich 483591-5G that was used as an internal standard for quantitative phase measurements. Deionised water was obtained from an Elga Purelab dispenser. Borosilicate glass capillaries (0.7 mm and 0.9 mm) for powder diffraction were purchased from Capillary Tube Supplies Ltd, Bodmin, UK. Crystal mounting loops were purchased from Hampton Research, USA.

2.2 Instrumentation

2.2.1 Powder X-ray Diffraction

All powder X-ray diffraction experiments were carried out on a Bruker D8 Advance diffractometer equipped with a copper tube and a Johansen monochromator, delivering K-alpha-1 radiation. The D8 was set in transmission capillary mode, with the incident radiation focussing on a LynxEye detector. The vast majority of samples were prepared in 0.7 mm borosilicate glass capillaries, this type of capillary exhibiting only a relatively small amorphous contribution to the observed X-ray diffraction data. 0.9 mm capillaries were occasionally used when sufficient sample was available and when stronger scattering, or easier sample loading, was required. The instrument was controlled by Bruker's XRD Commander software.

2.2.2 Single-crystal X-ray Diffraction

All single-crystal diffraction experiments, were carried out on a Rigaku Synergy diffractometer, equipped with a microfocus copper source, a HyPix 6000HE detector and an Oxford Cryosystems Cryostream device. Samples were mounted onto loops using Fomblin oil. The instrument was controlled by Rigaku's CrysalisPro software.

2.2.3 Nuclear Magnetic Resonance

NMR data were recorded using a Bruker Ascend 400 (400 MHz) spectrometer after dissolution of the samples in dimethyl sulphoxide (DMSO). 1H and 13C NMR spectra were both recorded at 400 MHz. Chemical shifts (δ) are quoted in parts per million (ppm) using the abbreviations described in Table 2. Integrations were done manually using MestreNova.

2.2.4 Optical Hot-stage Microscopy

All hot-stage microscopy experiments were carried out on a Mettler Toledo Hot Stage microscope.

2.2.5 Software

The table below lists the key software packages used in this work

Package	**Version**	**Use in this work**
DIFFRACplus EVA	14.0.0.0	Preliminary evaluation of PXRD diffraction data
CrysalisPro	V38.46	Evaluation of SX diffraction data
SHELX suite (XT, XL)	various	Solution and refinement of structures against SX diffraction data

TOPAS Academic	4.2	Rietveld refinement of structures from PXRD diffraction data, PXRD indexing, and quantitative phase analysis
Cambridge Structural Database System	5.24	Crystal structure search and retrieval; Mercury for structure visualisation and packing analysis; DASH for structure solution from PXRD data
MestReNova	14.3.0	Interpretation of NMR data
QuantumEspresso	6	DFT-D optimisation of periodic systems for crystal structure validation

2.3 Methods

As the majority of the work carried out involves analysis of diffraction data, the major elements of PXRD data analysis are outlined here.

2.3.1 Indexing

Powder indexing, using DASH and TOPAS, was typically carried out by selecting the first twenty low-angle peaks in a PXRD pattern, accurately recording their 2Θ positions and using these positions as input to the programs. Candidate unit cells were then evaluated firstly by comparison of the predicted reflection positions with observed PXRD peaks, and then confirmed by Pawley refinement.

2.3.2 Pawley Refinement

In Pawley refinement, reflection positions and reflection intensities were allowed to vary in a least square process that, coupled with a refinable peak shape, zero point and background parameterisation, attempt to fit the entire observed PXRD profile. These refinements were performed initially using DASH and then with TOPAS, the latter having the ability to fit the full observed PXRD data range more reliably than the former. At the end of a Pawley refinement, if the correct crystalline phase had been identified, there were no unaccounted-for peaks if the sample was a pure phase; unaccounted-for peaks may represent impurities, or an incorrect unit cell. Note that a Le Bail refinement performs essentially the same task as a Pawley refinement, though with an iterative (rather than a least-squares) approach. Pawley refinement was preferred in this work as it allows structure determination using DASH via the reflection intensities that are extracted during the Pawley refinement.

2.3.3 Space Group Determination

Pawley refinements were initially carried out in a space group that matched the indexed cell type, but which had no systematic absences. For example, a monoclinic cell would first be Pawley fitted in *P*2. Automatic space group determination was then usually performed using DASH, which determines the space group from sets of reflections for which there is no observed diffraction i.e. the systematic absences. For a correct space group determination, the fit to the observed data (as adjudged by R_{wp}) was essentially the same when fitted in a space group with no systematic absences (e.g. *P*2) and when fitted with the correct space group (e.g. *P*21/c).

2.3.4 Crystal Structure Determination

No structures were explicitly solved from PXRD data in this work; the interested reader is referred to one of the many articles on structure determination from powder data for full details of the basis of this technique and the use of DASH is strongly recommended for molecular materials.

2.3.5 Rietveld Refinement

Once a profile had been well fitted by the Pawley approach, most of the refinable parameters involved in the Pawley were fixed at their refined values e.g. zero point, peak shape, lattice parameters. Known atomic coordinates could then be refined against the observed data, again by a least-squares process in which the main variables were scale factor, background, atomic coordinates and isotropic temperature factors. As the quality of PXRD data was seldom good enough to permit the free refinement of large number of atomic positions (as is the case for the size of, and low symmetry of, the molecular materials of interest to this work) it was deemed good practice to use so-called rigid-body refinements in which the position, orientation and conformation of the molecule of interest is allowed to refine instead. Rigid-body methods are well implemented in TOPAS, which is strongly recommended for such refinements.

2.3.6 Quantitative Phase Analysis

A quantitative phase analysis (QPA) is essentially a multi-phase Rietveld refinement in which the scale factors of the two or more crystalline forms contributing to PXRD are allowed to vary. From these scale factors, the relative contributions of each phase can be calculated. In this work, QPA was performed using TOPAS and known single-crystal structures whenever possible, in order to maximise the accuracy of the approach. A sample QPA input file for TOPAS is shown in Chapter 4.8, Additional Information.

Chapter 3

Experimental Analysis of Powder Diffraction Data

Published as:

Spillman, M., Nicholls, D. & Shankland, K. (2020) Experimental analysis of powder diffraction data. Handbook on Big Data and Machine Learning in the Physical Sciences. 1-23

3 Experimental Analysis of Powder Diffraction Data

Foreword

During the course of this thesis, I was asked to participate in the writing of an invited chapter for a forthcoming book:

"Handbook on Big Data and Machine Learning in the Physical Sciences Vol. 2"

My participation was based on my experience with experimental powder diffraction, the various challenges I had encountered with powder indexing, structure determination, structure validation and quantitative phase analysis, and also my use of programs such as *TOPAS* for data analysis. My experience in collecting data at the Diamond Synchrotron for Pair Distribution Function analysis was also valuable. The aforementioned topics are the ones to which I contributed most strongly, though I was also involved in shaping the other sections, and the overall coverage of the chapter. It is not just a straightforward review; rather it is an overview that calls upon our collective experience in PXRD and aims to make recommendations for the use of PXRD that non-expert practitioners can call upon.

The book was published in May 2020 (ISBN: 978-981-120-444-9) as a two-volume set costing £860.

The chapter provides an excellent introduction to the capabilities of PXRD, with a strong focus on the general class of "molecular materials" that are relevant to my interests in this thesis.

Experimental Analysis of Powder Diffraction Data

Mark Spillman, Daniel Nicholls and Kenneth Shankland*

School of Chemistry, Food and Pharmacy,
University of Reading, Reading RG6 6AD, UK
**k.shankland@reading.ac.uk*

Powder diffraction is a core technique in materials analysis. Recent years have seen substantial developments in the methods available for the analysis of powder diffraction data, as rising computational power in particular has facilitated their implementation. This chapter aims to summarize some of these developments.

1 Introduction

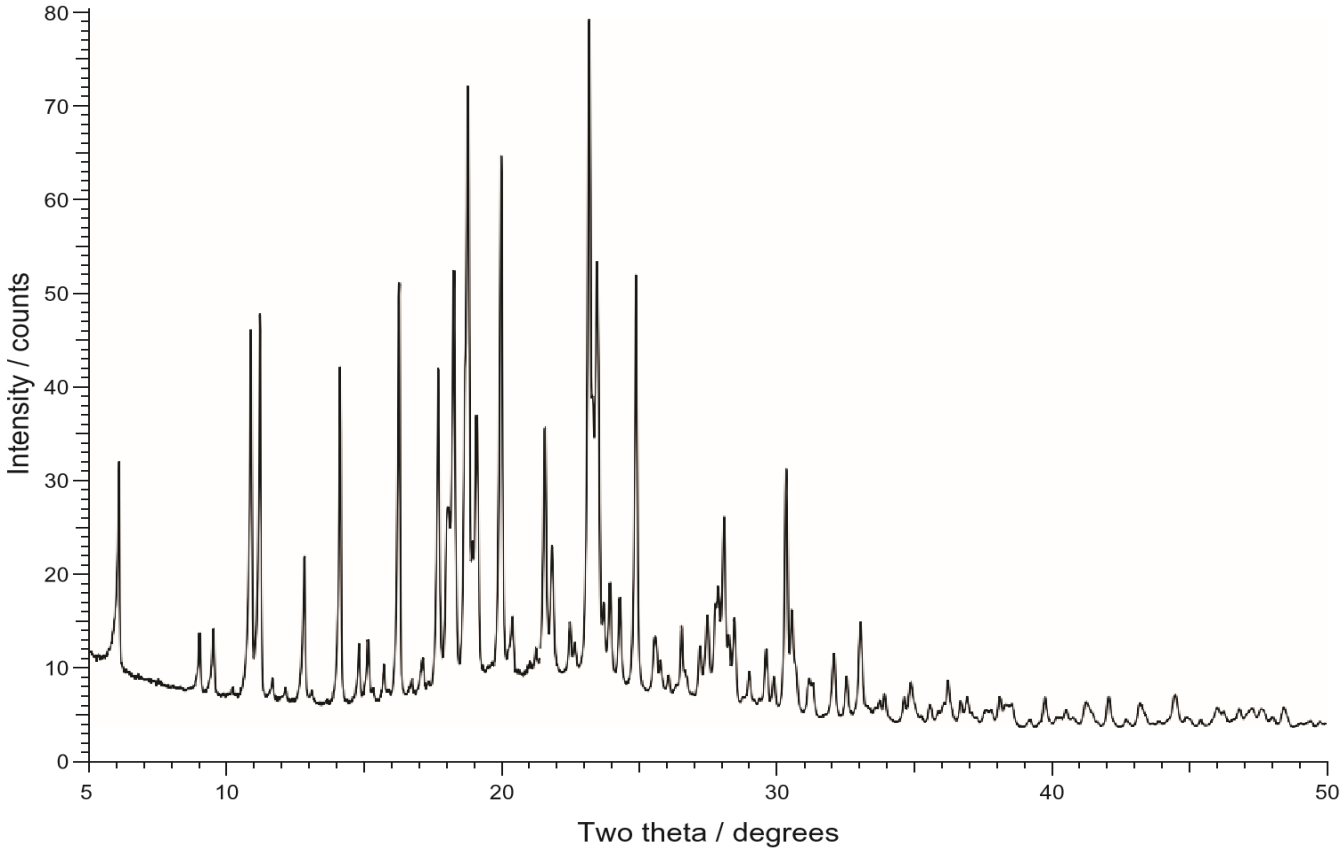
Diffraction of incident radiation is one of the most powerful methods for the analysis of crystalline materials in the solid state. The radiation most commonly takes the form of X-rays, less commonly neutrons and increasingly electrons. While single crystal diffraction is the "gold standard" for determining the three-dimensional structure of crystalline materials, the availability of a representative single crystal is a pre-requisite. There are many cases where obtaining a single crystal is difficult (e.g., materials formed as a result of mechanical grinding) and only polycrystalline materials are available as a result. In such cases, powder diffraction is a powerful alternative. While a powder diffraction pattern is intrinsically less "information rich" than its single-crystal equivalent, it can still yield a lot of valuable information about the crystallinity and structure of materials ranging from small molecules and inorganics, through to macromolecules and proteins. It finds particular utility in phase identification and quantification, because of the relative simplicity of sample preparation and presentation, and the relative simplicity of the instrumentation. Yet despite this simplicity, powder diffraction can yield much more information, including the complete crystal structure. Furthermore, as will be described later, powder diffraction experiments can even return useful information when the sample under study is not crystalline. There is no shortage of comprehensive articles and reviews covering many of the aspects of powders discussed here; our focus in this chapter (after a very brief introduction to instrumentation) is a method-based approach throughout, with a view to highlighting particular

steps in the analysis of powder data and acting as a starting point for those interested in exploring the topics further. Although diffraction from crystals can be observed using X-rays, neutrons and electrons, our focus here is on the use of X-rays, which are by far the most widely used in practice. Nevertheless, the approaches presented are largely applicable to the other radiations, where experimental methods permit. It is assumed that the reader is already familiar with the fundamental concepts and terminology of crystal structures, such as lattice parameters, space groups, reflections and fractional coordinates.

2 The X-ray Powder Diffraction Experiment

A powder X-ray diffraction (PXRD) pattern is collected by exposing a polycrystalline sample of the material under study to incident X-rays and detecting scattered X-rays. The radiation is typically, though not always, monochromatic and the modern detectors are typically zero-dimensional (so-called point detectors), one-dimensional (such as the Bruker LynxEye PSD, or the curved Mythen-2) or two-dimensional (so-called area detectors). The advantages of, disadvantages of, and reasons for choosing each type lie outside the scope of this chapter, but it is clear that as detector dimensionality increases, so does the volume of data collected and the complexity of reducing that data to an easily useable form. Many elements of data reduction are detector-specific, but if we consider the demanding, and increasingly common use of two-dimensional detectors, a number of software solutions are available for processing. *Fit2D*,[1] written in Fortran, has been a mainstay of data reduction for powder diffraction, but increasing experimental demands have led to the development of alternatives such as the Python-based *DIOPTAS*[2] and the Eclipse-based *DAWN*.[3] Important aspects of the latter, which is being developed at the Diamond Light Source, are: the ability to utilize multiple CPU cores to achieve on-fly-processing capability; the ability to support a wide range of data formats (including NeXus, f2d and ASCII); and the ability to incorporate new steps into the processing chain using Python programming.

Figure 1. A laboratory powder X-ray diffraction pattern collected from a molecular organic co-crystal.



Ultimately, irrespective of any proprietary or encoded format used, the reduced diffraction data take the familiar "intensity versus diffraction angle" form shown in Figure 1. We now turn our attention to some of the ways in which such data can be analyzed.

3 Pattern Matching

"Fingerprinting" of PXRD data is routinely performed automatically via user friendly GUI driven software such as *EVA*,[4] *QualX2*[5] *and Match!*.[6] These programs enable sophisticated peak-location-based search queries of various crystallographic databases (see below) to rapidly obtain likely candidates for the crystalline phases contributing to a PXRD pattern of interest. Searches can easily be constrained to ensure that the proposed candidates reflect the known properties of the sample (e.g., known elemental composition can constrain searches to include or exclude structures that contain certain elements).

Crystallographic databases, such as the Cambridge Structural Database (CSD),[7] the Powder Diffraction File,[8] the Inorganic Crystal Structure Database,[9] the Crystal Data for Metals Database

(CRYSTMET[10]), the Protein Data Bank (PDB[11]) or the Crystallography Open Database (COD[12]), contain hundreds of thousands of published crystal structures, and are constantly expanding; most newly determined crystal structures are deposited in the appropriate database. The availability of these databases provides a large body of crystallographic information enabling the previously mentioned pattern matching tasks to be carried out, as well as other detailed statistical studies of various properties of interest. One example relevant to crystal structure determination (see Section 6) is the modal torsion angle restraints provided by *MOGUL*[13] that can be employed by *DASH*.[14] Their use has been demonstrated to increase the success rate of global optimization searches by limiting the torsional search space to be explored. These restraints are derived from structure fragment searches of the CSD, and reflect the statistically most likely conformations a particular fragment will adopt.

The advent of high-throughput crystallography, typically seen in crystallization trials and polymorph screening, has necessitated the development of new tools to enable researchers to deal with the inevitably large data volumes generated in a short space of time. Methods of identifying and clustering similar datasets have been developed which allow users to rapidly identify new and unusual structures or polymorphs. For example, a 96-well plate of polycrystalline samples can now be rapidly scanned and the resultant data automatically analyzed to see how many "unique" patterns are present and if any of them are novel. One commonly applied technique in this context is principal component analysis (PCA), which decomposes an input data matrix into a given number of mutually orthogonal principal components (PCs) which are calculated such that the PCs maximally explain the variance in the observed data. PCA-decomposed PXRD patterns can be easily visualized and automatically clustered based on their relative positions within the PCA space. This and other related types of data decomposition allow PXRD patterns to be easily compared to a library of existing patterns from both current and previous experiments. Programs such as *PolySNAP*[15–18] and *RootProf* [19] have been developed to enable this kind of analysis, and feature sophisticated routines for data preparation, decomposition, visualization and clustering. These programs are not limited to working with PXRD data; unidimensional data from a variety of other sources such as FT-IR, Raman, X-ray absorption and NMR spectroscopy as well as DSC experiments are also amenable to this type of analysis. In addition, an extension to the *PolySNAP* software, entitled *dSNAP*[20] has been developed to enable similar classification and clustering exploration of structural fragments obtained from searches of the CSD. Such pattern matching and

classification programs are now an essential component of any pre-formulation pharmaceutical laboratory with powder diffraction facilities.

4 Indexing

In general, analytical work that goes beyond simple fingerprinting of a PXRD dataset requires knowledge of the lattice parameters that define the unit cell of the crystal. Without this information, many of the other techniques described in this chapter are rendered extremely challenging or impossible. The simplicity of the underlying equations that govern the relationship between the observed peak positions and the lattice parameters belies the complicated task of indexing, where neither the lattice parameters nor the *hkls* of the observed peaks are known *a priori*. Computer programs developed to tackle this problem generally return a list of possible unit cells, which are ranked according to a figure of merit (most commonly M_{20}[21] and F_N[22]) which provide a measure of the quality of a putative solution in terms of the difference between the calculated and observed peak positions, with a higher weighting given to unit cells of higher symmetry. An extremely useful "sense check" on the proposed volume of the indexed phase can be performed when the chemical formula of the material under study is known. The volume occupied by one molecule or formula unit of that material can be rapidly and quite accurately calculated using tabulated atomic volume contributions derived by Hofmann.[23] Dividing the volume of the indexed cell by the molecular volume should yield a value that lies close to a crystallographically sensible whole number (such as 1, 2, 4, etc.). This has been implemented in the *EXPO* software,[24,25] where it is used to guide the user on sensible unit cell choices. Values that deviate significantly from those expected need not necessarily indicate incorrect indexing — they may be indicative of a solvated crystal, or a crystalline structure with void space.

Where good quality diffraction data are available, the "classic" indexing programs, *DICVOL*,[26,27] *ITO*[28] and *TREOR*,[29] are usually able to obtain accurate results extremely rapidly from the 2θ values of the first 20 observed diffraction peaks. Convenient GUI-driven interfaces to these programs are included in a number of packages. However, there are several frequently encountered phenomena that can lead to difficulties.

One such problem (a dominant zone) arises when one unit cell edge is short relative to the other two edges. For example, the first 20 peaks of a PXRD pattern collected from a sample with a monoclinic unit cell of dimensions a = 16.0 Å, b = 4.40 Å, c = 17.50 Å, β = 95.0° all have k = 0 and thus provide no information on the length of the b-axis. Most indexing programs automatically test for dominant zone problems which ensures that a greater number of observed peaks are input such that sufficient information is available for lattice parameter determination — in the above case, peaks where k = 0 must be included. Even so, it is quite likely that few $k \neq 0$ reflections will manifest as clear, non-overlapping peaks and so whist estimates of a, c and β will likely be good, that of b will be poor. *TOPAS*[30] has useful functionality to deal with such scenarios. If a Pawley fit using the lattice parameters returned by indexing is poor, then the `continue_after_convergence` directive can be used to restart the Pawley refinement after convergence, using a new b lattice parameter, chosen at random from a specified range using the `val_on_continue` function. For example,

```
b @ 4.3 val_on_continue = Rand(4.2,4.5);

continue_after_convergence
```

allows the b lattice parameter to refine from different random start points in the range 4.2–4.5 Å, in the expectation that one or more of those start points will be sufficiently close to the true lattice parameter for the Pawley refinement to converge to a satisfactory fit.

The accidental overlap of reflections also presents particular problems for indexing. Multiple reflections may manifest as a single observed peak, with no way to identify the number of contributing reflections or their respective intensities. Exclusion of peaks from the indexing attempt can lead to overestimation of the symmetry of the predicted cell or simply failure to obtain a solution. This issue can be dealt with to some extent by indexing software; the SVD-Index algorithm[31] as implemented in the *TOPAS* software was developed to be robust against this effect, and shown to be highly successful against a wide range of synthetic data where one in three peaks was randomly removed from the input. The more recently developed Lp-Search[32] approach, also available through *TOPAS*, has shown remarkable resilience to overlapped or missing peaks, and was demonstrated to be able to rapidly index diffraction patterns of low-symmetry materials with the 30 lowest angle reflections excluded from the attempt. Where laboratory diffraction data are

insufficiently well resolved to be indexed, the higher angular resolution attainable at synchrotron facilities can also help mitigate such issues.

If the absence of input reflections is problematic, the presence of additional peaks in a PXRD pattern originating from crystalline impurities in the sample is potentially much worse. In such patterns, a single unit cell no longer accounts for all the observed reflections. If initial indexing attempts fail, PXRD patterns should be inspected for peaks that appear different from the majority, e.g., two or three very sharp peaks in the presence of a plethora of broader peaks may indicate the presence of a very crystalline impurity, whose peak positions can be safely eliminated from the indexing process. Similarly, the pattern can be easily checked for the presence of well-characterized, unreacted starting materials that may be present in low concentrations. Access to crystallographic databases, such as the Cambridge Structural Database or the ICDD Powder Diffraction File, can be invaluable in this respect.

Where impurity peaks are not readily identifiable, indexing may still be accomplished using programs such as *X-Cell*,[33] *SVD-Index*,[31] *McMaille*[34] and *N-TREOR*,[35] which have all been developed to enable indexing in the presence of impurities. The original *DICVOL* algorithm has also been updated[27] to improve performance in such situations. These programs make use of figures of merit that penalize solutions which do not index all of the observed input reflections, allowing proposed unit cells to be more effectively ranked. In practice, it is recommended that problematic PXRD datasets are indexed using multiple programs. Solutions should be checked carefully by performing a Pawley or Le Bail refinement of the indexed phase together with scale-factor-only Rietveld refinements of the identified impurity phases to ensure that all observed reflections are accounted for.

5 Space Group Determination

In principle, all crystal structure determination could take place in space group *P*1, and the correct symmetry assigned after the structure has been found. This is indeed the strategy employed when using charge flipping. However, in most other cases, the inclusion of symmetry information dramatically reduces the complexity and time taken for various calculations and therefore forms an important step in structural analysis. The assignment of a space group (or more accurately, an

extinction symbol[a] from systematic absences in observed diffraction data) is very well established. For example, absences in the class 0*k*0 when *k* is odd are indicative of a 2_1 screw axis along *b*. While identifying systematic absences from three-dimensional single-crystal diffraction data is generally straightforward, the one-dimensional PXRD case can be much more difficult. For example, in order to confidently assign a 2_1 screw axis, one would ideally look for the absence of the 010, 030, 050, etc. reflections. There is a good chance that the 010 reflection will manifest at a point in the pattern (at low 2θ) where there are few other reflections and so its presence or absence can be easily determined. On the other hand, there is a good chance that the 030, 050 and higher order reflections will manifest at points in the pattern (at higher 2θ) where there are many other reflections and thus be subject to accidental overlap, making it extremely challenging to say if they are present or absent.

[a] An extinction symbol summarizes the conditions that account for the observed reflections and absences, and multiple space groups may be associated with a given extinction symbol. For example, extinction symbol *P* 1 2_1 1 is consistent with space groups *P* 1 2_1 1 and P 1 2_1/*m* 1; unambiguous space group assignment must be determined through the use of additional information.

As such, statistical approaches have been developed[36–38] to determine the most probable extinction symbol given intensities extracted by a Pawley or Le Bail-type refinement, weighted by the expected reliability of their estimate. These approaches, implemented in the *ExtSym*[37] and *EXPO* programs, have been shown to be effective and are strongly recommended for routine use. Other programs such as *TOPAS* and *X-Cell* include space group determination steps directly into their indexing procedure and hence can also be used for extinction symbol assignment.

Where multiple space groups are consistent with the extinction symbol determined, the final choice of a space group is often informed by its frequency of occurrence in relevant crystallographic databases. For example, if diffraction from an organic structure is assigned the extinction symbol $P\ 1\ 2_1\ 1$, then knowledge derived from the CSD that $P2_1$ is observed approximately 230 times more often than $P2_1/m$ suggests that it is reasonable to first attempt further analysis in $P2_1$. Chirality and the presence or absence of a molecular center of symmetry are other relevant factors, e.g., an enantiopure chiral molecular crystal structure cannot be observed in a centrosymmetric space group such as $P2_1/m$.

As was the case with indexing, the space group assignment should be verified by a Pawley or Le Bail-type refinement over the full data range.

6 Crystal Structure Determination

The ability to solve crystal structures directly from PXRD data has advanced significantly in the last two decades. While PXRD instrumentation has undoubtedly advanced in that period, the advances are primarily attributable to algorithmic and computational developments that seek to compensate for the loss of accurate reflection intensity information, which is an inevitable consequence of the collapse of the three dimensions of reciprocal space onto the one dimension of a PXRD pattern. The main methodologies that are employed are direct-space methods, modified direct methods, and dual-space methods that employ elements of both.

6.1. Modified direct methods

Conventional direct methods, as developed for solving the phase problem in single-crystal crystallography, result in the generation of trial sets of phased structure factors that are used to generate

electron density maps using a Fourier transform. Such maps can then be interpreted for chemical sense and any partial structure thus obtained can be recycled to improve the phasing accuracy and hence iteratively locate the remainder of the structure using difference Fourier maps. Direct methods have been very successful in the single-crystal domain because in such experiments, one typically collects thousands of accurate reflection intensities to sub-ångström resolution. In the case of PXRD, one typically extracts only hundreds of reflection intensities from a pattern, the accuracy of which are compromised by accidental reflection overlap in the one-dimensional data. In addition, in the absence of any heavy elements in the material under study, PXRD patterns rarely extend to sub-ångström resolution. As such, conventional direct methods alone are of very limited value in crystal structure determination from powder diffraction data (SDPD) with successes confined largely to data collected on high-resolution synchrotron instruments.[39]

Fortunately, it is possible to incorporate partial structure information obtained from phasing back into the intensity extraction stage in order to improve the accuracy of the extracted reflection intensities. For example, consider a single observed peak in the PXRD pattern containing intensity contributions from two closely overlapping reflections. In the initial intensity extraction, only the summed intensity for these two reflections can be accurately extracted for use. In the initial direct methods solution attempt, the peak intensity is equally partitioned between the contributing reflections.

On the basis that the electron density map will have some element of truth, any partial structure identified is used to calculate a more probable and realistic assignment of the reflection intensities. By applying this approach across the pattern, the modified partitioning can be used as the starting point for another extraction and the intensities fed into another, hopefully improved, direct methods attempt.[40] Even so, the limitations of resolution still apply and the low-resolution Fourier maps can be extremely difficult to interpret. The development of three new approaches (wLSQ, resolution bias modification [RBM] and COVMAP) to directly addressing this issue has yielded impressive results.[41,42] As implemented in the *EXPO* software package,[41] they are applicable to a broad range of problems that include inorganic, organic and organometallic structures. Importantly, the focus on addressing the resolution issue means that the program is well adapted to dealing with diffraction data collected on standard laboratory instrumentation. It is worth noting that *EXPO* maintains the multi-solution approach characteristic of direct methods: it is possible to take multiple candidate solutions and "develop" them using RBM and COVMAP. This can be relatively time-consuming on a single processor (i.e., several hours required to reach a solution) and so a fast CPU is recommended.

Other direct methods approaches that show considerable promise are available. For example, the direct methods modulus sum function, *S*, is well suited to powder diffraction data and as implemented in the *XLENS* program of Rius[43] has been used to solve several complex structures.

6.2. Direct space approaches

One of the most important areas of development for SDPD has been the use of methods that operate in direct space only. In these methods, an initial model of a crystal structure is proposed and then adjusted in such a way as to maximize the agreement between observed and calculated reflection intensities. For high symmetry crystal systems, these structural variables to be adjusted may be individual atomic coordinates, but in lower symmetry systems (such as those commonly encountered when solving molecular crystal structures) it is more common to treat the structural model as a collection of connected atoms. For example, a molecule can be described in terms of its atom types, bond lengths and bond angles (both of which can be easily obtained from tabulated values or from known crystal structures) and torsion angles (many of which are known, though generally those that are free to rotate and determine the overall molecular conformation are unknown).

By assigning random values to unknown torsion angles, and random values to the position and orientation of the molecule within the unit cell, trial crystal structures can be easily generated using a relatively small number of structural variables.[44,45]

The problem posed by this approach equates to one of finding the lowest point (global minimum) on an agreement factor (such as R_{wp}) hypersurface, where the dimensionality of the hypersurface is equal to the number of structural variables. Given that the hypersurface is not smooth and contains a great many local minima (Figure 2), global optimization methods (including simulated annealing, genetic algorithms, evolutionary algorithms, random searches and grid searches) are normally employed in order to obtain a solution.

Taking the widely-used simulated annealing (SA) approach as an exemplar, it is important to realize that no single simulated annealing run is guaranteed to find the global minimum of the hypersurface, given a finite timescale. As such, it is conventional to run multiple SA runs in order to improve the chances of locating the global minimum. As each run is independent of any other and each run commences with a different set of random number seeds, this is a problem that is well-suited to

execution using a coarse-grained parallel approach in order to improve throughput. For example, the *DASH* structure determination package[14] can write out batch

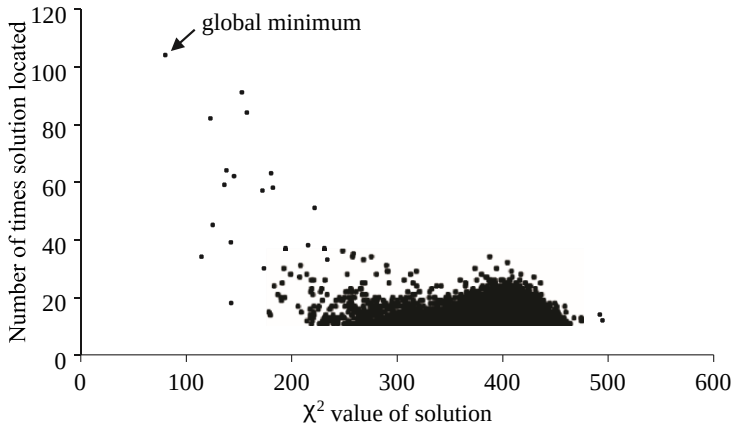


Figure 2. A frequency distribution of the χ^2 values on the χ^2 hypersurface of a molecular compound solved from powder diffraction data. Each point on the plot shows the frequency with which a particular χ^2 value was located by a local minimization algorithm, e.g., the global minimum at χ^2 = 80.12 was located by 104 minimizations. Only values with a frequency ≥ 11 are shown for clarity. There are 22,924 discrete χ^2 values totalling 302,080 hits and no binning has been applied.

files that can then be executed on either (a) the multiple cores of a modern CPU,[46] (b) a distributed computing network, such as *GridMP*[47] or (c) the Amazon Elastic Compute cloud.[48] Other programs that implement a similar approach include *EXPO* (which has a simulated annealing mode that utilizes Open MPI) and *FOX.Grid*,[49] which can run on multiple local CPU cores of a single personal computer, or can be distributed to multiple machines connected by a local area network. The ability to distribute calculations is not simply a matter of obtaining a result faster; more importantly, it allows one to compensate for the fall off in success rate in finding the global minimum as the complexity of structures under investigation increases. For example, 100 SA runs for a simple structure with only seven structural variables (three positional, three orientational and one torsional) might return a 100% success rate, i.e., every SA run finds the global minimum equating to the correct crystal structure. On the other hand, a problem involving 30 structural variables may return only a 0.1% success rate, meaning that 1000 SA runs are required in order to have a reasonable chance of finding the global minimum once. Faced with such odds, a parallel computing strategy becomes essential.

6.3. Other approaches

While direct methods and direct space methods represent the main SDPD approaches, there are also a number of other approaches that can be considered. These include the use of the Patterson-function,[43]

maximum entropy methods[50] and the increasingly used charge flipping method,[51] initially developed for single-crystal structure determination but now adapted for powder diffraction.[52] This method operates in real and reciprocal space. Briefly, a set of observed structure factor magnitudes extracted from a powder pattern is assigned random starting phases and an electron density map calculated. Any density below a certain user-definable threshold is "flipped" (i.e., has its sign changed) and a Fourier transform of the modified map returns an updated set of phases that are assigned to the observed reflection magnitudes. A new map is then generated and the whole process repeated until convergence (based on agreement between observed and calculated structure factor magnitudes) is achieved. At that point, the map is then subjected to interpretation and the space group symmetry derived from the structure. It is now the structure solution method of choice for many single-crystal crystallographers; for powder diffraction, it tends to perform best with high resolution synchrotron data as unsurprisingly, many of the issues related to low resolution that affect direct methods also affect charge flipping.

Since its initial publication, a number of modifications to the original algorithm have been proposed aimed at improving the performance of the algorithm on powder diffraction data. For example, charge flipping has been used in combination with histogram matching[53] as well as the direct methods tangent formula.[54] It is worth noting that several very challenging structures have since been solved by these modified approaches.

6.4. Massive structures: proteins and powders

Structure determination of proteins was, until relatively recently, considered the exclusive domain of single-crystal crystallographers. It was only with the realization that some polycrystalline proteins exhibited extremely sharp diffraction lines (consistent with *ca.* 1 μm, homogenous, virtually defect free crystals) that serious investigation of the potential of PXRD in protein structure characterization began. While a few studies have collected PXRD data in the laboratory, the vast majority of studies utilize synchrotron instrumentation, where the high instrumental resolution can fully exploit the intrinsically sharp diffraction patterns. Restrained Rietveld refinement methods can be used to refine structures containing more than 800 atoms against such data[55] and molecular replacement[56] has been used to solve protein structures such as the second SH3 domain of ponsin,[57] which contains 67 amino acid residues. In studying proteins using powders, it has been necessary to utilize both existing single-crystal codes such as *MOLREP*,[58] and modified Rietveld codes such as *GSAS-II*.[59] While existing studies may be

viewed as *tours de force*, it is unlikely that PXRD will play anything other than a niche role in the study of proteins in the longer term. Advances in single-crystal growth methodology, automated crystal screening, beamline design, detector sensitivity and the introduction of X-ray free electron lasers mean that tiny single crystals and soon even single *molecules* can ultimately yield sufficient diffraction information to allow structure solution to proceed.

