

# Pharmaceutical Cocrystal System: Experiment and Computational Orthogonal Correspondence

THESIS

Submitted in partial fulfilment  
of the requirements for the degree of  
**DOCTOR OF PHILOSOPHY**

by

**Raja Sekhar Voguri**  
**ID. No. 2016PHXF0109H**

Under the Supervision of  
**Prof. Manab Chakravarty**

and

Under the Co-Supervision of

**Dr. Archan Dey**

and

**Dr. Subhas Ghosal**



**BITS Pilani**  
Pilani | Dubai | Goa | Hyderabad

**BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI**

**2024**

# BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI

## CERTIFICATE

This is to certify that the thesis titled "Pharmaceutical Cocrystal System: Experiment and Computational Orthogonal Correspondence" submitted by Raja Sekhar Voguri ID No 2016PHXF0109H for award of Ph.D. of the Institute embodies original work done by him/her under my supervision.



**Signature of the Supervisor:**

**Name in capital letters: PROF. MANAB CHAKRAVARTY**

**Designation: Professor**

**Date: 06-12-2023**

# BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI

## CERTIFICATE

This is to certify that the thesis titled "Pharmaceutical Cocrystal System: Experiment and Computational Orthogonal Correspondence" submitted by Raja Sekhar Voguri ID No 2016PHXF0109H for award of Ph.D. of the Institute embodies original work done by him/her under my supervision.



Signature of the Supervisor:

Name in capital letters: PROF. SUBHAS GHOSAL

Designation: Professor

Date: 06-12-2023

Dr. Subhas Ghosal  
Associate Professor  
Department of Chemistry  
National Institute of Technology  
Durgapur-713209, W.B., INDIA



Signature of the Co-supervisor:

Name in capital letters: DR. ARCHAN DEY

Designation: Co-supervisor

Date: 06-12-2023

# BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI

## DECLARATION

I hereby declare that the research work embodied in the thesis entitled “Pharmaceutical Cocrystal System: Experiment and Computational Orthogonal Correspondence” has been carried out under the supervision of Prof. Manab Chakravarty, BITS Pilani, Dr. Subhas Ghosal, National institute of Technology, Hyderabad campus and Dr Archan Dey, IPDO, Dr. Reddy’s Lab, Hyderabad. The work is original and has not been submitted in part or full for any degree for this or to any other university.



**Raja Sekhar Voguri**

## Acknowledgments

*It is a pride moment to show the gratification with a sense of contentment at the long travelled Journey, to be able to recapture some of the fine moments, and to be able to thank some of whom were with me from the beginning, some who joined me at some stage during the journey, whose rally round kindness, love, and blessings have brought me to this day. I wish to thank each and every one of them with all my heart.*

*I want to express my deep and sincere gratitude to **Associate Professor Dr. Subhas Ghosal**. His guidance in applying computational studies has been of great value to me. His encouragement, personal guidance, and constant sharing of ideas in adopting computational methodologies have provided a good basis for the present thesis. I am deeply grateful to **Prof. Manab Chakravarty**, for his constant follow-up and guidance throughout this work. I am very thankful for his support in the NMR studies, which are of high value to any chemical research. His valuable guidance and rigorous review helped me in all the time of my thesis writing.*

*Completing this thesis work could not have been possible without the expertise and unwavering significant support of an exceptional person, **Dr. Archan Dey**. He is not only my co-supervisor but also a great mentor, guide, and friend who has given constant support, motivation, and shared research ideas. I owe immense thank you for holding a lot of patience while conveying complex theories and, from time to time, keeping me on the right track. I would like to thank him for spending his numerous early mornings, and invaluable personal time from his busy life.*

*I acknowledge my doctoral advisory committee (DAC) members, **Prof. Durba Roy** and **Prof. Krishnan Rangan**, for their support and encouragement during this period. I thank DRC convener **Prof. Tanmay Chatterjee** and former DRC convenor **Prof. Jayanti Subhalaxmi** for her valuable suggestions during course work. I thank the Head of the Department (HOD), **Prof. Sounak Roy**, and the chemistry department faculty members for their valuable interaction and discussions throughout the journey.*

*I am very thankful to Santosh Ranga, for his significant support at any time to access computational facilities and run computational jobs. It has always been an interesting journey working with Dr. Praveen Chappa. I thank him for his contribution and support towards this thesis work. I thank Abbagani Uppalaiah and Narishmha for their kind support for Single crystal-XRD and NMR studies. I sincerely thank **Assistant Professor Dr. Binoy Krishna Saha** (Pondicherry University) for his support in addressing complex disorders while solving the crystal structures of drug-polymer cocrystals. I acknowledge **Dr. Bankanidhi Sahoo's***

*contribution in applying AI/ML studies in this work. I sincerely thank Dr. Deepika (Dr. Reddy's Laboratories) and Dr. Banchhanidhi Prusti for their kind support.*

*I am grateful to Prof. V. Ramagopal Rao (Vice-Chancellor BITS), Prof. Souvik Bhattacharya (former Vice-Chancellor BITS) and Prof. G Sundar (Director BITS Pilani Hyderabad Campus) for giving me an opportunity to pursue and to carry out my Doctoral research work. I would like to express my sincere thanks to **Prof. V. V. Vamshi Krishna**, Dean, Academic Research Division (Ph.D. Programme), for their continuous support and encouragement during my research work. And I thank Mr. Praveen, ARD for his help and information in completing the course formalities in time.*

*It is impossible to extend enough thanks to my parents (**Voguri Saidaiah & Voguri Padmavathi**), for providing me with unfailing moral support and continuous encouragement throughout my years of study and financial support throughout this process. I am deeply indebted to Chandra Sekhar (brother), K Srinivasa Rao, K L Srinivasa Rao, Voguri Venkateswaralu (Grand Father) and all of my family members. I thank my Wife (Jyothi) and Daughter (Lakshmi Tejasa) for supporting and maintaining a comfortable environment at home to carry out my thesis work.*

Date: 12-12-2023



(Raja Sekhar Voguri)

## Abstract

This thesis deals with Pharmaceutical cocrystal Systems, aiming to fulfill the lagging areas where current experimental methodologies alone do not have the competence to achieve the objectives, which can be overcome by molecular simulations. Two aspects were discussed: 1) Application employing computational studies for molecular/atomic level structural characterization of cocrystals vs salt and 2) Virtual screening of drug–polymer cocrystals. These studies were conducted on pharmaceutical cocrystals composed of small molecule coformer and polymeric coformer.

**Chapter 1** briefly introduces pharmaceutical cocrystals and challenges in the characterization and traditional screening methods of cocrystals. It discusses molecular simulation tools and methodologies proven to overcome the above-mentioned challenges. DFT-D2 calculations overcame the challenges in the characterization of crystal structures. The adopted molecular simulation screening methodologies include tools on structural, thermodynamic, and kinetic-based descriptors combined with ML/AI tools. Overall, the simulation tools and corresponding methodologies employed to address the challenges in the characterization of cocrystals and in-silico screening of drug–polymer cocrystals were briefly discussed in this chapter.

**Chapter 2** is about materials and methods. It covers crystallization experimental procedures, characterization by X-ray diffraction, thermal, spectroscopy, microscopy techniques, etc. It also covers computational procedures and step-by-step methodologies applied to validate crystal structures, determine accurate proton position, address disorders in the crystal structures, and in-silico screening of drug–polymer cocrystals. Finally, this section includes simulation tools, computational procedures, and corresponding protocols and methodologies employed to overcome the challenges in the characterization of cocrystals and in-silico screening of drug–polymer cocrystals.

**Chapter 3** covers the characterization of dimorphic forms of agomelatine–phosphoric acid (1:1) molecular complexes using combined complementary experimental techniques and computational studies. The nature of these polymorphic two forms at room and high temperatures are enantiotropic and reversible. This system has thrown potential challenges due to the growth of a good diffraction-quality single crystal. Hence, combined ab initio powder XRD structure determination and dispersion-corrected density functional theory analysis were applied for structure characterization. The outcome of crystal structure results demonstrated. Thus, this chapter covers the application of experimental techniques combined with

computational studies to overcome the challenges in characterizing agomelatine–phosphoric acid (1:1) molecular complexes as a salt or cocrystal, which one is a burning question of the pharmaceutical industry from legal and regulatory perspective for given API with both the possibilities.

**Chapter 4** describes the characterization of dapsone–polyethylene glycol cocrystal, a drug–polymer cocrystal characterized by combining experimental and theoretical methodologies. The disorders in the crystal structure around the polyethylene glycol chain were addressed by employing DFT–D2 calculations. This work was extended toward in-silico screening to discover more drug–polymer cocrystals. Tailor-made computational screening protocol combined with artificial intelligence–machine learning tools were built with a success rate of 90%, discovering two more drug–polymer cocrystals with PEG as a cofomer. Finally, this chapter covers the characterization of dapsone–polyethylene glycol cocrystal by combining experimental and theoretical methodologies and in-silico screening studies for drug-polymer cocrystals.

**Chapter 5** is an extension of the previous chapter, where sulfanilamide–PEG cocrystal is one of the screened outcomes from the developed in-silico screening methodology for drug–polymer cocrystals. The material was successfully synthesized and further characterized by NMR spectroscopy techniques. Structural characterization of this material is highly challenging due to the intractable disordered polyethylene chain in the crystal structure. This complex disorder was addressed by combining experimental and computational tools, especially for solid-state NMR data interpretation. It has allowed the establishment of an interesting structure-property relationship phenomenon that can be connected to the contemporary fields of multi-domain single crystal and the emergence of chirality due to supramolecular structure.



## Table of Contents

Certificate	i
Declaration	iii
Acknowledgments	iv
Abstract	vi
Table of contents	viii
List of Tables	xi
List of Figures	xii
List of abbreviations	xvii
1.0. Introduction	1
1.1. Overview of Pharmaceutical Cocrystals	2
1.2. FDA and EMA guidelines	4
1.3. Migration of proton: Is it a Salt or Cocrystal?	5
1.4. Challenges with the characterisation of polycrystalline materials over single crystals	5
1.5. Recent advances in SDPD and treatment of disorders in crystal structures: An extension towards DFT-D2 calculations	6
1.6. Drug–polymer cocrystals: A promising and emerging alternative	8
1.7. In–silico screening of drug–polymer cocrystals	10
1.7.1. Molecular Simulation	11
1.7.2. QSPR (Quantitative structure-property relations)	16
1.7.2.1. Descriptors based approach	17
1.7.3. Conductor-Like Screening Model for Real solvents (COSMO-RS)	20
1.7.4. Crystal Structure Prediction	21
1.7.5. Knowledge and Database driven	24
1.8. Relevance of this work in Brief	27
1.9. References	29
2.0 Experimental and theoretical methods	38
2.1. Crystallization of materials	39
2.1.1. Agomelatine–phosphoric acid	39
2.2. Characterization	41
2.2.1. Powder-XRD analysis	41
2.2.2. Powder-XRD data collection for polymorphic form identification	42
2.2.3. Powder-XRD data collection for Indexing followed by Crystal structure determination	42
2.2.4. Variable Temperature Powder-XRD studies	43
2.2.5. Crystal structure determination from SCXRD	45

2.2.6.	Thermal analysis	43
2.2.7.	Hot Stage Microscopy	44
2.2.8.	Spectroscopy	44
2.2.9.	Theoretical Methods	46
2.2.9.1.	Geometric optimization of Molecular structure using Forcefield calculations	46
2.2.9.2.	Polymorphic purity from powder-XRD: Determination of Lattice parameters and their refinement by Pawley refinement	46
2.2.9.3.	Crystal Structure Determination from powder-XRD	46
2.2.9.4.	Comparison (overlay) of crystal structure CRYSTALCAMP	48
2.2.9.5.	Calculation of Molecular and crystal structure descriptors using QSAR Module	48
2.2.9.6.	Calculation of Thermodynamic descriptors	49
2.2.9.7.	Calculation of Supramolecular descriptors	50
2.2.9.8.	Calculation of NMR chemical shift calculations using GIPAW method	51
2.2.9.9.	Crystal Morphology simulation from crystal structure	54
2.3.	References	55
3.0	Agomelatine–phosphoric acid molecular complex	59
3.1.	Introduction	60
3.2.	Polymorphism and thermal behaviour of AGL–P polymorphic forms	66
3.3.	Crystal structure determination: Salt vs co–crystal confirmation	69
3.4.	Crystal structure analysis: Mechanism of phase transition	76
3.5.	Conclusions	84
3.6.	References	85
4.0	Dapsone–polyethylene glycol cocrystal and virtual screening of drug–polymer cocrystals	90
4.1.	Introduction	91
4.2.	Characterization of DAP–PEG complex	94
4.2.1.	Synthesis and Stoichiometry determination for DAP–PEG cocrystal	94
4.2.2.	Crystal structure analysis of DAP–PEG from Powder-XRD	95
4.2.3.	Crystal structure analysis of DAP–PEG complex from SXRD:	99
4.3.	Virtual Screening of drug–polymer cocrystals	101
4.4.	Summary	113
4.5.	References	113
5.0	Sulfanilamide–polyethylene glycol cocrystal	120
5.1.	Introduction	121