7 The Rietveld Method

Rietveld's method for the refinement of crystal structures against neutron powder diffraction data was published nearly 50 years ago.[60] While it is long-established as the *de-facto* standard for structure refinement using powder diffraction data, advancements in the application of the basic technique are still regularly being made. At its heart lies a least-squares minimization of the differences between the measured intensity and the calculated intensity at all points across a measured powder diffraction pattern. The calculated intensities are derived from the crystal structure(s) which contribute to the observed diffraction pattern. Atomic coordinates can be refined, modeling of the peak shape gives information about crystallite size and strain, refinement of peak positions gives information on unit cell dimensions and modeling of the background can give insights into any amorphous material present. Furthermore, when more than one crystalline phase is contributing, the scale factors for each phase can be used to quantify the relative amount (by mass) of each phase present in the sample. Refinement can be performed against one or more diffraction patterns, collected on one or more radiation sources, allowing one to combine the merits of (for example) X-ray and neutron powder patterns into a single refinement.

Advances in computing have been pivotal in enabling the Rietveld method to develop. By way of example, faster CPUs enabled more complex modeling of peak shapes, e.g., removing the need for the pseudo-Voigt approximation of the computationally expensive Voigt function. The ability to hold multiple (increasingly large) powder patterns in memory also greatly speeds up each iteration of the least-squares process. Numerous Rietveld codes are available; some of the more commonly used ones are listed in Table 1. Many of these have additional functionality; for example, some (such as *FOX* and *TOPAS*) also have the ability to solve structures.

Improved computing facilities have also facilitated the development of so-called parametric refinement. It is well known that collecting multiple powder diffraction datasets of the same material under different conditions (e.g., temperature, pressure) can yield a greater understanding of the structure being assessed, e.g., identifying any phase transitions it might undergo. Differential expansion or contraction of the lattice parameters under different conditions can also help to alleviate issues associated with the accidental overlap of reflections.[66] Furthermore, it is possible to derive important quantities such as activation energies and rate mechanisms from studying the evolution of structures as a function of such external variables. The classic approach to analyzing multiple data sets involves sequential refinement of parameters against each pattern individually, then deriving some function from studying the parameter evolution. It is preferable, however, to describe a parameter of interest as a function that can then be refined against all the data simultaneously. Utilizing physical formulas may allow for description of multiple parameters, under different conditions can also help to alleviate issues associated with the accidental overlap of reflections.[66] Trends in parameters which are too slight to observe when considered individually may well be observable with parametric refinement and treating the entire dataset as a whole helps avoid false minima. Methods for implementing parametric refinement are well described in the literature, such as its use with *TOPAS*, both manually[67] and through automated software systems that interact with the *TOPAS* kernel.[68]

Table 1. Some commonly used Rietveld refinement programs.

Program	Licensing	Reference
BGMN	Open source	61
FOX	Open source	49
FullProf	Academic	62
GSAS-II	Open source	59
Jade 9	Commercial	63
JANA2006	Academic	64
MAUD	Academic	65
TOPAS	Academic & Commercial	30

7.1. Structure verification

The question “How do I know that a crystal structure, solved and refined from powder data, is correct” is one frequently asked by those more familiar with structures solved from single-crystal diffraction

data. It is not sufficient to reply "the structure describes the observed data" because the data are (as mentioned several times already) of limited quality and resolution. The correct response is "the structure is chemically and crystallographically sensible, and it describes the observed data." Many tools are available to help one check the validity of a structure; the International Union of Crystallography's *checkCIF*[69,70] service (which calls upon the well-established *PLATON*[71] program) is the most widely used, performing a comprehensive range of checks not only on the structure but also on the data (if available) and reporting on errors and potential issues with both. Issues can often be minimized, particularly with molecular crystal structures obtained by global optimization methods, if good quality input models are used. Such models can be verified before use in a structure solution, using the *MOGUL*[13] program, which checks the geometry of the model against the hundreds of thousands of structures in the Cambridge Structural Database. If rigid-body type Rietveld refinement is used, this checked geometry (apart from conformational changes around rotatable bonds) is preserved throughout the refinement.

Periodic DFT-D calculations can now be performed on relatively modest computer hardware, in relatively short timeframes, and as such it is both possible and advisable to subject a refined crystal structure to a full energy minimization. The root-mean-square Cartesian displacement of the non-H atoms upon energy minimization should be small, typically around 0.1 Å, for a correct crystal structure.[72,73] This approach also has the advantage that it is able to locate hydrogen atoms with great accuracy, compensating for PXRD's weakness in this aspect.

7.2. Quantitative phase analysis

Quantitative phase analysis (QPA) is perhaps one of the biggest industrial application areas for PXRD data, being widely used in the analysis of geological materials. As mentioned earlier, when the crystal structures of the phases contributing to a measured PXRD pattern are known, Rietveld refinement of the scale factors for each phase against the data allow the various relative phase fractions (by mass) of the material to be determined. A recently developed method now also allows accurate QPA to be performed *without* knowledge of the contributing crystal structures.[74] Removal of the need for crystal structures particularly meets the requirements of pharmaceutical scientists who do not always have crystal structures, but still desire compositional information in the early stages of pharmaceutical materials development in order to guide their development processes.

Another increasingly important role for QPA is in the determination of *absolute* phase abundances as well as amorphous content, which is of particular import for the pharmaceutical and food industries. There are a variety of approaches but Rietveld-based analysis on samples spiked with a known amount of some internal standard (such as diamond powder) provides a straightforward method for determination of absolute phase abundances, with total amorphous form quantification obtained by difference. More sophisticated techniques such as PONKCS[75] allow individual contributing amorphous phases to be quantified.

8 Crystal Structure Prediction

It has long been a goal of the scientific community to be able to predict crystal structures. Superficially, the problem appears straightforward: construct a model of the material of interest and pack it regularly in such a way as to create a periodic structures whose potential energy can be calculated using well-known methods. The lowest energy (most thermodynamically stable) structure provides a prediction of a structure that is likely to be observed experimentally. In practice, crystal structure prediction (CSP) is far from straightforward, and faces some formidable obstacles. For example, the assumption that an observed crystal structure will be the most thermodynamically stable is manifestly wrong (the phenomenon of polymorphism is testament to that); the crystal energy landscape which is explored during CSP usually has a plurality of structures that are thermodynamically feasible and it is not obvious which (if any) will be observed; kinetic factors of crystallization are ignored; and calculations are performed nominally at 0 kelvin. Furthermore, the computational requirements are extremely demanding — high levels of theory are required to be able to correctly rank structures that may differ only slightly in energy.

Despite these challenges, significant successes have been achieved by many groups. In the field of molecular materials, the series of six CSP "blind tests" run by the Cambridge Crystallographic Data Centre have shown how the field has progressed in recent years. In the sixth blind test[76,77] in 2015, the five test systems (which included a polymorphic system, a salt and a co-crystal) were all predicted, with the exception of one $Z' = 2$ polymorph of one of the molecules. Interestingly, the total number of CPU hours used by the 21 teams was estimated to be >40 million, though that figure is heavily weighted by one team who used 30 million CPU hours alone. The most successful team used ~750,000 CPU hours in total. Dispersion-corrected density functional theory (DFT-D) calculations play an increasingly

important role in such calculations, but are currently too costly for use throughout the CSP process. The use of molecule-specific force fields,[78] derived from DFT-D calculations, has proven to be a significant advance, enabling good accuracy at the structure generation stage of CSP, at a greatly reduced computational cost relative to DFT-D.

As mentioned earlier, CSP frequently generates many crystal structures that lie close to the lowest energy structure in the crystal energy landscape. It is here that PXRD can play a valuable role. Given any hypothetical crystal structure, it is trivial to calculate a PXRD diffraction pattern, and so potential structures generated by CSP can be rapidly screened against observed PXRD data in order to match prediction to experiment. Some care must be taken when comparing data generated from a structure calculated at 0 kelvin with data measured (typically) at room temperature, but lattice parameters can always be refined using the Rietveld method. This approach is particularly valuable when the PXRD data are of low quality and not amenable to indexing — CSP provides yet another route to SDPD. If the observed PXRD data are sufficiently good, a unit cell and space group derived from the data can act as a very valuable constraint in the CSP process.

9 Pair Distribution Functions

Thus far, all the methods discussed (apart from CSP) have exploited the Bragg scattering that reflects the long-range, periodic ordering of atoms within a solid; no account has been taken of the diffuse scattering that reflects local deviations from the average structure. For nanocrystalline materials, where there is short- range order only, conventional crystallographic methods lose their power and a different approach is needed. One such approach is to describe the total scattering[79,80] (i.e., Bragg plus diffuse) in terms of a pair distribution function, $G(r)$, which is obtained via Fourier transform of diffraction data that has been modified to correct for factors such as background, Compton scattering and experimental geometry (Figure 3). The experimental requirements for PXRD data intended for use in PDF calculations are stringent: high-flux, high-energy, monochromatic synchrotron radiation is desirable (though Mo or Ag radiation can be used in the lab, albeit with longer collection times) as are large area detectors to facilitate rapid data collection. The maximum value of momentum transfer $Q = 4\pi/\lambda \sin\theta$ for the data should be as large as possible (at least 20 $^{\circ}A^{-1}$ to reduce termination ripples in the PDF and to maximize real space resolution in the PDF) and the instrument should have a low and stable background contribution. Initial analysis and transformation of the PXRD data requires the use of

specialized software such as *Gudrun*[81] or *PDFgetX3*[82]; further analysis can then be performed with a number of programs including *PDFgui*[83] and *TOPAS*. Of particular interest to those interested in molecular materials is the ability of PDF analysis to discriminate between truly amorphous materials and nanocrystalline materials; both display the typical "X-ray amorphous hump" when probed by conventional PXRD approaches, but may have significantly different physicochemical properties.[84] It is also possible to solve relatively simple molecular organic crystal structures from PDF data[85] using approaches

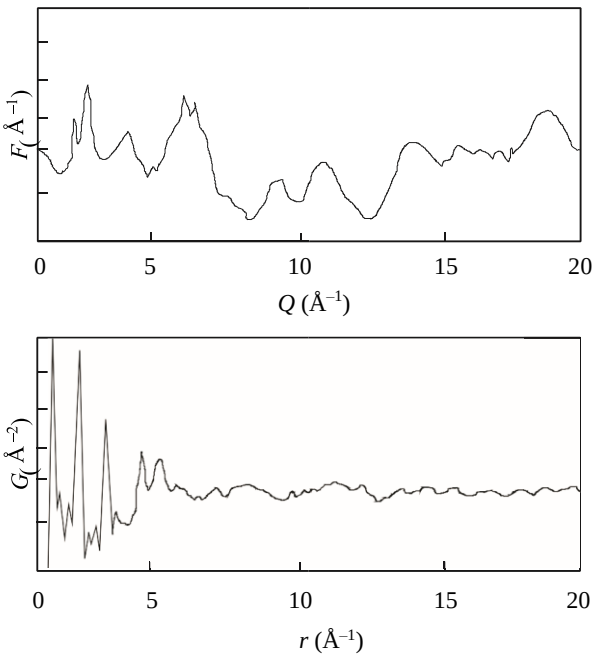


Figure 3. The total scattering reduced structure function, $F(Q)$, and the total scattering pair distribution function, $G(r)$, for data collected at a synchrotron (λ = 0.137 ˚A) from a molecular organic material.

analogous to those outlined in Section 6.2, allowing access to structures that are poorly crystalline or nanocrystalline.

10 References

1. Hammersley, A. P. *FIT2D V9.129 Reference Manual V3.1, ESRF Internal Report*, ESRF98HA01T, European Synchrotron Radiation Facility (1998).
2. Prescher, C. & Prakapenka, V. B. DIOPTAS: a program for reduction of twodimensional X-ray diffraction data and data exploration. *High Press. Res.* **35**, 223–230 (2015).
3. Basham, M. *et al*. Data Analysis WorkbeNch (*DAWN*). *J. Synchrotron Rad.* **22**, 853–858 (2015).
4. DIFFRAC.SUITE EVA. Available at: https://www.bruker.com/products/x-raydiffraction-and-elemental-analysis/x-ray-diffraction/xrd-software/eva/overview.html. (Accessed: 27 February 2018).
5. Altomare, A. *et al.* QUALX2.0: a qualitative phase analysis software using the freely available database POW COD. *J. Appl. Crystallogr.* **48**, 598– 603 (2015).
6. Match! Available at: http://www.crystalimpact.com/match/functions.htm (Accessed: 27 February 2018).
7. Groom, C. R., Bruno, I. J., Lightfoot, M. P. & Ward, S. C. The Cambridge structuraldatabase. *Acta Crystallogr. Sect. B* **72**, 171–179 (2016).
8. Fawcett, T. G., Kabekkodu, S. N., Blanton, J. R. & Blanton, T. N. Chemical analysisby diffraction: the Powder Diffraction File^{TM}. *Powder Diffr.* **32**, 63–71 (2017).
9. Hellenbrandt, M. The inorganic crystal structure database (ICSD) — Present andfuture. *Crystallogr. Rev.* **10**, 17–22 (2004).
10. White, P. S., Rodgers, J. R. & Le Page, Y. CRYSTMET: A database of the structuresand powder patterns of metals and intermetallics. *Acta Crystallogr. Sect. B* **58**, 343–348 (2002).
11. Burley, S. K. *et al.* Protein Data Bank (PDB): the single global macromolecular structure archive. In *Protein Crystallography Methods and Protocols*, eds. Wlodawer, A., Dauter, Z. & Jaskolski, M., Vol. 1607 (Springer, 2017), pp. 627–641.
12. Grazulis, S. *et al.* Crystallography Open Database (COD): an open-access collection of crystal structures and platform for world-wide collaboration. *Nucleic Acids Res.* **40**, D420–D427 (2012).
13. Bruno, I. J. *et al.* Retrieval of crystallographically-derived molecular geometry information. *J. Chem. Inf. Comput. Sci.* **44**, 2133–2144 (2004).
14. David, W. I. F. *et al.* DASH: a program for crystal structure determination from powder diffraction data. *J. Appl. Crystallogr.* **39**, 910–915 (2006).
15. Barr, G., Dong, W. & Gilmore, C. J. High-throughput powder diffraction. IV. Clustervalidation using silhouettes and fuzzy clustering. *J. Appl. Crystallogr.* **37**, 874–882 (2004).
16. Barr, G., Dong, W., Gilmore, C. & Faber, J. High-throughput powder diffraction. III.The application of full-profile pattern matching and multivariate statistical analysis to round-robin-type data sets. *J. Appl. Crystallogr.* **37**, 635–642 (2004).
17. Barr, G., Dong, W. & Gilmore, C. J. PolySNAP: a computer program for analysinghigh-throughput powder diffraction data. *J. Appl. Crystallogr.* **37**, 658–664 (2004).
18. Barr, G., Dong, W. & Gilmore, C. J. High-throughput powder diffraction. II. Applications of clustering methods and multivariate data analysis. *J. Appl. Crystallogr.* **37**, 243–252 (2004).

19. Caliandro, R. & Belviso, D. B. RootProf: software for multivariate analysis ofunidimensional profiles. *J. Appl. Crystallogr.* **47**, 1087–1096 (2014).
20. Barr, G., Dong, W., Gilmore, C. J., Parkin, A. & Wilson, C. C. dSNAP: a computerprogram to cluster and classify Cambridge Structural Database searches. *J. Appl. Crystallogr.* **38**, 833–841 (2005).
21. de Wolff, P. M. The definition of the indexing figure of merit M 20. *J. Appl. Crystallogr.* **5**, 243–243 (1972).
22. Smith, G. S. & Snyder, R. L. FN: a criterion for rating powder diffraction patternsand evaluating the reliability of powder-pattern indexing. *J. Appl. Crystallogr.* **12**, 60–65 (1979).
23. Hofmann, D. W. M. Fast estimation of crystal densities. *Acta Crystallogr. Sect. B* **58**, 489–493 (2002).
24. Altomare, A. *et al.* New techniques for indexing: N-TREOR in EXPO. *J. Appl. Crystallogr.* **33**, 1180–1186 (2000).
25. Altomare, A. *et al.* EXPO: a program for full powder pattern decomposition and crystal structure solution. *J. Appl. Crystallogr.* **32**, 339–340 (1999).
26. Boultif, A. & Louer, D. Indexing of powder diffraction patterns for low- symmetrylattices by the successive dichotomy method. *J. Appl. Crystallogr.* **24**, 987–993 (1991). 27. Louer, D. & Boultif, A. Indexing with the successive dichotomy method, DICVOL04.
Z. Krist. Suppl. **30**, 225–230 (2006).
28. Visser, J. W. & IUCr. A fully automatic program for finding the unit cell from powderdata. *J. Appl. Crystallogr.* **2**, 89–95 (1969).
29. Werner, P. E., Eriksson, L. & Westdahl, M. TREOR: a semi-exhaustive trial-and-errorpowder indexing program for all symmetries. *J. Appl. Crystallogr.* **18**, 367–370 (1985).
30. Coelho, A. A. TOPAS and TOPAS-Academic: an optimization program integratingcomputer algebra and crystallographic objects written in C plus. *J. Appl. Crystallogr.* **51**, 210–218 (2018).
31. Coelho, A. A. Indexing of powder diffraction patterns by iterative use of singular valuedecomposition. *J. Appl. Crystallogr.* **36**, 86–95 (2003).
32. Coelho, A. A. An indexing algorithm independent of peak position extraction for X-ray powder diffraction patterns. *J. Appl. Crystallogr.* **50**, 1323–1330 (2017).
33. Neumann, M. A. X-Cell: a novel indexing algorithm for routine tasks and difficultcases. *J. Appl. Crystallogr.* **36**, 356–365 (2003).
34. Le Bail, A. Monte Carlo indexing with McMaille. *Powder Diffr.* **19**, 249–254 (2004).
35. Altomare, A. *et al.* Advances in powder diffraction pattern indexing: N-TREOR09. *J. Appl. Crystallogr.* **42**, 768–775 (2009).
36. Markvardsen, A. J., David, W. I. F., Johnson, J. C. & Shankland, K. A probabilisticapproach to space-group determination from powder diffraction data. *Acta Crystallogr. Sect. A*. **57**, 47–54 (2001).
37. Markvardsen, A. J. *et al.* ExtSym: a program to aid space-group determination from powder diffraction data. *J. Appl. Crystallogr.* **41**, 1177–1181 (2008).
38. Altomare, A. *et al.* Advances in space-group determination from powder diffraction data. *J. Appl. Crystallogr.* **40**, 743–748 (2007).

39. Cernik, R. J. *et al.* The structure of cimetidine (C10H16N6S) solved from synchrotronradiation X-ray-powder diffraction data. *J. Appl. Crystallogr.* **24**, 222–226 (1991).
40. Altomare, A. *et al.* EXPO2009: structure solution by powder data in direct and reciprocal space. *J. Appl. Crystallogr.* **42**, 1197–1202 (2009).
41. Altomare, A. *et al.* EXPO2013: a kit of tools for phasing crystal structures from powder data. *J. Appl. Crystallogr.* **46**, 1231–1235 (2013).
42. Altomare, A., Cuocci, C., Giacovazzo, C., Moliterni, A. & Rizzi, R. COVMAP: a newalgorithm for structure model optimization in the EXPO package. *J. Appl. Crystallogr.* **45**, 789–797 (2012).
43. Rius, J. Patterson-function direct methods for structure determination of organiccompounds from powder diffraction data. XVI. *Acta Crystallogr. Sect. A.* **67**, 63–67 (2011).
44. David, W. I. F., Shankland, K., McCusker, L. B. & Baerlocher, C. *Structure Determination from Powder Diffraction Data* (Oxford University Press, 2002).
45. David, W. I. F. & Shankland, K. Structure determination from powder diffractiondata. *Acta Crystallogr. Sect. A* **64**, 52–64 (2008).
46. Griffin, T. A. N., Shankland, K., Van De Streek, J. & Cole, J. MDASH: a multicore-enabled program for structure solution from powder diffraction data. *J. Appl. Crystallogr.* **42**, 360–361 (2009).
47. Griffin, T. A. N., Shankland, K., Van De Streek, J. & Cole, J. GDASH: a grid-enabledprogram for structure solution from powder diffraction data. *J. Appl. Crystallogr.* **42**, 356–359 (2009).
48. Spillman, M. J., Shankland, K., Williams, A. C. & Cole, J. C. CDASH: a cloud-enabledprogram for structure solution from powder diffraction data. *J. Appl. Crystallogr.* **48**, 2033–2039 (2015).
49. Cerny, R., Favre-Nicolin, V., Rohlicek, J. & Husak, M. FOX, current state andpossibilities. *Crystals*. **7**(10), 322–331 (2017).
50. Gilmore, C. J., Shankland, K. & Bricogne, G. Applications of the maximum entropymethod to powder diffraction and electron crystallography. *Proc. R. Soc. London, A* **442**, 97–111 (1993).
51. Palatinus, L. The charge-flipping algorithm in crystallography. *Acta Crystallogr. Sect. B*. **69**, 1–16 (2013).
52. Sisak, D., Baerlocher, C., McCusker, L. B. & Gilmore, C. J. Optimizing the inputparameters for powder charge flipping. *J. Appl. Crystallogr.* **45**, 1125–1135 (2012).
53. Baerlocher, C., McCusker, L. B. & Palatinus, L. Charge flipping combined withhistogram matching to solve complex crystal structures from powder diffraction data. *Z. Krist.* **222**, 47–53 (2007).
54. Coelho, A. A. & IUCr. A charge-flipping algorithm incorporating the tangent formulafor solving difficult structures. *Acta Crystallogr. Sect. A* **63**, 400–406 (2007).
55. Von Dreele, R. B., Stephens, P. W., Smith, G. D. & Blessing, R. H. The first proteincrystal structure determined from high-resolution X-ray powder diffraction data: a variant of T3R3 human insulin-zinc complex produced by grinding. *Acta Crystallogr. Sect. D*. **56**, 1549–1553 (2000).
56. Doebbler, J. A. & von Dreele, R. B. Macromolecular powder diffraction: Structuresolution via molecular replacement. *Z. Krist. Suppl.* **30**, 33–37 (2009).
57. Margiolaki, I., Wright, J. P., Wilmanns, M., Fitch, A. N. & Pinotsis, N. Second SH3 domain of ponsin solved from powder diffraction. *J. Am. Chem. Soc.* **129**, 11865–11871 (2007).
58. Vagin, A. & Teplyakov, A. MOLREP: an automated program for molecular replacement. *J. Appl. Crystallogr.* **30**, 1022–1025 (1997).

59. Toby, B. H., Von Dreele, R. B. & IUCr. *GSAS-II*?: the genesis of a modern opensource all purpose crystallography software package. *J. Appl. Crystallogr.* **46**, 544–549 (2013).
60. Rietveld, H. M. A profile refinement method for nuclear and magnetic structures. *J. Appl. Crystallogr.* **2**, 65–71 (1969).
61. Bergmann, J., Friedel, P. & Kleeberg, R. BGMN—a new fundamental parametersbased Rietveld program for laboratory X-ray sources, its use in quantitative analysis and structure investigations. *CPD Newsl.* **20**, 5–8 (1998).
62. Rodriguez-Carvajal, J. Recent advances in magnetic-structure determination byneutron powder diffraction. *Phys. B.* **192**, 55–69 (1993).
63. KS Analytical Systems — Jade 9. Available at: http://ksanalytical.com/jade-9/.(Accessed: 1st March 2018).
64. Petr´ıcek, V., Du˘sek, M. & Palatinus, L. Crystallographic computing systemJANA2006: general features. *Z. Krist.* **229**, 345–352 (2014).
65. Lutterotti, L., Vasin, R. & Wenk, H.-R. Rietveld texture analysis from synchrotrondiffraction images. I. Calibration and basic analysis. *Powder Diffr.* **29**, 76–84 (2014).
66. Shankland, K., David, W. I. F. & Sivia, D. S. Routine ab initio structure determinationof chlorothiazide by X-ray powder diffraction using optimised data collection and analysis strategies. *J. Mater. Chem.* **7**, 569–572 (1997).
67. Stinton, G. W. & Evans, J. S. O. Parametric Rietveld refinement. *J. Appl. Crystallogr.* **40**, 87–95 (2007).
68. Rajiv, P., Dinnebier, R. E. & Jansen, M. 'Powder 3D parametric' — a program forautomated sequential and parametric rietveld refinement using topas. *Mater. Sci. Forum* **651**, 97–104 (2010).
69. Spek, A. L. Single-crystal structure validation with the program PLATON. *J. Appl. Crystallogr.* **36**, 7–13 (2003).
70. Spek, A. L. Structure validation in chemical crystallography. *Acta Crystallogr. Sect.D.* **65**, 148–155 (2009).
71. Spek, A. L. PLATON: An integrated tool for the analysis of the results of a singlecrystal structure determination. *Acta Crystallogr. Sect. A.* **46**, 34 (1990).
72. van de Streek, J. & Neumann, M. A. Validation of experimental molecular crystalstructures with dispersion-corrected density functional theory calculations. *Acta Crystallogr. Sect. B.* **66**, 544–558 (2010).
73. van de Streek, J. & Neumann, M. A. Validation of molecular crystal structures frompowder diffraction data with dispersion-corrected density functional theory (DFT-D). *Acta Crystallogr. Sect. B.* **70**, 1020–1032 (2014).
74. Toraya, H. A new method for quantitative phase analysis using X-ray powderdiffraction: direct derivation of weight fractions from observed integrated intensities and chemical compositions of individual phases. *J. Appl. Crystallogr.* **49**, 1508–1516 (2016).
75. Scarlett, N. V. Y. & Madsen, I. C. Quantification of phases with partial or no knowncrystal structures. *Powder Diffr.* **21**, 278–284 (2006).
76. Groom, C. R. & Reilly, A. M. Sixth blind test of organic crystal-structure predictionmethods. *Acta Crystallogr. Sect. B.* **70**, 776–777 (2014).

77. Reilly, A. M. *et al.* Report on the sixth blind test of organic crystal structure prediction methods. *Acta Crystallogr. Sect. B* **72**, 439–459 (2016).
78. Neumann, M. A. Tailor-made force fields for crystal-structure prediction. *J. Phys. Chem. B.* **112**, 9810–9829 (2008).
79. Egami, T. & Billinge, S. J. L. *Underneath the Bragg Peaks: Structural Analysis of Complex Materials* (Elsevier Science, 2012).
80. Masadeh, A. S. Total scattering atomic pair distribution function: new methodologyfor nanostructure determination. *J. Exp. Nanosci.* **11**, 951– 974 (2016).
81. McLain, S. E., Bowron, D. T., Hannon, A. C. & Soper, A. K. *Gudrun: A Computer Program Developed for Analysis of Neutron Diffraction Data* (STFC Rutherford Appleton Laboratory, 2012).
82. Juhas, P., Davis, T., Farrow, C. L. & Billinge, S. J. L. PDFgetX3: a rapid and highlyautomatable program for processing powder diffraction data into total scattering pair distribution functions. *J. Appl. Crystallogr.* **46**, 560–566 (2013).
83. Farrow, C. L. *et al.* PDFfit2 and PDFgui: computer programs for studying nanostructure in crystals. *J. Phys.: Condens. Matter*. **19**, 335219–335225 (2007).
84. Billinge, S. J. L. *et al.* Characterisation of amorphous and nanocrystalline molecular materials by total scattering. *CrystEngComm* **12**, 1366–1368 (2010).
85. Prill, D., Juhas, P., Billinge, S. J. L. & Schmidt, M. U. Towards solution and refinementof organic crystal structures by fitting to the atomic pair distribution function. *Acta Crystallogr. Sect. A.* **72**, 62–72 (2016).

Chapter 4

Rietveld-Based Quantitative Phase Analysis of Sugars in Confectionary

Published as:

Nicholls, D. et al. (2018) Rietveld-Based Quantitative Phase Analysis of Sugars in Confectionery. Food Analytical Methods. 11 (10), 2673–2681.

4 Rietveld-Based Quantitative Phase Analysis of Sugars in Confectionary

Foreword

The following paper was published in *Food Analytical Methods* in 2018. It contains the key findings of a substantive investigation into develop a PXRD-based method for the quantification of crystalline and amorphous sugars in chocolate crumb and chocolate product itself.

I certify that I performed all of the experimental work that is reported, performed all of the data analysis and wrote the bulk of the final submitted manuscript. Mark Spillman (MS) and Kenneth Shankland (KS) contributed to development of the QPA approach as follows: MS advised on the generation of the amorphous lactose and the practicalities of the diamond spiking; KS contributed to the development of the practicalities of PXRD data collection, developing suitable multiphase TOPAS input files and the approach to the quantification of amorphous content. Carole Elleman (CE) supplied samples of interest for analysis.

All named authors contributed to the writing of the final manuscript.

Rietveld-Based Quantitative Phase Analysis of Sugars in Confectionery

Daniel Nicholls[a] & Kenneth Shankland[a] & Mark Spillman[b] & Carole Elleman[c]

[a]School of Pharmacy, University of Reading, Whiteknights, Reading RG6 6AD, UK

[b]School of Chemistry, University of Reading, Whiteknights, Reading RG6 6AD, UK

[c]Reading Scientific Services Ltd, Reading Science Centre (Mondelez), Whiteknights, Reading RG6 6LA, UK

k.shankland@reading.ac.uk

Electronic supplementary material:

The online version of this article (https://doi.org/10.1007/s12161-018-1243-9) contains supplementary material, which is available to authorized users.

1 Abstract

Sugars are a near-ubiquitous ingredient in food products, yet rising rates of obesity and related illnesses have prompted a drive to reduce their content. The use of amorphous sugars in confectionery may be one way of achieving this by providing a similarly sweet sensation due to increased dissolution rate. However, accurate amorphous and crystalline form characterisation and quantification of complex foodstuffs can be difficult. In this study, a method for the quantification of crystalline and amorphous sugars in chocolate precursors, using powder X-ray powder diffraction, is presented. The method was first validated by the use of known compositions of mixtures of amorphous and crystalline sugars, then employed in assessing two chocolate crumb samples. The results show that the method can reliably determine the absolute quantity of amorphous and crystalline components in a confectionery sample, whilst maintaining sample integrity, apart from the addition of an inert internal standard. As such, it is a valuable addition to other techniques currently used.

Keywords: Quantitative phase analysis, Chocolate, Crystalline sugars, X-ray powder diffraction, Amorphous form quantification

2 Introduction

Gradual changes in human lifestyle and behaviour over the past 100 years have led to a dramatic rise in the incidence of diabetes, obesity and related illnesses (coined “diabesity”) across the globe (Amos et al. 1997; Astrup and Finer 2000; King et al. 1998). Consequently, governments have responded by introducing regulations (such as required information on the fat and sugar content of foodstuffs) and taxations (for example, the recently announced Soft Drinks Industry Levy in the UK) which are likely to become more restrictive on the food industry over the coming decades (Zimmet et al. 2001). As a result, there is heightened interest in research and development of food products that contain reduced sugar and fat content.

There are many options available to confectionery manufacturers to reduce the total quantity of sugars in foods. Perhaps the most widely used approach in the market is the use of sugar replacements (such as sugar alcohols) in chewing gums and carbonated soft drinks. However, these ingredients are more costly, do not necessarily provide the same taste profile and may produce undesirable gastric effects in some consumers (Nabors 2011). Another option is the use of sugars in their amorphous forms; the lack of long-range order in amorphous solids allows for increased dissolution rate within the mouth, providing a sweet sensation at reduced gross sugar content (Hartel et al. 2011).

However, achieving the amorphisation of the sugars in products is not as straightforward as simply replacing previously crystalline powders with their amorphous counterparts. For example, in chocolate manufacture, the product is exposed to a range of physical manipulations, including several heating and cooling steps, mixing within a conch, and the tempering process (Beckett 1994; Beckett 2008). Consequently, components of the mixture that were amorphous at the outset of processing may have been given enough molecular mobility to recrystallise during processing, nullifying the intended impact on the final product. Hence, modifications to the manufacturing process may be required, and prior to this, a method of characterisation and quantification of the amorphous and crystalline components within a product at each stage of preparation is potentially advantageous in ensuring the desired final product is achieved. Quantitative phase analysis using powder X-ray diffraction (PXRD) is one method that can be employed for this purpose; other commonly used methods for quantification of amorphous content include DSC, DVS and solution calorimetry.

PXRD has been used extensively as an analytical technique since its initial application by Debye and Scherrer a century ago (Debye and Scherrer 1916).In recent years, there has been great interest in the application of the Rietveld method (Rietveld 1969) to determine the relative quantities of crystalline components in a powder mixture, a process known as quantitative phase analysis (QPA). Though QPA has been extensively utilised within the cements, mining and ceramics industry, and to some extent within the pharmaceutical industry, its application in other areas is much less common (Aranda et al. 2012). Typically, QPA reports the percentage abundance of each crystalline phase present in a mixture, giving the sum of crystalline phases as 100% of the powder that is irradiated in the X-ray beam. However, by use of an appropriate internal standard, the contribution of non-crystalline[2] content can also be calculated (De La Torre et al. 2001). This is done by scaling each crystalline phase's contribution using a scale factor determined by the calculated and known concentrations of the standard, as shown in Eq. 1:

$$W_{\alpha(abs)} = W_{\alpha(calc)} \cdot \frac{W_{s(known)}}{W_{s(calc)}}$$

Equation 1

where $W_{\alpha(abs)}$ is the absolute weight fraction of phase α, $W_{\alpha(calc)}$ is the calculated relative weight fraction of phase α obtained from the QPA, $W_{s(calc)}$ is the calculated relative weight fraction of the standard material obtained from the QPA and $W_{s(known)}$ is the known weight fraction of the standard material.

In order to apply QPA to a mixture, prior knowledge of the crystallographic form of each crystalline component is required. In the case of chocolate and chocolate precursors, these components include (but are not limited to) sucrose, α-lactose monohydrate and β-lactose (see Electronic Supplementary Material for sample PXRD patterns). Furthermore, a well-characterised and highly crystalline internal standard, with an elemental composition comparable to the samples of interest, is also required.

[2] This includes X-ray amorphous solids (solids that do not contribute any visible Bragg diffraction to the PXRD pattern) and any liquid phases present, e.g. liquid fats. For simplicity, these non-crystalline entities are henceforth referred to as the amorphous content of the samples.

Here, the accuracy and precision of QPA, as applied to known compositions of crystalline and amorphous sugars using synthetic diamond powder as an internal standard, is assessed and the resultant methodology applied to two chocolate crumb samples.

3 Materials and Methods

3.1. Obtaining Pure Phases

Crystalline sucrose (CAS 57-70-1), α-lactose monohydrate (CAS 5989-81-1) and synthetic crystalline diamond powder (CAS 7782-40-3) were obtained from Sigma Aldrich, UK. Amorphous lactose was prepared by dissolving α-lactose monohydrate in water and freeze-drying small individual aliquots of the resultant 10% w/v aqueous solution as follows: pre-freezing at −80 °C, a primary drying step of 72 h at −50 ° C under a vacuum pressure of 0.06 mbar, then finally a secondary drying step over P_2O_5 for 48 h at 25 °C. For more information on this process, see Jawad (2012). The resultant powders, which were sealed and stored in desiccators, exhibited no Bragg diffraction peaks when PXRD data were collected on a laboratory diffractometer.

3.2. Preparation of Powder Mixtures of Known Composition

Five mixtures comprising varying proportions of sucrose, amorphous lactose (checked by PXRD to ensure it was X-ray amorphous) and diamond were prepared by manually mixing accurately weighed quantities of the aforementioned powders using a pestle and mortar. The composition of each mixture is shown in Table 1. Mixtures were immediately loaded into capillaries (see the “Crystallographic Phase Information” section) and the capillaries sealed with wax to prevent ingress of water from the atmosphere.

3.3. Chocolate Crumbs

Two generic chocolate crumbs[3] (henceforth referred to as crumb A and crumb B) of notionally identical composition, but made using different drying methods (roller drying under vacuum or vacuum oven drying), were supplied for analysis by QPA by Mondelez UK R&D Ltd. An amount of crumb was first weighed accurately, and then spiked with a known mass of diamond powder. Each mixture was then mixed manually using a pestle and mortar to ensure an even distribution mass of diamond powder. of the diamond powder within the sample. The composition of each mixture is shown in Table 2.

Table 1. The known composition of the powder mixtures generated for QPA

Mixture	Sucrose		Amorphous lactose		Diamond	
	Grams	%	Grams	%	Grams	%
1	0.0402	44.87	0.0393	43.86	0.0101	11.27
2	0.0329	35.84	0.0336	36.60	0.0253	27.56
3	0.0702	79.50	0.0099	11.21	0.0082	9.29
4	0.0183	19.32	0.0660	69.69	0.0104	10.98
5	0.0449	46.38	0.0468	48.35	0.0051	5.27

3.4. Crystallographic Phase Information

Relevant crystallographic information for all phases present in the mixtures and likely to be present in the crumbs was sourced from the literature and is given in Table 3.

3.5. Powder X-Ray Diffraction Data Collection and Analysis

Each sample was loaded into a 0.7-mm borosilicate capillary. For the mixtures of known sugar composition, three capillaries of each mixture were prepared, whilst subsequently, a single capillary of each of the crumb mixtures was prepared. Diffraction data were collected under ambient conditions on a Bruker D8 ADVANCE powder diffractometer configured in capillary-transmission geometry using

[3] Information on the general composition of chocolate crumb, and its role in chocolate manufacture, can be found in the following links: http://bit.ly/2rke10t and http://bit.ly/2smT1Vo

monochromatic Cu $K_{\alpha 1}$ radiation and a LynxEye detector. Data were collected in the range 3.5–80° 2θ, with a 0.017° step for 10 s per step, which equates to a total time data collection time of approximately 12.5 h per dataset. As the focus of this study was the accurate quantification of crystalline and amorphous sugars, crumb A was also collected at 47 °C[4] in order to melt the cocoa butter present and remove its crystalline contribution to the pattern. The six known crystalline forms of cocoa butter all have relatively large unit cells, of relatively low symmetry. As a result, these crystalline phases produce a large number of Bragg diffraction peaks that can overlap with those produced by the sugar phases in crumb and chocolate samples, potentially reducing the accuracy and precision of the composition values determined by QPA. As such, this simple heating procedure provides a convenient method for removing the crystalline cocoa butter contribution to the diffraction pattern. Heated samples of crumb A were subsequently cooled to ambient temperature and re-measured after ca. 16 h to observe whether or not the cocoa butter had recrystallised to its original form(s). All PXRD data analysis was performed using TOPAS (Coelho 2003), and a sample TOPAS QPA input file is available as electronic supplementary information. For a brief and effective introduction to QPA as a technique in PXRD, please see Madsen and Scarlett (2008).

Table 2. The composition of the crumb mixtures generated for QPA

Sample	Crumb		Diamond	
	Grams	%	Grams	%
Crumb A	0.0858	89.84	0.0097	10.16
Crumb B	0.0906	90.69	0.0093	9.31

4 Materials and Methods

4.1. QPA of Sugar Mixtures of Known Composition

The results of the QPA conducted on the mixtures of known sugar composition are shown in Table 4, and a representative fit to a PXRD dataset (mixture 1) is shown in Fig. 1.

[4] Temperature control was achieved using an Oxford Cryosystems Cryostream Compact device, mounted co-axially with the capillary.

4.2. QPA of Crumb Samples

The results of the QPA conducted on the crumb samples are shown in Table 5, and a representative fit to a PXRD dataset (crumb B) is shown in Fig. 2.

5 Discussion

5.1. QPA of Known Mixtures

There is very good agreement between the percentage composition values returned by QPA and the known composition across all samples, with the greatest difference occurring in mixture 5 ($\Delta_{QPA\text{-}measured}$ = 1.57%). The proportion of diamond internal standard in the analysed mixture affects the accuracy of the technique, with mixtures containing smaller diamond concentrations (mixtures 3 and 5) displaying poorer agreement with the measured quantities than those containing higher diamond concentrations (mixtures 1, 2 and 4). It also affects precision, as evidenced by the higher standard deviations obtained for mixtures 3 and 5 compared to mixtures 1, 2 and 4. As the determination of the absolute content of each phase within a mixture relies heavily on the scaling of the internal standard, small changes in the QPA-determined diamond percentages can have a significant impact on the final determined percentages of each phase. Obtaining a perfectly homogenous dispersion of the diamond standard throughout each mixture is difficult, subject as it is to inconsistencies in user technique, potential loss of powders during mixing or differences in adhesion of each phase to the pestle and mortar during mixing. Introducing a larger percentage of diamond powder in the final mix helps to reduce the relative impact of the aforementioned factors on the resultant powder pattern. The results suggest a diamond concentration of at least 10% w/w is important in QPA accuracy when dealing with samples of this nature. It is likely that the use of 0.9-mm capillaries, exposing a larger volume of the sample in the incident X-ray beam, would also help mitigate any mixing issues, at relatively small cost to the resolution of the collected PXRD pattern.

Results are precise within a sample set, with the largest standard deviation values occurring in mixtures with low diamond concentrations: mixtures 3 and 5 have a standard deviation of 1.4 and 1.3% respectively). The repeatability coefficient is defined as

$$Repeatability = \sqrt{\frac{\sum_{i=1}^{n}(sd^2 \cdot Degrees\ of\ freedom)}{(M-n)}}$$ Equation 2

where M is the total number of measurements, n is the number of mixtures analysed, degrees of freedom is the number of repeat analyses minus one and sd is the standard deviation for a set of repeats.

Table 3. Published crystallographic information for each phase used in the QPA process

	α-Lactose monohydrate	β-Lactose	Sucrose	Cocoa butter (form V)	Diamond
Space group	$P2_1$	$P2_1$	$P2_1$	Cc	$Fd\bar{3}m$
a (Å)	7.937	10.839	10.863	5.442	3.567
b (Å)	21.568	13.349	8.704	127.638	3.567
c (Å)	4.815	4.954	7.762	8.214	3.567
β (°)	109.77	91.31	102.94	88.69	-
Volume ($Å^3$)	775.673	716.606	715.354	5703.967	45.385
Reference	Fries et al. (1971)	Ken and Akira (1974)	Hynes and Le Page (1991)	van Mechelen et al. (2006)	Fayos (1999)
Temperature factors	Anisotropic[a]	Anisotropic[a]	Isotropic[b]	Isotropic[b]	Isotropic[c]

[a] Anisotropic displacement parameters used for non-hydrogen atoms and isotropic temperature factors for hydrogen atoms

[b] Isotropic temperature factors used for all atoms

[c] Isotropic temperature factor allowed to refine in QPA

For the QPA measurements performed on the sugar mixtures, the coefficient was calculated to be 0.78 (M = 15, n = 5, degrees of freedom = 2), indicating that for two repeat measurements on the same mixture, there is a 95% probability that the results of the QPA will differ by less than 0.78%. The results suggest that multiple repeat measurements of a mixture are not necessary to obtain quantitative data within 1% accuracy.

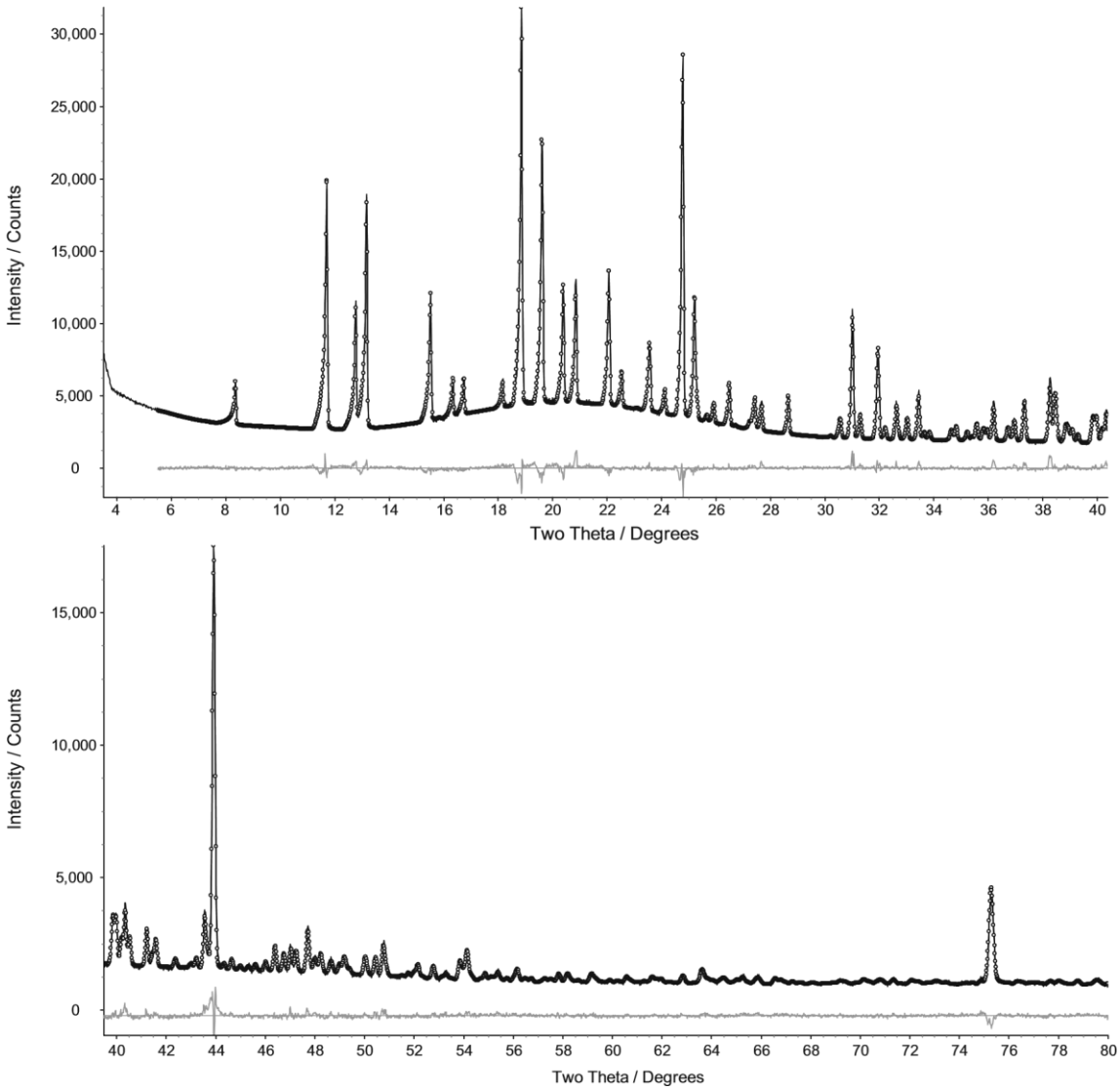


Figure 1. The Rietveld fit obtained for PXRD data collected from mixture 1 in the 2θ ranges 3.5–40° (upper plot) and 40–80° (lower plot). Observed data (points), calculated data (solid black line) and the difference profile (solid grey line) are shown. Diffraction from the diamond internal standard is clearly visible in the form of the very strong peaks at ca. 44° and ca. 75° 2θ.

The goodness of fit between calculated and measured diffraction patterns is indicated by R_{wp} (Young 1993). In all cases, the fits to the data are good, with low R_{wp} values and difference plots (see, for example, Fig. 1) consistent with the input crystalline phases providing a very good description of the observed PXRD data. It is important to consider not only the R_{wp} values when evaluating fits but also the visual fit to the data—samples containing large amounts of amorphous material will have systematically lower R_{wp} values than those containing smaller amounts (e.g. mixture 4 has 78.29%

amorphous lactose and R_{wp} = 3.18 whilst mixture 3 has 87.64% crystalline sucrose and R_{wp} = 5.56) as a consequence of the way in which R_{wp} is calculated. Whilst the background-subtracted R_{wp}' could be used to eliminate this systematic difference, the combination of R_{wp} plus close visual inspection of the difference plot is an effective way of ensuring that a good fit has been obtained. Furthermore, when dealing with samples where the exact compositions are not known a priori, this method allows the identification of unfitted features that may indicate the presence of additional crystalline phases that are not currently included in the Rietveld fit, or reveal inadequacies in the existing models.

Table 4. Results of the QPA conducted on mixtures of known composition

Mixture	Known percentages (exc. diamond)		QPA Percentages (n = 3)		R_{wp}
	Sucrose	Amorphous Lactose	Sucrose	Amorphous Lactose	
1	50.57	49.43	50.68 ± 0.6	49.32 ± 0.6	4.14
2	49.48	50.52	49.42 ± 0.5	50.58 ± 0.5	4.43
3	87.64	12.36	86.66 ± 1.4	13.34 ± 1.4	5.56
4	21.71	78.29	20.97 ± 0.8	79.03 ± 0.8	3.18
5	48.96	51.04	47.12 ± 1.3	52.88 ± 1.3	3.74

The QPA percentage is reported as a mean ± s.d. of three separate analyses. The goodness of fit between the calculated and measured powder X-ray diffraction patterns at the end of the QPA is presented as R_{wp}

Table 5. Results of the QPA conducted on mixtures of crumb and diamond

Crumb	QPA Composition (%)					R_{wp}
	Sucrose	α-Lactose monohydrate	β-Lactose	Cocoa butter form V	Amorphous content	
A (ambient)	41.44	2.75	4.11	2.15	49.55	5.39
A (47 °C)	42.41	3.22	4.37	0.00	50.00	4.74
A (recooled)	46.45	2.75	4.62	0.12	46.06	4.92
B (ambient)	51.31	0.00	5.82	2.01	40.86	3.59

The QPA percentage is reported as the result of a single analysis. The goodness of fit between the calculated and measured powder X-ray diffraction patterns at the end of the QPA is presented as R_{wp}

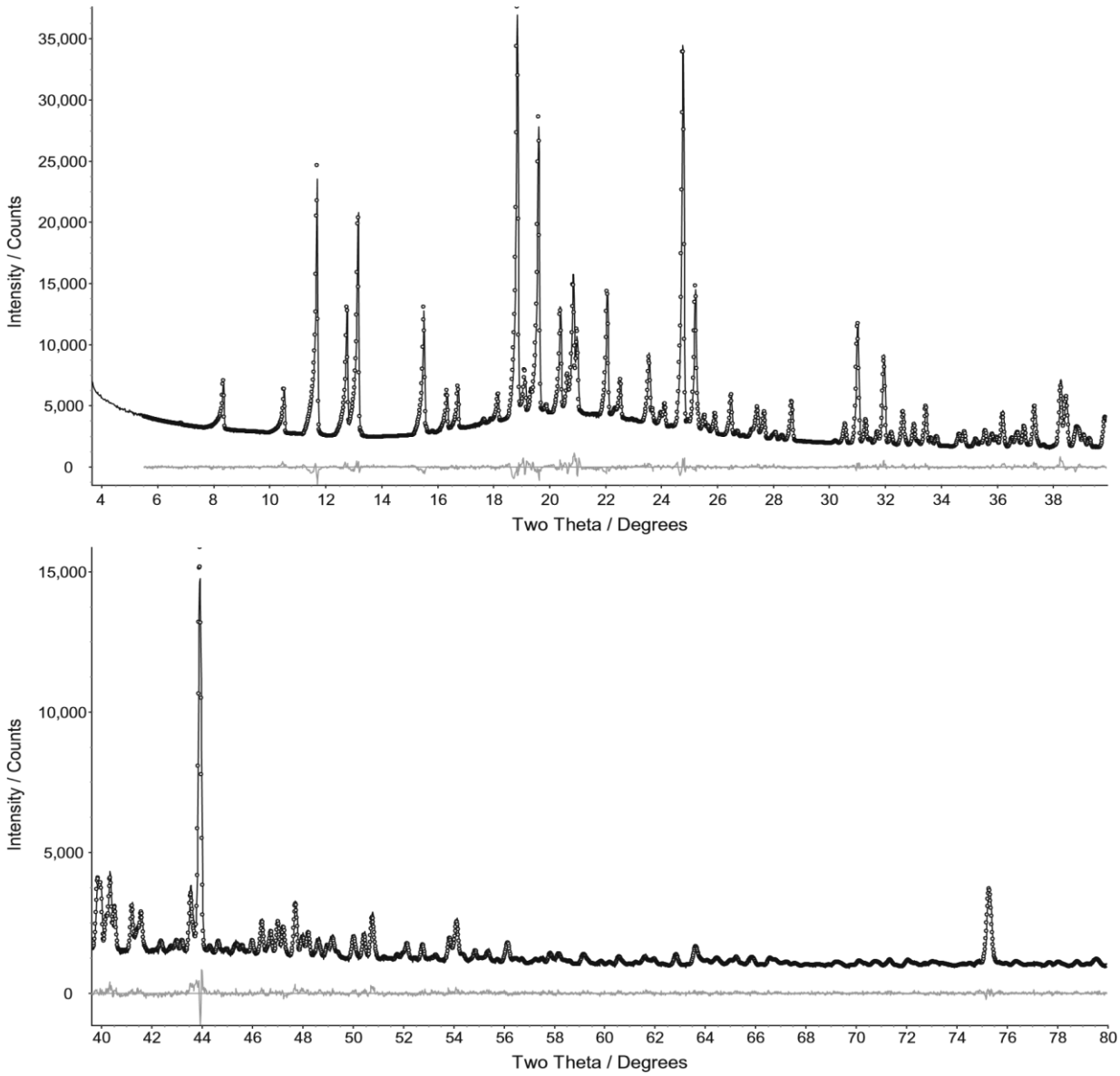


Figure 2. The Rietveld fit obtained for PXRD data collected from crumb B in the 2θ ranges 3.5–40° (upper plot) and 40–80° (lower plot). Observed data (points), calculated data (solid black line) and the difference profile (solid grey line) are shown. Diffraction from the diamond internal standard is clearly visible in the form of the very strong peaks at ca. 44° and ca. 75° 2θ

5.2. QPA of Chocolate Crumb

The results shown in Table 5 illustrate the power and the flexibility of PXRD for QPA of phase mixtures, such as the chocolate crumbs examined in this work. The method is able to quantify the crystalline sugars and form V cocoa butter at ambient temperatures, and Fig. 3 shows that upon heating to 47 °C, the diffraction peak at 19.4° 2θ (corresponding to the strongest diffraction peak of form V cocoa butter) disappears, indicative of cocoa butter melting and the removal of its crystalline contribution to the diffraction pattern. The slight lateral shifting of peaks, also clearly visible in Fig. 3, is attributable to unit cell expansion of the crystalline phases at elevated temperature.

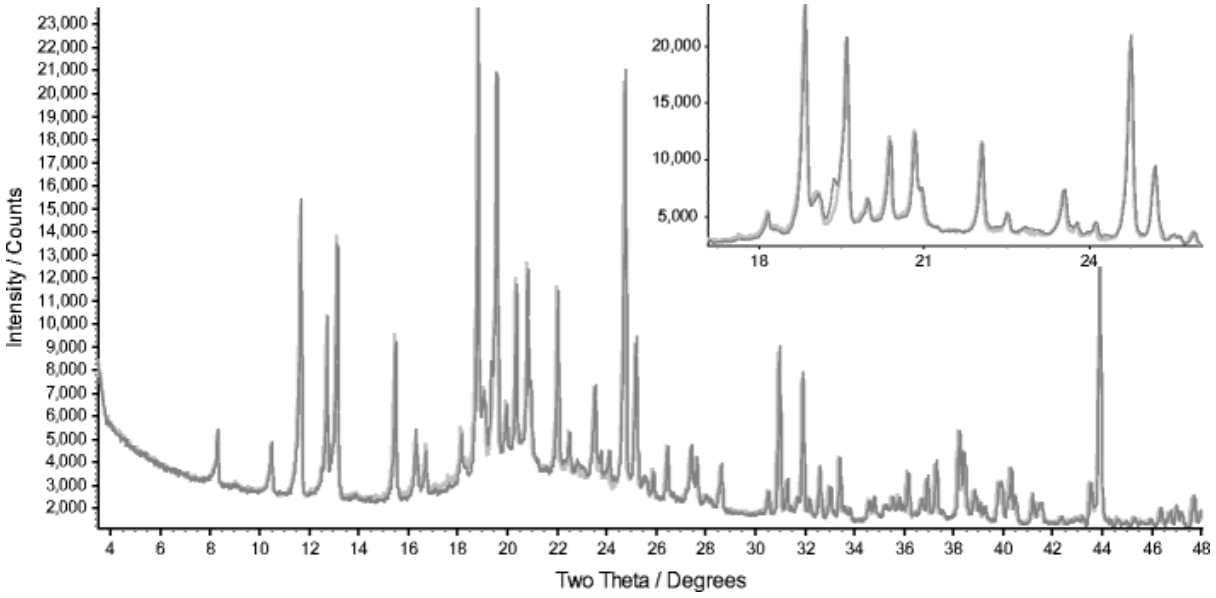


Figure 3. Overlay of heated (light grey) and ambient (dark grey) measurements of Chocolate Crumb A. The inset shows the 17 – 26 ° 2θ region, where the loss of form V cocoa butter peaks in the heated crumb pattern is most evident

Upon cooling of the heated sample back to ambient conditions, the peak at 19.4° 2θ does not immediately reappear, and there are no additional visible diffraction features, even after ca. 16 h at ambient conditions, which cannot be accounted for by the crystalline sugars. This suggests that, owing to the uncontrolled nature of the cooling back to ambient temperature, the cocoa butter may have recrystallised into several polymorphs whose relatively low abundance makes their quantification impractical or has solidified as an amorphous or nanocrystalline solid. QPA performed on the heated and cooled sample reported 0.12% of form V cocoa butter, a negligible amount that reflects correlations between phases included in the calculation as opposed to an accurate value; the same R_{wp} for the QPA can be achieved when the form V contribution is

excluded. Furthermore, the heated and cooled sample showed an increased percentage of crystalline sucrose within the mixture, suggesting that the increased molecular mobility introduced by the heating stage has allowed some amorphous sucrose to recrystallise.

Chocolate crumb B displays a much different pattern to that of crumb A. Firstly, no crystalline alpha-lactose monohydrate was observed in this pattern; rather, there is an increased quantity of the beta anomer. Furthermore, a much higher quantity of crystalline sucrose, and reduced total amorphous content, is present.

It is important to remember that QPA cannot distinguish between different amorphous components within the crumb mixture. Therefore, whilst it is possible to state with good accuracy what percentage of the mixture is amorphous, it is not possible to directly determine the amorphous percentage of each individual phase within the crumb, be it lactose, sucrose, cocoa butter, non-fat cocoa solids or various milk solids. However, evaluation of the amorphous percentages of each phase is possible if the crumb recipe is known. For example, if the recipe states 60% sucrose and QPA returns a crystalline value of 50%, then 10% of the crumb is amorphous sucrose. In this work, Mondelez has confirmed that the results obtained make compositional sense for the chocolate crumbs (whose recipes were not disclosed to us) that they supplied.

This clearly demonstrates the value of QPA in this context: evaluating changes in crystalline composition as a function of sample processing can help to explain differences that may be detected in taste, texture and manufacturability.

5.3. Internal Standard and Diffraction Geometry

In developing a QPA method, it is important that a suitable internal standard is chosen. For molecular organic materials, widely used QPA internal standards such as Al_2O_3 and ZnO are not suitable as microabsorption (Klug and Alexander 1974) becomes an issue. The choice of diamond powder as an internal standard aimed to minimise microabsorption whilst providing sharp diffraction peaks that did not markedly interfere with the regions of strong diffraction observed from the crystalline components of the samples under examination. Both synthetic and naturally sourced diamond powders were tested. It was found that the naturally sourced diamond powder contained trace quantities of zirconia, which produced significant contributions to the diffraction patterns of the spiked mixtures. Whilst these contributions are

easily modelled, microabsorption and incidental peak overlap are complications that are best avoided and so synthetic diamond powder was chosen as the internal standard. The internal standard method assumes that the percentage crystallinity of the standard is known; in this work, based on laboratory PXRD and using a hard-wearing standard, 100% crystallinity is assumed. The validity of this assumption is illustrated by the excellent agreement between measured and calculated phase percentages in the known mixtures. Whilst relatively expensive as a standard,[5] the amount of diamond powder required for each individual analysis is sufficiently small that the cost per sample prepared is small. The internal standard method avoids the need for additional corrections to account for scattering from the capillary.

This QPA work has been carried out using a laboratory-based Bruker D8 diffractometer operating in capillary-transmission geometry; the monochromatic incident X-ray beam passes through the sample, contained in a thin-walled borosilicate glass capillary, and diffracted X-rays are collected using a position-sensitive detector. The sample is rotated along the axis of the diffractometer to reduce the detrimental effects of preferred orientation of crystallites within the sample. The use of reflection-based PXRD[6] for QPA of molecular materials is not recommended; sample presentation issues frequently lead to preferred orientation and sample transparency effects, and these will significantly affect the accuracy of values returned by QPA analysis.

6 Conclusions

Rietveld-based QPA, with a suitable internal standard, allows quick and easy identification and quantification of crystalline phases in samples related to chocolate manufacture. Furthermore, it allows for accurate quantification of total amorphous content by difference. The presence or absence of detectable amounts of a crystalline phase can be easily determined by incorporation of that phase into the Rietveld calculation and assessing the impact upon the R_{wp} of the resultant fit to the data.

Temperature control of the sample allows for the removal of crystalline cocoa butter contributions to the pattern, if desired. The method, whilst straightforward, doesrequire the use

[5] Approximately £160 for 5 g, at the time of writing

[6] Within industry, reflection-based PXRD is commonly used, as it lends itself to rapid sample preparation and presentation, rapid data collection and automated sample changing.

of a diffractometer operating in capillary-transmission X-ray diffraction geometry to ensure that good quality diffraction data, largely free from the effects of preferred orientation, are obtained. Suitable software for QPA is also required; in this work, TOPAS has been used, but many other alternatives are available and are equally suitable.

6.1. Acknowledgements

KS and CE gratefully acknowledge the Biotechnology and Biological Sciences Research Council and Mondelez UK R&D Ltd. for funding for DN through the iCASE scheme. MJS is grateful to the Defence Academy of the United Kingdom for allocating time for continued collaboration and academic development.

Funding This study was funded by the Biotechnology and Biological Sciences Research Council (grant number BB/L015730/1) and Mondelez UK R&D Ltd.

6.2. Compliance with Ethical Standards

Daniel Nicholls declares that he has no conflict of interest.

Kenneth Shankland declares that he has no conflict of interest.

Mark Spillman declares that he has no conflict of interest.

Carole Elleman declares that she has no conflict of interest.

This article does not contain any studies with human participants or animals performed by any of the authors.

7 References

Amos AF, McCarty DJ, Zimmet P (1997) The rising global burden of diabetes and its complications: estimates and projections to the year 2010. Diabet Med 14 Suppl 5:S1–85

Aranda MAG, De la Torre ÁG, León-Reina L (2012) Rietveld quantitative phase analysis of OPC clinkers, cements and hydration. Products. Rev Mineral Geochem 74:169–209. https://doi.org/10. 2138/rmg.2012.74.5

Astrup A, Finer N (2000) Redefining type 2 diabetes: 'diabesity' or 'obesity dependent diabetes mellitus'? Obes Rev 1:57–59

Beckett S (1994) Industrial chocolate manufacture and use. Springer US

Beckett S (2008) The science of chocolate. Royal Society of Chemistry, UK

Coelho A (2003) TOPAS user manual. Bruker AXS GmbH, Karlsruhe, Germany

De La Torre AG, Bruque S, Aranda MAG (2001) Rietveld quantitative amorphous content analysis. J Appl Crystallogr 34:196–202. https:// doi.org/10.1107/S0021889801002485

Debye P, Scherrer P (1916) X-ray interference by disorder particles. Physik Z 17:277–283

Fayos J (1999) Possible 3D carbon structures as progressive intermediates in graphite to diamond phase transition. J Solid State Chem 148: 278–285. https://doi.org/10.1006/jssc.1999.8448

Fries DC, Rao ST, Sundaralingam M (1971) Structural chemistry of carbohydrates. III. Crystal and molecular structure of 4-O-[beta]-dgalactopyranosyl-[alpha]-d-glucopyranose monohydrate ([alpha]lactose monohydrate). Acta Crystallogr B 27:994–1005. https:// doi.org/10.1107/S0567740871003364

Hartel RW, Ergun R, Vogel S (2011) Phase/state transitions of confectionery sweeteners: thermodynamic and kinetic aspects. Compr Rev Food Sci Food Saf 10:17–32. https://doi.org/10.1111/j.1541-4337. 2010.00136.x

Hynes RC, Le Page Y (1991) Sucrose, a convenient test crystal for absolute structures. J Appl Crystallogr 24:352–354. https://doi.org/10. 1107/S0021889891002492

Jawad R (2012) A study of phase transitions of chocolate sugars during manufacture. PhD thesis. In: King's College London

Ken H, Akira S (1974) The crystal and molecular structure of βlactose. Bull Chem Soc Jpn 47:1872–1879. https://doi.org/10. 1246/bcsj.47.1872

King H, Aubert RE, Herman WH (1998) Global burden of diabetes, 1995-2025: prevalence, numerical estimates, and projections. Diabetes Care 21:1414–1431

Klug H, Alexander LE (1974) X-ray diffraction procedures: for polycrystalline and amorphous materials, 2nd edn. Wiley-Interscience, USA

Madsen IC, Scarlett NVY (2008) Quantitative phase analysis. In: Dinnabier RE (ed) Powder diffraction: theory and practice. Royal Society of Chemistry

Nabors LB (2011) Alternative sweeteners. In: Alternative sweeteners, fourth edition. CRC Press, pp 1–10. https://doi.org/10.1201/ b11242-2

Rietveld H (1969) A profile refinement method for nuclear and magnetic structures. J Appl Crystallogr 2:65–71. https://doi.org/10.1107/ S0021889869006558

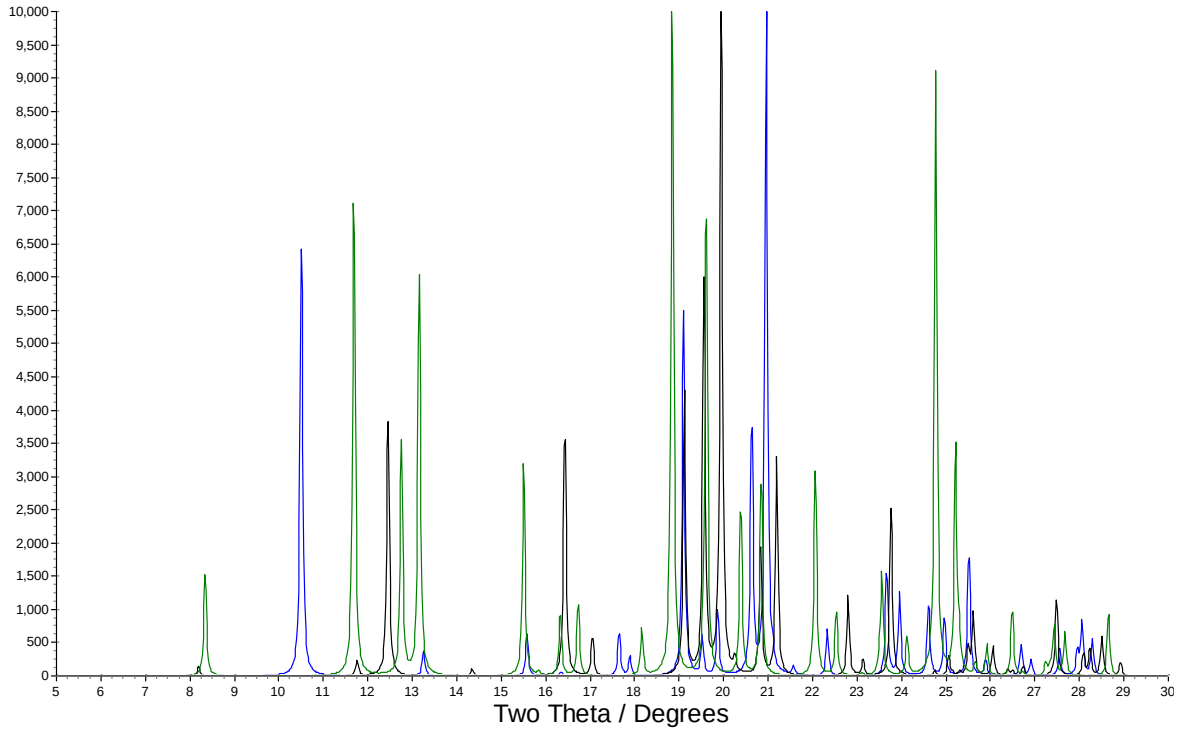
van Mechelen JB, Peschar R, Schenk H (2006) Structures of monounsaturated triacylglycerols. II. The [beta]2 polymorph. Acta Crystallogr B 62:1131–1138. https://doi.org/10.1107/

S0108768106037086

Young R (1993) The Rietveld method. Oxford University Press, UK

Zimmet P, Alberti KGMM, Shaw J (2001) Global and societal implications of the diabetes epidemic. Nature 414:782–787

8 Additional Information and Comments



An overlay of simulated PXRD data (λ=1.5406Å) for β-lactose (blue; CSD refcode BLACTO), α-lactose monohydrate (black; CSD refcode LACTOS10) and sucrose (green; CSD refcode SUCROS08).