5.2.	Characterisation of SUL–PEG cocrystal:	122
5.3.	Crystal structure analysis and insights into structure–property relationship:	127
5.4.	Summary	140
5.5.	References	141
	Future perspective	144
	Annexures	145
	List of publications	219
	Biography	
	Biography of the Candidate	220
	Biography of the Supervisor	221
	Biography of the Co-Supervisors	222

## List of Tables

Table No.	Caption	Page No.
1.1	Commercially available pharmaceutical cocrystals	4
2.1	List of Screening experiments for API-PEG complexes	40
3.1	Crystal structure survey of agomelatine M-APIs (Multi-component Active Pharmaceutical Ingredients)	63
3.2	Comparison between RT-form and HT-form from Rietveld refinement and DFT optimized with respect to crystal structure similarity and Amide group bond lengths and their ratio.	70
3.3	Lattice parameters of RT-form and HT-form	72
3.4	Chemical Shift of Amidic C in the Corresponding SS-NMR and Compared to AGL Form-II	76
4.1	Crystal structure analysis of drug-polymer co-crystal/inclusion complexes available in the literature	92
4.2	Intermolecular hydrogen-bonding interactions (Å, °)	96
4.3	Comparison of crystallographic data obtained from SXRD and SDPD of DAP-PEG co-crystal	100
4.4	Intermolecular hydrogen-bonding interactions (Å, °)	101
4.5	Comparison of $E_{\text{bind}}$ , $\chi$ and $E_{\text{mix}}$ values calculated for Dapsone and PEG	104
4.6	Master datasheet for ML modelling of various descriptors along with experimental and prediction result summary	110
5.1	Table illustrates the impact of different PEG molecular weights and process on the crystallinity of DAP-PEG and SUL-PEG materials	125
5.2	Crystallographic and DFT-D2 optimization details of SUL-PEG cocrystal	128
5.3	Summary of second harmonic generator (SHG) results	131
5.4	Composition of hydrogen bonding interaction and $^{13}\text{C}$ SS-NMR spectra (Geometry Optimized Structure)	134
5.5	Comparison table on Helix parameters of PEG (As such), SUL-PEG ( <i>Pbca</i> ), SUL-PEG ( <i>Pca2<sub>1</sub></i> ) and DAP-PEG	138

## List of Figures

Figure No.	Caption	Page No.
1.1	Flow chart of crystal structure determination from Powder-XRD Methodology	7
1.2	Flow chart of methodology to address disorders in Crystal structures using dispersion corrected DFT calculations	8
1.3	Drug–polymer co-crystal: A subclass between Amorphous solid dispersions and cocrystals	9
1.4	Overview of in-silico screening of drug–polymer co-crystals	11
1.5	Overview of in-silico screening of drug–polymer co-crystals using descriptors from QSPR calculations	16
1.6	Flow chart of CSP protocol	22
1.7	Overview of in-silico screening of drug–polymer co-crystals using descriptors from ML based calculations	25
1.8	Overview on ML based in-silico screening method adopted for drug–polymer co-crystals	27
2.1	Schematic illustration of Bragg’s Law	42
2.2	Flow chart showing the overview of validation of crystal structure by using powder/single crystal-XRD and SS–NMR as complementary techniques	51
2.3	Flow chart showing the overview of validation of crystal structure by using powder-XRD and SS–NMR as complementary techniques	53
2.4	Flow chart showing the methodology for morphology prediction	55
3.1	Outlines of FDA and EMA guidelines for cocrystals conformation	60
3.2	Schematic presentation of AGL–P in the form of (a) cocrystal (b) continuum and (c) salt of AGL–P via C=O...H-O-(P) interaction	62
3.3	Schematic presentation of AGL with torsion angle indicated	62
3.4	(A) Overlay of DSC and TGA thermograms of AGL–P with DSC curve showing two endothermic events at ~42°C indicates phase transition and melting at ~118°C. TGA indicates no weight loss. (B) HSM images of AGL–P at 25°C and 50°C representative of RT-form and HT-form, respectively.	66
3.5	Overlay of VT–XRPD powder-XRD data collected in the temperature range of 30°C to 60°C at each 10°C increment and at 30°C after cooling	67
3.6	DSC thermogram for heat–cool cycle of AGL–P with phase transition and recrystallization events are indicated	68
3.7	Overlay of VT–XRPD patterns of AGL–P RT–form and HT–form during heating and cooling monitoring	69
3.8	Data collection strategy for SDPD by optimizing the factors: preferred orientation and rate of HT–form to RT–form conversion. Please note that peaks (marked with arrows) corresponding to the RT–form in HT–form data were trimmed. The broad beak peak indicated with dotted box corresponds to	70

	the background from foil used in foil transmission XRPD data collection	
3.9	Overlay of RT-form and HT-form crystal structures Rietveld-refined (red) and DFT-optimized (blue)	71
3.10	Representative molecular diagram showing proton shift between agomelatine and phosphoric molecules in RT-form and HT-form and the associated bond distances and angle matrix	73
3.11	Overlay of solid-state <sup>13</sup> C NMR spectrum of agomelatine form-II and AGL-P RT-form	74
3.12	Overlay of experimental FT-IR spectrum of AGL, phosphoric acid (PA), and RT-form. Wavenumbers for Amide-I and Amide-II band are listed in the table	75
3.13	Representative plot between $\Delta pK_a$ and $d_{C-O}/d_{C-N}$ . M-APIs of AGL reported as salts are indicated by red color triangles (▲) and M-APIs of AGL reports as cocrystal are indicated by black color circles (●). RT and HT forms are marked by a blue diamond symbol (◆). For eye guidance, green and brown dotted lines represent the median for salt and cocrystal $d_{C-O}/d_{C-N}$ ratio respectively	76
3.14	Crystal packing diagram showing possible hydrogen bonding interactions for RT-form (A) and HT-form (B)	77
3.15	Image analysis of HSM pictures RT-form (A) and HT-form (B). (C) Table for crystal size measured along with long the major axis. (D) Simulated crystal morphology for RT and HT forms	78
3.16	(A) Crystal structure showing trilayers building block of RT-form and HT-form with representative distance matrix. (B) Temperature-induced proton migration and change in torsion angle (C) Change in Naphthyl-Naphthyl stacking angle between consecutive trilayers and associated distances to support negative thermal expansion along the b-axis	79
3.17	(A) Molecular structure of agomelatine showing four possible torsion angles. (B) All torsion angle values for RT and HT-forms are tabulated. (C) Plot between $\tau_3$ and $d_{C-O}/d_{C-N}$ from the crystal structures reported with AGL M-APIs as one of its components. Circles (●) indicates co-crystal, red triangles (▲) indicate salt and green circles (●) indicates pure AGL polymorph. Data points corresponding to 1: HT-form and 2: RT-form are indicated	80
3.18	Comparison of Hirshfeld surface analysis and molecular interactions for RT and HT-forms. (A) Surface is normalized with respect to internal and external interactions [ $d_{norm}$ with colour scale: -0.8 (red) to 1.4 (blue)] (B) Surface with respect to external interactions only [ $d_e$ colour scale: 0.5 (red) to 2.5 (blue)].	83
3.19	Relative contributions (%) of intermolecular interactions to the Hirshfeld surface areas of the molecular constituents in RT and HT-forms	83

3.20	(A) Comparison of crystallization schematics for optimized patented process and modified current process. Differentiating steps were highlighted, namely H <sub>3</sub> PO <sub>4</sub> reagent grade (red) and drying step (blue). (B) Overlay of powder-XRD data of three experimental trials showing concomitant formation of RT and HT-forms. However, a trend of increasing in HT-form content with H <sub>3</sub> PO <sub>4</sub> was noticed	84
4.1	Molecular structures of (A) dapsone and (B) polyethylene glycol	91
4.2	Schematic presentation of working framework for Virtual Screening of polymeric cocrystals	93
4.3	Powder-XRD overlay of (A) Dapsone, (B) PEG 4K and (C) DAP-PEG 4K	94
4.4	<sup>1</sup> H NMR of DAP-PEG (1:4) using CDCl <sub>3</sub> as solvent. <sup>1</sup> H NMR 400 MHz (CDCl <sub>3</sub> ): δ = 7.4 (d, 4H), 6.5 (d, 4H), 5.9 (s, 4H), 3.5 (s, 16H)	95
4.5	Experimental, calculated and the difference profile of DAP-PEG600 co-crystal obtained from Rietveld refinement	96
4.6	Overall close packing showing interdigitated chains of PEG and hydrogen bonded DAP molecules and Hydrogen bonding interactions in DAP-PEG crystal structures solved from (I) XRPD & (II) SXRD data and (III) Overall close packing showing interdigitated chains of PEG crystal structures	98
4.7	(A) Overlay of experimental <sup>13</sup> C ss-NMR spectra of dapsone and DAP-PEG 4K experimental along with simulated <sup>13</sup> C SS-NMR spectra of DAP-PEG 600 (SDPD) and DAP-PEG 600 (SXRD); (B) Plots between experimental <sup>13</sup> C chemical shifts and simulated <sup>13</sup> C chemical shifts for DAP-PEG SDPD and DAP-PEG SCXRD. The R <sup>2</sup> fit represents the correlation of chemical shift value of experimental and simulated; and (C) DAP-PEG molecular structure with C-atoms assignment	99
4.8	Flowchart showing protocol for virtual screening. *Amorphous cell was employed to simulate Hydrogen bonding propensity and drug-PEG interactions; Also, to mimics kinetic information.	102
4.9	Plot between calculated E <sub>bind</sub> (Kcal/Mole) between DAP-PEG verses change in PEG chain length (n)	104
4.10	Plot between descriptors and their coefficients. The coefficient values close to 1 or -1 indicate the corresponding better descriptors that fulfill classification model	107
4.11	3D plots-I (ΔlogP, Prediction Probability and API-PEG <sub>af</sub> ) and 3D plots-II (ΔlogP, V <sub>M</sub> /V <sub>Mpm</sub> and E <sub>bind</sub> ) for (A) Training dataset and (B) Screening dataset	109
5.1	Molecular structures of (A) SUL-PEG 400, SUL-PEG 600, SUL-PEG 400/4K (1:1, solid solution), SUL-PEG 4K and (B) DAP-PEG 400 (1:4), DAP-PEG 600, DAP-PEG 400/4K (1:1, solid solution), DAP-PEG 4K	122
5.2	Powder-XRD overlay of (A) Sulfanilamide, (B) PEG 4K and (C) SUL-PEG 4K	123

5.3	Pawley profile fitting of SUL-PEG 600 cocrystal with lattice parameters embedded in it. Red data points indicate experimental data, blue pattern indicates calculated pattern and black pattern indicates the difference plot between experimental and calculated pattern. Green colour lines indicate <i>h k l</i> reflections.	123
5.4	NMR study of the SUL-PEG 4K cocrystal: (A) <sup>1</sup> H NMR spectra in DMSO-D6 and (B) <sup>13</sup> C NMR spectra in DMSO-D6. Structural assignments for both spectra were indicated by labelling with assigned Atom numbers. The signals integration in <sup>1</sup> H NMR spectra for Atom no. 1 (2H, SUL) and Atom no. 5 & 6 (4H, PEG) is in the ratio of 2.0: 9.9 = 1:2.5, confirming the existence of both species with desired stoichiometry.	124
5.5	Powder-XRD overlay of (A) SUL-PEG 4K and (B) DAP-PEG 4K cocrystals synthesized using fast solvent (FS) evaporation process	124
5.6	XRPD overlay of (A) DAP-PEG 600 and DAP-PEG 4K (2 Preparations) and (B) SUL-PEG 600 and SUL-PEG 4K (2 Preparations). Here SS indicates that sample was prepared by slow solvent evaporation method and FS indicates that samples were prepared by fast solvent evaporation method labelled with %crystallinity	126
5.7	XRPD overlay of (A) DAP-PEG 600 and DAP-PEG 4K (2 Preparations) and (B) SUL-PEG 600 and SUL-PEG 4K (2 Preparations). Here SS indicates that sample was prepared by slow solvent evaporation method and ASD indicates that samples were prepared by Anti-Solvent Addition method labelled with %crystallinity.	126
5.8	XRPD overlay of (A) DAP-PEG 400, DAP-PEG 400/4K (Solid Solution), DAP-PEG 4K and (B) SUL-PEG 600, SUL-PEG 400/4K (Solid Solution), SUL-PEG 4K. Here ASD indicates that samples were prepared by Anti-Solvent Addition method labelled with %crystallinity. Since SUL and PEG 400 did not form cocrystals, hence %crystallinity was not mentioned	127
5.9	Asymmetric units (indicated with dotted boxes) of (A) SUL-PEG ( <i>Pbca</i> ), (B) SUL-PEG ( <i>Pca2<sub>1</sub></i> ) and (C) DAP-PEG crystal structures showing hydrogen bonding interactions between drug and PEG chain. Overall packing of interdigitated chains of PEG & drug molecules with hydrogen bonding interactions for (D) SUL-PEG ( <i>Pbca</i> ), (E) SUL-PEG ( <i>Pca2<sub>1</sub></i> ) and (F) DAP-PEG	129
5.10	(A) Reported and (B) Simulated SS <sup>13</sup> C NMR patterns of form- $\alpha$ , form- $\beta$ and form- $\gamma$ SUL crystal structures	132
5.11	(A) Experimental and (B) Simulated <sup>13</sup> C SS-NMR patterns of form- $\beta$ SUL crystal structures, (C) sulfanilamide molecular structure with C-atoms assignment and (D) Plots between experimental <sup>13</sup> C chemical shifts and simulated <sup>13</sup> C chemical shifts for form- $\beta$ SUL. The R <sup>2</sup> fit represents the correlation of chemical shift value of experimental and simulated	132



5.12	(A) SUL ( $\beta$ -polymorph) experimental $^{13}\text{C}$ SS-NMR spectra with peak assignment and read arrow indicate a spinning side band (B) SUL-PEG experimental $^{13}\text{C}$ SS-NMR spectra with peak assignment and read arrow indicate a spinning side band; (C) SUL-PEG simulated $^{13}\text{C}$ SS-NMR spectra of local crystal structure that solved in $Pca2_1$ space group. (D) SUL-PEG simulated $^{13}\text{C}$ SS-NMR spectra of average crystal structure that solved in $Pbca$ space group	133
5.13	Figure showing top view and side view of PEG chains in (A) PEG ( <i>As such</i> ), (B) SUL-PEG ( $Pbca$ ), (C) SUL-PEG ( $Pca2_1$ ) and (D) DAP-PEG. In top view, right-handed helices were indicated in blue colour and left-handed helices were indicated in red colour. PEG chain helix parameters such as helix Pitch, helix pitch Angle and helix diameter were indicated. For DAP-PEG, diameter of pseudo helix PEG chain was not calculated	137

### List of Abbreviation

API	Active Pharmaceutical Ingredients
GRAS	Generally Recognized As Safe
EMA	European Medical Agency
CSD	Cambridge Structural Database
NDA	New Drug Application
ANDA	Abbreviated New Drug Application
FDA	Food and Drug Administration
XRD	X-ray diffraction
XRPD	Powder X-ray diffraction
SXRD	Single X-ray diffraction
ED	Electron Diffraction
SDPD	Structure determination from powder diffraction
DFT	Density Functional Theory
ASD	Amorphous solid dispersions
MD	Molecular Dynamics
COMPASS	condensed-phase optimized molecular potentials for atomistic simulation
CED	Cohesive Energy Density
QSPR	Quantitative structure-property relations
MEPS	Molecular Electrostatic Potential Surface
CCDC	Cambridge Crystallographic Data Centre
HBP	Hydrogen Bond Propensity
CC	Cocrystal
SSIP	Surface site interaction points
COSMO-RS	Conductor-Like Screening Model for Real solvents
QM	Quantum mechanics
PCM	Polarizable continuum
ASC	apparent surface charge
CSP	Crystal Structure Prediction
ML	Machine Learning
M-APIs	Multicomponent Active Pharmaceutical Ingredients
AGL-P	Agomelatine phosphoric Acid
VT-XRPD	Variable Temperature Powder X-ray diffraction
RT-form	Room Temperature form
HT-form	Hight Temperature form
DAP-PEG	Dapsone-polyethylene glycol cocrystal system
SUL-PEG	Sulfanilamide-polyethylene glycol cocrystal system
$\lambda$	Wavelength
d	d-spacing
n	Integer
Cu	Copper

Å	Angstrom
DSC	Differential Scanning Calorimetry
TGA	Thermo Gravimetric Analysis
°C	Centigrade
min	Minutes
mg	Milligram
HSM	Hot Stage Microscopy
IR	Infrared
cm <sup>-1</sup>	Centimetre inverse
FT-IR	Fourier Transform Infrared
KBr	Potassium Bromide
NMR	Nuclear Magnetic Resonance
MHz	Mega Hertz
TMS	Tetra methyl Silane
ppm	Parts Per Million
DMSO	Dimethyl sulfoxide
CP/MAS	Cross-Polarization Magic Angle Spinning
eV	Electron Volt
GGA	Generalized Gradient Approximation
PBE	Perdew, Burke and Ernzerhof
CASTEP	Cambridge Serial Total Energy Package
QSAR	Quantitative structure-activity relationships
GIPAW	Gauge Including Projector Augmented Waves
EFG	electric field gradient
δ <sub>iso</sub>	Isotropic chemical shift
σ <sub>ref</sub>	Reference Shielding
BFDH	Bravais-Friedel Donnay-Harker
$R_{hkl}$	Relative growth rate
$d_{hkl}$	interplanar distance
IUPAC	International Union of Pure and Applied Chemistry
EA	Ethyl Acetate
PA	Phosphoric Acid

# Chapter 1

## Introduction

---

*This chapter starts with a brief introduction to pharmaceutical cocrystals followed by the application of molecular simulation tools and methodologies to overcome the challenges in the structural characterization and traditional screening methods for cocrystals.*

---

### 1.1. Overview of Pharmaceutical Cocrystals

The foundation for the approach of Pharmaceutical cocrystals has a hereditary relationship with the concepts of crystal engineering, which were first introduced by Pepinsky in 1955 and later extended by G. M. J. Schmidt<sup>1</sup>. It took over 100 years to evolve from early stages to commercial applications, starting from the discovery of the X-ray diffraction phenomenon using Bragg's Law in 1913<sup>2</sup>, followed by concepts of crystal engineering, molecular packing, the study of organic supramolecular compounds (synthons, motifs and graph sets)<sup>3</sup>, application of supramolecular concepts in pharmaceutical cocrystals, the study of Physicochemical properties & patenting and formulation<sup>4,5</sup>.

The concepts of supramolecular packing were shaped by Gautam R. DesiRaju to bridge with applications in pharmaceuticals by establishing reliable connections between molecular and supramolecular assemblies, which are governed by chemical and geometrical factors<sup>6</sup>. He defined the concept of crystal engineering as "*the understanding of intermolecular interactions in the context of crystal packing and in the utilization of such understanding in the design of new solids with desired physical and chemical properties*"<sup>7</sup>. The molecular arrangement and intermolecular interactions between the components in a crystallized material depend on various factors, such as molecular structural properties (size, shape, and topology), number and nature of functional groups, molecular flexibility, etc.

Generally, screening experiments are designed based on these crystal engineering concepts, molecular properties, and CSD database analysis to achieve novel polymorphs or cocrystals<sup>8</sup>. The complexity of these factors, which dictate the supramolecular assemblies in the crystals, will multiply with an increase in the number of components in the mixed crystals<sup>9</sup>. Also, predicting based on simple previous knowledge is difficult since it depends on many factors. Pharmaceutical cocrystals are multi-component neutral species formed by cocrystallization of an API (Active Pharmaceutical Ingredient) with more than one Generally Recognized As Safe (GRAS) conformer molecule interacting by means of non-covalent or nonionic intermolecular interactions without impacting the therapeutic activity of the Drug molecule. These non-covalent intermolecular interactions in the solid forms between API and co-formers can be hydrogen bonding interactions, Van der Waal's forces of attraction,  $\pi$  -  $\pi$  interactions, and halogen interactions<sup>10,11</sup>. A co-former molecule should be pharmaceutically acceptable under the category of Generally Recognized As Safe (GRAS)<sup>11</sup>. Although EMA (European Medical Agency) depicts the cocrystals as a subclass of solvates, solvents cannot be considered as co-formers to form cocrystals, implying that a co-former molecule should exist as a solid material under room temperature conditions<sup>12</sup>. The physicochemical properties of crystalline material

---

depend on the molecular arrangement in its crystal structure. During cocrystallization, the crystal structure of the native API polymorphic form will change through the incorporation of another molecule (co-former) in its crystal lattice<sup>12</sup>. This phenomenon should occur without making or breaking bonds and must retain the chemistry of the API<sup>3</sup>. As a result, the material physiochemical properties such as colour, solubility, bioavailability, dissolution rate, stability (chemical and physical), hygroscopicity, mechanical properties, morphological properties, and dose–response relationship may be altered<sup>11</sup>.

Statistically, it is well known that about 90% of new chemical entities, which are under drug development pipeline are getting rejected and 40% of the drugs already in the market are suffering due to poor solubility and permeability<sup>13</sup>. In general, these molecules will fall either under BCS (Biopharmaceutical classification system) class–II (Low solubility and High permeability) or BCS class–IV category (High solubility and Low permeability)<sup>11</sup>. Traditionally, for APIs exhibiting poor physical properties, well-known alternatives preferred are salts, polymorphs, hydrates, solvates, and amorphous forms<sup>12,14</sup>. These forms have their own limitations, such as 1) drug molecules without ionizable functional groups cannot form salts, 2) molecules having limited scope to show polymorphism, 3) molecules having low affinity to bind or encapsulate solvents or forming solvates with low stability and 4) many molecules will form only unstable or high energy amorphous materials with low glass transition temperatures resulting with poor stability<sup>12</sup>. These limitations have made drug agencies leap for an alternative, materializing concepts of pharmaceutical cocrystals from academics to industrial applications<sup>13,15</sup>.

After FDA and EMA drug agencies published guidelines for pharmaceutical cocrystals in 2013<sup>16</sup> and 2014<sup>17</sup>, it opened broader opportunities for both innovator and generic industries, accelerating the scope for product development followed by filing for NDA (New Drug Application) and ANDA (Abbreviated New Drug Application) Even then, the innovator must prove superior properties for cocrystal forms compared to its native polymorph. The challenge lies for generic companies to demonstrate safety and efficacy equivalency to the innovator product<sup>12</sup> (**Table 1.1**).

**Table 1.1:** Commercially available pharmaceutical cocrystals<sup>12</sup>.

Brand Name	Year of approval	API	Co-former	Improved property
<b>Depakote</b>	1989	Valproic acid	Valproate sodium	Stability and hygroscopicity
<b>Lexapro</b>	2002	Escitalopram Chloral hydrate	Oxalate Betaine	Stability
<b>Suglat</b>	2014	Ipragliflozin	L-proline	Stability
<b>Entresto</b>	2015	Valsartan	Sacubitril	Pharmacokinetics and bioavailability
<b>Odomzo</b>	2015	Sonidegib monophosphate	Phosphoric acid	Improved hygroscopic behaviour
<b>Steglatro</b>	2017	Ertugliflozin	Z-Pyroglutamic acid	Stability
<b>Siponimod–fumaric acid</b>	2019	Siponimod	Fumaric acid	Stability and hygroscopicity

### 1.2. FDA and EMA guidelines

As per recent FDA guidelines for pharmaceutical cocrystals, to predict whether an API and a coformer forms cocrystal or salt:

*"Generally speaking, if the API and its coformer have a  $\Delta pK_a$  ( $pK_a$  (base) -  $pK_a$  (acid)) > 1, there will be substantial proton transfer resulting in ionization and potential formation of salt as opposed to a cocrystal. On the other hand, if the API and its coformer have a  $\Delta pK_a$  ( $pK_a$  (base) -  $pK_a$  (acid)) < 1, there will be less than substantial proton transfer. If this criterion is met, the API–coformer entity should be classified as a co-crystal<sup>18</sup>".*

As per FDA guidelines, API and co-former should be neutral in the crystal structure, interacting with each other via non-covalent or non-ionic interactions. They must satisfy the  $\Delta pK_a$  ( $pK_a$  (base) –  $pK_a$  (acid)) rule. API and co-former should dissociate before reaching the site of pharmacological activity<sup>19</sup>. But as per inferences from research articles<sup>20</sup>, a salt will be obtained if the  $\Delta pK_a$  between two components is greater than 3, and a cocrystal is obtained if the  $\Delta pK_a$  is less than -1<sup>20</sup>. In general, a salt is formed through an acid–base reaction between an API and an acid or a base, and proton transfer happens from an acid to a base. For most of the cocrystals,  $\Delta pK_a$  value falls between -1 and 3; in such cases, depicting the proton position between acid and base is challenging. Hence, thorough characterization is needed for these adducts falling in this intermediate  $\Delta pK_a$  range<sup>21</sup>.