It is clear that the contributions of each phase are easily identified and Rietveld refinement explicitly deals with any overlapping contributions seen in an observed diffraction pattern.

8.1 Sample TOPAS QPA file

```
/*
  A template for quantitative multiphase Rietveld refinement using TOPAS

  Associated with publication "Rietveld-based Quantitative Phase Analysis
  of Sugars in Confectionary" by Nicholls et al.

  Please note the following TOPAS input file conventions

  Text contained between lines starting /* and ending */ are comments

  Text after lines starting with ' are comments

  Any parameter flagged with an @ symbol is one that is refined during the Rietveld
  fit e.g. a phase scale factor

  Any parameter not flagged with an @ symbol is treated as fixed e.g. the LP_Factor
  parameter value

*/

' Profile fit parameters (inc. background)
         r_exp  1.621
         r_wp  3.545

' Profile fit parameters (exc. background)
         r exp dash  5.329
         r_wp_dash  11.655

' Tell TOPAS what diffraction datafile to work with
' For example, xdd "sucrose.raw"
         xdd <name_of_diffraction_data_file>

' Tell TOPAS to model the background with 18th order Chebyshev polynomial
         bkg @  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

' Define the range of diffraction data you wish to fit, values in two-theta
         start_X       5.5000
         finish X      80.0000

' Allow the zero point for the diffractometer to refine
         Zero_Error(@, 0.01067`)

' The argument to the LP Factor macro is the monochromator take-off angle.
' The value of 27.28 is appropriate for the D8 Advance diffactometers
' at Uni. of Reading
```

```
         LP_Factor( 27.28)

' Model the low angle asymmetry in the pattern attributable to axial divergence
         Simple_Axial_Model(@, 11.12570)

' Specify the incident wavelength and the step size of the diffraction data
         lam
         ymin_on_ymax 0.001
         la 1 lo  1.540600
         x_calculation_step 0.017

' Allow peak widths to be described by crystallite size and strain terms
' (Lorentzian and Gaussian). These parameters are normally first refined
' in a Pawley-type refinement, then used in the Rietveld refinement. The
' values below are simply initial values

         CS_L(@, 1000)
         CS_G(@, 1000)
         Strain_L(@, 0.1)
         Strain_G(@, 0.1)

'       ?????
         PV(pva1, pvxo1, pvfwhm1, pvg1, 8952.37120, 20, 7.48076, 1.00000)


/*
  From this point onwards, there is a series of blocks that define the
  various possible crystalline contributions to the diffraction pattern.
  The presence of the keyword str marks the start of a new crystalline
  phase. The phase is defined by a space group, lattice parameters and
  atomic coordinates + temperature factors.
  The overall contribution of any phase to the observed diffraction
  pattern is evidenced by the value of the refineable scale factor
  for that phase.  It is from these scale factors that the percentage
  contributions to the pattern are ultimately determined.
*/

str
         phase_name "Diamond"
         space_group "Fd-3m"
         Cubic(@  3.55449)
         volume  45.3272472
         scale @  0.01
         site C      x 0.00000       y 0.00000       z 0.00000     occ C 1.0    beq @ 0.09

str
         phase_name "Sucrose"
         ' COORDINATES AND BISO VALUES TAKEN DIRECTLY FROM HYNES AND LE PAGE
         ' J APPL CRYST. (1991) 24, 352-354
```

```
' "Sucrose, a convenient test crystal for absolute structures"
' Note that this single-crystal structure was solved from data collected at 300K
' Structure available in CSD as SUCROS08
volume  713.97933`
space_group "P21"
a  @  10.85769
b  @  8.69932
c  @  7.75628
al   90
be  @ 102.95101
ga   90
scale @ 0.001

site O1     x 0.1706        y 0.34690       z 0.3910   occ O   1   beq   1.61
site O2     x 0.2291        y 0.4361        z 0.7475   occ O   1   beq   2.45
site O3     x 0.3086        y 0.7479        z 0.7023   occ O   1   beq   3.17
site O4     x 0.3484        y 0.8137        z 0.3549   occ O   1   beq   3.85
site O5     x 0.3772        y 0.3988        z 0.3688   occ O   1   beq   1.78
site O6     x 0.5817        y 0.5461        z 0.2855   occ O   1   beq   2.81
site O7     x 0.0296        y 0.2351        z 0.6205   occ O   1   beq   2.42
site O8     x 0.2123        y 0.0945        z 0.3162   occ O   1   beq   1.81
site O9     x -0.0745       y 0.3179        z 0.2044   occ O   1   beq   2.22
site O10    x -0.0222       y 0.0985        z -0.0890  occ O   1   beq   2.45
site O11    x 0.3265        y 0.2385        z 0.0397   occ O   1   beq   2.74
site C1     x 0.2993        y 0.3585        z 0.4857   occ C   1   beq   1.69
site C2     x 0.3134        y 0.4766        z 0.6342   occ C   1   beq   1.80
site C3     x 0.2852        y 0.6373        z 0.5644   occ C   1   beq   1.89
site C4     x 0.3744        y 0.6713        z 0.4426   occ C   1   beq   2.10
site C5     x 0.3592        y 0.5523        z 0.2960   occ C   1   beq   2.03
site C6     x 0.4576        y 0.5718        z 0.1837   occ C   1   beq   2.55
site C7     x 0.1030        y 0.1316        z 0.5430   occ C   1   beq   2.02
site C8     x 0.1247        y 0.1934        z 0.3692   occ C   1   beq   1.49
site C9     x 0.0073        y 0.1910        z 0.2149   occ C   1   beq   1.59
site C10    x 0.0653        y 0.1670        z 0.0555   occ C   1   beq   1.63
site C11    x 0.1767        y 0.0624        z 0.1278   occ C   1   beq   1.99
site C12    x 0.2886        y 0.0811        z 0.0472   occ C   1   beq   2.48
site H1     x 0.269         y 0.375         z 0.836    occ H   1   beq   4.5
site H2     x 0.244         y 0.762         z 0.737    occ H   1   beq   3.2
site H3     x 0.337         y 0.892         z 0.431    occ H   1   beq   5.5
site H4     x 0.603         y 0.419         z 0.296    occ H   1   beq   10.6
site H5     x 0.078         y 0.318         z 0.651    occ H   1   beq   2.6
site H6     x -0.033        y 0.387         z 0.186    occ H   1   beq   2.1
site H7     x -0.008        y 0.146         z -0.177   occ H   1   beq   6.4
site H8     x 0.345         y 0.286         z 0.147    occ H   1   beq   9.0
site H9     x 0.328         y 0.264         z 0.526    occ H   1   beq   2.8
site H10    x 0.403         y 0.475         z 0.706    occ H   1   beq   1.4
site H11    x 0.197         y 0.645         z 0.494    occ H   1   beq   2.0
site H12    x 0.459         y 0.677         z 0.516    occ H   1   beq   1.6
site H13    x 0.280         y 0.556         z 0.223    occ H   1   beq   1.6
```

```
site H14    x 0.442         y 0.674         z 0.130    occ H   1   beq   2.8
site H15    x 0.434         y 0.493         z 0.066    occ H   1   beq   3.1
site H16    x 0.054         y 0.033         z 0.512    occ H   1   beq   1.6
site H17    x 0.183         y 0.107         z 0.624    occ H   1   beq   0.9
site H18    x -0.045        y 0.092         z 0.230    occ H   1   beq   1.3
site H19    x 0.089         y 0.260         z 0.026    occ H   1   beq   0.9
site H20    x 0.151         y -0.027        z 0.112    occ H   1   beq   2.6
site H21    x 0.365         y 0.015         z 0.117    occ H   1   beq   2.4
site H22    x 0.270         y 0.036         z -0.078   occ H   1   beq   2.0

str
        phase_name "Beta-lactose"
        ' COORDINATES AND BISO VALUES TAKEN DIRECTLY FROM HIROTSU AND SHIMADA
        ' BULL CHEM SOC JAP 47(8) 1872-1879 [1974]
        ' "The crystal and molecular structure of beta-lactose"
        ' Note that this single-crystal structure was solved from data collected at RT
        ' Structure available in CSD as BLACTO
        ' Note that no Biso values are reported for the H atoms and so they should be refined
        volume  714.775148`
        space_group "P21"
        a  @ 10.81964
        b  @ 13.33442
        c  @ 4.95634
        al 90
        be @ 91.64360
        ga 90
        scale @ 0.001
        site HC1          x 0.813         y 0.585         z 0.21          occ H    1    beq 6.0
        site HC2          x 0.67          y 0.703         z -0.19         occ H    1    beq 6.0
        site HC3          x 0.721         y 0.737         z 0.36          occ H    1    beq 6.0
        site HC4          x 0.51          y 0.718         z 0.416         occ H    1    beq 6.0
        site HC5          x 0.632         y 0.571         z 0.444         occ H    1    beq 6.0
        site HC6-1        x 0.399         y 0.533         z 0.157         occ H    1    beq 6.0
        site HC6-2        x 0.418         y 0.545         z 0.495         occ H    1    beq 6.0
        site HC1p         x 0.94          y 0.309         z -0.676        occ H    1    beq 6.0
        site HC2p         x 0.832         y 0.233         z -0.215        occ H    1    beq 6.0
        site HC3p         x 0.754         y 0.41          z -0.543        occ H    1    beq 6.0
        site HC4p         x 0.852         y 0.422         z 0.018         occ H    1    beq 6.0
        site HC5p         x 0.979         y 0.473         z -0.48         occ H    1    beq 6.0
        site HC6p-1       x 1.099         y 0.458         z 0.08          occ H    1    beq 6.0
        site HC6p-2       x 1.04          y 0.566         z -0.054        occ H    1    beq 6.0
        site HO2          x 0.882         y 0.788         z 0.013         occ H    1    beq 6.0
        site HO3          x 0.575         y 0.865         z 0.05          occ H    1    beq 6.0
        site HO4          x 0.361         y 0.735         z -0.025        occ H    1    beq 6.0
        site HO6          x 0.535         y 0.4           z 0.492         occ H    1    beq 6.0
        site HO1p         x 1.077         y 0.191         z -0.588        occ H    1    beq 6.0
        site HO2p         x 0.794         y 0.171         z -0.645        occ H    1    beq 6.0
        site HO6p         x 1.22          y 0.455         z -0.246        occ H    1    beq 6.0
        site HO3p         x 0.652         y 0.439         z -0.093        occ H    1    beq 6.0
```

```
        site C1           x 0.7544        y 0.5881        z 0.0177        occ C    1    ADPs { 0.000877795      0.023471512
        0.000937221                0.000000000                0.000401199                -0.000474673 }
        site C2           x 0.7328        y 0.6984        z -0.0391       occ C    1    ADPs { 0.000926562      0.022568762
        0.000758946                0.000148364                0.000312044                0.000000000 }
        site C3           x 0.666         y 0.746         z 0.1955        occ C    1    ADPs { 0.001121627      0.017152259
        0.000886285               -0.000148364                0.000546077                0.001288397 }
        site C4           x 0.5466        y 0.688         z 0.2375        occ C    1    ADPs { 0.001072861      0.01986051
        0.00104928                -0.000593457                0.000300899                0.001627449 }
        site C5           x 0.5744        y 0.5768        z 0.2829        occ C    1    ADPs { 0.000999711      0.020763261
        0.00085063                 0.000593457                0.000791254                0.001152776 }
        site C6           x 0.4573        y 0.5163        z 0.306         occ C    1    ADPs { 0.001389843      0.022568762
        0.001808226                0.002077099                0.001092153                0.001356207 }
        site C1p          x 0.9553        y 0.2796        z -0.478        occ C    1    ADPs { 0.001048478      0.021666011
        0.001207182               -0.001632007                0.000434633                0.001695259 }
        site C2p          x 0.8242        y 0.2694        z -0.376        occ C    1    ADPs { 0.001146011      0.023471512
        0.00150261                 0.000741821                0.000936132                0.002305552 }
        site C3p          x 0.7684        y 0.3742        z -0.3392       occ C    1    ADPs { 0.000829029      0.027082514
        0.001207182                0.001186914                0.000813543                0.000745914 }
        site C4p          x 0.8557        y 0.4444        z -0.18         occ C    1    ADPs { 0.000877795      0.014444008
        0.001156246               -0.000890185                0.000501499               -0.001559638 }
        site C5p          x 0.9857        y 0.4431        z -0.2897       occ C    1    ADPs { 0.000950945      0.01895776
        0.00117662                -0.001335278                0.000412344               -0.001424018 }
        site C6p          x 1.0788        y 0.5001        z -0.1179       occ C    1    ADPs { 0.000902179      0.028888015
        0.001736916                0.000741821                0.000557221                0.000813724 }
        site O1           x 0.8086        y 0.5456        z -0.2101       occ O    1    ADPs { 0.001121627      0.018055009
        0.000927034               -0.001632007                0.00057951                -0.000881535 }
        site O2           x 0.8442        y 0.7471        z -0.1082       occ O    1    ADPs { 0.001487376      0.027082514
        0.001242837                0.004005835                0.000880409                0.002780225 }
        site O3           x 0.6449        y 0.85          z 0.1506        occ O    1    ADPs { 0.001462992      0.014444008
        0.0014313         -0.000593457                0.000111444                0.001491828 }
        site O4           x 0.4682        y 0.698         z 0.0058        occ O    1    ADPs { 0.001024095      0.031596267
        0.001466955               -0.002077099               -0.000300899               -0.000474673 }
        site O5           x 0.641         y 0.5386        z 0.0566        occ O    1    ADPs { 0.000902179      0.01895776
        0.00097797                 0.000741821                0.000835832                0.002848035 }
        site O6           x 0.4802        y 0.4118        z 0.3041        occ O    1    ADPs { 0.001926273      0.01895776
        0.001726729                0.003412378                0.000156022               -0.001559638 }
        site O1p          x 1.0099        y 0.1856        z -0.4838       occ O    1    ADPs { 0.001170394      0.024374263
        0.001640137               -0.002077099                0.000690954                0.002915845 }
        site O2p          x 0.7472        y 0.2183        z -0.5615       occ O    1    ADPs { 0.001316693      0.03791552
        0.003127467               -0.002670556               -0.00067981                 0.011731192 }
        site O3p          x 0.6516        y 0.3647        z -0.223        occ O    1    ADPs { 0.001048478      0.033401768
        0.002516236                0.002373828                0.001526786                0.007933812 }
        site O5p          x 1.028         y 0.3412        z -0.3          occ O    1    ADPs { 0.000829029      0.01895776
        0.001492423               -0.002225464                0.000222888                0.001559638 }
        site O6p          x 1.191         y 0.5153        z -0.2549       occ O    1    ADPs { 0.001048478      0.027985265
        0.002949191                0.000741821                0.000668665               -0.004678915 }

str
        phase_name "alpha-Lactose.H2O"
```

```
        ' COORDINATES AND TEMPERATURE FACTOR VALUES TAKEN DIRECTLY FROM RIES, RAO AND SUNDARALINGAM
        ' ACTA CRYST B27, 994 [1971]
        ' "STRUCTURAL CHEMISTRY OF CARBOHYDRATES III"
        ' Note that this single-crystal structure was solved from data collected at RT
        ' Structure available in CSD as LACTOS10
        volume  785.799556
        space_group "P21"
        a @ 7.96440
        b @ 21.84359
        c @ 4.78483
        al 90
        be @ 109.85201
        ga 90
        scale @ 0.001
        site C1     x 0.22747      y  0.34031      z  0.39038   occ  C  1  ADPs { 0.004020904 0.030619049        0.004806453
0.000984161        0.001408338        0.001387836 }
        site C2     x 0.36079      y  0.39241      z  0.42648   occ  C  1  ADPs { 0.004660593 0.037684983        0.00485652        -
0.001312214        0.001242651        0.000198262 }
        site C3     x 0.30803      y  0.44897      z  0.56702   occ  C  1  ADPs { 0.005437359 0.035329672        0.005323814       -
0.003280535        0.001739711       -0.001189574 }
        site C4     x 0.11435      y  0.46720      z  0.40301   occ  C  1  ADPs { 0.006396893 0.028263737        0.006141578       -
0.000656107        0.002374844       -0.001982624 }
        site C5     x -0.00462     y  0.41041      z  0.38683   occ  C  1  ADPs { 0.004569209 0.028263737        0.006325158
0.000656107        0.002071085       -0.001982624 }
        site C6     x -0.19777     y  0.42372      z  0.21408   occ  C  1  ADPs { 0.004889054 0.040040295        0.008728384
0.001640268        0.001960627       -0.000991312 }
        site O1     x 0.26711      y  0.29441      z  0.22056   occ  O  1  ADPs { 0.004843361 0.025908426        0.005957998       -
0.000328054        0.002512916        0.000396525 }
        site O2     x 0.53208      y  0.37081      z  0.60255   occ  O  1  ADPs { 0.00328983  0.058882786        0.006658939
0.000328054        0.000607518       -0.000198262 }
        site O3     x 0.42755      y  0.49844      z  0.56556   occ  O  1  ADPs { 0.008224576 0.042395606        0.00754346        -
0.008201338        0.002816675       -0.005947871 }
        site O4     x 0.09457      y  0.48604      z  0.11142   occ  O  1  ADPs { 0.0080875   0.030619049        0.006358536       -
0.001312214        0.001850169        0.001189574 }
        site O5     x 0.05091      y  0.36080      z  0.23479   occ  O  1  ADPs { 0.003335523 0.025908426        0.00664225        -
0.000328054        0.001076964       -0.002180886 }
        site O6     x -0.30237     y  0.36980      z  0.19878   occ  O  1  ADPs { 0.004889054 0.054172163        0.007576838       -
0.001968321        0.001933012       -0.000594787 }
        site C1p    x 0.19181      y  0.10398      z  0.09644   occ  C  1  ADPs { 0.005071822 0.037684983        0.006441982
0.002624428        0.000856048        0.000594787 }
        site C2p    x 0.03239      y  0.14025      z  0.12064   occ  C  1  ADPs { 0.004934746 0.030619049        0.006008066
0.000656107        0.001215036        0.000793049 }
        site C3p    x 0.05416      y  0.20917      z  0.08064   occ  C  1  ADPs { 0.004020904 0.028263737        0.00737657
0.000656107        0.002015856        0.001387836 }
        site C4p    x 0.23610      y  0.23129      z  0.28504   occ  C  1  ADPs { 0.004614901 0.030619049        0.005540772
0.000656107        0.001684482        0.000594787 }
        site C5p    x 0.38571      y  0.19233      z  0.24441   occ  C  1  ADPs { 0.00502613  0.03297436         0.006024755
0.002296375        0.001905398        0.001387836 }
```

```
         site C6p    x 0.56833     y  0.20924     z  0.45416   occ  C  1  ADPs { 0.004066596 0.044750917        0.00934588
0.002296375        0.001159807       -0.001387836 }
         site O1p    x 0.19004     y  0.10699     z  -0.19173  occ  O  1  ADPs { 0.006214124 0.063593409        0.005474015
0.004592749        0.001739711       -0.00337046  }
         site O2p    x -0.13082    y  0.11634     z  -0.07738  occ  O  1  ADPs { 0.005757203 0.028263737        0.007943998        -
0.000984161        0.001629253        0.000991312 }
         site O3p    x -0.08848    y  0.23986     z  0.14357   occ  O  1  ADPs { 0.004066596 0.030619049        0.013384635
0.000656107        0.00334135        -0.000594787 }
         site O5p    x 0.35420     y  0.12918     z  0.29925   occ  O  1  ADPs { 0.006168432 0.03297436         0.006742385
0.003280535        0.001215036        0.001784361 }
         site O6p    x 0.70387     y  0.17312     z  0.39816   occ  O  1  ADPs { 0.005163206 0.056527475        0.011331879
0.004920803        0.003313736        0.007533969 }
         site O12    x -0.47115    y  0.05777     z  -0.08802  occ  O  1  ADPs { 0.006671045 0.044750917        0.008578183
0.002296375        0.002761446        0.002577411 }
         site HC1    x 0.23400     y  0.32700     z  0.57200   occ  H  1  beq 1.8
         site HC2    x 0.36700     y  0.40200     z  0.22100   occ  H  1  beq 2.8
         site HC3    x 0.31900     y  0.43500     z  0.77800   occ  H  1  beq 3.8
         site HC4    x 0.08300     y  0.50200     z  0.51300   occ  H  1  beq 2.4
         site HC5    x 0.00900     y  0.39300     z  0.60100   occ  H  1  beq 2.9
         site HC6A   x -0.20800    y  0.44200     z  0.01800   occ  H  1  beq 2.9
         site HC6B   x -0.22900    y  0.45800     z  0.32400   occ  H  1  beq 3.6
         site HO2    x 0.58600     y  0.36700     z  0.49900   occ  H  1  beq 3.3
         site HO3    x 0.43900     y  0.51500     z  0.71400   occ  H  1  beq 4.2
         site HO4    x 0.10100     y  0.52000     z  0.11900   occ  H  1  beq 4.1
         site HO6    x -0.35100    y  0.36300     z  0.04500   occ  H  1  beq 4.3
         site HC1p   x 0.19200     y  0.05500     z  0.15900   occ  H  1  beq 1.0
         site HC2p   x 0.02400     y  0.13600     z  0.32000   occ  H  1  beq 2.3
         site HC3p   x 0.05300     y  0.21900     z  -0.1280   occ  H  1  beq 3.8
         site HC4p   x 0.24700     y  0.22900     z  0.49100   occ  H  1  beq 1.4
         site HC5p   x 0.38900     y  0.19700     z  0.03500   occ  H  1  beq 2.1
         site HC6pA  x 0.58400     y  0.20000     z  0.67900   occ  H  1  beq 3.3
         site HC6pB  x 0.59600     y  0.25700     z  0.47800   occ  H  1  beq 4.3
         site HO1p   x 0.28200     y  0.09600     z  -0.1950   occ  H  1  beq 5.0
         site HO2p   x -0.13000    y  0.12800     z  -0.2310   occ  H  1  beq 5.6
         site HCO3p  x -0.07200    y  0.29000     z  0.17100   occ  H  1  beq 6.8
         site HCO6p  x 0.74600     y  0.19000     z  0.28500   occ  H  1  beq 9.8
         site HO12A  x -0.39900    y  0.07600     z  -0.0330   occ  H  1  beq 3.1
         site HO12B  x -0.47300    y  0.03900     z  -0.2540   occ  H  1  beq 6.3

str
                  phase_name "Cocoa-butter-V"
         ' COORDINATES TAKEN DIRECTLY FROM VAN MECHELEN, PESCHAR, SCHENK
         ' ACTA CRYST B62, 1131-1138 [2006)
         ' "Structures of mono-unsaturated triacylglycerols. II. The [beta]2 polymorph"
         ' Note that this synchrotron powder structure was solved from data collected at 280K
         ' Structure available in CSD as JEMSAW01
         ' Note that Uiso values from the paper were converted to Biso values using B=U*8*pi^2
                  volume 5703.967
                  space_group "Cc"
```

```
a @ 5.44219
b @ 127.638
c @ 8.21365
al 90
be @ 88.6936
ga 90
scale @ 0.0001
site O1     x 0.603        y 0.348190     z 0.598        occ O    1     beq  2
site C1     x 0.68371      y 0.357354     z 0.6518       occ C    1     beq  2
site C2     x 0.47928      y 0.365219     z 0.66312      occ C    1     beq  2
site O2     x 0.8931       y 0.359118     z 0.6857       occ O    1     beq  2
site C3     x 0.51784      y 0.372484     z 0.80685      occ C    1     beq  2
site C4     x 0.33235      y 0.381372     z 0.80412      occ C    1     beq  2
site C5     x 0.38527      y 0.389169     z 0.93863      occ C    1     beq  2
site C6     x 0.19655      y 0.397947     z 0.93812      occ C    1     beq  2
site C7     x 0.24505      y 0.405566     z 1.07587      occ C    1     beq  2
site C8     x 0.05385      y 0.414243     z 1.07763      occ C    1     beq  2
site C9     x 0.09800      y 0.421613     z 1.21924      occ C    1     beq  2
site C10    x -0.08692     y 0.430528     z 1.21841      occ C    1     beq  2
site C11    x -0.03768     y 0.437987     z 1.35816      occ C    1     beq  2
site C12    x -0.22171     y 0.446943     z 1.35846      occ C    1     beq  2
site C13    x -0.16787     y 0.454479     z 1.49637      occ C    1     beq  2
site C14    x -0.35402     y 0.463352     z 1.49917      occ C    1     beq  2
site C15    x -0.30580     y 0.470626     z 1.64139      occ C    1     beq  2
site C16    x -0.49226     y 0.479490     z 1.64469      occ C    1     beq  2
site C17    x -0.4343      y 0.486961     z 1.78273      occ C    1     beq  2
site C18    x -0.62471     y 0.495636     z 1.7726       occ C    1     beq  2
site C19    x 0.78567      y 0.339971     z 0.57037      occ C    1     beq  2
site C20    x 0.70291      y 0.332615     z 0.43986      occ C    1     beq  2
site O3     x 0.89898      y 0.325175     z 0.3990       occ O    1     beq  2
site C21    x 0.83165      y 0.315526     z 0.3576       occ C    1     beq  2
site C22    x 1.05153      y 0.308982     z 0.31152      occ C    1     beq  2
site O4     x 0.6261       y 0.312191     z 0.3720       occ O    1     beq  2
site C23    x 0.97101      y 0.299144     z 0.22300      occ C    1     beq  2
site C24    x 1.19162      y 0.292113     z 0.18856      occ C    1     beq  2
site C25    x 1.11046      y 0.282154     z 0.10353      occ C    1     beq  2
site C26    x 1.33170      y 0.275214     z 0.06556      occ C    1     beq  2
site C27    x 1.24974      y 0.265248     z -0.01885     occ C    1     beq  2
site C28    x 1.46423      y 0.257677     z -0.03679     occ C    1     beq  2
site C29    x 1.4086       y 0.249516     z -0.16220     occ C    1     beq  2
site C30    x 1.3619       y 0.23946      z -0.13942     occ C    1     beq  2
site C31    x 1.36144      y 0.233769     z 0.02042      occ C    1     beq  2
site C32    x 1.16602      y 0.22525      z 0.02199      occ C    1     beq  2
site C33    x 1.23767      y 0.216439     z 0.13622      occ C    1     beq  2
site C34    x 1.03439      y 0.20827      z 0.14562      occ C    1     beq  2
site C35    x 1.10309      y 0.199572     z 0.26258      occ C    1     beq  2
site C36    x 0.89819      y 0.191482     z 0.27325      occ C    1     beq  2
site C37    x 0.96811      y 0.18271      z 0.38848      occ C    1     beq  2
site C38    x 0.77258      y 0.174212     z 0.38886      occ C    1     beq  2
```

```
site C39    x 0.63406       y 0.338670      z 0.29087       occ C    1      beq  2
site O5     x 0.8047        y 0.347361      z 0.26366       occ O    1      beq  2
site C40    x 0.84143       y 0.350474      z 0.11047       occ C    1      beq  2
site C41    x 1.06842       y 0.357090      z 0.09409       occ C    1      beq  2
site O6     x 0.7151        y 0.34798       z -0.00092      occ O    1      beq  2
site C42    x 1.09178       y 0.364061      z 0.24293       occ C    1      beq  2
site C43    x 0.91126       y 0.373126      z 0.23479       occ C    1      beq  2
site C44    x 0.94676       y 0.380427      z 0.37855       occ C    1      beq  2
site C45    x 0.77279       y 0.389709      z 0.36857       occ C    1      beq  2
site C46    x 0.81318       y 0.397047      z 0.51121       occ C    1      beq  2
site C47    x 0.63532       y 0.406206      z 0.50522       occ C    1      beq  2
site C48    x 0.67719       y 0.413473      z 0.64844       occ C    1      beq  2
site C49    x 0.49873       y 0.422616      z 0.64408       occ C    1      beq  2
site C50    x 0.53955       y 0.42979       z 0.78862       occ C    1      beq  2
site C51    x 0.35856       y 0.43885       z 0.78623       occ C    1      beq  2
site C52    x 0.39903       y 0.446002      z 0.93110       occ C    1      beq  2
site C53    x 0.21113       y 0.454811      z 0.93262       occ C    1      beq  2
site C54    x 0.25885       y 0.462247      z 1.07290       occ C    1      beq  2
site C55    x 0.07097       y 0.471058      z 1.07434       occ C    1      beq  2
site C56    x 0.11072       y 0.478197      z 1.21948       occ C    1      beq  2
site C57    x -0.07932      y 0.486927      z 1.22207       occ C    1           beq  2
site H1     x -0.6029       y 0.49943       z 1.6667        occ H    1           beq  4
site H2     x -0.7934       y 0.49253       z 1.7798        occ H    1      beq  4
site H3     x -0.6030       y 0.50066       z 1.8646        occ H    1      beq  4
site H4     x -0.2663       y 0.48997       z 1.7644        occ H    1      beq  4
site H5     x -0.4946       y 0.48145       z 1.8604        occ H    1      beq  4
site H6     x -0.6611       y 0.47654       z 1.6622        occ H    1      beq  4
site H7     x -0.4837       y 0.48335       z 1.5387        occ H    1      beq  4
site H8     x -0.1365       y 0.47361       z 1.6293        occ H    1      beq  4
site H9     x -0.3197       y 0.46659       z 1.7455        occ H    1      beq  4
site H10    x -0.5235       y 0.46038       z 1.5110        occ H    1      beq  4
site H11    x -0.3396       y 0.46739       z 1.3950        occ H    1      beq  4
site H12    x -0.1774       y 0.45063       z 1.6024        occ H    1      beq  4
site H13    x 0.0011        y 0.45744       z 1.4800        occ H    1      beq  4
site H14    x -0.3917       y 0.44407       z 1.3738        occ H    1      beq  4
site H15    x -0.2091       y 0.45075       z 1.2522        occ H    1      beq  4
site H16    x 0.1324        y 0.44088       z 1.3455        occ H    1      beq  4
site H17    x -0.0529       y 0.43410       z 1.4636        occ H    1      beq  4
site H18    x -0.2568       y 0.42761       z 1.2312        occ H    1      beq  4
site H19    x -0.0718       y 0.43440       z 1.1128        occ H    1      beq  4
site H20    x 0.2681        y 0.42454       z 1.2092        occ H    1      beq  4
site H21    x 0.0801        y 0.41765       z 1.3239        occ H    1      beq  4
site H22    x -0.1140       y 0.41111       z 1.0897        occ H    1      beq  4
site H23    x 0.0661        y 0.41825       z 0.9731        occ H    1           beq  4
site H24    x 0.4123        y 0.40870       z 1.0599        occ H    1      beq  4
site H25    x 0.2370        y 0.40174       z 1.1820        occ H    1      beq  4
site H26    x 0.0281        y 0.39492       z 0.9529        occ H    1      beq  4
site H27    x 0.2079        y 0.40175       z 0.8318        occ H    1      beq  4
site H28    x 0.5533        y 0.39218       z 0.9204        occ H    1      beq  4
```

```
site H29    x 0.3775       y 0.38551      z 1.0462       occ H    1      beq  4
site H30    x 0.1631       y 0.37846      z 0.8210       occ H    1      beq  4
site H31    x 0.3436       y 0.38500      z 0.6963       occ H    1      beq  4
site H32    x 0.4963       y 0.36845      z 0.9106       occ H    1      beq  4
site H33    x 0.6879       y 0.37544      z 0.8001       occ H    1      beq  4
site H34    x 0.4769       y 0.36943      z 0.5605       occ H    1      beq  4
site H35    x 0.3189       y 0.36148      z 0.6781       occ H    1      beq  4
site H36    x 0.9457       y 0.34323      z 0.5360       occ H    1      beq  4
site H37    x 0.8079       y 0.33597      z 0.6738       occ H    1      beq  4
site H38    x 0.6100       y 0.17732      z 0.4209       occ H    1      beq  4
site H39    x 0.7629       y 0.17110      z 0.2775       occ H    1      beq  4
site H40    x 0.8164       y 0.16863      z 0.4684       occ H    1      beq  4
site H41    x 0.9828       y 0.18559      z 0.5010       occ H    1      beq  4
site H42    x 1.1294       y 0.17965      z 0.3521       occ H    1      beq  4
site H43    x 0.8703       y 0.18850      z 0.1626       occ H    1      beq  4
site H44    x 0.7438       y 0.19491      z 0.3146       occ H    1      beq  4
site H45    x 1.1294       y 0.20262      z 0.3730       occ H    1          beq  4
site H46    x 1.2581       y 0.19613      z 0.2226       occ H    1      beq  4
site H47    x 1.0101       y 0.20523      z 0.0350       occ H    1      beq  4
site H48    x 0.8783       y 0.21166      z 0.1849       occ H    1      beq  4
site H49    x 1.2646       y 0.21936      z 0.2473       occ H    1      beq  4
site H50    x 1.3928       y 0.21309      z 0.0945       occ H    1      beq  4
site H51    x 1.1494       y 0.22240      z -0.0906      occ H    1      beq  4
site H52    x 1.0053       y 0.22830      z 0.0597       occ H    1      beq  4
site H53    x 1.5263       y 0.23052      z 0.0363       occ H    1      beq  4
site H54    x 1.3256       y 0.23884      z 0.1107       occ H    1      beq  4
site H55    x 1.3302       y 0.23615      z -0.2472      occ H    1      beq  4
site H56    x 1.3973       y 0.25085      z -0.2818      occ H    1      beq  4
site H57    x 1.4927       y 0.25419      z 0.0702       occ H    1      beq  4
site H58    x 1.6153       y 0.26163      z -0.0714      occ H    1      beq  4
site H59    x 1.1879       y 0.26703      z -0.1291      occ H    1      beq  4
site H60    x 1.1150       y 0.26187      z 0.0472       occ H    1      beq  4
site H61    x 1.4140       y 0.27335      z 0.1692       occ H    1      beq  4
site H62    x 1.4503       y 0.27908      z -0.0073      occ H    1      beq  4
site H63    x 1.0285       y 0.28403      z -0.0002      occ H    1      beq  4
site H64    x 0.9917       y 0.27827      z 0.1761       occ H    1      beq  4
site H65    x 1.3131       y 0.29592      z 0.1171       occ H    1      beq  4
site H66    x 1.2701       y 0.29025      z 0.2936       occ H    1      beq  4
site H67    x 0.8476       y 0.29528      z 0.2923       occ H    1          beq  4
site H68    x 0.8950       y 0.30115      z 0.1178       occ H    1      beq  4
site H69    x 1.1648       y 0.31313      z 0.2387       occ H    1      beq  4
site H70    x 1.1395       y 0.30695      z 0.4123       occ H    1      beq  4
site H71    x 0.5557       y 0.32866      z 0.4813       occ H    1      beq  4
site H72    x -0.0698      y 0.49086      z 1.1168       occ H    1      beq  4
site H73    x -0.2473      y 0.48385      z 1.2371       occ H    1      beq  4
site H74    x -0.0461      y 0.49182      z 1.3140       occ H    1      beq  4
site H75    x 0.2792       y 0.48131      z 1.2112       occ H    1      beq  4
site H76    x 0.0945       y 0.47403      z 1.3222       occ H    1      beq  4
site H77    x -0.0982      y 0.46802      z 1.0818       occ H    1      beq  4
```

```
site H78    x 0.0894        y 0.47521       z 0.9715        occ H    1      beq  4
site H79    x 0.4276        y 0.46527       z 1.0596        occ H    1      beq  4
site H80    x 0.2463        y 0.45832       z 1.1780        occ H    1      beq  4
site H81    x 0.0423        y 0.45179       z 0.9458        occ H    1      beq  4
site H82    x 0.2238        y 0.45874       z 0.8275        occ H    1      beq  4
site H83    x 0.5682        y 0.44905       z 0.9232        occ H    1      beq  4
site H84    x 0.3811        y 0.44186       z 1.0340        occ H    1      beq  4
site H85    x 0.1870        y 0.43604       z 0.7918        occ H    1      beq  4
site H86    x 0.3834        y 0.44292       z 0.6831        occ H    1      beq  4
site H87    x 0.7111        y 0.43260       z 0.7828        occ H    1      beq  4
site H88    x 0.5152        y 0.42573       z 0.8919        occ H    1      beq  4
site H89    x 0.3264        y 0.41990       z 0.6488        occ H    1          beq  4
site H90    x 0.5261        y 0.42664       z 0.5408        occ H    1      beq  4
site H91    x 0.8494        y 0.41620       z 0.6424        occ H    1      beq  4
site H92    x 0.6514        y 0.40951       z 0.7524        occ H    1      beq  4
site H93    x 0.4628        y 0.40351       z 0.5108        occ H    1      beq  4
site H94    x 0.6626        y 0.41015       z 0.4012        occ H    1      beq  4
site H95    x 0.9855        y 0.39976       z 0.5057        occ H    1      beq  4
site H96    x 0.7861        y 0.39313       z 0.6155        occ H    1      beq  4
site H97    x 0.5992        y 0.38714       z 0.3722        occ H    1      beq  4
site H98    x 0.8048        y 0.39356       z 0.2642        occ H    1      beq  4
site H99    x 1.1202        y 0.38301       z 0.3773        occ H    1      beq  4
site H100   x 0.9125        y 0.37652       z 0.4821        occ H    1      beq  4
site H101   x 0.7395        y 0.37033       z 0.2383        occ H    1      beq  4
site H102   x 0.9400        y 0.37708       z 0.1310        occ H    1      beq  4
site H103   x 1.0558        y 0.35986       z 0.3434        occ H    1      beq  4
site H104   x 1.2630        y 0.36688       z 0.2468        occ H    1      beq  4
site H105   x 1.2159        y 0.35244       z 0.0842        occ H    1      beq  4
site H106   x 1.0582        y 0.36155       z -0.0056       occ H    1      beq  4
site H107   x 0.4634        y 0.34148       z 0.3058        occ H    1      beq  4
site H108   x 0.6406        y 0.33391       z 0.1941        occ H    1      beq  4
```

Chapter 5

Physical Characterisation of Dried Lactose Syrups: An Investigation

Draft planned for submission to "Food Analytical Methods":

Nicholls, D.; Shankland, K.