### 1.3. Migration of proton: Is it a Salt or Cocrystal?

Hydrogen, being a weak scattering atom, it is very challenging to identify the location of the proton between acid and base<sup>22</sup>. Such protons are bound to be labile in between the path of acid–base interacting sites. The position of the proton can be influenced and may be varied due to various factors such as crystal environment and temperature, etc. Usually, if a good quality crystal was grown, then it might be possible to determine position of proton between acid and base by using X-ray diffraction or Neutron diffraction techniques. There are other techniques or methods, such as Infrared spectroscopy, Raman spectroscopy, X-ray photoelectron, photoluminescence and solid-state NMR, which can be orthogonally used in this context<sup>21</sup>.

### 1.4. Challenges with the characterisation of polycrystalline materials over single crystals

Although single crystal XRD (SCXRD) technique is superior, many pharmaceutical materials are likely to form polycrystalline materials rather than grow as good quality single crystals even after numerous efforts. Determination of crystal structure for such materials may not be possible by single crystal diffraction techniques, and even after attempts by more extended periods for collecting the diffraction data, may end up with poor quality data. In such cases, there is no choice but to rely on other techniques such as 3D electron diffraction or powder-XRD diffraction studies.

3D electron diffraction (ED)/ micro ED has extended the limits of crystallography by determining crystal structures from microns and sub-micron crystals<sup>23,24</sup>. Primarily, this technique has become a boon to those who want to determine the crystal structure for materials having challenges to grow them into single crystals. Also, determination of accurate positions of H atoms in the small compounds becomes the most important<sup>25,26</sup>. The major challenges with these techniques are 1) Its requirement for experts both from electron diffraction and crystallography domains<sup>23</sup> and 2) Since Electrons strongly interact with materials, it can lead to sample degradation for organic compounds. Traditional single X-ray diffraction techniques cannot be helpful when getting quality single crystals is a major challenge and X-rays being weak scatters of H atoms cannot be helpful to accurately determine the positions of H atoms<sup>27</sup>. Even using sophisticated techniques such as synchrotron radiation, it is difficult to measure H-atoms positions<sup>23</sup>. Advanced techniques such as neutron diffraction studies can be helpful, but it require completely deuterated materials<sup>26</sup>.

On the other hand, although it is feasible to determine crystal structure from powder-XRD data, there are certain limitations that should be noted. SC-XRD gives three-dimensional data; whereas powder-XRD provides one-dimensional data, and it is more challenging to solve crystal structure from powder-XRD compared to single crystal X-ray diffraction data. Usually,

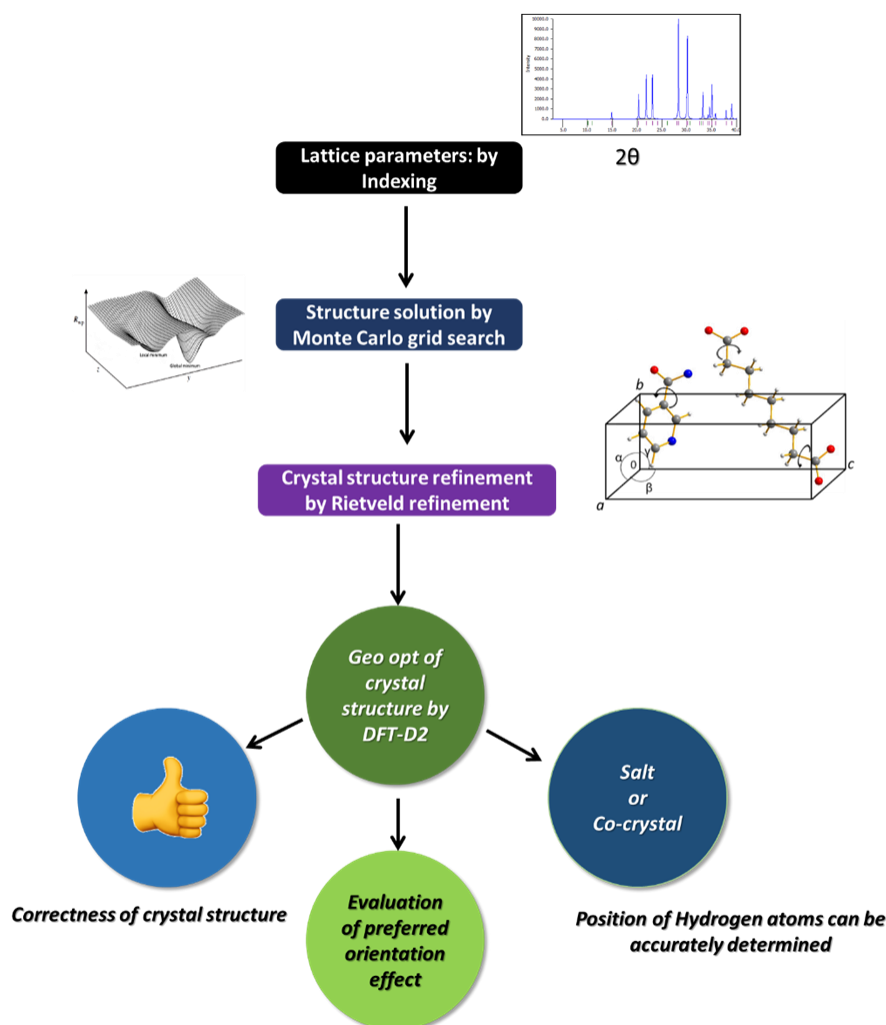


SC-XRD gives rise to thousands of individual unique reflections. All reflections are measured individually. In XRPD data, all symmetry equivalent reflections will have the same *d*-*spacings*, implying that the three-dimensional data from SXRD is compressed to one dimensional data. So, even for a powder-XRD data with well resolved peaks, only hundreds of reflections can be found and most of them are partially overlapped. Apart from these limitations, sample preparation and data collection procedures are crucial for this technique. In the present context, the next chapters in this thesis will focus on characterization and its challenges using powder-XRD studies.

### **1.5. Recent advances in structure determination from powder data (SDPD) and treatment of disorders in crystal structures: An extension towards DFT-D2 calculations**

A powder-XRD pattern is a plot between  $2\theta$  positions and peak intensities. The lattice parameters, crystal system and space group of the crystal structure can be determined from  $2\theta$  positions by indexing procedure. The atomic positions distribution in the unit cell can be determined from the peak intensities in the powder-XRD by using Monte Carlo direct space methods. In direct space methods, a landscape of trial crystal structures was generated by applying translational, rotational, and conformational movements to the molecular models in the unit cell. Monte Carlo grid search methodologies are applied to find the global minima. Upon reaching a reasonable approximate crystal structure in the structure solution step, it is further refined. The refinement of the approximate crystal structure obtained in the structure solution step is taken care by the Rietveld refinement technique. By applying bond lengths and bond angles constraints and restraints to the molecular models, the powder-XRD pattern is simulated and compared to the experimental powder-XRD pattern point by point. A flow chart on crystal structure determination using powder-XRD is presented in Fig. 1.1. After refinement of crystal structure, the following questions may be raised due to ambiguities on the structural arrangement and quality of the powder-XRD used:

- Position of hydrogen atoms
- Preferred orientation effect
- Any ambiguous discrepancies in the crystal structure

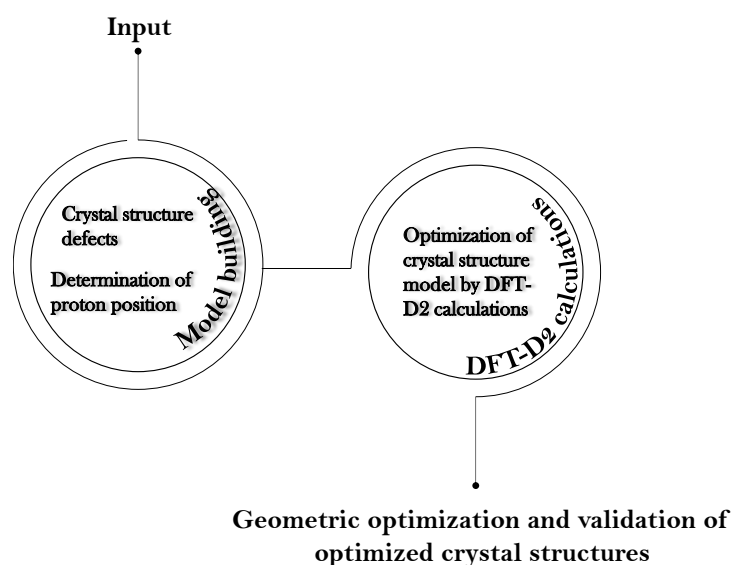


**Figure 1.1:** Flow chart of crystal structure determination from powder-XRD methodology.

Jacco van de Streek and Marcus A Neumann has published research papers on improving the quality of crystal structures determined from SXRD and XRPD data. He has demonstrated by performing geometric and energy optimization of crystal structures by using DFT-D2 calculations, the following challenges can be improved (**Figure 1.1**).

1. Correctness of crystal structure
2. Evaluation of preferred orientation effect
3. Position of Hydrogen atoms can be accurately determined.

Among the above three capabilities, 'determination of hydrogen atom positions' was effectively applied to improve the correctness of the crystal structure and to determine cocrystal or salt by accurately calculating the proton position between drug molecule and coformer.



**Figure 1.2:** Flow chart of methodology to address disorders in crystal structures using dispersion corrected DFT calculations.

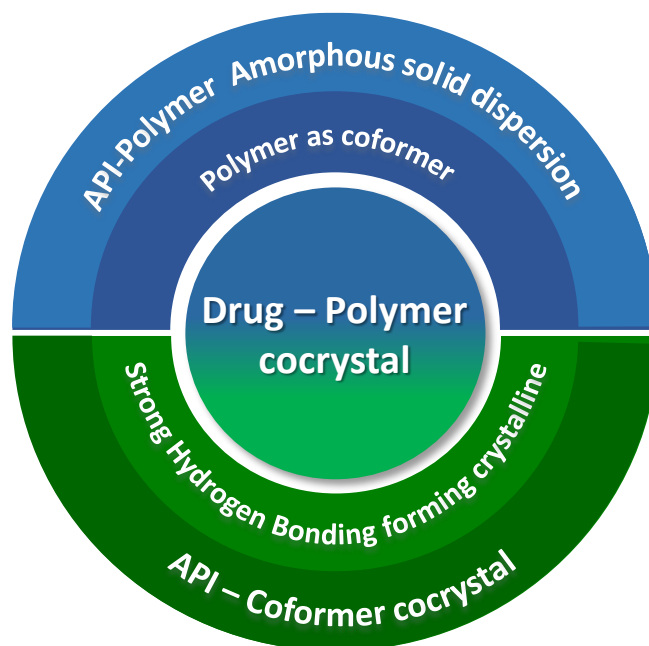
Similarly, dispersion corrected DFT calculations can also be adopted in the treatment of disordered crystal structures (**Figure 1.2**). Adequate literature has been reported on the application of similar methodologies to address a wide range of challenges while solving crystal structures. Some of them are listed below:

- Confirmation of Structural features or order-disorder challenges during single crystal growth or due to thermal phase transitions<sup>28</sup>.
- Order-Disorder phase transition triggered due to proton transfer between donor and acceptor adducts (ionic salt–neutral complexes)<sup>29–31</sup>.
- Conformational disorders due to temperature effects or phase transitions<sup>32</sup>.
- Accurate determination of H atoms positions in disordered structures<sup>30</sup>.

### 1.6. Drug–polymer cocrystals: A promising and emerging alternative

Traditionally, it is well known that other than solvates (hydrate), salts, and amorphous materials, cocrystals have emerged as a promising strategy where co-formers are small molecules that modulate physicochemical, mechanical, and pharmacokinetic properties. Under multi-component systems, amorphous solid dispersions come under another subclass where the coformer is a polymer excipient. Amorphous solid dispersion is also known as a frequently chosen multi-component system by generic industries to overcome either inferior dissolution properties or Intellectual property requirements. However, these complexes are known to suffer from inferior physical properties (stability). Now, API–polymer cocrystals are multi-

component systems like cocrystals, where coformer is a polymer (**Figure 1.3**). Urea-PEG and Thiourea-PEG complexes are the first API-polymer complexes reported in 1961. There are a few more API-polymer complexes (Resorcinol, p-Nitrophenol, 2-Methyl resorcinol, p-dihalobenzene, Hydroquinone, Mavacoxib, Griseofulvin, Diflunisal, and Nevirapine) reported which comes under either cocrystals or inclusion complexes.



**Figure 1.3:** Drug-polymer cocrystal: A subclass between amorphous solid dispersions and cocrystals.

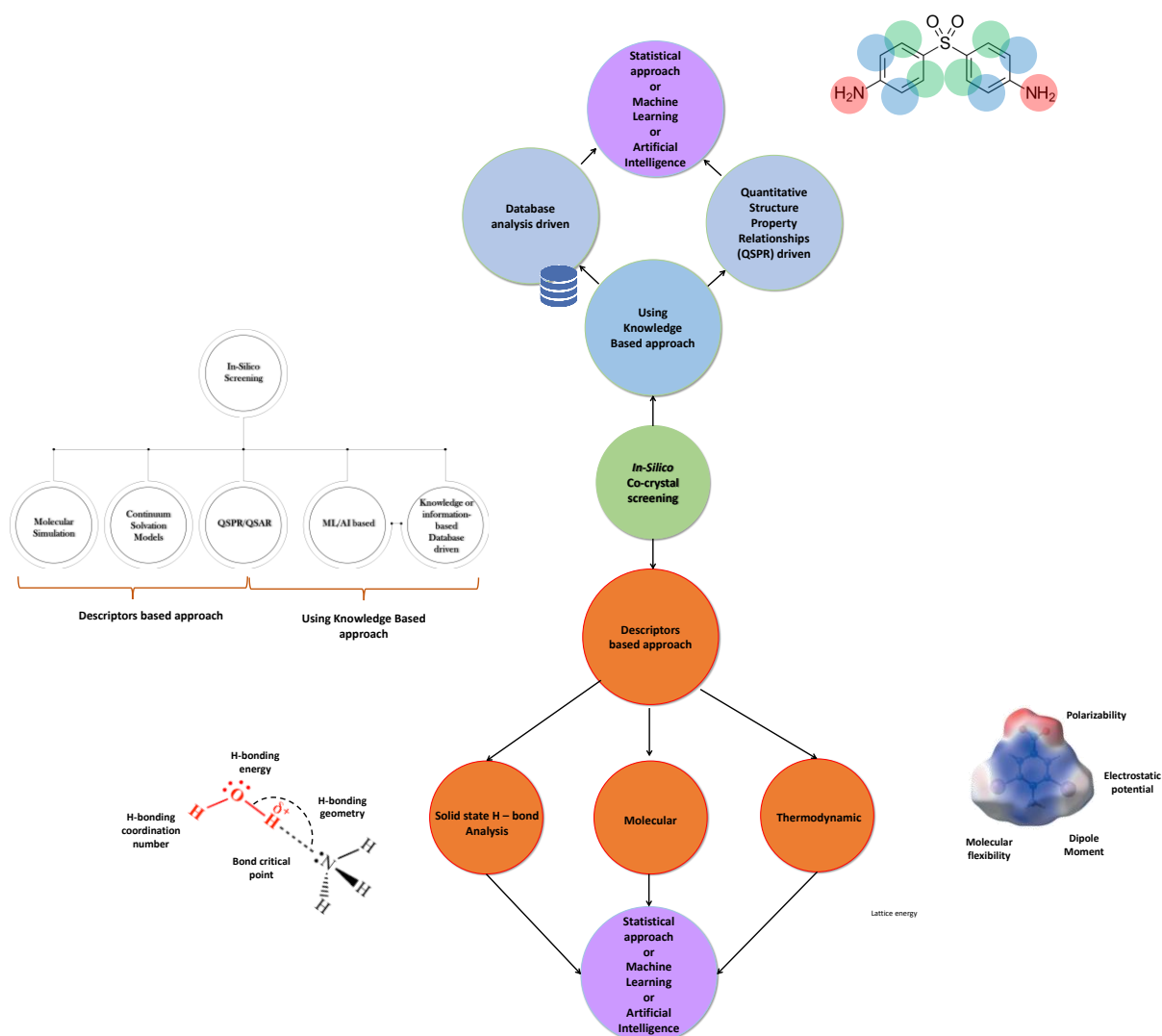
Although this can be a promising and futuristic strategy, the efforts required to find suitable polymer that can act as a co-former is experimentally challenging. Not only, due to its laborious, tedious, and time-consuming process but also the requirement to explore vast number of coformers and possibilities of crystallization methodologies. Such a huge number of variables demands rigorous and accurate balance to attain possible routes to the target of interest. This concern has grabbed the attention on the application of computational tools towards screening suitable co-formers for the selected API molecule of interest. Although this is a promising strategy, the main challenge that needs to be overcome for this class of complexes is a sensible design of in-silico methodologies and selection of suitable descriptors.

### 1.7. In-silico screening of Drug-polymer cocrystals

---

Cocrystallization is considered an economically feasible approach due to the availability of many molecules suitable as coformers, leading to the possibility of vast combinations of drug-coformers<sup>33</sup>. The exploration of drug-coformer landscape to identify the best possible combination that leads to the one with desired physicochemical properties is generally carried out by cocrystal screening experimental procedures. It is a known fact that the cocrystal screening for identifying suitable co-formers as an experimentally driven procedure is a time-consuming, expensive and tedious process<sup>33</sup>. To mitigate such issues, pharmaceutical companies are willing to adopt *priori* routes, wherein emerging computational capabilities come in. In-silico tools have been playing a major role in drug discovery (docking studies), innovating novel drug molecules in the pharmaceutical field. These capabilities are being extended by pushing boundaries towards polymorph screening based on fundamental principles postulated on the development of supramolecular concepts<sup>34</sup>. Building rational design of multi-component crystals based on supramolecular concepts are well introduced by Gautam Desiraju<sup>35</sup>. Researchers such as Aakeroy and Zawartoko have utilized these concepts to lay the foundation in developing knowledge-based cocrystal design strategies based on complementary and H-bond descriptors<sup>36</sup>. From past few years, significant research is going on extending these concepts to develop in-silico methodologies or computer aided methods for screening cocrystals. These concepts have inspired the development of various in-silico methodologies or computer aided methods for screening cocrystals (**Figure 1.4**).

1. Descriptors driven approach
2. Knowledge-based or database driven approach



3. **Figure 1.4:** Overview of in-silico screening of drug-polymer cocrystals.

### 1.7.1. Molecular Simulation

Molecular dynamics is a powerful tool to design molecules and mimic their behaviour at atomistic level to predict key properties<sup>37</sup>. These tools have been well explored for Amorphous solid dispersions (ASDs) extensively, where API and polymers are the key components of these systems. Similar concepts can also be applied for drug-polymer cocrystals, where drug-drug, drug-polymer and polymer-polymer interactions play key role in both ASDs and drug-polymer cocrystals, except that remarkable difference is observed in the order of the molecular arrangement, thereby differing impact on physical properties of the material<sup>38</sup>. The forces between the molecules of different molecules are referred as adhesive forces and whereas forces between different molecules are referred as cohesive forces. In other words, the free energy of mixing calculated from Flory-Huggin's interactions indicate strength of drug-polymer, drug-drug and polymer-polymer interactions. As per reported observations, it is well

reported in the literature that negative values of free energy indicate strong adhesive forces and positive values indicate strong cohesive interactions<sup>36,39–41</sup>. In general, the feasibility for the formation of ASDs is assessed by calculating compatibility or affinity between polymer and drug. Affinity between the drug and polymer molecules can be estimated theoretically by calculating the key parameters under structural and thermodynamical descriptors<sup>42</sup>.

These molecular simulations can help build complex models wherein such calculations can give privilege for calculating energy and interaction parameters<sup>37</sup>. Energy descriptors include free energy of mixing, binding energy, mixing energy and cohesive energy density<sup>38,43,44</sup>. Structural descriptors include non-bonding interactions, interaction parameters and hydrogen bond propensity<sup>45</sup>.

MD simulations are performed by using force field calculations, where forces are based on molecular mechanics and empirical forcefields which are vital to achieve the accuracy of the molecular modelling simulations. Forcefield files are summarized with interactions of bonded and non-bonds molecular interactions.

MD simulations are performed by using force field calculations, where forces are based on molecular mechanics and empirical forcefields which are vital to achieve the accuracy of the molecular modelling simulations. Forcefields are generally classified as three class (Class I, Class II and Class III). Class I forcefields include potential energies ( $U$ ) of bond, bond angle, torsional angles, LJ (Lennard–Jones) and electrostatic (ES) interactions. Examples of class I include AMBER, GAFF, OPLS and DRIEDING.

$$U_{Class\ I} = U_{bond} + U_{angle} + U_{torsion} + U_{LJ} + U_{ES}$$

In additions to the Class I, Class II forcefield include bond–bond and bond–angle coupling terms. Example for class II forcefield is PCFF.

$$U_{Class\ II} = U_{Class\ I} + U_{bond-bond} + U_{bond-angle}$$

In addition to the Class I and Class II, Class III forcefields include polarization effects. These calculations are very sensitive to the chosen forcefield. In the current work, COMPASS force field were used throughout the MD studies<sup>46</sup>.

COMPASS forcefield calculated potential energy of a system by considering the condensed state of a material as a result of  $E_{bond-bond\ interactions}$  and  $E_{non-bond\ interactions}$ .  $E_{bond-bond}$  interactions consists of stretch-stretch, stretch–bend, bend–bend, stretch–torsion, bend–torsion and bend–

bend–torsion terms. The non-bond interactions consist of electrostatic interactions ( $E_{\text{Cou}}$ ), Van der Waals interactions ( $E_{\text{vdW}}$ ) and Hydrogen bonding interactions.  $E_{\text{vdW}}$  is described by the Lennard-Jones 6-12 potential and the electrostatic energy is obtained from the partial charges of atoms in the system as estimated by the charge-equilibration method<sup>37</sup>.

$$E_{\text{pot}} = E_{\text{b}} + E_{\theta} + E_{\varphi} + E_{\chi} + E_{\text{cross}} + E_{\text{Cou}} + E_{\text{vdW}}$$

Where  $E_{\text{b}}$  = bond stretching energy,  $E_{\theta}$  = valence angle bending energy,  $E_{\varphi}$  = dihedral torsion energy,  $E_{\chi}$  = out-of-plane energy,  $E_{\text{cross}}$  = cross-term interaction energy,  $E_{\text{Cou}}$  = Coulombic interaction energy and  $E_{\text{vdW}}$  = van der Waals interaction energy.

These calculations are executed by either modelling a three-dimensional periodic cell (amorphous cell) or Flory-Huggins theory of Mixing. The application of amorphous cell in the molecular simulations involves series of steps such as (i) construction or packing, (ii) energy minimization, (iii) equilibrium phase and (iv) production phase. Amorphous cell is a periodic cell in which molecules are packed as per given composition, density ( $\text{g}/\text{cm}^3$ ) and molar volume ( $\text{cm}^3/\text{mol}$ ). Then, amorphous cells are minimized and refined by molecular dynamics calculations. In the Equilibrium phase, cell is relaxed to reach system stabilization or equilibrated density. In the production phase, equilibrated cell is subjected to molecular dynamics calculations<sup>42,47</sup>.

The combined application of Flory-Huggins model and molecular simulation techniques is another way of finding the compatibility of binary mixtures i.e., thermodynamics of mixing and phase separation. Flory-Huggins follows regular lattice arrangement of molecules. Whereas extended Flory Huggins models involve off-lattice calculation involves generation of clusters where configurations are generated by packing number of molecules of one type around another molecule<sup>48-50</sup>.

The combined Flory-Huggins and amorphous cell methodologies calculate the following parameters<sup>48-50</sup>.

2. Cohesive energy density
2. Solubility
2. Binding energy
2. Free energy of Mixing
2. Interaction energy



## Chapter 1: Introduction

---

Understanding on the solubility and cohesive energy between drug and polymer is important to calculate the physical properties of the materials. It can give insight on miscibility, non-bonding interactions and non-bonding interaction parameters.

Solubility ( $\delta$ ) is calculated by square root of Cohesive Energy Density (CED).

$$\delta = \sqrt{CED}$$

The solubility parameter can be calculated in two ways.

1. Hildebrand and Scott
2. Hansen

According to Hildebrand and Scott theory, Cohesive energy density (CED) refers to total attractive forces within a condensed state material. It is defined as increase in energy per unit volume.

$$CED = \frac{E_{coh}}{V}$$

Energy of Mixing can be related to the cohesive energy of pure components and corresponding blend as follows,

$$\Delta E_{mix} = \phi_A \left( \frac{E_{coh}}{V} \right)_A + \phi_B \left( \frac{E_{coh}}{V} \right)_B - \left( \frac{E_{coh}}{V} \right)_{mix}$$

$\phi_A$  and  $\phi_B$  indicates volume fractions of components in Blend.