5 Physical Characterisation of Dried Lactose Syrups: An Investigation

Foreword

The following paper is planned for submission to *Food Analytical Methods*. It consists of the investigations into the properties of an unidentified crystalline form of lactose formed by oven-drying concentrated aqueous solutions of lactose. These include the thermal characterisation of the dried syrup in contrast to α-lactose monohydrate by differential scanning calorimetry and hot-stage microscopy; the determination of the chemical composition of the dried syrup to be a mixture of the α- and β- anomers of lactose by ^{1}H-, ^{13}C, and ^{13}C-DEPT NMR; and the analysis of powder X-ray diffraction (PXRD) data collected of the sample in contrast to currently published lactose crystal structures. The use of a random peak selection script to attempt indexing of the powder pattern gave interesting results but the presence of multiple crystal forms of lactose in the sample prevented further progress towards structure solution.

I certify that I performed all the experimental work that is reported, performed all the data analysis and wrote the majority of the manuscript. Kenneth Shankland was consulted in the practicalities of collecting the PXRD data, aided in the development of *TOPAS* input files that screened powder patterns against known crystal forms, and reviewed and edited the manuscript for submission.

Physical Characterisation of Dried Lactose Syrups: An Investigation

Daniel Nicholls and Kenneth Shankland

School of Pharmacy, University of Reading, Whiteknights, Reading RG6 6AD, UK

1 Abstract

Concentrated (approximately 40 % w/v) aqueous solutions of D-lactose were oven dried producing a solid cake of mixed crystalline phases of lactose, one of which exhibits a previously unidentified powder X-ray diffraction pattern. The sample was confirmed to be a mixture of alpha and beta anomers of lactose and its physical properties were characterised by hot-stage microscopy, differential scanning calorimetry, nuclear magnetic resonance and powder X-ray diffraction. High quality patterns obtained from the mixture using the latter technique were collected, but difficulties in powder indexing prevented a full explanation of the observed X-ray data.

Supplemental NMR spectra are included at the end of this paper.

2 Introduction

Chocolate exists as a semi-solid dispersion of fine sugars and cocoa solids throughout a continuous fat phase. The primary fat in chocolate, cocoa butter, is relatively well understood and characterised – its importance in tempering has led to a significant amount of research into controlling polymorphism of cocoa butter during cooling (Tewkesbury et al., 2000). However, the sugars (i.e. lactose and sucrose) in chocolate are much less well understood. The physical

form of the sugars in chocolate confectionary can affect both the processability of the product and its final mouth feel and texture. High levels of amorphous sugar, for example, result in reduced amounts of free fat in the matrix which influences how much extra fat is required for processing in downstream stages. Furthermore, amorphous sugars require less energy to solubilise them (Yff et al., 2004), which means that on initial contact they will be detected as sweeter in the mouth than their crystalline counterparts.

The processes and material developments that chocolate products undergo during manufacture control the final crystal forms of the ingredients of the confectionary (Fryer and Pinschower, 2000). One particular method of chocolate production involves an intermediary material referred to as "chocolate crumb", a vacuum dried and crystalline mixture of milk, cocoa butter and sugar. Chocolate crumb has been shown to have a substantial impact on the flavour profile of the resultant chocolate due to the caramelisation of sugars during processing (Afoakwa et al., 2007; Beckett, 2011). Additionally, the reduced water activity of chocolate crumb confers greater shelf life when compared with milk powders, an advantage that was historically beneficial in allowing for year-round chocolate production (Beckett, 2011).

Due to the importance of the processing in chocolate crumb in flavour development, manufacturing processes are mostly proprietary. However most production methods follow a similar schematic. Pasteurised milk is concentrated to approximately 40% solids in an evaporator, where sugar is then added to the mixture and further evaporated to 90% solids. This intermediary is commonly known as sweetened condensed milk (SCM). Finally, cocoa butter and other cocoa solids are then added to the SCM and further drying processes bring the total solid concentration up to 99% (Edmondson et al., 2005) Typically, drying processes occur at temperatures greater than 60 °C; however the exact temperatures are often held secret by chocolate manufacturers (Muresan et al., 2000).

Previous unpublished work conducted by Nicholls has shown that mixtures of sucrose and lactose behave differently in mixtures than if they were analysed independently. Differential scanning calorimetry (DSC) of sucrose/lactose mixtures shows a shift in the melting temperature of alpha lactose monohydrate due to an interaction with melting sucrose. Most investigations into sugar interactions during food manufacture are concentrated on the Maillard reaction and its importance in flavour development (Labuza et al., 1998; O'Brien et al., 1998) or the rheological behaviour of chocolate with different sugar crystalline forms (Babin, 2005; Aguilar and Ziegler, 1995). Our focus is on the changes in the sugar crystalline

forms and their relative amounts during chocolate crumb and chocolate manufacture, noting that amorphous sugars may also be present (Nicholls et al., 2018).

This body of work began as an investigation into sugar interactions that occur during the sweetened condensed milk drying process. However, early on in this work, it was discovered that powders obtained from drying pure lactose syrups gave powder X-ray diffraction (PXRD) patterns that did not conform to any currently known crystal forms of lactose in the literature or in the Cambridge Structural Database (CSD; see table 1) Here, we report the physical characterisation of the powders obtained by drying.

Table 1. Reported crystalline forms of D-lactose for which crystal structures are available.

Form	Abbreviation	SpGrp	Z	Z'	T (K)	Refcode	Type	Ref.
α-lactose monohydrate	α-L.H_2O	$P2_1$	2	1	150	LACTOS11	SX	(Smith et al., 2005)
β-lactose	β-L	$P2_1$	2	1	293	BLACTO	SX	(Hirotsu and Shimada, 1974)
α-lactose (hygroscopic)	α-LH	$P2_1$	2	1	293	EYOCUQ	PXRD	(Platteau et al., 2004)
α-lactose (stable anhydrous)	α-LS	$P1$	2	2	293	EYOCUQ01	PXRD	(Platteau et al., 2005)
αβ-lactose	αβ-L_T	$P1$	2	2	120	LAKKEO01	SX	(Guiry et al., 2008)

SpGrp = Space group; Refcode = Cambridge Structural Database refcode; SX = structure obtained from single-crystal X-ray diffraction; PXRD = structure obtained from powder X-ray diffraction

3 Experimental

3.1. Sample preparation

Batches of the dried syrup cake were generated as follows. 40 grams of α-lactose monohydrate (Sigma, BN: SLBQ8461V) were dissolved in 100 mL of purified water at 45 °C in a 250 mL Pyrex beaker using a stirring hotplate. The lactose had visibly dissolved after approximately

40 minutes but stirring was continued on the hotplate for a further hour in order to ensure complete dissolution. The beaker was sealed with Parafilm to reduce water loss by evaporation during the dissolution process. 6 mL aliquots of the syrup were measured using a glass pipette and poured into 5.5 cm Petri dishes before being placed in a Lenton Thermal Designs fan oven at 120 °C for 90 minutes to ensure complete drying of the samples. Samples were then removed from the oven and allowed to cool, the off-white cake produced removed from the dishes using a spatula, and stored in glass vials prior to analysis. Where appropriate, samples were lightly ground using an agate mortar and pestle in order to create free-flowing powders suitable for PXRD measurements.

3.2. Hot-stage microscopy (HSM)

A small quantity of α-lactose monohydrate and lightly ground dried syrup was placed onto separate slides. A second glass slide was used to distribute each powder over a small area in a thin layer. A slide cover was then placed over the area containing the sample.

The slides were then placed onto the hot stage of a Mettler Toledo Hot Stage microscope. The light source of the microscope was adjusted such that the image being recorded was not overly bright or dim. The microscope was initially focused at 5 times magnification by use of the microscope and PC display, then gradually increased to 20 times, with further focusing at each magnification.

The hot stage was programmed to run from 20 to 250 °C, at a ramp rate of 10 °C / minute. Camera recording was started on the PC, followed by initiating the hot stage ramp.

3.3. Differential Scanning Calorimetry (DSC)

Between 5 and 10 mg of each sample (α-lactose monohydrate and ground dried syrup) were accurately weighed and loaded into pans of identical weight and thickness (calibrated for the DSC), and each sample pan then placed in a Mettler Toledo DSC 1. Samples were run from 20 °C to 350 °C at a ramp rate of 10 °C/min. Data were analysed and integrated where appropriate using the Mettler Toledo STARe software.

3.4. Nuclear Magnetic Resonance (NMR)

NMR data were recorded using a Bruker Ascend 400 (400 MHz) spectrometer after dissolution of the samples in dimethyl sulphoxide (DMSO). ^{1}H and ^{13}C NMR spectra were both recorded at 400 MHz. Chemical shifts (δ) are quoted in parts per million (ppm) using the abbreviations described in Table 2. Integrations were done manually using MestreNova.

Table 2. NMR Abbreviations

Abbreviation	
s	singlet
d	doublet
dd	doublet of doublets
t	triplet
dt	doublet of triplets
ddt	doublet of doublet of triplets
td	triplet of doublets
q	quartet
m	multiplet
br	broad

3.5. Powder X-Ray Diffraction (PXRD)

Both α-lactose monohydrate and lightly ground dried syrup were loaded into a separate 0.7 mm borosilicate glass capillaries and mounted on a Bruker D8 Advance Diffractometer operating in transmission capillary geometry, with a LynxEye detector and monochromatic incident radiation of wavelength 1.54056 Å. Samples were scanned over the 4 to 45° 2θ range over a 4-hour data collection time. Data were analysed using the EVA (Bruker AXS, 2018), TOPAS (Coelho, 2018) and DASH (David et al., 2006) packages.

4 Results

4.1. Hot-stage Microscopy

The hot-stage microscope camera records an image of the powder over the course of the ramping experiment - Figures 1 and 2 show several snapshots of the recorded images of α-lactose monohydrate and the dried lactose syrup over the course of the experiment.

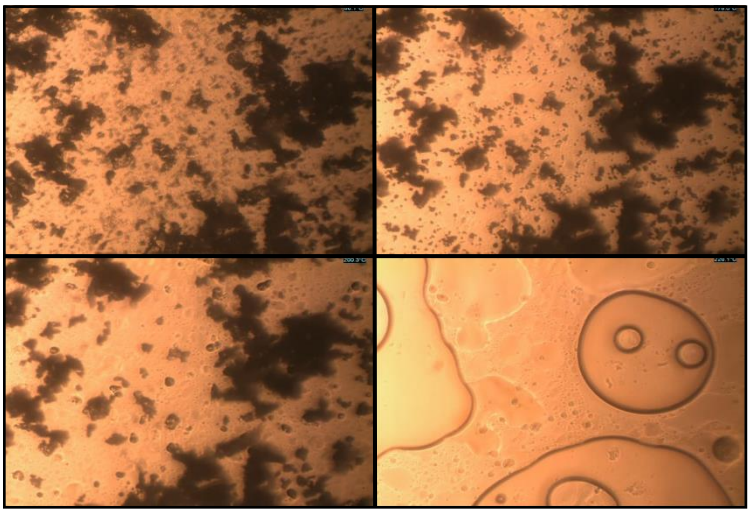


Figure 1. 20 x Magnification images of the α-lactose monohydrate sample during the course of the heating. Top-left: 50 °C ; Top-right: 175 °C ; Bottom-left: 200.3 °C ; Bottom-right: 220.1 °C.

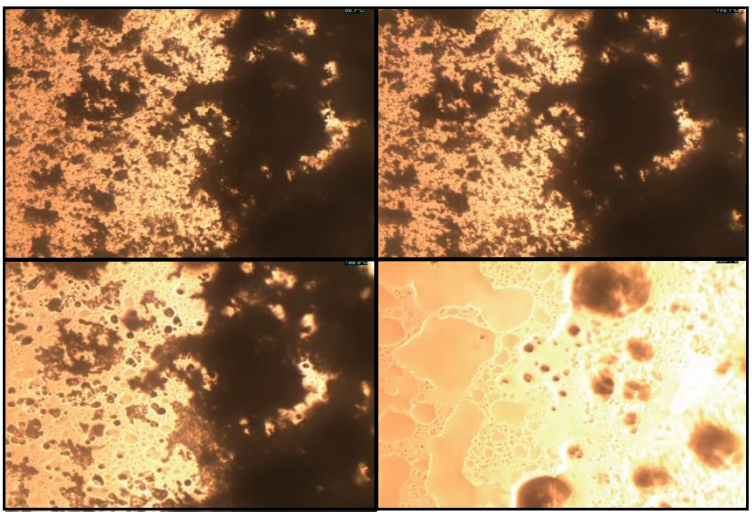


Figure 2. 20 x Magnification images of the new batch dried lactose syrup sample during the course of the heating. Top-left: 50.7 °C ; Top-right: 175.7 °C ; Bottom-left: 199.9 °C ; Bottom-right: 220.1 °C.

The light microscope also records the brightness of the light passing through the sample and reports the data as the derivative i.e. the change in intensity (arbitrary units) over time (seconds) at the current point in the temperature ramp. Figures 3 and 4 show these data plotted for each of the samples assessed.

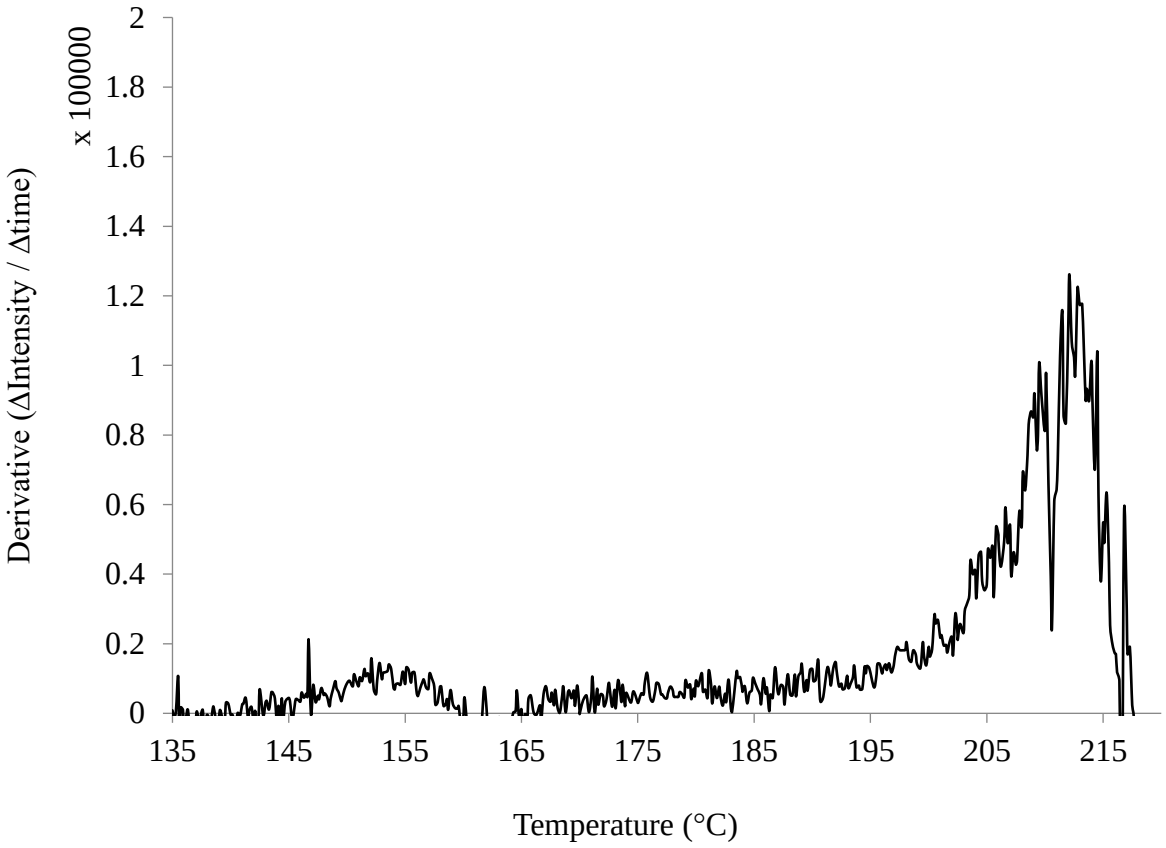


Figure 3. A plot of the change in light intensity over time against temperature for alpha-lactose monohydrate during the temperature ramp between 135 and 220 °C.

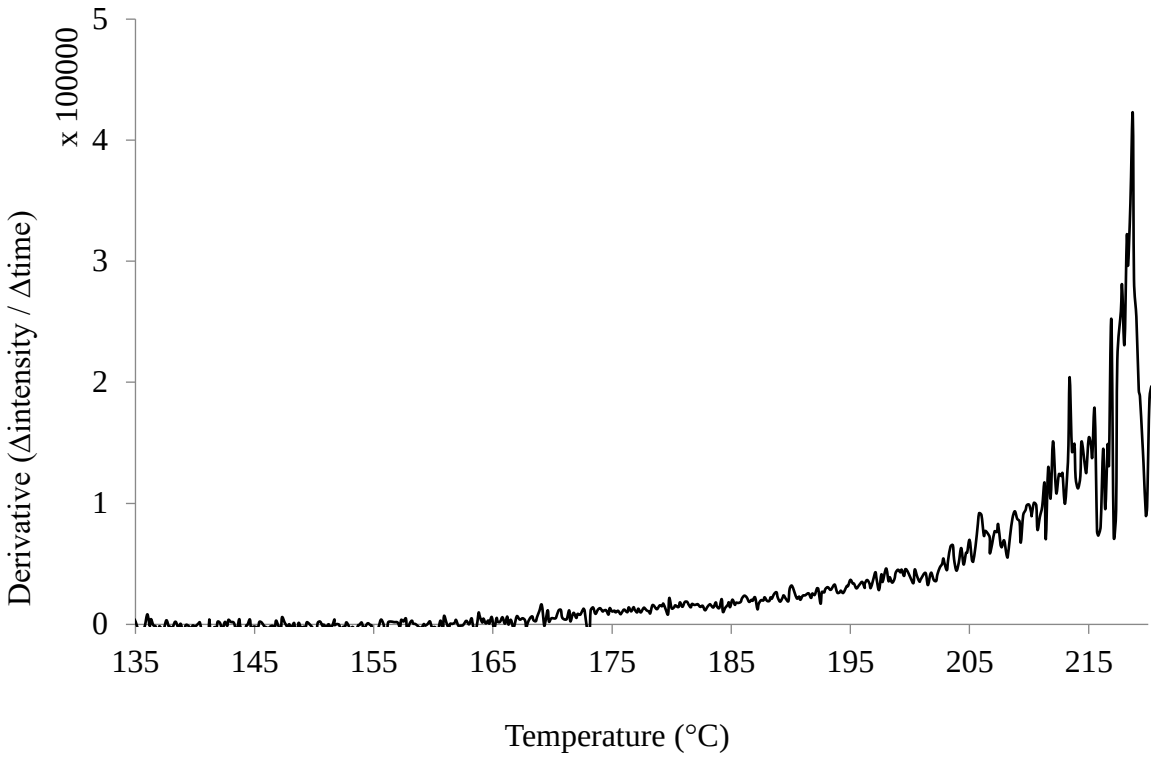


Figure 4. A plot of the change in light intensity over time against temperature for the dried lactose syrup during the temperature ramp between 135 and 220 °C.

4.2. Differential Scanning Calorimetry

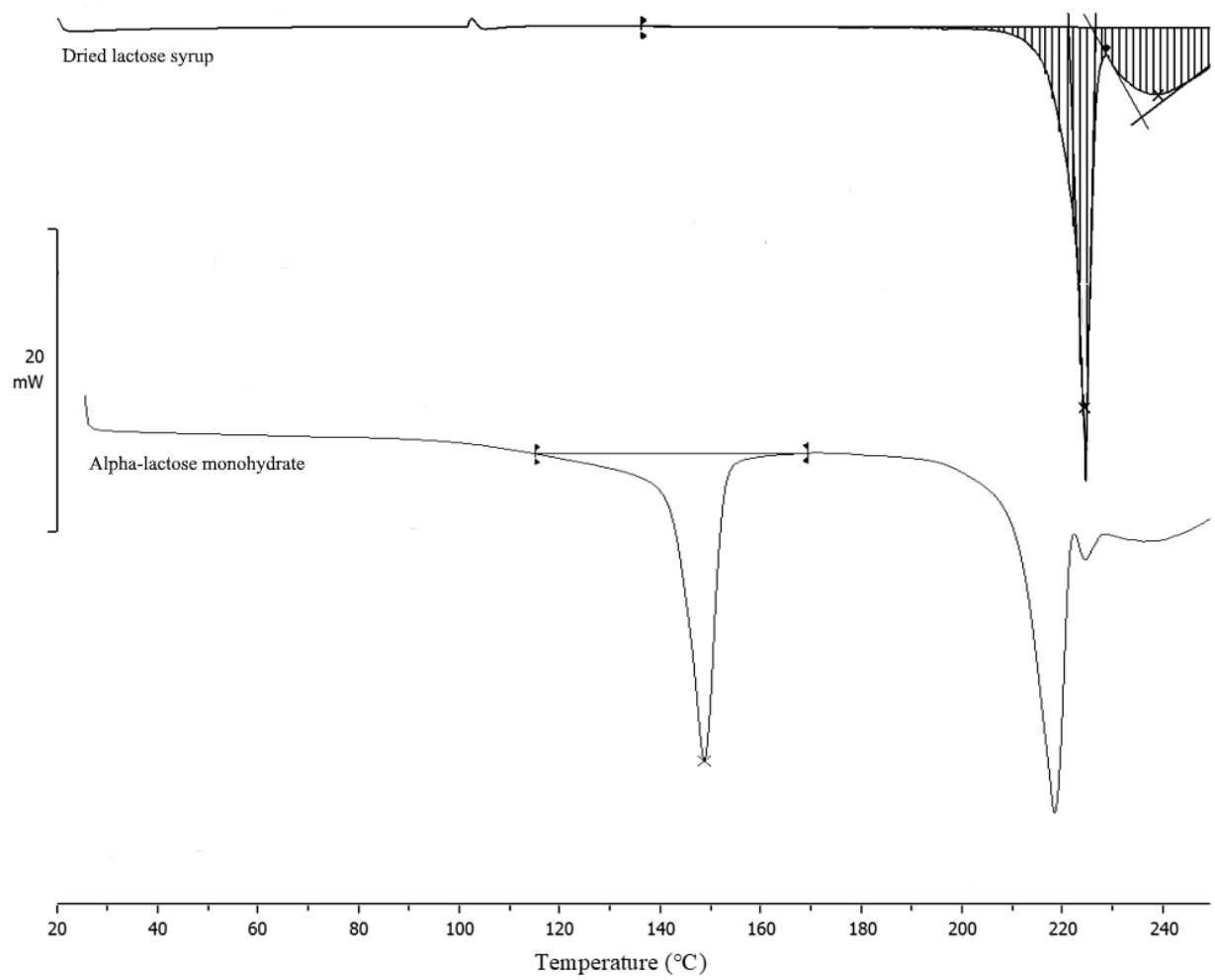


Figure 5. Enthalpy changes during alpha lactose monohydrate and dried lactose syrup DSC.

Table 3. Data from the integrated peaks of the dried syrup sample and α-lactose monohydrate

Sample	Mass of sample (mg)	Integral (mJ)	Normalised Integral (J/g)	Peak (°C)
Dried lactose syrup melt	5.232	-811.54	-155.11	222.8
α-lactose monohydrate dehydration	5.812	-810.91	-139.52	148.2
α-lactose monohydrate melt	5.812	n/a	n/a	219.2

4.3. Nuclear Magnetic Resonance

^{1}H, ^{13}C DEPT and ^{13}C NMR spectra recorded are shown across Figures S1 through S6. Table 4 details the chemical shifts corresponding to the labelled hydrogens in Figure 6 for alpha lactose monohydrate.

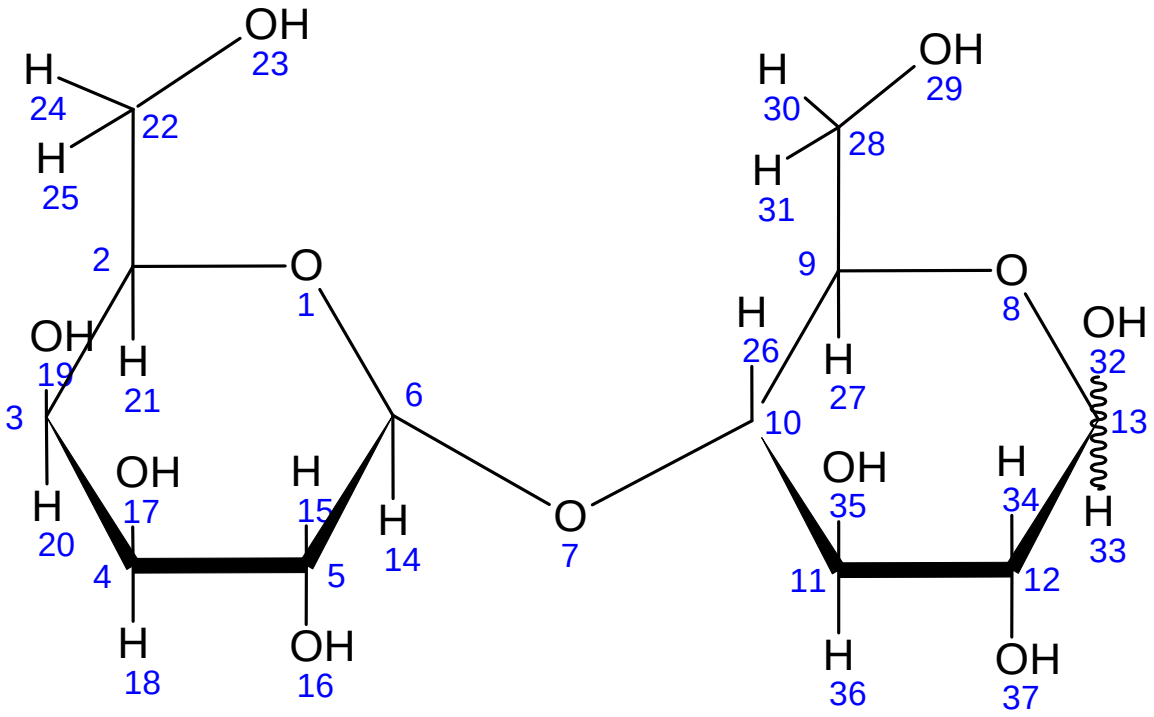


Figure 6. Labelled atoms of D-lactose, the anomeric centre that dictates whether the lactose is the alpha or beta polymorph occurs at the O32-C13-H33 stereocentre.

Table 4. Hydrogen assignments for ^{1}H-NMR of α-lactose monohydrate

Chemical shift / δ (ppm)	Integral	Hydrogen assignment	Multiplet
6.36	1.00	O32-H	d
5.12	1.03	O17-H	d
4.91	0.99	H33	t
4.81	1.02	O16-H	d
4.68[a]	2.06	O23-H; O37-H	m
4.55	1.03	O19-H	d
4.49[a]	2.02	O29-H; O35-H	m
4.19	1.00	H14	d
3.72	1.10	H27	dt
3.64[a]	3.00	H20; H30; H31	tt
3.53[a]	4.20	H22; H23; H36; H21	m
3.31[a]	3.10	H15; H18; H26	m
3.18	1.02	H34	ddd

[a] significant overlap between chemical shift peaks, individual integration not possible

4.4. Powder X-Ray Diffraction

The dried lactose syrup sample when scanned using PXRD exhibited a complicated pattern suggesting a mixture of multiple crystalline phases (Figure 7). Even when one such phase previously characterised (β-lactose) was identified by its characteristic peak at ~10.5 ° 2θ, the extent of peak overlap between phases made manual selection of a set of peaks that could be confidently said to belong to a single crystalline phase a near-impossible task.

To attempt to overcome this large degree of peak overlap, and in particular the presence of confounding peaks from other phases within the dried syrup powder, a Python script was written to randomly select 20 peak positions from the first 25 non-β phase peaks, and then index them using likely space groups in *TOPAS*. The cells generated by the indexing were assessed for goodness of fit to the full PXRD pattern and whether all randomly selected peaks were described by the indexed cell. Of these results, the space group $P2_1$ was found to have the best goodness of fit whilst still indexing all peaks selected – a sampling of the top ten fits is shown in table 5.

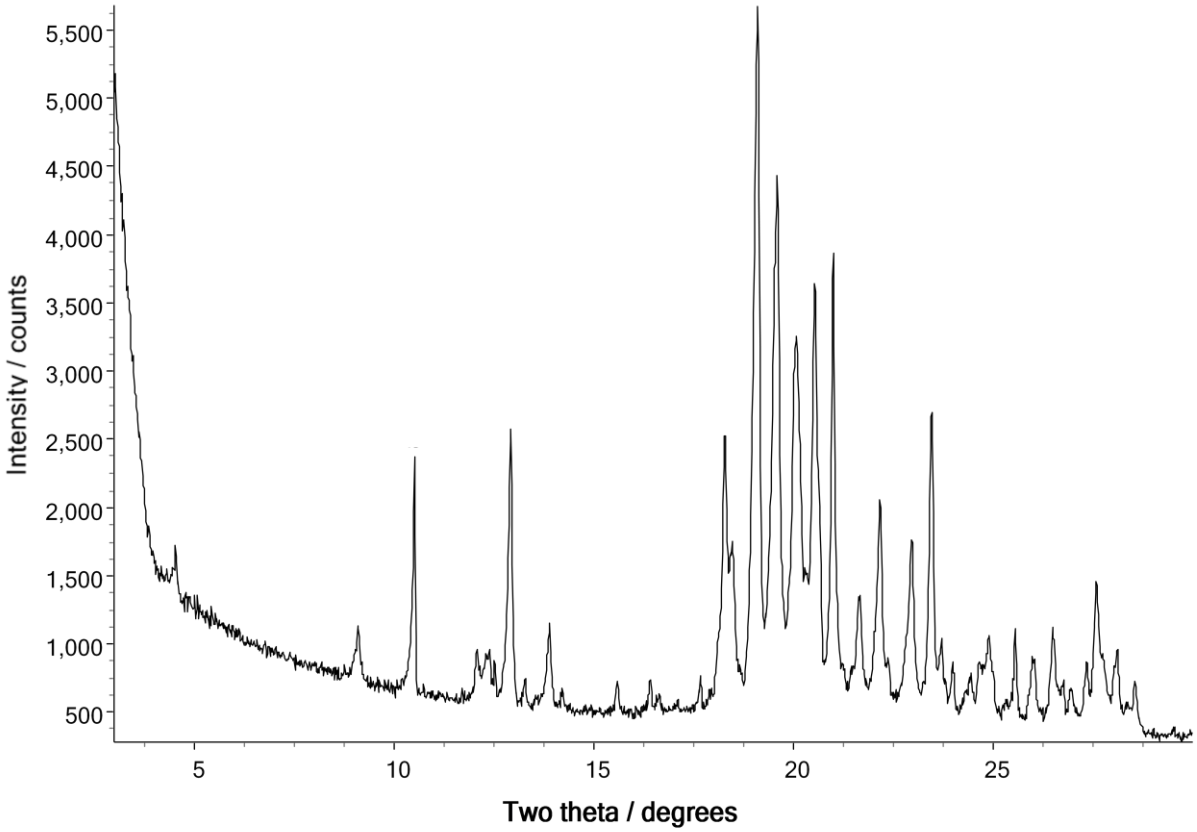


Figure 7. Powder X-ray diffraction pattern in the range 3-30° 2θ of the dried lactose syrup sample.