The interaction parameter ( $\chi$ ) can be calculated from  $\Delta E_{mix}$  as follows.

$$\chi = \left[ \frac{\Delta E_{mix}}{RT\phi_A\phi_B} \right] V_m$$

According to Hansen theory, CED is defined as energy per unit volume required to vaporize in a condensed phase.

$$CED = \sqrt{\frac{E_{vaporization}}{V}} = \sqrt{\frac{\Delta H_{vap} - RT}{V}}$$

$$\Delta H_{mix} = V_t(\delta_A - \delta_B)^2\phi_A\phi_B$$

Hansen has distinguished solubility parameter ( $\delta$ ) into polar, hydrogen bonding and dispersive component.

## Chapter 1: Introduction

---

$$\delta^2 = \delta_h^2 + \delta_p^2 + \delta_d^2$$

Where  $\delta_h$ ,  $\delta_p$  and  $\delta_d$  share the contributions from the dispersive forces, polar forces, and hydrogen bonding, respectively.

$$\delta_h = \sqrt{\frac{\Sigma F_h}{V}}$$

$$\delta_p = \sqrt{\frac{\Sigma F_p^2}{V}}$$

$$\delta_d = \sqrt{\frac{\Sigma E_d}{V}}$$

$F_h$  is the molar attraction constant due to dispersive component;  $F_p$  is the molar attraction constant due to polar component;  $E_h$  is the hydrogen bonding energy and  $V$  is the molar volume. Other General expression for the free energy of Mixing of binary energy is

$$\frac{\Delta G}{RT} = \frac{\phi_b}{n_b} \ln \phi_b + \frac{\phi_s}{n_s} \ln \phi_s + \chi \phi_b \phi_s$$

$\Delta G$  is the free energy of mixing (per mole),  $\phi_i$  is the volume fraction of component,  $n_i$  is the degree of polymerization of component  $i$ ,  $\chi$  is the interaction parameter,  $T$  is the absolute temperature, and  $R$  is the gas constant.

Blends Module:

$$\Delta E_{mix} = \frac{1}{2} [Z_{bs}(E_{bs})_T + Z_{sb}(E_{sb})_T - Z_{bb}(E_{bb})_T - Z_{ss}(E_{ss})_T]$$

$Z$  = coordination number

$E$  = Binding energy

$S$  = screen molecule

$B$  = base molecule

$$(E_{bs})_T = \frac{\int dE_{ij} E_{ij} P(E_{ij}) e^{\frac{-E_{ij}}{RT}}}{\int dE_{ij} P(E_{ij}) e^{\frac{-E_{ij}}{RT}}}$$

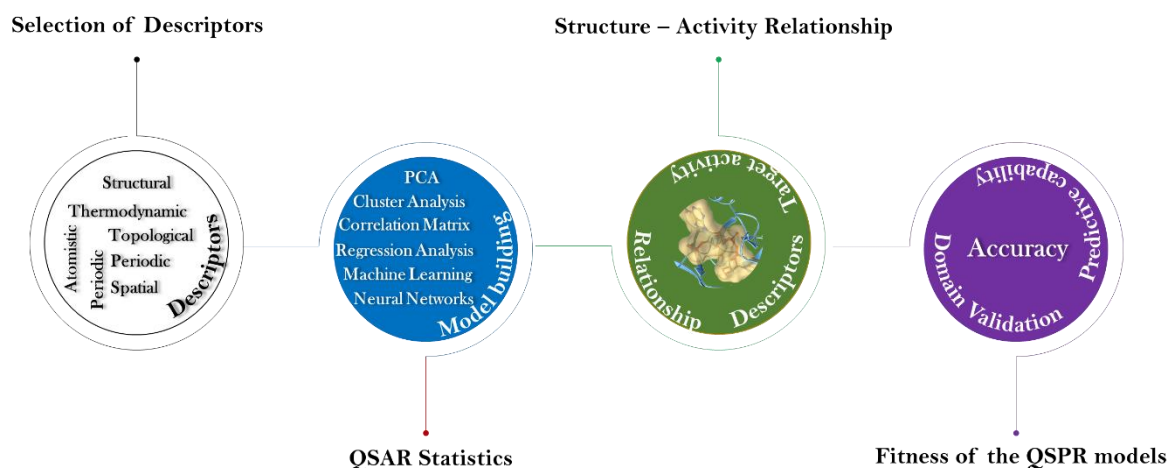
$i$  = one molecule,  $j$  =  $Z_{ij}$  molecules

The binding energy is a measure of the energy of interaction between two components. Together with the coordination numbers, it enables generation of the mixing energy and the chi parameter and of phase diagrams.

### 1.7.2. QSPR (Quantitative structure-property relations)

The foundation for relating structure-property relations was first laid by Brown and Fraser in 1868. It has laid a roadmap for the application in the drug design. In general, these methodologies employ Molecular modelling methods, and multivariate non-linear functions are employed to quantify the relationship. Descriptors play a major role in the establishment of accurate model. Topological, geometrical, quantum chemical, conformational, morphology and molecular surface properties are few known descriptors developed which are well explored in the building QSPR models.

To derive accurate correlations between structural descriptors and corresponding properties, selecting right representation descriptors is the most crucial step. It is essential to consider true representative molecular descriptors rather than computing a large number of descriptors, which could lead to overfitting. This is why QSPR is either recognized as one of the best methodologies or worst misleading calculations. Such situations would arise when models were built without proper validation or utilized beyond their applicability. The property cliff pertaining to the structure-property landscapes should be adequately rough i.e., the core data should contain both positive and negative results, covering all extremes adequately to produce unbiased results. Among many, this section is further discussed on the Molecular Structural, Thermodynamic, Molecular Electrostatic Potential Surface (MEPS), Conductor-Like Screening Model for Real solvents (COSMO-RS) and Crystal Structure Prediction (**Figure 1.5**).



**Figure 1.5:** Overview of in-silico screening of drug–polymer cocrystals using descriptors from QSPR calculations.

### 1.7.2.1. Descriptors based approach

A molecular descriptor is the mathematical representation of molecular chemical information in the form of a symbolic representation. To build the In-silico tools, it is necessary to establish the mathematical relationship between molecular structures and properties. Molecular descriptors play an important role in the establishment of QSPR to map structure–property relationship on the calculated descriptor results. Descriptors are mainly classified into experimental and theoretical descriptors. Only theoretical descriptors were considered in the current work to predict molecular properties prior to the experimental work. Among, thousands of descriptors known, selection of appropriate descriptors is very important. Since the current work aims to screen drug–polymer cocrystals, relevant selective descriptors come under the categories of structural, thermodynamic, topological, and crystal structure prediction were discussed.

#### 1.7.2.1.1. Molecular and Supramolecular descriptors

##### *Structural descriptors*

*Number of rotatable bonds:*

Among many structural descriptors, the number of rotatable bonds in the molecular structure is an important descriptor indicating the molecular flexibility of the molecule. Molecular flexibility is one of the many indicators for the property of polymorphism.

##### *Spatial descriptors*

*Molecular Volume:*

Molecular volume, dipole moment and polar surface are the three most important descriptors under spatial descriptors. The molecular van der Waals volume descriptor is a well-known descriptor that can relate the packing density of molecules in crystal structure with the native molecular volume. Volume can differ based on the molecular geometry of the molecule. This parameter is calculated by considering van der Waals spheres of the atoms having radius equal to the van der Waals radius of the atom.

*Dipole Moment:*

Dipole moment is calculated based on partial charges on the atoms in the molecule. Solvent accessible surface area of polar atoms in the molecule is calculated as polar surface area.

##### **Supramolecular descriptors**

*Molecular Shape:*

Molecular shape descriptors play a crucial role in the supramolecular assembly of molecules in the crystals<sup>51</sup>. Kitaigorodskii's postulates on crystal packing for optimistic utilization of space in the crystal structure and minimum energy. Many factors govern the crystal packing in materials; among them, molecular shape and volume are important. A molecular structure with an awkward shape may not have tendency to show efficient crystal packing. Flat and elongated shaped molecules will have tendency to form cocrystals with molecules exhibiting similar shapes. Globularity is another molecular shape descriptor which relates molecular surface to its molecular volume<sup>52</sup>. The globularity of a molecule with smooth surface will be smaller for molecular structures with smooth surfaces compared to the molecular surfaces with hollow surfaces and bumps. The extent of close packing in general is calculated as below<sup>53</sup>:

$$\text{Packing coefficient } (k) = Z \frac{V_0}{V}$$

Z = number of molecules in the unit cell; V = volume of unit cell; V<sub>0</sub> = molecular volume

*Hydrogen Bond Propensity (HBP):*

HBP is calculated to derive the probability of hydrogen bond formation between the multi-components. These methods come under the knowledge-based prediction category where crystal structures from the CCDC database are taken as input<sup>45</sup>. A statistical model is derived and incorporating the information from the test subject's propensity of hydrogen bond formation can be estimated from the equation below.

$$\Delta HBP = HBP_{API-Coformer} - HBP_{API-API} - HBP_{Coformer-Coformer}$$

Here, it is assumed that the formation of cocrystal is ruled by strongest donor-acceptor hydrogen bonds among all pairs<sup>54,55</sup>. If  $\Delta HBP > 0$ , then the probability for the formation of cocrystal is high; if  $\Delta HBP < 0$ , then the probability of cocrystal formation is low<sup>45</sup>.

### 1.7.2.1.2. Thermodynamic descriptors approach

Experimentally, the process of crystallization comprises of both thermodynamic and kinetic aspects<sup>56</sup>. If a cocrystal formation between two components is thermodynamically driven rather than structurally driven, then they are likely to form cocrystals. Also, thermodynamic parameters are directly proportional to the strength and directionality of hydrogen bonding network of the crystalline materials. Thermodynamic concepts are taken as basis for the screening cocrystals. In general, the Gibbs free energy for cocrystals is always less than corresponding individual components<sup>33</sup>. However, due to huge computational requirements, computational methods can accommodate thermodynamic calculations over kinetic calculations. These kinetic aspects which are structurally driven by intermolecular interactions,

---

cannot be described by simple descriptors. These intermolecular interactions include formation of synthons and motifs between the molecules in the crystal structures.

Hence, most of the computational methodologies relevant to in-silico screening of cocrystals are mainly based on thermodynamic descriptors. These thermodynamic descriptors include parameters such as free energy, enthalpy etc. There are several such methodologies for in-silico screening. Among them, few popular methodologies are listed below:

- (i) Thermodynamic characteristics (free energy, enthalpy, electrochemical potentials etc.) of pure compounds vs cocrystals (CC).
- (ii) The supramolecular synthons were also considered as a driving force for cocrystallization; they play an important role while planning the cocrystallization strategy.

### 1.7.2.1.3. Molecular Electrostatic Potential Surface (MEPS)

MEPS is a potential tool which is widely used to range the hydrogen bond donor and acceptor sites in the molecules representing relative electrostatic surface of the molecules. At any point on the surface of a molecule, MEP is a gas phase calculation<sup>57</sup>, calculated as the ratio of net electrostatic effect to the total charge distribution of the molecule<sup>58</sup>. Charge distribution on the molecular surface is generally indicated in the range of red to blue colour scale. Red indicates region of negative charge and whereas blue indicates region of positive charge<sup>58</sup>. These discrete points on the molecular surface are called as surface site interaction points (SSIPs). SSIPs will help to define hydrogen bond donor and acceptor sites. These sites even represent weak polar sites of the molecule which can eventually forms weak electrostatic interactions. These properties will mimic the molecular surface of the molecules with its environment<sup>57</sup>. Thereby, MEPS can be related to the dipole moment, electronegativity, partial charges, and chemical reactivity of the molecule<sup>58</sup>. MEPS are generally calculated through DFT-D2 geometric optimization calculations which are generally performed by CASTEP, GAUSSIAN 09 or DMol3 etc algorithms by using B3LYP/6-31G basis set<sup>59,60</sup>. HOMO–LUMO analysis gives energy gap between maximum positive value and maximum negative value of nucleophile and electrophile sites respectively. Its magnitude indicates reactivity of the molecule<sup>60</sup>.

MEPS is a potential tool which is widely used to range the hydrogen bond donor and acceptor sites in the molecules representing relative electrostatic surface of the molecules. To evaluate the formation for the scope of cocrystal formation between API and cofomer, MEPs gives an accurate estimation over hydrogen bond propensity (HBP)<sup>33</sup>. HBP calculations are based on best donor–acceptor pair between the molecules, Whereas MEPS considers all possible

intermolecular interactions between the molecules. The interaction energy between the molecules can be calculated by ranking the pairs of interaction sites as per strength of interactions<sup>61</sup>.

$$E = - \sum_i \alpha_i * \beta_i$$

Here  $\alpha$  represents hydrogen bond donor sites and  $\beta$  represents hydrogen bond acceptor sites.