Table 5. Best results from the randomised peak indexing of PXRD data with space group $P2_1$

Number	Goodness of fit	Volume ($Å^3$)	a (Å)	b (Å)	c (Å)	β (°)
1	5.32	922.88	7.44	19.46	6.48	100.14
2	4.71	655.30	10.02	8.64	7.81	104.56
3	4.13	907.60	7.32	19.44	6.38	89.52
4	4.08	1365.30	7.33	38.84	4.80	92.53
5	3.98	725.40	15.65	4.77	10.03	75.61
6	3.85	1451.99	7.73	38.88	5.09	71.51
7	3.77	1404.44	7.54	38.88	4.93	103.62
8	3.65	1434.27	7.63	38.90	5.03	73.74
9	3.59	1409.39	15.11	9.73	10.99	119.26
10	3.57	916.95	15.65	6.03	10.03	75.61

5 Discussion

The alpha lactose monohydrate sample (Figure 1) shows a change in light intensity across the range of 143 - 164 °C; upon reviewal of the footage recorded, it is clear that the powder agitates at these temperatures. This is most likely due to the loss of water from the alpha lactose monohydrate crystal structure, reported to occur at 145 °C (Raut et al, 2011). The intensity procedes to rapidly increase as the temperature ramps, from 1.2 x 10^8 at 166 °C, to 1.75 x 10^8 at 216 °C. Visually this appears as the powder melting into a transparent liquid, permitting more light to pass through. Beyond 216 °C, the molten lactose decomposes into a brown liquid which moves across the slide, causing rapid fluctuations in the intensity of light passing through the sample.

The dried lactose syrup exhibits no significant visual or intensity variation events prior to 170 °C, suggesting that the syrup does not contain any hydrated crystalline forms. However, similarly to the alpha lactose monohydrate sample, the intensity increases as the temperature ramps after 170 °C as the sample melts and subsequently degrades. These observations suggest that there is no signficant amount of alpha lactose monohydrate present in the dried syrup, and

that it may well contain a crystalline form (or forms) the same as the one(s) that are formed by the solid-state dehydration of alpha lactose monohydrate

DSC data support this supposition. There are no significant energy changes in the dried syrup sample prior to the melting at 222.8 °C, whilst in the alpha lactose monohydrate sample a dehydration occurs at 148.2 °C with subsequent melting at 219.2 °C. Both samples degrade after melting, showed by the broad energy changes of both systems following melting.

Hydrogen NMR chemical shifts of alpha lactose monohydrate were relatively simple to assign with the exception of the clusters of multiplets seen at 3.64, 3.53 and 3.31 ppm. The dried lactose syrup, however, was much more challenging to assign. Many of the multiplets assigned in the alpha lactose monohydrate sample are still visible on the dried syrup spectrum although several shift integrals are now fractional. For example, the O32-H doublet assigned with a integral value of 1.00 on the alpha lactose monohydrate spectrum is integrated to 0.47 on the dried syrup spectrum as the same chemical shift (6.36 ppm). A new doublet downfield of this at 6.70 ppm is observed on the dried syrup spectrum – this is likely to be the anomer of the O32-H hydrogen as the sum of their integration values is close to one total hydrogen (1.01). Hence, the sample in solution appears to consist of both alpha and beta lactose anomers, suggesting that during the syrup dissolution and cooking process anomeric rotation has occurred and the dried syrup consists of both alpha and beta lactose. These “split” shifts occur throughout the NMR spectrum, with other shfits being split in a similar ratio. ^{13}C NMR and ^{13}C DEPT corroborate these findings, with both samples showing near identical shifts across the spectra, only varying in intensity.

As the combination of HSM, DSC and NMR strongly indicate that dried syrup consists of a mixture of the alpha and beta anomers, all known lactose crystal structures deposited in the Cambridge Structural Database (CSD) were fitted against the measured PXRD pattern in a multiphase Rietveld refinement. The scale factors for each crystalline phase was allowed to vary in the refinement in order to achieve the best possible fit to the data based on the known structure; the resultant 'best fit' is shown in Figure 8. It is clear from this figure that the observed PXRD cannot be explained on the basis the currently known crystal structures; there are regions in the pattern where the calculated intensity is significantly underestimated (e.g. the peak at 13.9° 2θ) and regions where the calculated intensity is significantly overestimated (e.g. the peak at 12.2° 2θ). Clearly the majority crystalline phase of the dried lactose syrup does not consist of any of the currently known structures deposited in the CSD.

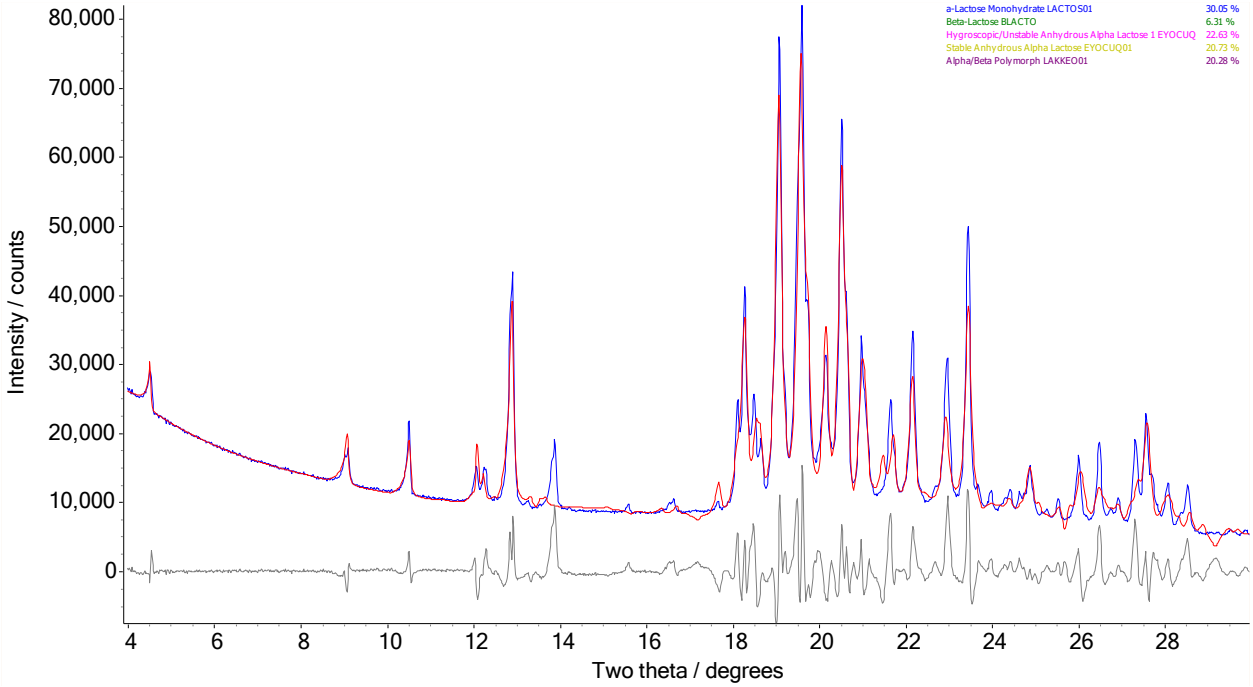


Figure 8: Attempted fit of dried lactose syrup powder pattern with currently known lactose crystal structures in the CSD. The calculated pattern (red) shows a poor fit with the observed pattern (blue), as can be seen by the large difference between the two patterns shown in light grey.

Randomly selecting peaks and indexing with *TOPAS* in order to identify the unknown component of the mixture gave relatively good fits to the data. However, knowing that the crystal structure does consist of lactose, few of the reported volumes corroborated well with the estimated molecular volume (403.49 $Å^3$, obtained by the method of Hoffman) or multiples thereof. The indexed cells with long *b* axes (i.e. numbers 4, 6, 7, 8) have cell volumes that are similar to lactose structures deposited in the CSD. For example, given that β-lactose has a cell volume of 705.285 $Å^3$ and a $Z = 2$, the novel crystalline phase may contain four molecules of lactose ($Z = 4$, suggesting Z'=2 for $P2_1$) leading to volumes of approximately 1400 $Å^3$, a value in reasonable agreement with, for example indexing solution 7 in Table 5 Unfortunately, none of the indexing solutions (combined with other combinations of known phases) ultimately allowed a complete description the powder pattern; the complexity and extent of overlap was simply too great and a likely long *b* axis was another confounding factor.

6 Conclusion

The dried lactose syrup was found to have similar physical properties to alpha lactose monohydrate, both melting and degrading at similar temperatures but with no enthalpy changes due to dehydration as seen in HSM and DSC. ^{1}H, ^{13}C and ^{13}C DEPT NMR showed that the substance consists of a mixture of the alpha and beta anomers of lactose when dissolved in a solution of DMSO, which was not observed in the spectra of alpha lactose monohydrate. Hence, the dried syrup consists of one or more crystal structures containing both alpha and beta lactose.

The powder data collected was unable to be fitted by any combination of the currently known crystal structures of lactose. Though randomly selecting peaks for indexing gave some potential structures for refinement, these were unable to be convincingly refined due to the mixture of crystal structures present in the powder pattern causing too much peak overlap and preventing the accurate extraction of intensities from the pattern. For this unidentified crystal structure of lactose to be solved from PXRD data alone, production of a phase-pure sample would be a significant step forward; however, this is not possible with the current method of sample production.

It is however, highly informative to look back at some of the putative cells in Table 5 in light of subsequent work in which a microcrystal, representative of the novel bulk phase, *was* eventually obtained from the dried lactose syrup (Nicholls et al, 2019). This single crystal allowed the new phase to be definitively characterised as crystallising in space group $P2_1$ with cell dimensions of 5.00 x 38.64 x 7.60, β=106.2°, V=1411.26Å at 150K; values which are in very good agreement with solution 7 of Table 5. It seems thus likely that with the application of an automated, systematic multi-phase combined Rietveld / Pawley refinement approach to full-pattern fitting, a convincing fit to the dried lactose syrup data *would* have eventually been obtained, though at considerable computational cost due to the problem complexity.

7 References

Afoakwa, E. O. et al. (2007) Factors influencing rheological and textural qualities in chocolate - a review. *Trends in Food Science and Technology*. [Online] 18 (6), 290–298.

Aguilar, C. A. & Ziegler, G. R. (1995) Viscosity of molten milk chocolate with lactose from spray-dried whole-milk powders. *Journal of food science*. 60 (1), 120–124.

Babin, H. (2005) *Colloidal properties of sugar particle dispersions in food oils with relevance to chocolate processing*.

Beckett, S. T. (2011) *Industrial chocolate manufacture and use*. John Wiley & Sons.

Edmondson, P. T. et al. (2005) Modelling of heat transfer, mass transfer and flavour development in chocolate crumb. *Food and Bioproducts Processing*. [Online] 83 (2 C), 89–98.

Fryer, P. & Pinschower, K. (2000) The materials science of chocolate. *Mrs Bulletin*. 25 (12), 25–29.

Guiry, K. P. et al. (2008) Effect of 1-Deoxy-d-lactose upon the Crystallization of d-Lactose. *Crystal Growth and Design*. 8 (11), 3927–3934.

Hirotsu, K. & Shimada, A. (1974) The crystal and molecular structure of β-lactose. *Bulletin of the Chemical Society of Japan*. 47 (8), 1872–1879.

Labuza, T. P. et al. (1998) *Maillard reactions in chemistry, food and health*. Elsevier.

Muresan, S. et al. (2000) Aroma profile development of intermediate chocolate productsI. Volatile constituents of block-milk. *Food chemistry*. 68 (2), 167–174.

Nicholls, D. et al. (2018) Rietveld-Based Quantitative Phase Analysis of Sugars in Confectionery. *Food Analytical Methods*. 11 (10), 2673–2681.

Nicholls, D. et al. (2019) A new crystalline form of *αβ*-D-lactose prepared by oven drying a concentrated aqueous solution of D-lactose. *Acta Crystallographica Section C*. 75 (7), 904-909.

O'Brien, J. et al. (1998) *The Maillard reaction in foods and medicine*. Vol. 223. Woodhead Publishing.

Platteau, C. et al. (2004) Ab initio structure determination of the hygroscopic anhydrous form of α-lactose by powder X-ray diffraction. *Acta Crystallographica Section B: Structural Science*. 60 (4), 453–460.

Platteau, C. et al. (2005) Structure determination of the stable anhydrous phase of α-lactose from X-ray powder diffraction. *Acta Crystallographica Section B: Structural Science*. 61 (2), 185–191.

Smith, J. H. et al. (2005) α-lactose monohydrate: a redetermination at 150 K. *Acta Crystallographica Section E: Structure Reports Online*. 61 (8), o2499–o2501.

Tewkesbury, H. et al. (2000) Modelling temperature distributions in cooling chocolate moulds. *Chemical Engineering Science*. 55 (16), 3123–3132.

Yff, B. T. S. et al. (2004) An investigation of calibration methods for solution calorimetry. *International journal of pharmaceutics*. 269 (2), 361–372.

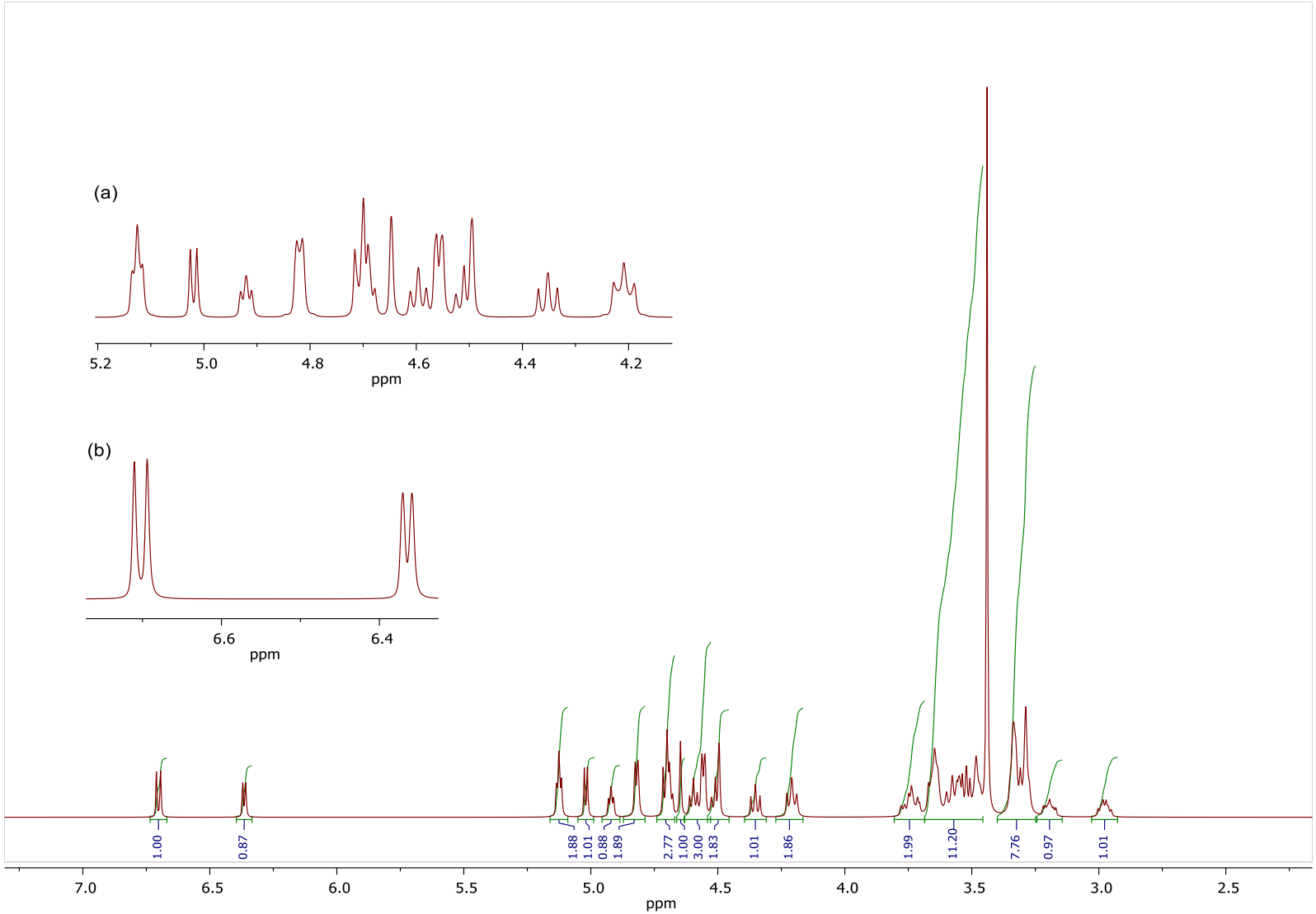


Figure S1. ^{1}H NMR of dried lactose syrup at 400 MHz. Inlays show (a) the region between 5.2 and 4.2 ppm; (b) the region from 6.8 to 5.3 ppm for clarity.

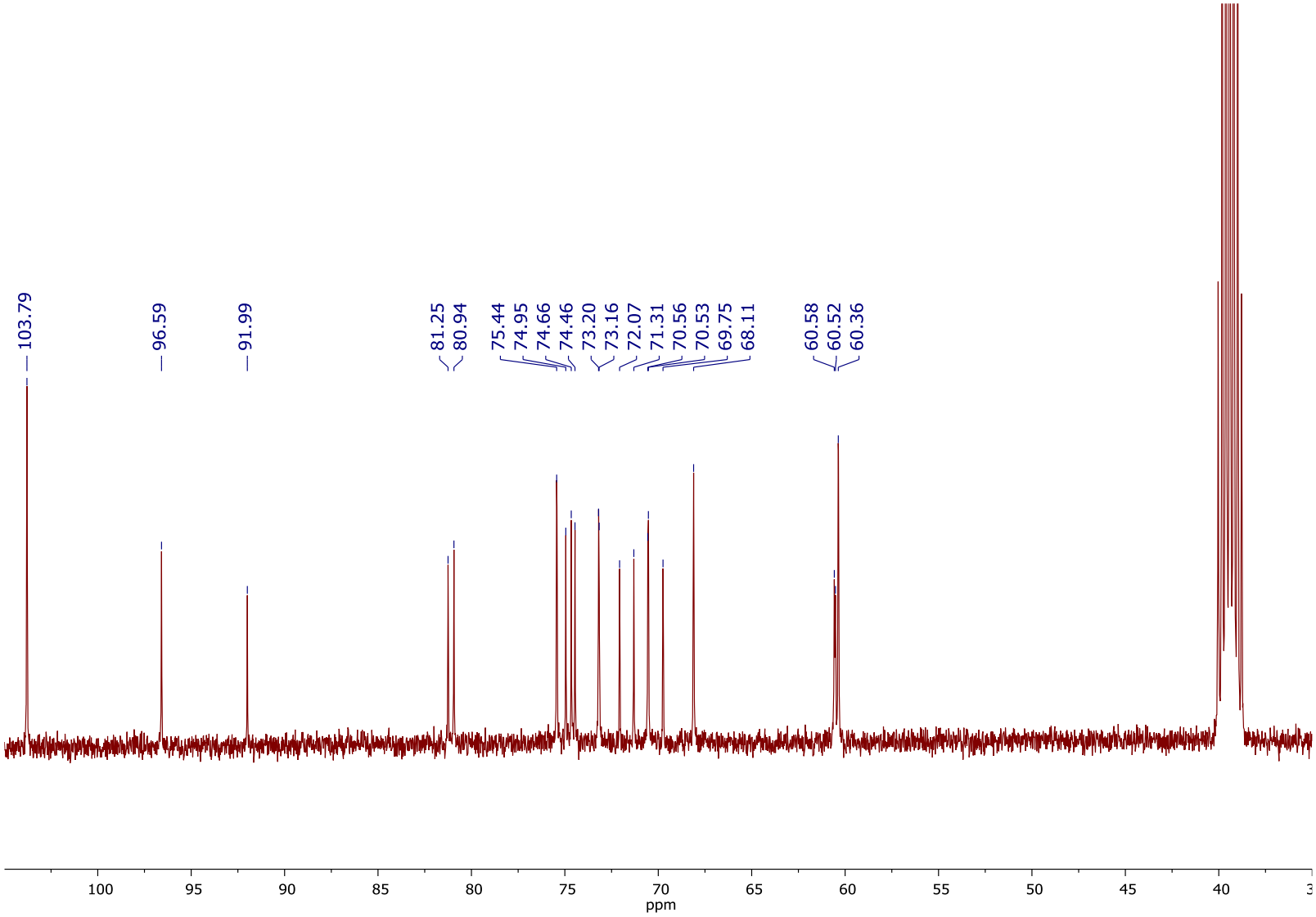


Figure S2. ^{13}C NMR of dried lactose syrup at 400 MHz. The septet at 39.41 ppm is the solvent DMSO.

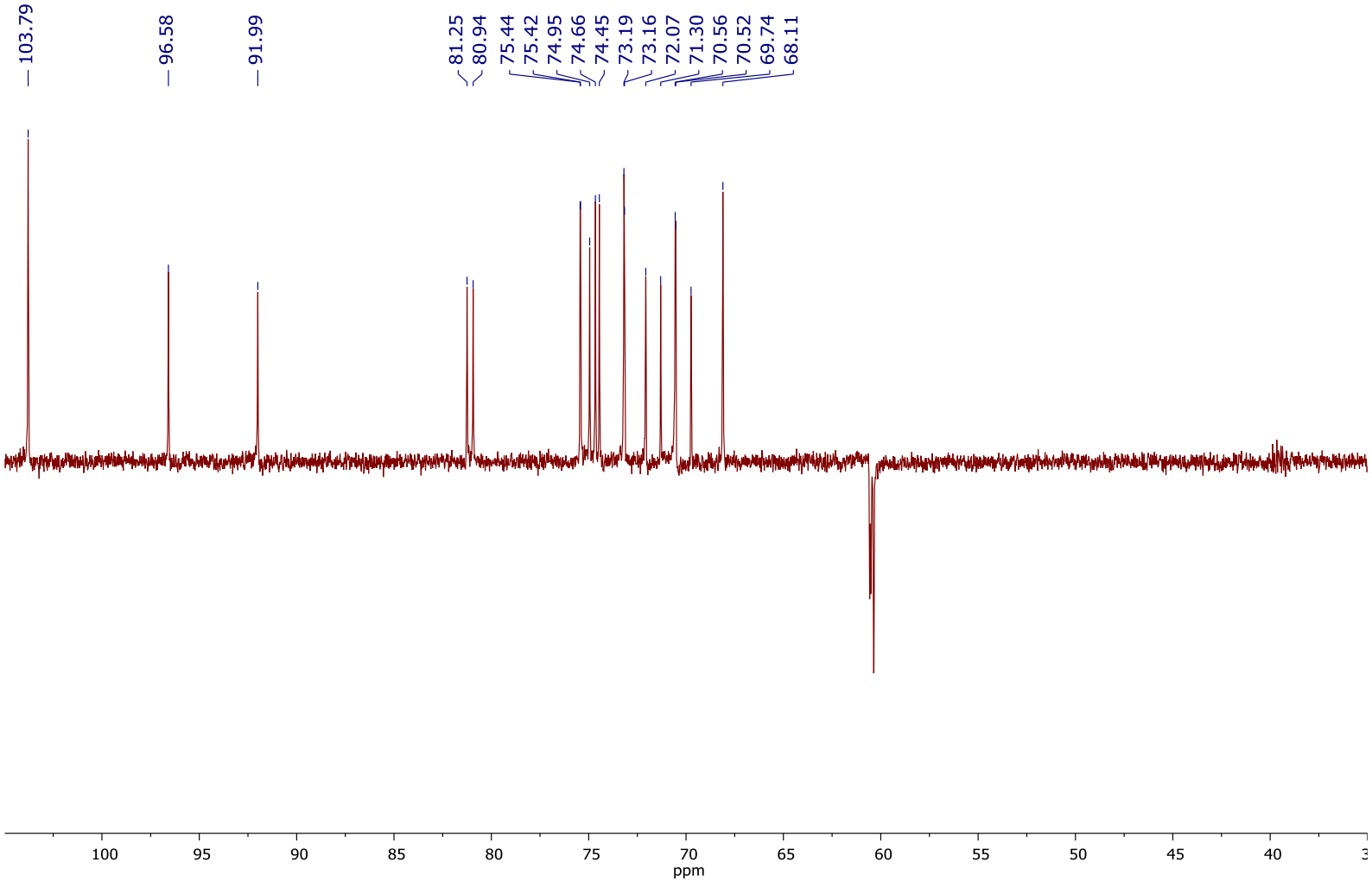


Figure S3. ^{13}C DEPT NMR of dried lactose syrup at 400 MHz. The inverted peaks represent -CH_2 inversions

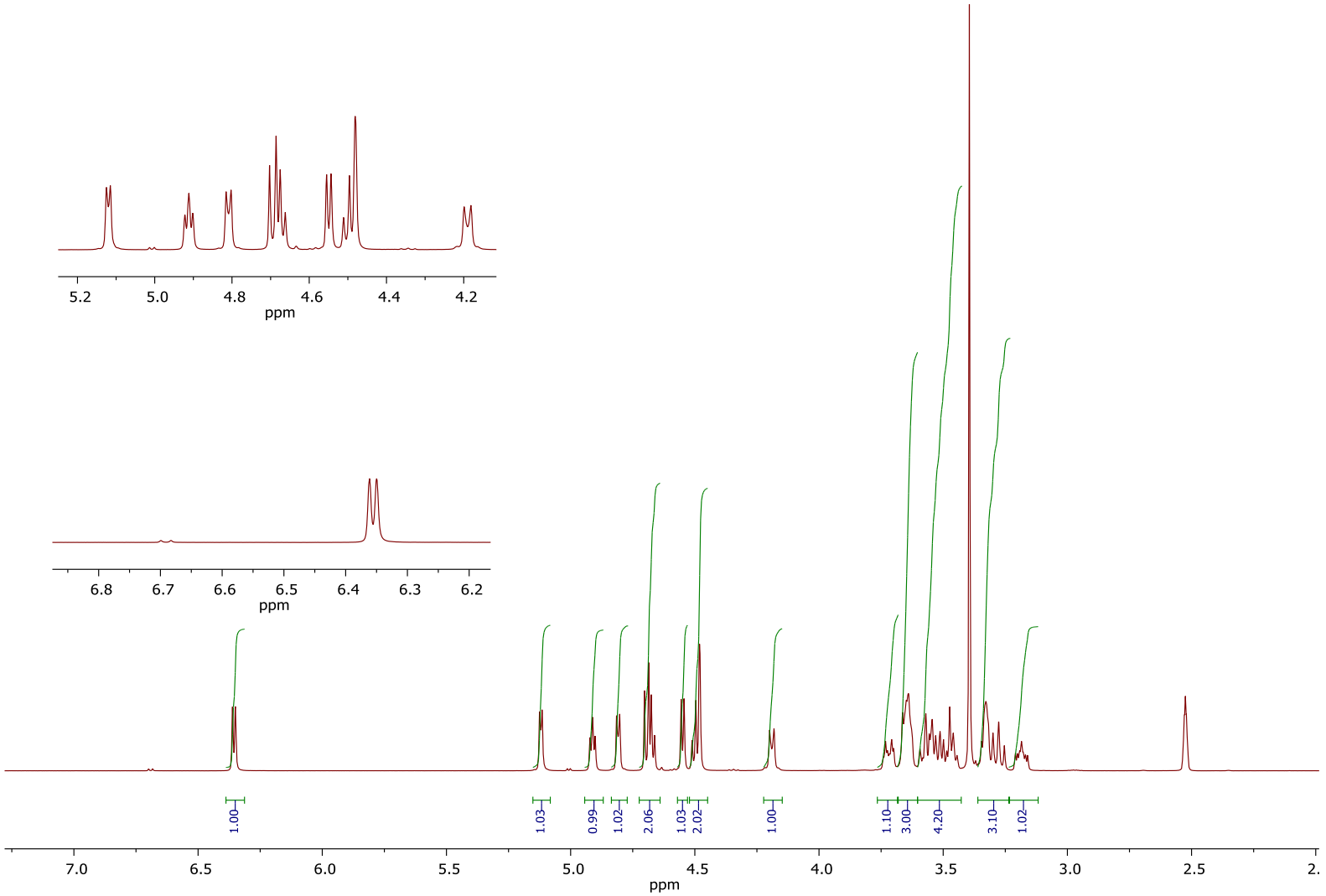


Figure S4. ^{1}H NMR of α-lactose monohydrate at 400 MHz. Inlays show (a) the region between 5.2 and 4.1 ppm; (b) the region from 6.8 to 6.3 ppm for clarity.

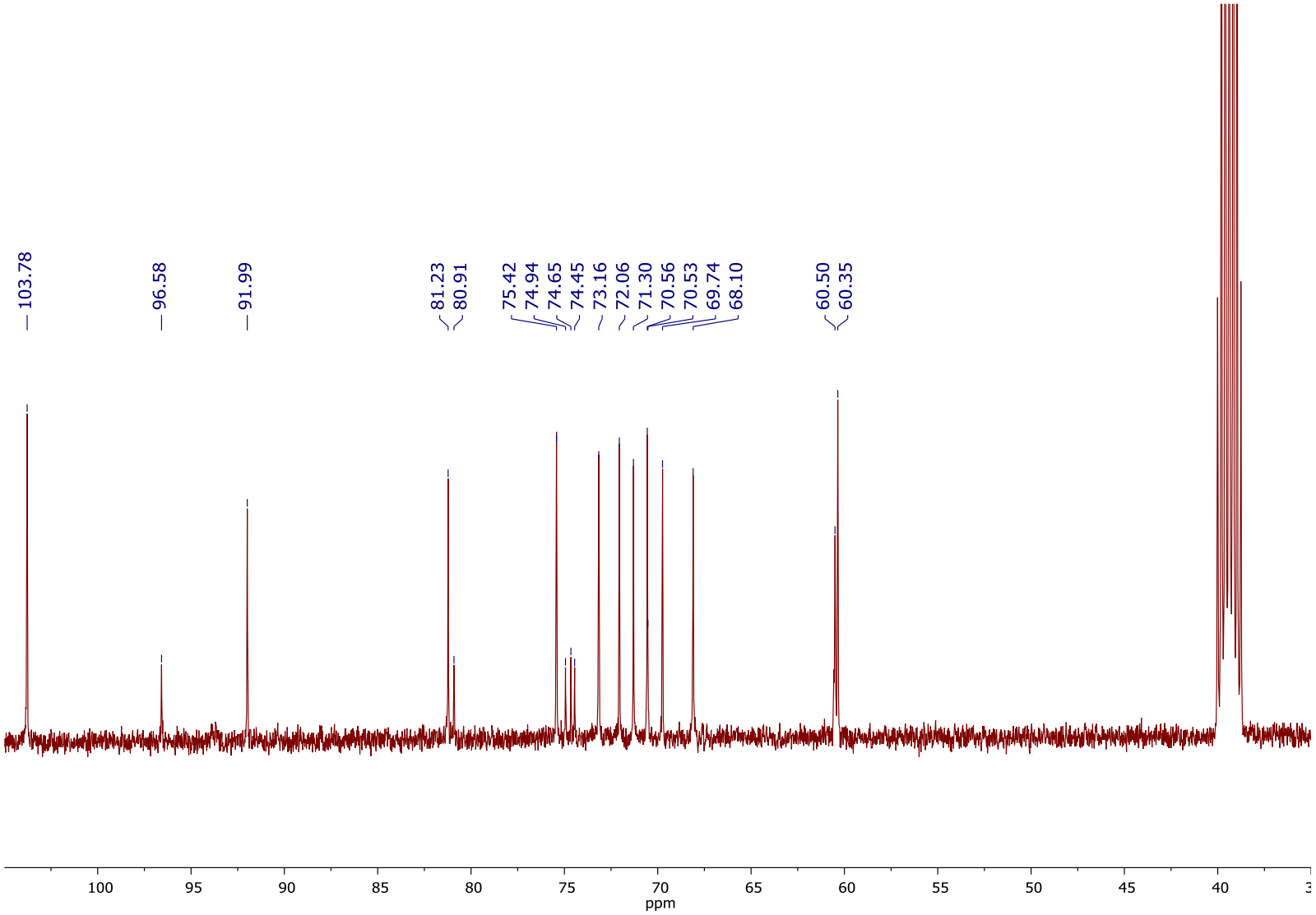


Figure S5. ^{13}C NMR of α-lactose monohydrate at 400 MHz. The septet at 39.41 ppm is the solvent DMSO.

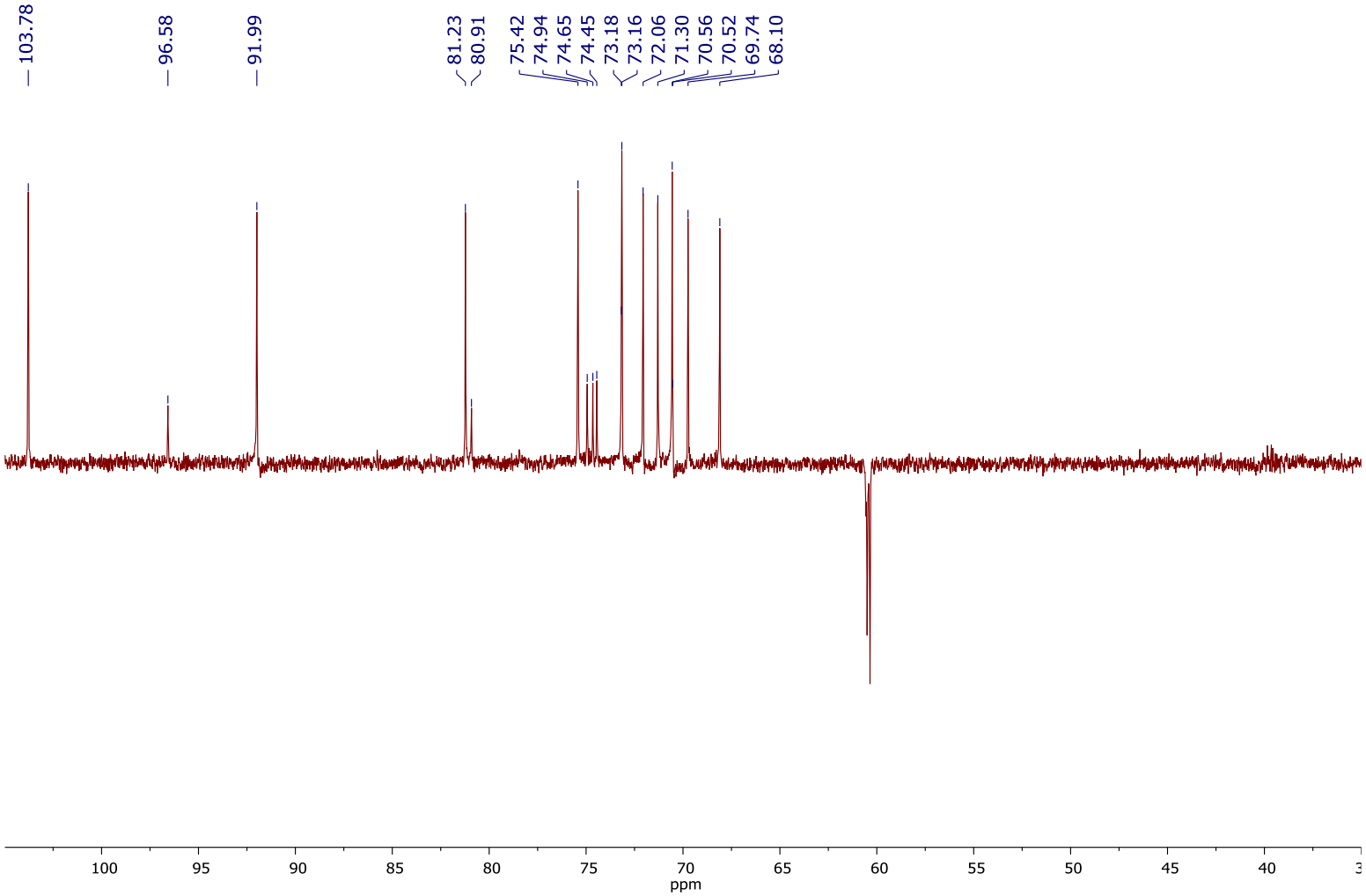


Figure S6. ^{13}C DEPT NMR of α-lactose monohydrate at 400 MHz. The inverted peaks represent -CH_2 inversions

7 Additional Information and Comments

7.1 Unpublished DSC Data

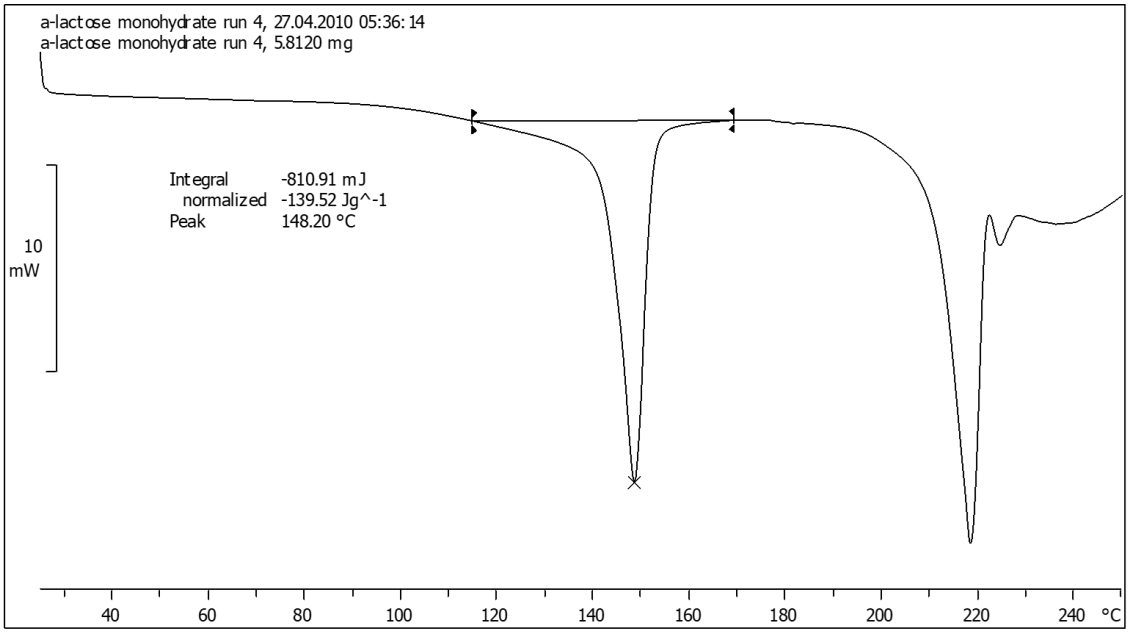


Figure 9: DSC trace of alpha lactose monohydrate, showing loss of water of hydration peak at 148 °C, the two peaks to the right at 220 °C and 225 °C show melting of the anhydrous α and β lactose anomers respectively.

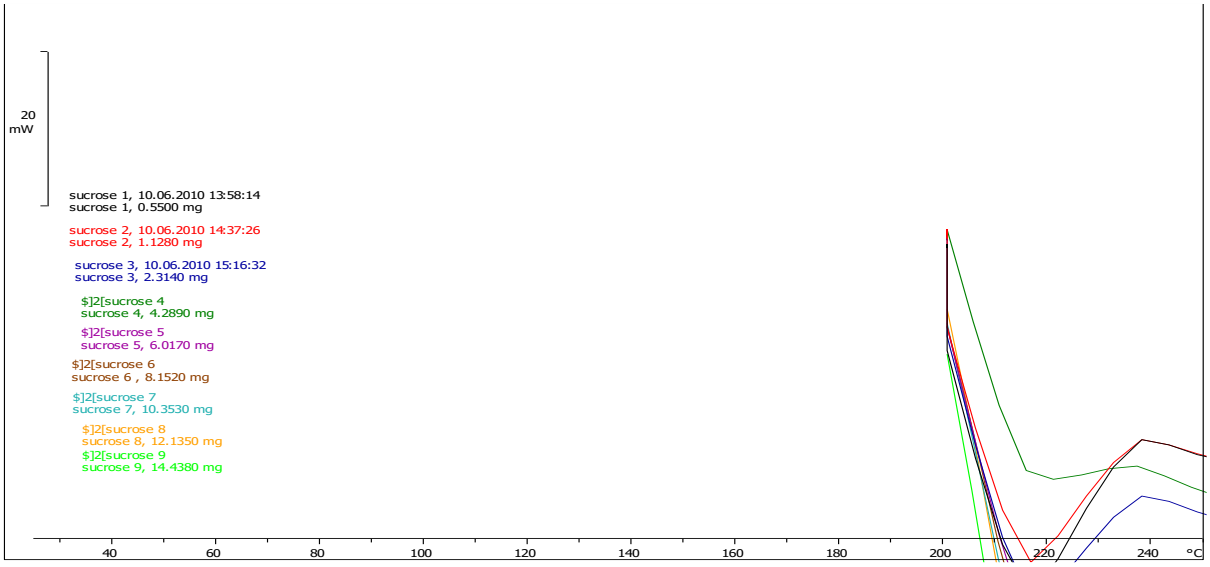


Figure 10: DSC traces of sucrose showing melting at 195 °C and degradation at ca. 220 °C.

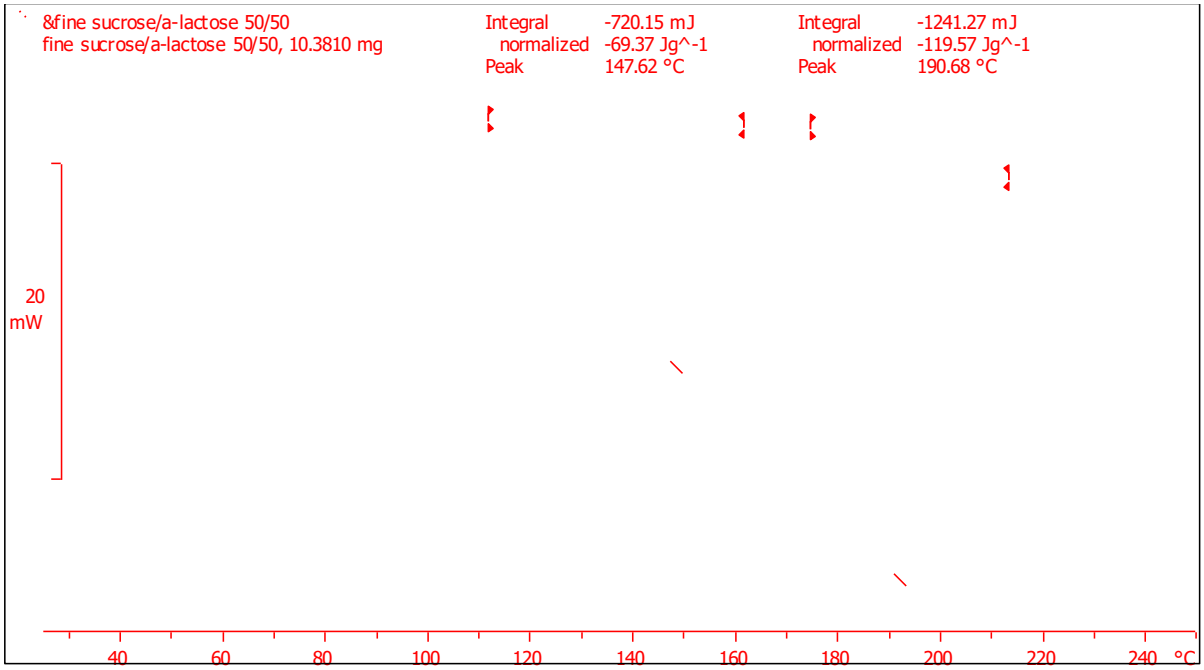


Figure 11: DSC trace showing melting peaks for a mixture of sucrose and lactose.

The DSC traces of alpha lactose monohydrate, sucrose, and a mixture of alpha lactose monohydrate and sucrose are shown in Figures 9, 10, and 11 respectively. Though the loss of water of hydration peak for alpha lactose monohydrate in Figure 11 is consistent with the compound in isolation, the melting peak for alpha lactose monohydrate has shifted to a lower temperature and is perhaps the cause for the distinct shoulder on the sucrose melt at approximately 200 °C. This is likely due to an interaction with the melting sucrose, as melting points of mixtures tend to occur over a larger range of lower temperatures, resulting in a broader peak that has shifted left on the DSC trace. Furthermore, the melt peak of the lactose anomers is overlaid with the decomposition noise associated with sucrose, making accurate integration unachievable. This combination of factors makes accurate quantification of sugar mixtures by DSC impossible, hence the need for PXRD QPA in Chapter 4.

7.2 Identification of new polymorph in Table 5

Following the identification of the new αβ-L_M structure, review of the data in Table 5 shows that the random peak selection indexing has some potential for indexing of complex mixtures. Result number 7 of the $P2_1$ indexing shows great agreeability with the structure found later down the line. Table 6 below shows the two datasets for ease of comparison.

Table 6. Comparison of random-peak indexing result number 7 with the single-crystal diffraction determined structure of alpha-beta lactose monohydrate.

Structure	Goodness of fit	Volume ($Å^3$)	a (Å)	b (Å)	c (Å)	β (°)
Number 7	3.77	1404.44	7.54	38.88	4.93	103.62
αβ-L_M	n/a	1411.26	7.6006	38.6364	5.0044	106.200

The similarities between the two cells highlights the potential for random peak selection indexing in complex mixtures. In samples where attainment of a single crystal of sufficient quality for single crystal diffraction is not possible, the search may give some headway towards identification of an unknown substance.

However, there are limitations to this approach. Firstly, the indexing was done with a selection of the most common space groups observed in organic compounds. Hence if all 230 space groups were to be ran through this method the quantity of data would be incredibly time consuming to analyse. Furthermore, the goodness of fit is not an effective indicator of the accuracy of the cell determined – in the $P2_1$ space group alone there were six other cells with a higher goodness of fit. Hence, analysis of many of the potential cells is required which is expensive in computation time. Finally, when this indexing result was returned to following identification of the new polymorph and structure determination was attempted with the random peak indexing cell, sensical results were unattainable. Regardless, this approach does present a potential avenue for gleaning more information on an uncharacterizable pattern within a multi-phase mixture if a cautious approach is taken to reduce the vast quantity of data produced.

Chapter 6

A new crystalline form of αβ-D-lactose prepared by oven drying a concentrated aqueous solution of D-lactose

Published as:

Nicholls, D. et al. (2019) A new crystalline form of *αβ*-D-lactose prepared by oven drying a concentrated aqueous solution of D-lactose. *Acta Crystallographica Section C*. 75 (7), 904-909.

6 A new crystalline form of αβ-D-lactose prepared by oven drying a concentrated aqueous solution of D-lactose

Foreword

The following paper was published in *Acta Crystallographic Section C* in 2019. It consists of the crystal structure determination of a monoclinic polymorph of αβ-D-lactose from a dried aqueous lactose solution through single crystal X-ray diffraction. Periodic density functional theory calculations (DFD-D) were used to produce an energy-minimised structure that informed the orientation of a hydrogen atom in single crystal structure solution. The solved structure was used in quantification of the powder X-ray diffraction data collected on the dried syrups.

I certify that I performed all the experimental work that is reported, conducted the bulk of data analysis, and wrote the majority of the manuscript. Kenneth Shankland aided in running the DFD-D calculations and the practicalities of single crystal extraction and collection. Norman Shankland helped in the comparison between DFD-D energy minimized and experimental structures in *Mercury* and aided in the production of packing diagrams. Carole Elleman was consulted on the importance and context of the research in chocolate crumb manufacture.

All named authors contributed to the writing of the final manuscript.

A new crystalline form of αβ-D-lactose prepared by oven drying a concentrated aqueous solution of D-lactose

Daniel Nicholls[a], Carole Elleman[b], Norman Shankland[c] and Kenneth Shankland[a*]

[a] School of Pharmacy, University of Reading, Whiteknights, Reading RG6 6AD, UK

[b] Reading Scientific Services Ltd, Reading Science Centre (Mondelez), Whiteknights, Reading RG6 6LA, UK

c CrystallografX Limited, 2 Stewart Street, Glasgow, Strathclyde, G62 6BW, UK

[*]*k.shankland@reading.ac.uk*

1 Synopsis

A monoclinic polymorph of αβ-D-lactose has been identified from crystalline phase mixtures obtained by oven drying concentrated aqueous solutions of D-lactose. This is the second crystalline form of αβ-D-lactose to be identified and it has a high degree of structural 3D similarity to the existing triclinic form.

2 Abstract

A new crystalline form of αβ-D-lactose has been prepared by rapid drying of an approximately 40 % w/v syrup of D-lactose in water at 120 °C in a fan oven. Initially identified from its novel powder X-ray diffraction pattern, the monoclinic crystal structure was solved from a microcrystal recovered from the generally polycrystalline, mixed-phase residue obtained at the end of the drying step. This is the second crystalline form of αβ-D-lactose to be identified and it has a high degree of structural 3D similarity to the existing triclinic form.

3 Introduction

D-Lactose (henceforth referred to simply as lactose) is a disaccharide that plays a very important role in the food and pharmaceutical industries (Booij, 1985; Audic *et al.*, 2003; Gohel, 2005; Lifran *et al.*, 2000) and whose solid state properties in both the crystalline and

amorphous states have been studied extensively for decades (Hockett & Hudson, 1931; Bushill *et al.*, 1965; Kirk *et al.*, 2007; Terban *et al.*, 2016). It exists in two anomeric forms (Figure 1) and several crystalline forms have been reported and fully characterised crystallographically (Table 1). Methods of the preparation for the various forms have been described in considerable detail in the publications cited in Table 1 and elsewhere (Simpson *et al.*, 1982; Simone *et al.*, 2019). Whilst other forms with varying mixes of the two anomers have been reported in the literature (e.g. an α:β ratio of 5:3, Hockett & Hudson, 1931), these reports lack compelling crystallographic evidence to back them up. Indeed, the multidisciplinary nature of much of the research on lactose crystals and crystallisation means that crystallography rarely plays the central role in the evaluation of experimental outcomes.

Our own interest in the solid-state properties of lactose stems from its significance in chocolate crumb manufacture. Chocolate crumb is typically manufactured by the plate drying of a sweetened condensed milk (produced by either dissolving sucrose in milk or adding water to a sucrose and milk powder mixture) to ~80-90 % solids. Cocoa liquor is then mixed into this material which is subsequently dried under vacuum to leave a chocolate crumb product (Wells, 2009). The crumb method of chocolate manufacture is widely used by larger manufacturers, in the main due to the characteristic and brand-defining flavours that crumb chocolates possess. The vacuum drying stage provides the perfect conditions for Maillard reactions to occur, allowing for the development of complex, cooked flavours that can vary significantly depending on the exact conditions of the drying. This dependence of flavour upon conditions means that crumb production methodologies are often maintained secret or patented (Wells, 2009).

We have previously accurately quantified the relative proportions of crystalline and amorphous sugars in both chocolate and chocolate crumb using powder X-ray diffraction (PXRD) based quantitative phase analysis (QPA) (Nicholls *et al.*, 2018). Under the assumption that the majority of solid-state transformations of, and interactions between, the sugars in crumb manufacture occur during the initial plate drying of the sweetened condensed milk, we have explored the phase composition of the materials that result from drying of the somewhat simpler systems of concentrated sucrose and lactose syrups. Here, we report the results of drying syrups that contain only lactose.

Table 1. Reported crystalline forms of D-lactose for which crystal structures are available.

Form	Abbreviation	SpGrp	*Z*	*Z′*	*T* (K)	Refcode	Type	Ref.
α-lactose monohydrate	α-L.H_2O	$P2_1$	2	1	150	LACTOS11	SX	Smith *et al.* (2005)
β-lactose	β-L	$P2_1$	2	1	293	BLACTO	SX	Hirotsu and Shimada (1974)
α-lactose (hygroscopic)	α-LH	$P2_1$	2	1	293	EYOCUQ	PXRD	Platteau *et al.* (2004)
α-lactose (stable anhydrous)	α-LS	*P*1	2	2	293	EYOCUQ01	PXRD	Platteau *et al.* (2005)
αβ-lactose	αβ-L_T	*P*1	2	2	120	LAKKEO01	SX	Guiry *et al.* (2008)

SpGrp = Space group; Refcode = Cambridge Structural Database refcode; SX = structure obtained from single-crystal X-ray diffraction; PXRD = structure obtained from powder X-ray diffraction

4 Experimental

4.1. Sample preparation

40 grams of α-lactose monohydrate (Sigma, BN: SLBQ8461V) were dissolved in 100 mL of purified water at 45 °C on a stirring hotplate. Stirring was continued for approximately 1 hour after the disappearance of the last dispersed solid to ensure complete dissolution. Aliquots (3 mL or 6 mL) were then measured out and poured into 5.5 cm Petri dishes before being placed in a Lenton Thermal Designs fan oven at a number of different temperatures (Table 2). The temperature at the level of the tray inside the oven was confirmed using an RS52 Digital Thermometer equipped with a K-type thermocouple. Preliminary experiments showed that 30 minutes of drying appeared sufficient to completely dry a 3 mL sample, but 90 minutes of drying was chosen for use throughout, in order to account for the use of 6 mL samples. Upon removal from the fan oven, samples were allowed to cool and the crystallised powders, which had hardened as off-white crusts in the Petri dishes, were removed using a spatula and stored

in glass vials ready for analysis. Samples were ground in an agate mortar and pestle in order to create free-flowing powders suitable for PXRD measurements.

Table 2.Sample drying conditions for the lactose syrups

Sample number	Temperature (°C)	Syrup volume (mL)
1	100	3
2	100	6
3	110	3
4	110	6
5	120	3
6	120	6
7	140	3
8	140	6

4.2. Powder X-ray diffraction

Each powder was loaded into a 0.7 mm borosilicate glass capillary and mounted on a Bruker D8 Advance Diffractometer operating in transmission capillary geometry, with a LynxEye detector and monochromatic incident radiation of wavelength 1.54056 Å. Samples were typically scanned in the relatively narrow range 5-35° 2θ with a 1 hr data collection time, for phase identification purposes. All PXRD runs were performed within a few hours of the samples being removed from the oven. Data were analysed using the EVA (Bruker AXS, 2018), TOPAS (Coelho, 2018) and DASH (David et al., 2006) packages.

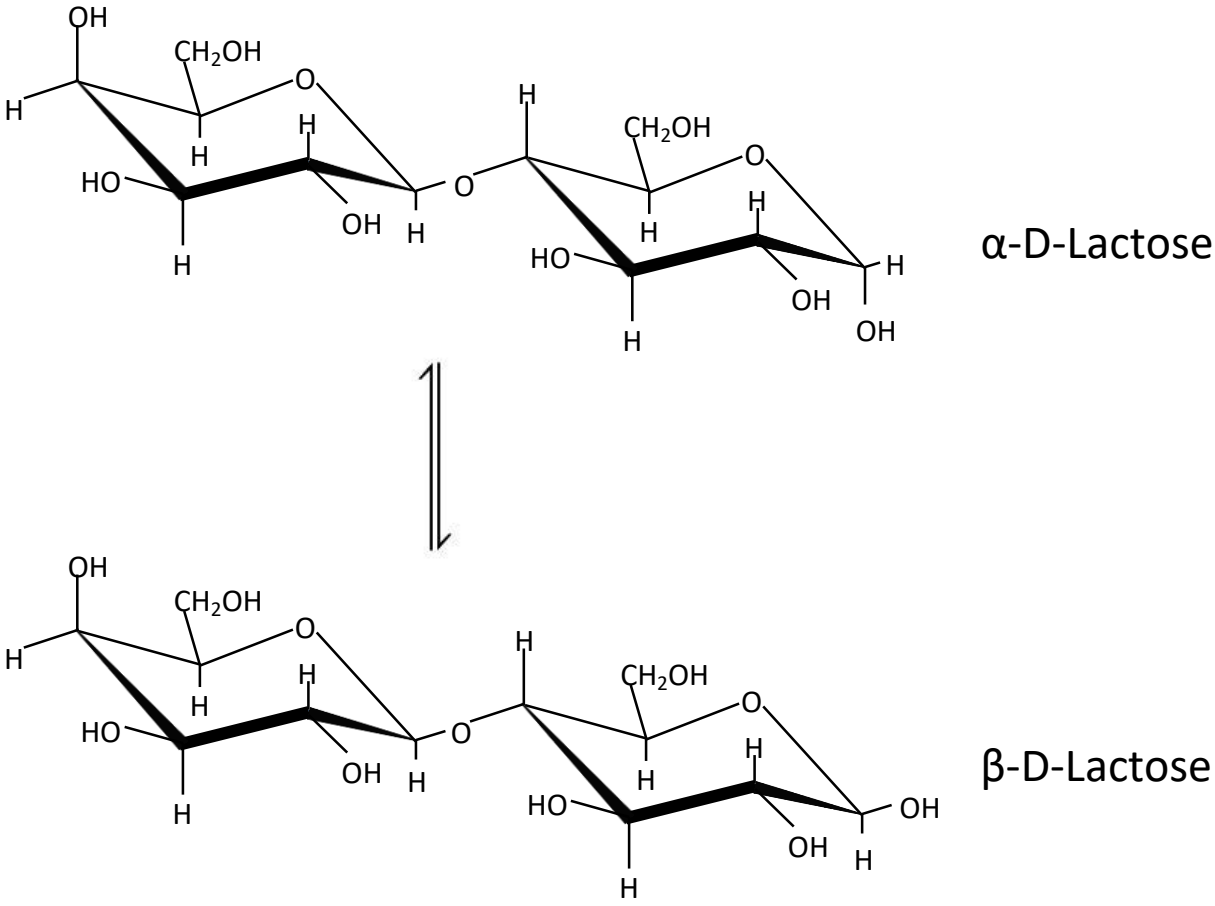


Figure 1. The molecular structures of the α and β anomers of D-lactose.

4.3. Single-crystal diffraction

Dried samples were examined under a polarising microscope to check for the possible occurrence of single crystals. Any potential single crystals were carefully removed and mounted on a Rigaku Synergy single-crystal diffractometer equipped with a microfocus copper X-ray source, a Hypix 3000 single-photon counting detector and an Oxford Cryosystems Cryostream cooling device. After data collection at 100 K, structures were solved by *SHELXT* (Sheldrick, 2015a) and refined by *SHELXL* (Sheldrick, 2015b), using the overarching *OLEX2* application (Dolomanov *et al.*, 2009).

Table 3. Experimental Details

Crystal Data	
Chemical formula	$C_{12}H_{22}O_{11}$
M_r	342.29
Crystal system, space group	Monoclinic, $P2_1$
Temperature (K)	100
a, b, c (Å)	5.0044 (3), 38.6364 (14), 7.6006 (4)
β (°)	106.200 (5)
V (Å^3)	1411.26 (13)
Z	4
Radiation type	Cu Kα
μ (mm^{-1})	1.26
Crystal size (mm)	0.08 × 0.02 × 0.02
Data collection	
Diffractometer	XtaLAB Synergy Dualflex HyPix
Absorption correction	Multi-scan (*CrysAlis PRO*; Rigaku OD, 2019)
T_{min}, T_{max}	0.804, 1.000
No. of measured, independent and observed [$I > 2\sigma(I)$] reflections	8132, 3861, 3316
R_{int}	0.057
(sin $\theta/\lambda)_{max}$ (Å^{-1})	0.597
Refinement	
$R[F^2 > 2\sigma(F^2)]$, $wR(F^2)$, S	0.048, 0.120, 1.02
No. of reflections	3861
No. of parameters	431
No. of restraints	2
H-atom treatment	H-atom parameters constrained
$\Delta\rho_{max}$, $\Delta\rho_{min}$ (e Å^{-3})	0.38, -0.25

Computer programs: *CrysAlis PRO* (Rigaku OD, 2019), *SHELXT* (Sheldrick, 2015*a*), *SHELXL2017* (Sheldrick, 2015*b*) and *OLEX2* (Dolomanov *et al*, 2009)

4.4. Periodic density functional theory calculations

Periodic density functional theory with van der Waals dispersion corrections (DFT-D) was used for geometry optimisation of crystal structures of interest. The PBE functional was used with PAW pseudopotentials and the Grimme D3 correction, as implemented in the *pw.x*

executable of the *QuantumEspresso* program (Giannozzi *et al.*, 2009; Giannozzi *et al.*, 2017). The lengths of bonds involving H atoms were normalised using *Mercury CSD* (Macrae *et al.*, 2008) and input files for *pw.x* were then created from these normalised CIFs using the *cif2qe* script of *QuantumEspresso*. Automatic *k*-point sampling was used; the kinetic energy cutoffs for wavefunctions and charge density were 50 and 400 Ry respectively. The convergence thresholds on total energy and forces were set to 0.0001 and 0.001 a.u. respectively. Initial geometry optimisations were carried out with lattice parameters fixed at their crystallographic values, with subsequent variable cell geometry optimisations starting from the endpoint of the fixed-cell calculations. All calculations were carried out on a Dell Precision T7810 workstation equipped with two 2.40 GHz 8-core Intel Xeon E5-2630 v3 CPUs, running the Microsoft Windows 10 operating system, and using the Windows Subsystem for Linux (WSL) feature to allow the Linux-compiled, MPI-enabled *pw.x* executable to utilise multiple cores.

5 Results

All of the dried samples exhibited PXRD patterns consistent with phase mixtures of crystalline lactose and, with the exception of those samples dried at 100 °C, all possessed a strong contribution from a crystalline phase that did not correspond to any of the known phases listed in Table 1 (Figure 2). Even after identifying one of the contributing phases as β-L, attempts to index the unknown contribution(s) using *DASH* and *TOPAS* were unsuccessful. Careful examination of the 110 °C sample under a polarising microscope revealed that whilst the bulk of the sample was obviously polycrystalline, there were also a few very small single crystals present. Single crystal diffraction showed that most of these were β-L, but one very small (76 × 24 × 18 μm) single crystal indexed to a monoclinic cell that did not match any of the known forms. After careful Pawley refinement, to account for the ~200 K temperature difference between the single-crystal measurement and the PXRD experiments, it was found that this cell could explain the peak positions of the unknown phase in the PXRD pattern. A full single crystal data collection resulted in the crystal structure of a new monoclinic form of αβ-D-lactose (henceforth referred to as αβ-L_M), whose crystallographic details are summarised in Table 4

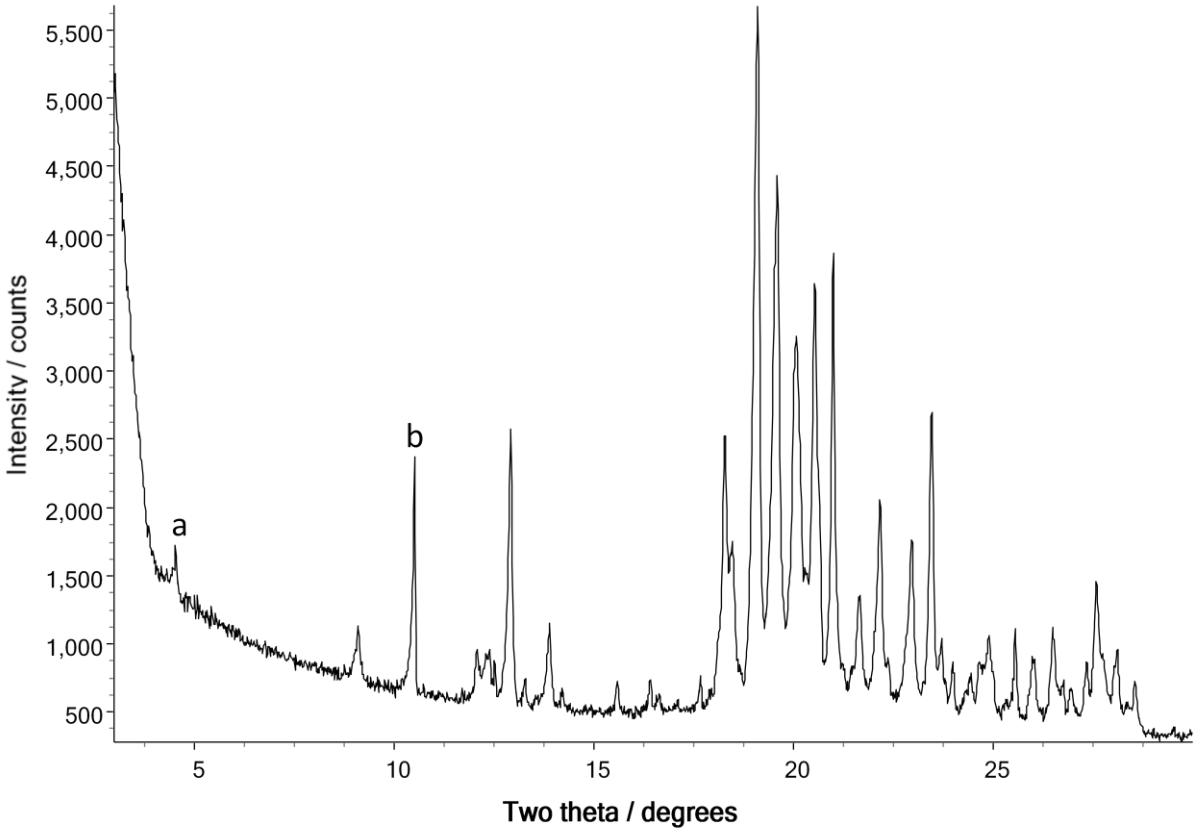


Figure 2. Powder X-ray diffraction data in the range 3-30° 2θ collected from dried sample 3. The data cannot be fully explained by the combination of any of the previously report crystal forms of lactose. The presence of β-L in the sample is evident from the diagnostic peak labelled *b*, whilst the low angle peak labelled *a* may be attributable to α-LS, αβ-L_T, or the unknown form of lactose (subsequently identified as monoclinic αβ-L_M) that appears to comprise the bulk of the sample.

Knowing the structure of αβ-L_M, it was then possible to perform Rietveld-based quantitative phase analysis (QPA) on the PXRD patterns in order to quantify the amounts of the various forms of crystalline lactose present in the samples (Table 3, Figure 3).

Table 4. Relative abundance of crystalline lactose phases (% w/w) in each of the samples listed in Table 2, as obtained by Rietveld-based QPA.

The estimates are approximate, since the PXRD data had not been collected with the intention of performing QPA; more accurate and precise quantification would require longer data collections over a wider 2θ range.

Sample number	$\alpha\beta$-L_M	$\alpha\beta$-L_T	β-L	α-LS	α-L·H_2O
1	0	0	41	0	59
2	0	0	27	0	73
3	57	8	23	12	0
4	80	3	10	7	0
5	81	19	0	0	0
6	91	4	0	5	0
7	90	10	0	0	0
8	92	8	0	0	0

6 Discussion

Despite the small size of the crystal retrieved from the phase mixture, the crystal structure determination of αβ-L_M was relatively straightforward. Whilst the single-crystal diffraction data were of sufficient quality to allow the positioning of the majority of hydroxyl H atoms directly from the difference Fourier maps, the location of a few such hydrogens were less well determined. Geometry optimisation in the solid state using DFT-D has been shown to be a powerful method for crystal structure verification (van de Streek & Neumann, 2010, 2014) and it was employed here to check the hydroxyl group orientations of the single-crystal structure. Fixed-cell geometry optimisation of the refined crystal structure of αβ-L_M gave a structure in which the position of H16 in particular moved significantly: specifically, 0.92 Å, corresponding to a clockwise rotation of approximately 59° around bond O9-C9 to form a hydrogen bond [O9-H16···O11; symmetry code: (i) $x - 1, y, z - 1$] that was not evident in the original single crystal structure refinement. Whilst this optimisation also improved the linearity of some other

hydrogen bonds (e.g. those involving H22, H36 and H38) only the position change of H16 was deemed to be significant enough to merit revisiting the single-crystal structure refinement. A DFIX command was therefore used in order to hold H16 in the correct (*i.e.* DFT-D) hydrogen bonding position in the final SHELX refinement. The asymmetric unit of the crystal structure of αβ-L_M, with the stereochemistry of the chiral centres highlighted, is shown in Figure 4.