The energy gain of cocrystal formation is calculated as

$$\Delta E = E_{Cocrystal} - (E_{API} + E_{Coformer})$$

### 1.7.3. Conductor-Like Screening Model for Real solvents (COSMO-RS)

COSMO is a well-known approach extensively used for predicting thermodynamic properties and constructing phase diagrams. COSMO was originally developed by Andreas Klamt based theory of continuum models. The molecules (solute or solvent) in the liquid are always influenced by their neighbour surrounding molecules. In molecular modelling, defining discrete microscopic description of solvent molecules is computationally expensive. Combining continuum solvation models and quantum chemical methods is an effective methodology to define solvent effects in QM approach. In other words, polarization in the dielectric solute molecules due to neighbour solvent molecules can be translated into QM approach.

Polarizable continuum (PCM) model was developed by Tomasi and co-workers to indicate the ASC (apparent surface charge). The polarization of the solvent is induced by the solute and solute itself is polarized back by the solvent, an iterative self-consistent field approach is the straightforward solution to the problem. Continuum PCM methods contemporaneously account for long-range electrostatic interactions and polarization effects. These aspects are not easily available in discrete methods. Continuum models automatically give configurationally samples solvent effect.

To calculate the accurate calculation of such molecular properties, efficient sampling of all possible arrangements of solute and solvent molecules and their mean thermodynamics is required. Later, it was further developed to COSMO-RS. It was built by integrating dielectric continuum model with statistical thermodynamics. COSMO-RS treats all molecules (solute and solvent) in a system by imparting quantum chemical and statistical thermodynamics equivalently rather than solely dielectric field. This combination will help to treat interacting surfaces. This is useful for fast screening of cocrystals and many more applications. It gives miscibility of cocrystal components in supercooled melt phase mimic cocrystal solid state.

$$H_{Excess\ Enthalpy} = H_{AB} - x_m H_{pure,A} - x_n H_{pure,B}$$

$H_A$  and  $H_B$  are enthalpies of pure components A and B.  $x$  is mole fraction.  $H_{AB}$  = enthalpy stoichiometry mixture of component A and component B.

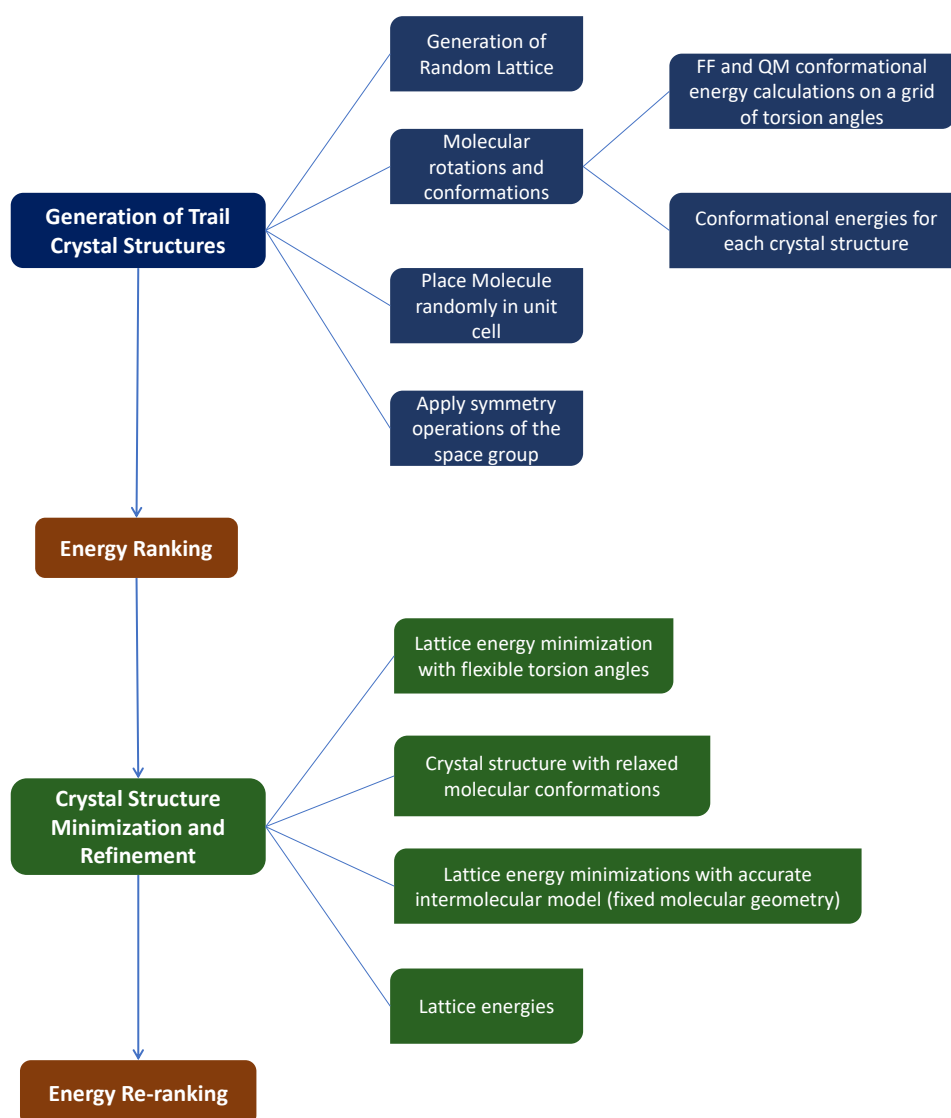
In general, excess free energy is considered as an indicator for theoretical screening of cocrystals compared to enthalpy. As small modifications in the molecular systems can show higher impact on enthalpy values and their impact would be low on free energy values.

### 1.7.4. Crystal Structure Prediction

From many years, crystal structure prediction has been one of the most fascinating dreams of crystal engineers and computational researchers to predict polymorphs with reliable accuracy and desired physico-chemical properties. This is evident from the substantially increasing efforts and progress over years on extensive exploration of new CSP methodologies are increasing on polymorph prediction<sup>62-65</sup>. The crystal structure prediction methodologies can be fundamentally categorized into three steps<sup>36</sup> (**Figure 1.6**).

1. Generation of trail crystal structures by exploration of preferential molecular conformational landscape and crystal packing arrangements of target molecules
2. Minimization and optimization of trail crystal structures
3. Ranking of resulting crystal structures forming using some form of scoring and fitness solution.





**Figure 1.6:** Flow chart of CSP protocol<sup>66–69</sup>.

Trail crystal structures will be generated using various sampling methods and conformational and crystal packing parameters. Among reported CSP protocols so far, most successful protocol was developed by Neumann et al as sixth blind test, where space sampling step for generation of trail crystal structures is followed by lattice energy minimization. It incorporates forcefields for sampling step and DFT–D2 calculations for the minimization step. There are few developments on utilization of neural network also reported<sup>66</sup>. The magnitude of the structural landscape of a molecule depends on the molecular flexibility. With increase in number of rotational bonds and torsional bonds would increase the flexibility of the molecule. By following sampling procedures based on molecular conformations, intermolecular orientations, space groups, unit cell volumes and intermolecular interactions, vast number of crystal structures are generated. These crystal structures are ranked based on calculated free energies<sup>70</sup>.

## Chapter 1: Introduction

---

Free energy ( $\Delta G_{sub}$ ) can be calculated by either computing from (i) Thermodynamic ( $\Delta G_{sub} = \Delta H_{sub} - T\Delta S_{sub}$ ) and another from (ii) molecular simulation<sup>71</sup>.

The input  $\Delta H_{sub}$  and  $\Delta S_{sub}$  for the Gibbs equation can be calculated as follows

Enthalpy of sublimation can be calculated from the following equation.

$$\Delta H_{sub}^o = -U_{lattice}^{\square} - 2RT$$
$$U_{lattice}^{\square} = U_{intermolecular}^{\square} + \Delta E_{intramolecular}^{\square}$$

The sublimation entropy can be calculated as difference between the entropy of an ideal gas and the entropy of the crystal at a given temperature and pressure. The entropy of the crystal is the sum of intra and intermolecular contributions. By considering the intra-molecular contributions towards entropy as zero, the entropy of the rigid molecules can be calculated by the following equation.

$$\Delta S_{sub}^o = S_{rot,gas}^{\square} + S_{trans,gas}^{\square} - S_{ext,cryst}^{\square}$$

$S_{trans,gas}$  and  $S_{rot,gas}$  are the transitional contributions towards entropy of gas at temperature T which can be calculated from statistical thermodynamics.  $S_{ext,cryst}$  is the intermolecular vibrational contribution towards entropy of the crystal at T which can be computed by rigid molecule lattice dynamics.

In general, for organic molecules sublimation free energy is reported in the gas phase ( $\Delta G_{sub}^o$ ) at 1 atm pressure. Hydration free energies ( $\Delta G_{hyd}^*$ ) from  $\Delta G_{sub}^o$  can be calculated from the below equation.

$$\Delta G_{hyd}^* = \Delta G_{sub}^o - RT \ln \left( \frac{V_m P_o}{RT} \right)$$

$V_m$  is the molar volume of the crystal, and  $P_o$  is the standard atmospheric pressure (1 atm=101.325kPa).

Free energy and lattice energy calculated from the above procedure can also be used to screen conformers for the most stable cocrystal<sup>72</sup>. The free energy difference between most stable cocrystal and neat crystal structures of API and conformer can be calculated as below.

$$\Delta \Delta G_c = G_{[API_n c_m]} - (nG_{[API]} + mG_{[c]})$$

Where  $\Delta \Delta G_c$ =lattice energy difference,  $G_{[API_n c_m]}$ = lattice energy of most stable cocrystal,  $nG_{[API]}$  and  $mG_{[c]}$  are lattice energies of API and cocrystal neat crystal structures; API = Active

Pharmaceutical Ingredient,  $c$  = conformer,  $m$  and  $n$  are stoichiometry between API and ingredient.

If conformer is liquid leading to formation of solvate, then  $mG_{[c]}$  should be replaced with conformer's enthalpy of fusion ( $\Delta G_{[c]fus}$ ).

$$\Delta\Delta G_c = G_{[API_n c_m]} - (nG_{[API]} + m(G_{[c]} + \Delta G_{[c]fus}))$$

Here, enthalpy of fusion at  $T_{c,fus}$  can be calculated as below

$$\Delta G_{[c]fus} = \Delta H_{[c]fus} \left( 1 - \frac{T}{T_{c,fus}} \right)$$

Where  $H_{[c]fus}$  = conformer's enthalpy of fusion.

Similarly, the lattice energy difference between most stable cocrystal and neat crystal structures of API and conformer can be calculated as below.

$$\Delta\Delta U_c = U_{[API_n c_m]} - (nU_{[API]} + mU_{[c]})$$

Where  $\Delta\Delta U_c$ =lattice energy difference,  $U_{[API_n c_m]}$ = lattice energy of most stable cocrystal,  $nU_{[API]}$  and  $mU_{[c]}$  are lattice energies of API and cocrystal neat crystal structures.

If conformer is liquid leading to formation of solvate, then  $mU_{[c]}$  should be replaced with conformer's enthalpy of fusion ( $\Delta G_{[c]fus}$ ).