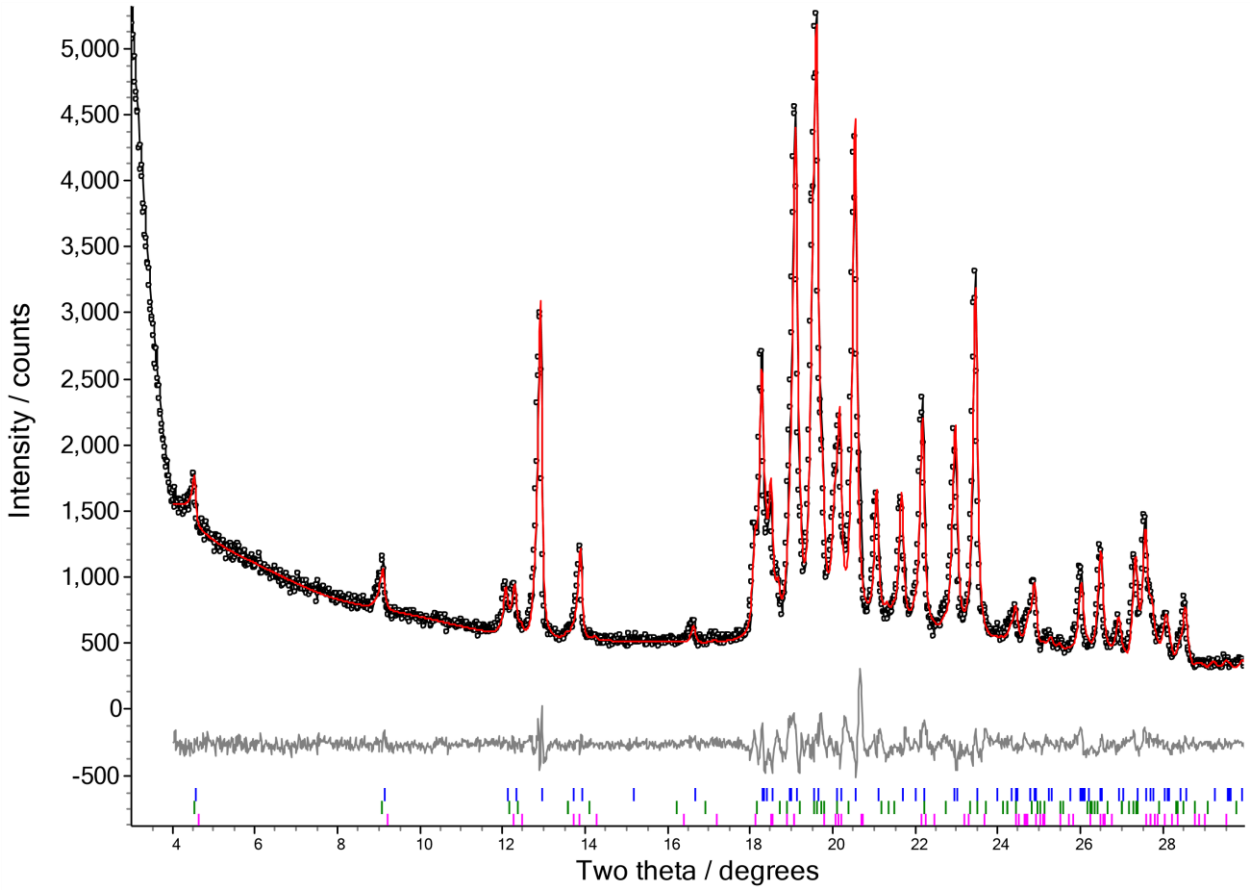


Figure 3. A three-phase Rietveld fit to the powder diffraction data in the range 3-30° 2θ collected from dried sample 6. Observed data are shown as black circles, the fit to the data is shown as a red line and the difference plot is shown in grey. Blue tick marks indicate reflection positions corresponding to αβ-L_M (91 % w/w); green tick marks indicate reflection positions corresponding to α-LS (5 % w/w); pink tick marks indicate reflection positions corresponding to the αβ-L_T (4 % w/w).

Differences between the experimental and energy-minimized αβ-L_M crystal structures are relatively small – a 15 molecule overlay in Mercury, returned RMSD = 0.054 Å, a favourable value that supports the correctness of experimental αβ-L_M (van de Streek & Neumann, 2010). The αβ-L_M and αβ-L_T polymorphs each comprise alternating two-dimensional layers of α and β anomers (Figure 5) and αβ-L_M shows a high degree of three-dimensional similarity with αβ-L_T (Figure 6).

Unsurprisingly, the DFT-D total energy difference between the asymmetric units of the polymorphs is too small to infer the rank order of stability: $\Delta E_{(\alpha\beta\text{-LM} - \alpha\beta\text{-LT})} = 1.8$ kJ mol^{-1}, with the energy for αβ-L_T calculated using the corrected and energy-minimised αβ-L_T crystal structure of van de Streek and Neumann (2014).

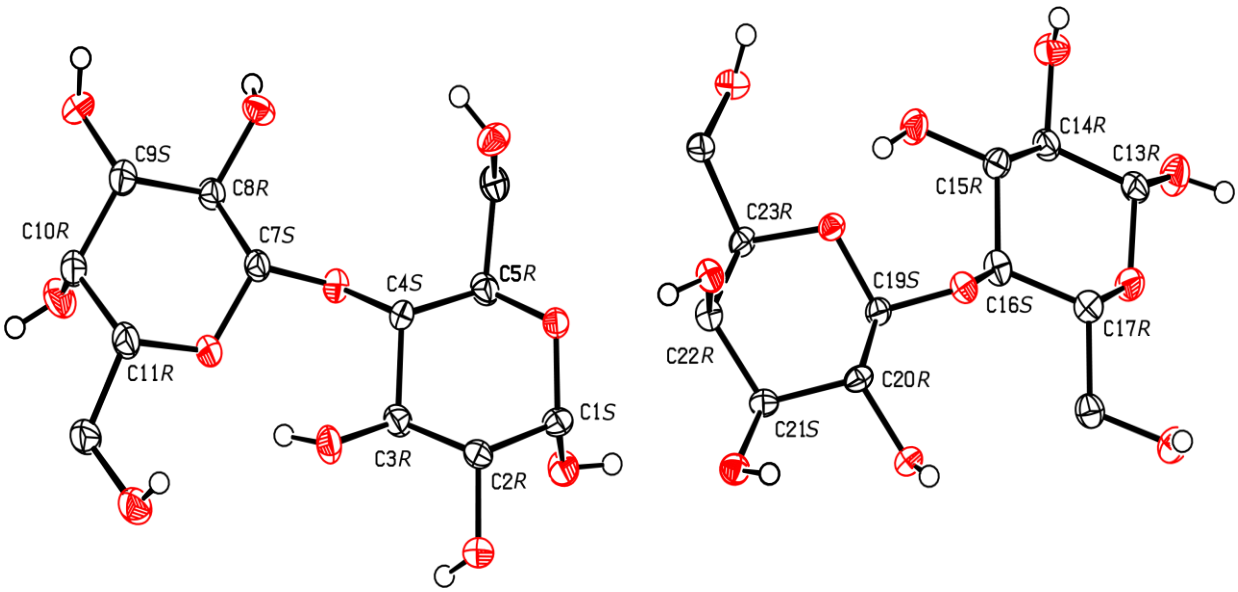


Figure 4. The contents of the asymmetric unit of the αβ-L_M crystal structure, with the stereochemistry of the chiral centres shown. The α anomer (left) has C1 S, whilst the β anomer (right) has C13 R.

The use of PXRD was crucial in identifying the presence of a new crystalline form of lactose. However, the complexity of the phase mixtures returned in the crystallisations, coupled with the relatively large unit cell size of αβ-L_M meant that we were unable to index it from the PXRD alone.

Appearing to comprise the bulk of the material recovered from crystallisations occurring at temperatures of 110 °C or above, the αβ-L_M phase was clearly polycrystalline in nature; from all the samples prepared, only one very small crystal of αβ-L_M suitable for single-crystal diffraction was ever recovered. In spite of its diminutive size, the crystal diffracted sufficiently well to yield a crystal structure that, as evidenced by Rietveld-based QPA, proved to constitute the majority phase in the bulk samples 3 to 8. Whilst we did not attempt any quantification of amorphous content, the relatively low backgrounds of all the collected PXRD data sets suggest that there is little in the way of amorphous material in the recrystallised samples and we observed no changes in the PXRD of samples as a function of time, over a timescale of 12 months.

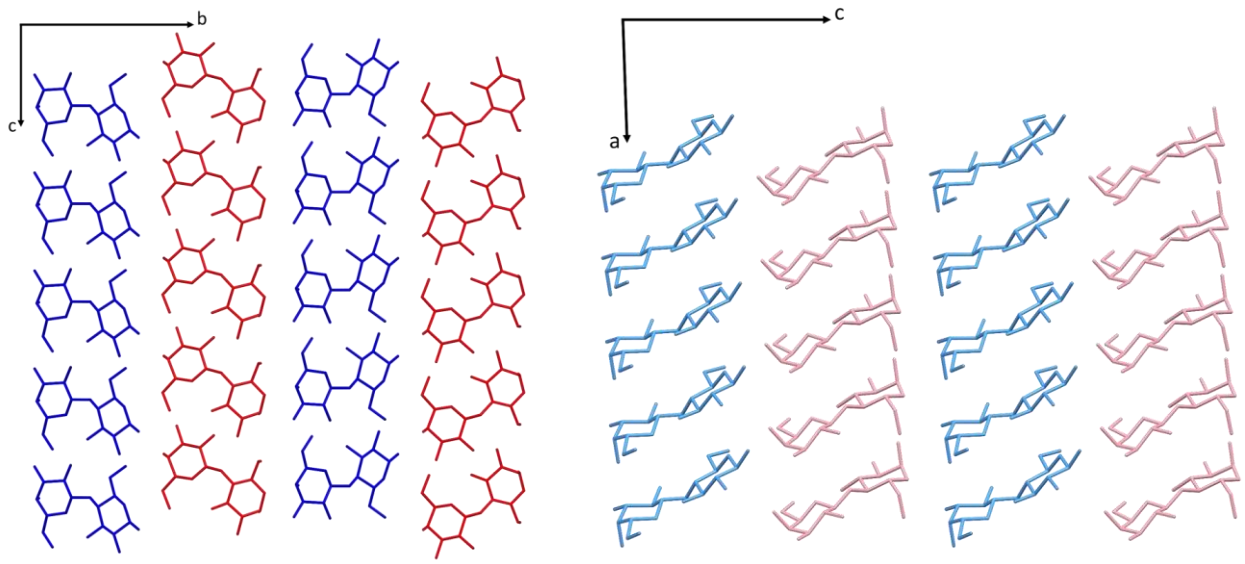


Figure 5. Packing diagrams of αβ-L_M (left) and αβ-L_T (right). Both crystal structures comprise alternating two-dimensional layers of the α anomer (red, salmon) and the β anomer (blue, light blue).

The QPA analysis shows that the composition of the phase mixture returned from the syrup drying experiments exhibit a temperature dependence. It also shows clearly that the previously identified form αβ-L_T crystallises alongside αβ-L_M in the syrup drying process, adding to the existing methods (Lefebvre *et al.*, 2005; Guiry *et al.*, 2008) for the crystallisation of αβ-L_T.

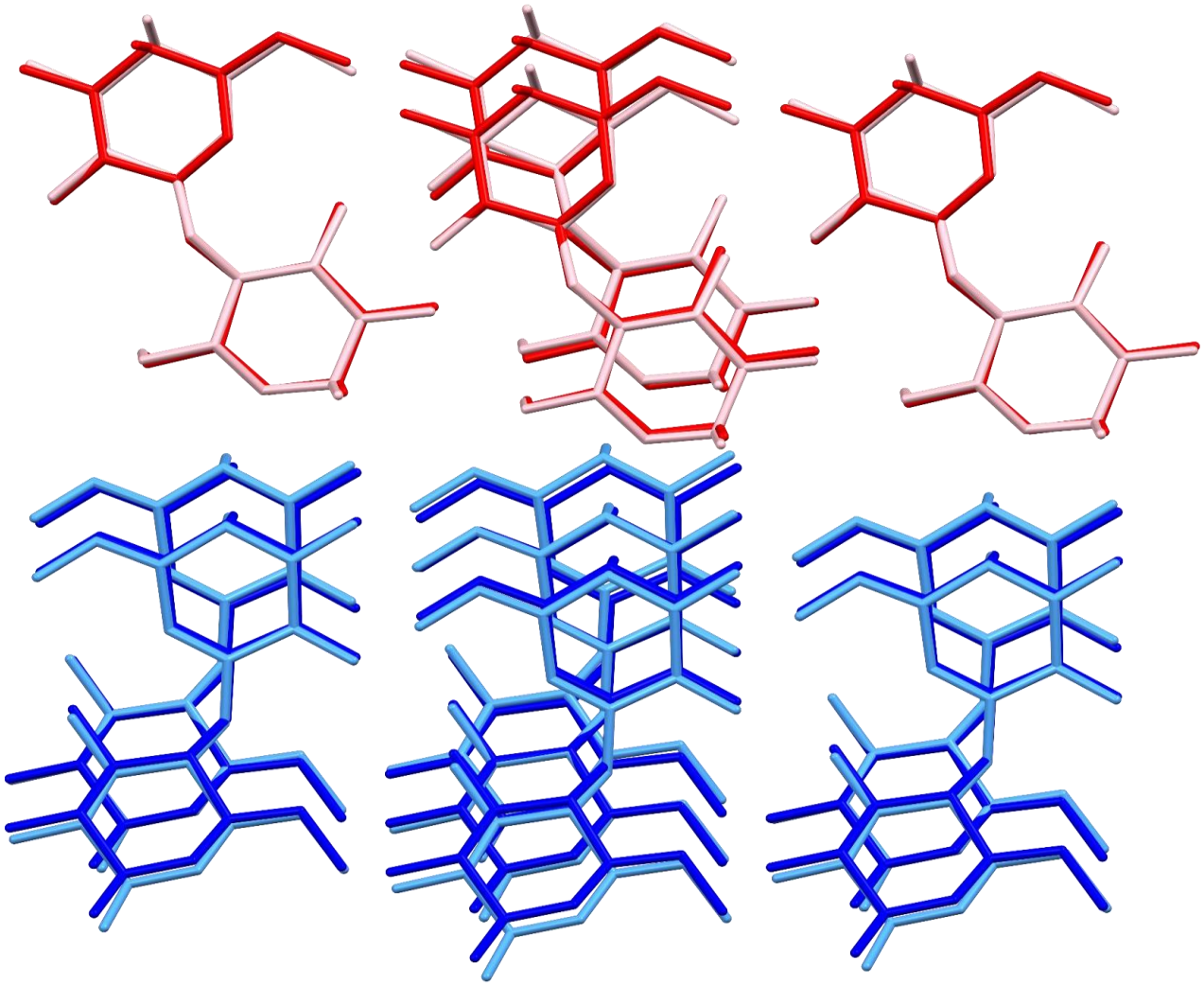


Figure 6. The three-dimensional similarity of αβ-L_M and αβ-L_T, illustrated by overlaying molecules in each structure using the 'Crystal Packing Similarity' feature of Mercury. The RMSD for the overlay is 0.336 Å and the colour coding used is the same as in Figure 5.

7 Conclusion

The use of PXRD, single-crystal diffraction and periodic DFT-D has allowed the identification and characterisation of a new monoclinic crystalline form of αβ-D-Lactose. The structure adds to the five known, well-characterised forms of D-Lactose. Many recent publications (Fan & Roos, 2015; Haque & Roos, 2005; Jouppila *et al.*, 1998; Nijdam *et al.*, 2007; Saffari *et al.*, 2015; Yazdanpanah & Langrish, 2011) continue to cite literature whose lactose phase identification methods are sub-optimal (i.e. based upon tabulated PXRD line positions and estimated intensities), and which continue to refer to mixed phase forms (such as α:β = 5:3 and α:β = 4:1) that lack compelling crystallographic evidence for their existence. It is recommended

that where PXRD is to play a role in phase identification in lactose mixtures, the analysis should be based on whole-pattern fitting methods using verified crystal structures, and ideally employ transmission (as opposed to the more widely used reflection) instrument geometry, in order to minimise the confounding effects of preferred orientation in samples.

8 Acknowledgements

We gratefully acknowledge the University of Reading's Chemical Analysis Facility for X-ray diffraction facilities and Nick Spencer for technical support with those facilities.

8.1. Funding information

We gratefully acknowledge the BBSRC and Mondelez Ltd. (Case Award BB/L015730/1) for funding for Dan Nicholls.

9 References

Audic, J.-L., Chaufer, B. & Daufin, G. (2003). *Lait* **83**, 417-438.

Booij, C. J. (1985). *J. Soc. Dairy Tech.* **38**, 105-109.

Bruker AXS (2018). *https://www.bruker.com/products/x-ray-diffraction-and-elemental-analysis/x*https://www.bruker.com/products/x-ray-diffraction-and-elemental-analysis/x-ray-diffraction/xrd-software/eva/overview.html*ray-diffraction/xrd-software/eva/overview.html.*

Bushill, J. H., Wright, W. B., Fuller, C. H. F. & Bell, A. V. (1965). *J. Sci. Food Agri.* **16**, 622-628.

Coelho, A. A. (2018). *J. Appl. Cryst.* **51**, 210-218.

David, W. I. F., Shankland, K., van de Streek, J., Pidcock, E., Motherwell, W. D. S. & Cole, J. C.

(2006). *J. Appl. Cryst.* **39**, 910-915.

Dolomanov, O. V., Bourhis, L. J., Gildea, R. J., Howard, J. A. K. & Puschmann, H. (2009). *J. Appl.*

Cryst. **42**, 339-341.

Fan, F. H. & Roos, Y. H. (2015). *Food Research International* **67**, 1-11.

Giannozzi, P., Andreussi, O., Brumme, T., Bunau, O., Nardelli, M. B., Calandra, M., Car, R.,

Cavazzoni, C., Ceresoli, D., Cococcioni, M., Colonna, N., Carnimeo, I., Dal Corso, A., de Gironcoli,

S., Delugas, P., DiStasio, R. A., Ferretti, A., Floris, A., Fratesi, G., Fugallo, G., Gebauer, R.,

Gerstmann, U., Giustino, F., Gorni, T., Jia, J., Kawamura, M., Ko, H. Y., Kokalj, A., Kucukbenli, E.,

Lazzeri, M., Marsili, M., Marzari, N., Mauri, F., Nguyen, N. L., Nguyen, H. V., Otero-de-la-Roza, A.,

Paulatto, L., Ponce, S., Rocca, D., Sabatini, R., Santra, B., Schlipf, M., Seitsonen, A. P., Smogunov, A., Timrov, I., Thonhauser, T., Umari, P., Vast, N., Wu, X. & Baroni, S. (2017). *J. Phys. Condens.*

Matter **29**.

Giannozzi, P., Baroni, S., Bonini, N., Calandra, M., Car, R., Cavazzoni, C., Ceresoli, D., Chiarotti, G.

L., Cococcioni, M., Dabo, I., Dal Corso, A., de Gironcoli, S., Fabris, S., Fratesi, G., Gebauer, R.,

Gerstmann, U., Gougoussis, C., Kokalj, A., Lazzeri, M., Martin-Samos, L., Marzari, N., Mauri, F.,

Mazzarello, R., Paolini, S., Pasquarello, A., Paulatto, L., Sbraccia, C., Scandolo, S., Sclauzero, G., Seitsonen, A. P., Smogunov, A., Umari, P. & Wentzcovitch, R. M. (2009). *J. Phys. Condens. Matter*

21.

Gohel, M. C. J., P.D. (2005). *J. Pharm. and Pharmaceut. Sci.* **8**, 76-93.

Guiry, K. P., Coles, S. J., Moynihan, H. A. & Lawrence, S. E. (2008). *Crystal Growth & Design* **8**, 3927-3934.

Haque, K. & Roos, Y. H. (2005). *Carbohydrate Research* **340**, 293-301.

Hirotsu, K. & Shimada, A. (1974). *Bull. Chem. Soc. Japan* **47**, 1872-1879.

Hockett, R. C. & Hudson, C. S. (1931). *J. Am. Chem. Soc.* **53**, 4455-4456.

Jouppila, K., Kansikas, J. & Roos, Y. H. (1998). *Biotechnology Progress* **14**, 347-350.

Kirk, J. H., Dann, S. E. & Blatchford, C. G. (2007). *Int. J. Pharm.* **334**, 103-114.

Lefebvre, J., Willart, J.-F., Caron, V., Lefort, R., Affouard, F. & Danede, F. (2005). *Acta Cryst.* **B61**, 455-463.

Lifran, E., Hourigan, J., Sleigh, R. & L Johnson, R. (2000). *Food Australia* **52**.

Macrae, C. F., Bruno, I. J., Chisholm, J. A., Edgington, P. R., McCabe, P., Pidcock, E., RodriguezMonge, L., Taylor, R., van de Streek, J. & Wood, P. A. (2008). *J. Appl. Cryst.* **41**, 466-470. Nicholls, D., Shankland, K., Spillman, M. & Elleman, C. (2018). *Food Analytical Methods* **11**, 26732681.

Nijdam, J., Ibach, A., Eichhorn, K. & Kind, M. (2007). *Carbohydrate Research* **342**, 2354-2364.

Platteau, C., Lefebvre, J., Affouard, F. & Derollez, P. (2004). *Acta Cryst.* **B60**, 453-460.

Platteau, C., Lefebvre, J., Affouard, F., Willart, J. F., Derollez, P. & Mallet, F. (2005). *Acta Cryst.*

B61, 185-191.

Rigaku OD (2019). *CrysAlis PRO 1.171.40.43a*

Saffari, M., Ebrahimi, A. & Langrish, T. (2015). *J. Food. Eng.* **164**, 1-9.

Sheldrick, G. M. (2015a). *Acta Cryst.* **A71**, 3-8.

Sheldrick, G. M. (2015b). *Acta Cryst.* **C71**, 3-8.

Simone, E., Tyler, A. I. I., Kuah, D., Bao, X. F., Ries, M. E. & Baker, D. (2019). *Org. Process Res.*

Dev. **23**, 220-233.

Simpson, T. D., Parrish, F. W. & Nelson, M. L. (1982). *J. Food. Sci.* **47**, 1948-1951.

Smith, J. H., Dann, S. E., Elsegood, M. R. J., Dale, S. H. & Blatchford, C. G. (2005). *Acta Cryst.* **E61**, O2499-O2501.

Terban, M. W., Cheung, E. Y., Krolikowski, P. & Billinge, S. J. L. (2016). *Crystal Growth & Design* **16**, 210-220. van de Streek, J. & Neumann, M. A. (2010). *Acta Cryst.* **B66**, 544-558. van de Streek, J. & Neumann, M. A. (2014). *Acta Cryst.* **B70**, 1020-1032.

Wells, M. A. (2009). *Beckett's Industrial Chocolate Manufacture and Use*, edited by S. Beckett. UK:

Wiley-Blackwell

Yazdanpanah, N. & Langrish, T. A. G. (2011). *Drying Technology* **29**, 1046-1057.

10. Additional Information and Comments

The absolute stereochemistry of the structure was not established by experiment; the data were not collected with that aim (indeed, the data collection strategy had the "Friedel mates are equivalent" option checked to reduce data collection time, for fear that the crystal would be unstable in the beam) and the 'nonsense' Flack parameter of -0.3(2) reflects this. However, the absolute stereochemistry of all centres is well established from the known chemistry of the molecule.

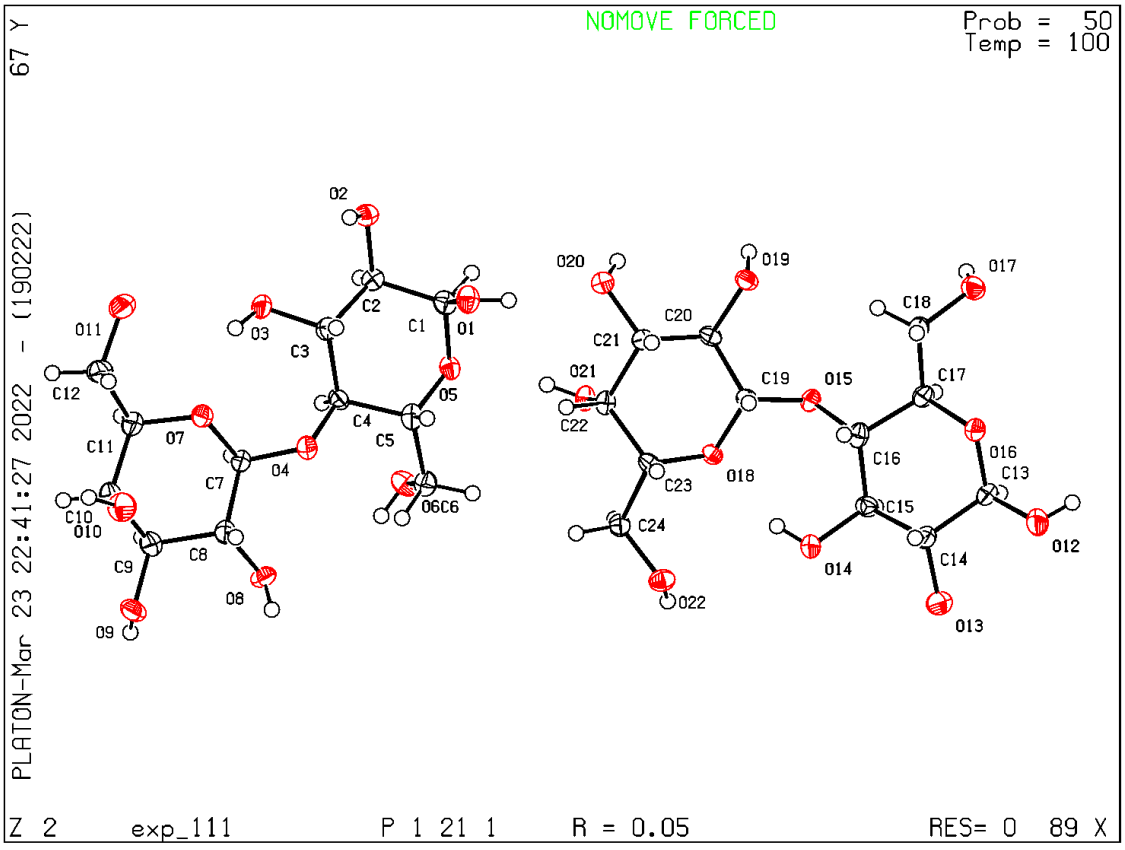


Figure 7. ORTEP plot of αβ-D-lactose.

Chapter 7

Conclusions and Future Work

7 Conclusions

It was stated at the outset that X-ray diffraction was, with the exception of studies on fats in chocolate, both (a) seriously under-represented as a technique for studying the behaviour of crystalline sugars and (b) used in a somewhat basic manner for the characterisation of chocolate and its various components. This work set out to change that status quo and demonstrate that X-ray diffraction could in fact help resolve some of the issues affecting chocolate manufacture. Referring back to the objectives of this body of work, we can attempt to quantify the level of success achieved.

Objective one: "To apply PXRD to sugar raw materials, chocolate intermediates and chocolate samples. In particular, to establish whether or not conventional laboratory-based powder X-ray diffractometers (of a type that could be easily used by industry) are sufficiently powerful to enable reliable and informative characterisation of these materials and samples."

This has certainly been achieved - the industrial partner in this PhD programme of work, Reading Scientific Services Ltd. (RSSL), has recognised the value of PXRD and QPA in particular. The methodology of PXRD characterisation and QPA has been employed extensively during the PhD project to assist RSSL-led investigations into form changes of chocolate and its related products across various processing steps. For example, the amorphicity of spray-dried milk and sucrose solutions was assessed by use of QPA to evaluate spray-dried sweetened milk as a potential avenue for introduction of the amorphous form of sucrose into chocolate formulations. The milk powders were found to consist small quantities of crystalline lactose polymorphs mixed with entirely X-ray amorphous sucrose. The stability of various crystalline and amorphous forms of lactose throughout chocolate processing was also investigated. Milk powders containing differing proportions of solid-state lactose forms were made into chocolate bars and it was found that the relative proportions of these forms were maintained throughout chocolate processing, allowing for caution to be taken in selection of milk powders to be used in chocolate products.

Furthermore, my developments in the applications of QPA and PXRD have led to the employment of post-doctoral research assistants to collect and analyse further chocolate product PXRD data, as well as the extension of the technique into other food products requiring

analysis. My academic supervisor has also continued to use my methodology developments and apply them to other phase mixtures outside of the food science research space.

The level of interest in the application of PXRD to foodstuffs from RSSL has remained consistently high since the initial experiments conducted, so much so that the physical sciences department on-site is looking to invest in a powder diffractometer. Unfortunately, the proprietary nature of their products and manufacturing processes means a lot of this research cannot be published to the wider scientific community. Nevertheless, PXRD has proven itself to be an incredibly powerful technique when correctly utilised in the food industry.

Objective two: "To develop a reliable methodology for the accurate quantification of crystalline sugars in chocolate intermediates and chocolate samples."

This objective has also clearly been achieved, as demonstrated by the results presented in the QPA paper (Chapter 4). It remains to be seen whether or not the methodology will gain traction in the food community. One factor limiting its adoption is that whilst the equipment required to do high-quality PXRD work on organic materials is not especially expensive (ca. £250k), many companies shy clear of the transmission geometry needed for accurate quantitative and structural work, preferring instead to go with reflection geometry for sample throughput and fingerprinting. Regardless, the expertise I have gained from QPA methodology development and research has allowed me to consult on alternate approaches to crystalline phase quantification that RSSL considered to ultimately improve research outcomes and chocolate product processability.

Objective three: "To develop a reliable methodology for the accurate quantification of amorphous sugars in chocolate intermediates and chocolate samples."

Similarly to objective two, this has been achieved as evidenced by the results presented in the QPA paper (Chapter 4). Though quantification of amorphous content in simple mixtures is achievable by direct methods, the complex composition of chocolate and its products means that accurate quantification relies on the introduction of a standard and furthermore the knowledge of the chemical composition of the sample being assessed (by which the amorphous quantity of each component can be ascertained through difference calculations). This is not a serious limitation for in-house, proprietary research, but when analysing competitors' products without knowing the formulation composition only the total amorphous content can obtained by the internal standard QPA method. Alternative QPA approaches such as the partial or no

known crystalline structure (PONKCS) direct method may be able to elucidate the amorphous quantity of every non-crystalline component to some degree. However, the significant peak overlap between broad amorphous diffraction "humps" will make extraction of individual phase amorphous intensities challenging if at all possible.

Objective four: "To use diffraction to investigate the phase behaviour of sugars upon drying of syrups under conditions much simpler than, but closely related to, those experienced during the manufacture of chocolate crumb."

This was perhaps the most challenging aspect of the work undertaken. The full unravelling of the multiphase dried lactose syrup took, on and off, a year to accomplish. Multiple PXRD collections of samples generated over a variety of temperatures and heating durations failed to give produce a powder pattern consisting of a single phase. It took the arrival of a powerful new microfocus SX diffractometer to re-kindle interest in the problem. Knowing its capabilities relative to the old SX diffractometer, the powder sample of the dried syrup was re-examined very closely for single crystals, and a tiny crystal was found which turned out to be a new αβ-lactose polymorph that then allowed the full PXRD pattern to be fitted. Some element of good fortune has to be acknowledged; this was the *only* single crystal of the new polymorph that was found – the rest was strictly polycrystalline.

Multiple other dried syrups were investigated but not included within this thesis, including syrups consisting of mixtures of lactose and sucrose. It was observed that as the concentration of sucrose in the syrup increased, the resultant solid became glassier in appearance and more difficult to analyse by PXRD. Removing the high sucrose concentration dried syrups from the watch glass was challenging due to the sticky properties of the solid. Furthermore, grinding the substance into a powder of particle size appropriate for PXRD analysis was not possible without prior treatment (such as freezing) of the dried syrup, which in turn potentially alters the crystallographic composition of the material. Some sucrose-containing samples that still contained a high concentration of lactose in solution were analysable with difficulty and produced powder diffraction patterns similar to that observed in the published work in chapter 6, with the addition of a background amorphous hump of sucrose and a minor presence of crystalline sucrose.

In order to truly achieve this objective and begin to understand the peculiar thermal behaviour of lactose and sucrose in chocolate crumb, the crystalline properties and thermal

characteristics of dried lactose and sucrose syrups needs to be more extensively assessed. If the behaviour of simple dried sucrose and lactose syrups can be understood, the relative complexity of the mixture and drying procedure can be subsequently changed to more closely match chocolate crumb production. For example, sucrose could be added to milk and the resultant sweetened milk dried to form conditions that more closely match the drying of sweetened condensed milk in crumb manufacture.

7.1 Future work

One application of PXRD in the context of amorphous materials not utilised in this thesis is the pair distribution function approach (PDF, see Chapter 3 for more information). PDF is an incredibly powerful in that it is capable of looking at short range order of amorphous materials, and so is a technique that could prove invaluable in determining whether the sugars that exist in the amorphous state in chocolate and chocolate crumb are truly amorphous or merely nanocrystalline. PDF analysis of high-resolution powder diffraction data could be leveraged to understand the impact of different methods of amorphous generation on the amorphous material produced. High resolution data of amorphous sugars generated by different amorphization techniques (melt-quenching, freeze-drying, ball-milling) was collected at the Diamond XPDF (I15-1) beamline during this thesis. However, due to time constraints related to ill health and the lack of adequate expertise in data processing, the analysis of these materials was not completed. Given the time, it is likely that significant insights into the structure and behaviour of amorphous sugars could be garnered from that data.

Additionally, it would be interesting to observe the crystallographic behaviour of sugar substitutes in chocolate products. The use of artificial sweeteners is steadily growing in chocolate manufacture and the profound difference in physicochemical properties between artificial sweeteners and sugars is certain to have a significant impact on the processability and quality of the final product beyond that which is currently reported in the literature. An investigation into the degree of crystallinity of these compounds in sugar-substituted chocolates could aid in the development of a product that is as agreeable and manufacturable as traditional milk powder and crumb-based chocolates.