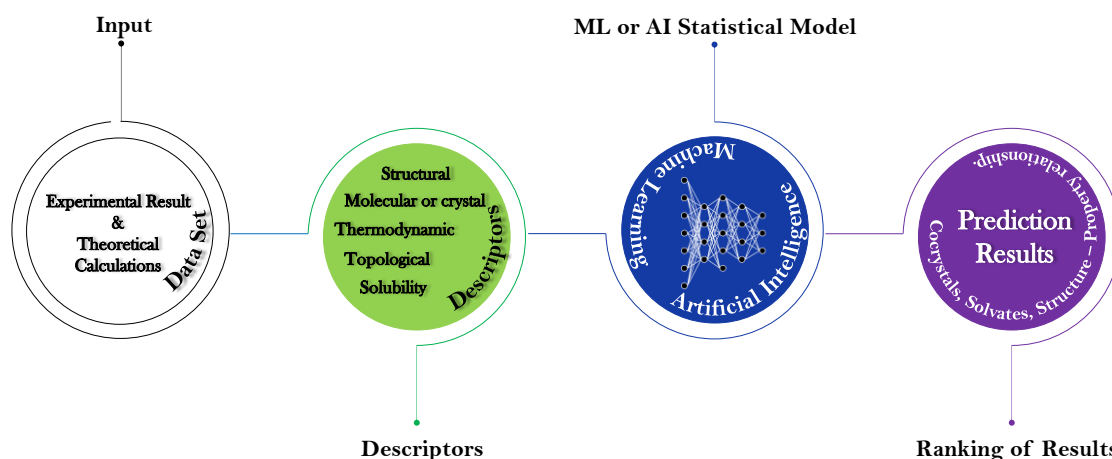
$$\Delta\Delta U_c = U_{[API_n c_m],min} - (nU_{[API]} + mU_{[c]}) - m\Delta H_{c,fus} \left( 1 - \frac{T}{T_{c,fus}} \right)$$

### 1.7.5. Knowledge and Database driven

Screening of pharmaceutical cocrystals is experimentally very time-consuming and expensive<sup>73</sup>. In general, the experimental strategies followed for the identification of suitable conformers for an API cocrystallization are designed based on concepts of supramolecular synthons by Gautam Desiraju<sup>45</sup> and  $\Delta pK_a$  rule<sup>74</sup>. Although the possible intermolecular interactions between the molecules rationalize the possibility of nucleation, practically such attempts may not hold true all the time<sup>73</sup>. Alternatively, CSP (crystal structure prediction) methods are increasingly gaining attention for organic molecules, where structure–energy landscapes are combined with property predictions for cocrystal screening by means of molecular descriptors<sup>75</sup>. The selection of descriptors is a critical step and must always be based

on unbiased, sensible and universal<sup>76</sup>. In these predictive tools, molecular descriptors such as Hydrogen bond propensity<sup>77</sup>, polarity, molecular shape, lattice energy, MEPS, Hansen solubility, COSMO-RS and statistical analysis of reported cocrystals in the CSD database<sup>78</sup> are generally used for the design and discovery of cocrystals. But, these CSP tools may not be fully efficient due to drawbacks such as (i) Contribution of irrelevant crystal structure in the data sets, (ii) Datasets containing large number of crystal structures differing by small lattice energy and (iii) Lack of transferability to design crystallization experiments to obtain targeted products<sup>78</sup>. As per literature, despite the attempts with rationale thought process based on in-silico methods, most of the attempts to find a suitable conformer for an API end with either unsuccessful or eutectic mixtures (prediction of cocrystal formation between drug and conformer by simple structural parameters). To deduct more accurate output from in-silico methods, there is requirement for more sophisticated and, rigorous, accurate, and robust thought process together.

The virtual screening tools utilized for the cocrystal screening are generally divided into three categories: (i) Knowledge-based, (ii) Physics-based and (iii) Machine learning or AI approaches. The above discussion, the CSP from databases such as CCDC and molecular descriptors comes under knowledge-based category. Physics-based methods involve COSMO and MEPS calculations; the third category belongs to ML and AI applications<sup>54</sup>. The new requirements have demanded a new direction in machine learning applications, and it is increasingly grabbing attention in crystal structure prediction studies. Including CSP parameters (molecular descriptors, structural descriptors, energy descriptors, etc.) with ML framework concepts makes it possible to establish a structure-property relationship for the predicted crystal structures<sup>75</sup>. Researchers in this area have made significant progress in the application of ML and AI successfully that can help to design cocrystals with specific required properties<sup>73,79</sup>. Since descriptors are the key input components for CSP by ML calculations, it is very important to have prior knowledge on selected descriptors, crystallization process, molecular structure and crystal packing<sup>76</sup>. The calculations can be computationally expensive, if they demand large datasets and quantum mechanical calculations (DFT)<sup>76</sup>. In general, the overview of protocol involving machine learning models will start with a dataset as input. In general, the input dataset is either experimental and simulation results. This dataset is fed to various molecular descriptors and the obtained results are ranked in a master sheet and a statistical model is derived from it. By feeding test subjects to the statistical models to obtain to predict the target class (**Figure 1.7**).

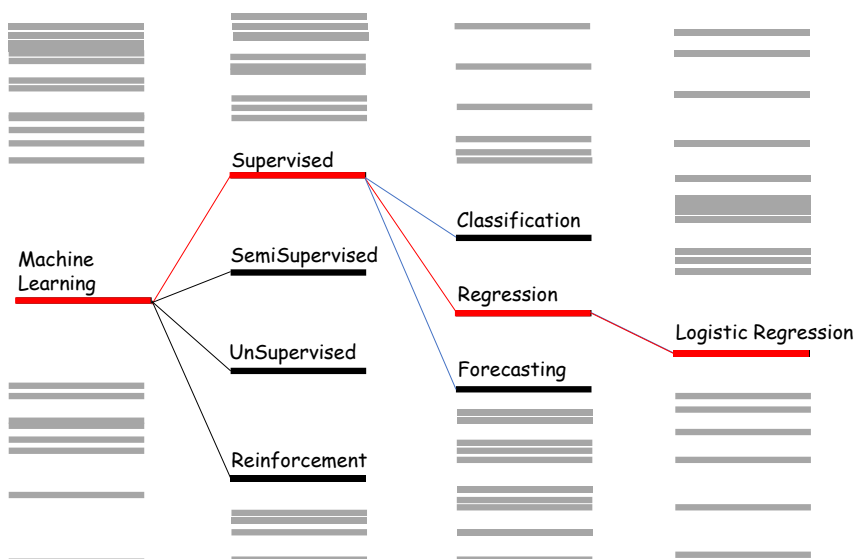


**Figure 1.7:** Overview of in-silico screening of drug–polymer cocrystals using descriptors from ML based calculations.

ML is classified as four categories (i) Supervised learning, (ii) Semi Supervised learning, (iii) Unsupervised learning and (iv) Reinforcement learning. In this thesis, only specific topics under ML were discussed confined as per scope of application in the present work. Supervised learning utilizes labelled dataset having both input and output parameters and unsupervised learning can find patterns in the unlabelled dataset. Supervised learning is categorized into Classification, Regression and Forecasting. Among them, Regression is statistical analysis applied to find correlation between dependent variables and resultant multiple variables. Logistic regression provides probability of an event based on input data i.e., the outcome of the analysis will be either 0 or 1. The logistic regression can be written in the form of equation as below<sup>74</sup> (**Figure 1.8**).

$$p(x) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_n x_n)}}$$

Where  $\beta$  is intercept,  $x$  is dependent or response variable.



**Figure 1.8:** Overview on ML based in-silico screening method adopted for drug–polymer cocrystals.

### 1.8. Relevance of this work in Brief

Recent increase in research papers indicates the growing interests of academics and industries in the area of pharmaceuticals to incorporate in their product development procedure with theoretical studies such that development time and resources can be reduced, thereby minimizing product development costs. Especially significant work is going with similar objectives in the area of pharmaceuticals cocrystals. Literature review in this prospective has indicated that significant efforts were being kept where experimental and computational studies go hand in and hand to develop effective in-silico studies. The current studies focus on structural characterisation and elucidation, addressing disorders in cocrystals and in-silico screening studies. It includes employing theoretical studies for the determination of proton migration between cocrystal and salt for small molecule cocrystals. Refinement of crystal structures by DFT-D2 calculations. Addressing challenges in disorders in crystal structures in drug-polymer cocrystals. Virtual screening of drug-polymer cocrystals. Structural studies to establish structure-property relationship.

**Chapter 2** focuses on experimental, characterisation, and theoretical methodologies employed in this work

**Chapter 3** concentrates on the solid-state phase transition of a multi-component active pharmaceutical ingredient (M-APIs), composed of (1:1) agomelatine and phosphoric acid (AGL-P), is characterized using a set of complementary techniques: in situ variable

temperature X-ray powder diffraction (VT-XRPD), thermal analysis, spectroscopic techniques, and hot stage microscopy (HSM). It is observed that these dimorphic forms (RT-form and HT-form) are enantiotropic and reversible in nature. The salt-cocrystal continuum of this system is demonstrated using ab initio powder-XRD structure determination (SDPD) and dispersion-corrected density functional theory (DFT-D2) analysis. Furthermore, this solid-state phase transition can be inferred as a martensitic-like transformation, where simultaneous proton migration and small conformational switching trigger the concerted molecular displacements of entire layers, leading to microscopic crystal contraction.

**Chapter 4** covers characterization of dapsone-polyethylene glycol by combined experimental and computational studies followed by development of virtual screening of drug-polymer cocrystals. In the initial stage, crystal structure was solved from powder-XRD data and later it was solved from Single crystal-XRD. Upon solving crystal structure, the crystal structure was subjected to geometric optimization by DFT-D2 calculations. Being motivated to discover more such drug-polymer complexes with PEG as coformer, this study was also extended further to develop potential computational methodologies together aimed for virtual screening of drug-PEG (drug-polyethylene glycol) cocrystal complexes. These methodologies are incorporated with molecular descriptors, thermodynamic descriptors and supramolecular descriptors combined with ML/AI studies. Experimental outcome has confirmed that the success rate of the protocol is about 90% discovering two drug-polymer cocrystals. Among possible outcomes from screening of suitable APIs to form drug-polymer cocrystals with PEG, sulfanilamide – PEG (SUL-PEG) was considered as an extended for further characterization in **Chapter 5**.

**Chapter 5** is an extended work of outcomes from Virtual Screening of drug-polymer cocrystals from **Chapter 4**. SUL-PEG cocrystal was subjected through thorough characterization and comparison of its structural and physicochemical properties using techniques such as X-ray diffraction, differential scanning calorimetry, Fourier-transform infrared spectroscopy, dynamic vapour sorption, and dissolution testing on five isostructural polymeric cocrystals with different grades of PEG. The crystal structure of SUL-PEG 600 is determined using single crystal X-ray diffraction. Additionally, solid-state nuclear magnetic spectroscopy and dispersion corrected density functional theory analysis are used to shed light on drug-polymer interactions and chiral amplification via PEG helical conformation that results in novel order-disorder phenomena and symmetry breaking at crystal's short-range order (multi-domain single crystal). The gained structural insight helps to understand the impact of

---

the synthetic procedure on physical properties such as crystallinity and thermal stability. The study further indicates that polymeric cocrystals can be a promising option for addressing pharmaceutical material properties and can provide a deeper molecular-level understanding of drug-polymer interactions that can be leveraged to design innovative and effective drug delivery systems.

### 1.9. References

1. Shan, N. & Zaworotko, M. J. The role of cocrystals in pharmaceutical science. *Drug Discov. Today* **13**, 440–446 (2008).
2. Bolla, G. & Nangia, A. Pharmaceutical cocrystals: walking the talk. *Chem. Commun.* **52**, 8342–8360 (2016).
3. Karimi-Jafari, M., Padrela, L., Walker, G. M., & Croker, D. M. Creating cocrystals: A review of pharmaceutical cocrystal preparation routes and applications. *Cryst. Growth Des.* **18**, 6370–6387 (2018).
4. Nangia, A. K. & Desiraju, G. R. Heterosynthons, Solid Form Design and Enhanced Drug Bioavailability. *Angew. Chemie Int. Ed.* **61**, e202207484 (2022).
5. Nangia, A. Conformational polymorphism in organic crystals. *Acc. Chem. Res.* **41**, 595–604 (2008).
6. Bolla, G., Sarma, B., & Nangia, A. K. Crystal Engineering of Pharmaceutical Cocrystals in the Discovery and Development of Improved Drugs. *Chem. Rev.* **122**, 11514–11603 (2022).
7. Nangia, A. K. & Desiraju, G. R. Crystal Engineering: An Outlook for the Future. *Angew. Chemie Int. Ed.* **58**, 4100–4107 (2019).
8. Nangia, A. & Desiraju, G. R. Supramolecular Synthons and Pattern Recognition. in *Design of Organic Solids* vol. 198 57–95 (1998).
9. Vishweshwar, P., McMahon, J. A., Bis, J. A., & Zaworotko, M. J. Pharmaceutical Co-Crystals. *J. Pharm. Sci.* **95**, 499–516 (2006).
10. Cappuccino, C., Cusack, D., Flanagan, J., Harrison, C., Holohan, C., Lestari, M.,



- Walsh, G., & Lusi, M. How Many Cocrystals Are We Missing? Assessing Two Crystal Engineering Approaches to Pharmaceutical Cocrystal Screening. *Cryst. Growth Des.* **22**, 1390–1397 (2022).
11. Xia, M. Y., Zhu, B. Q., Wang, J.-R., Yang, Z. E., & Mei, X. F. Superior Dissolution Behavior and Bioavailability of Pharmaceutical Cocrystals and Recent Regulatory Issues. *ACS Med. Chem. Lett.* **13**, 29–37 (2022).
  12. Kumar Bandaru, R., Rout, S. R., Kenguva, G., Gorain, B., Alhakamy, N. A., Kesharwani, P., & Dandela, R. Recent Advances in Pharmaceutical Cocrystals: From Bench to Market. *Front. Pharmacol.* **12**, (2021).
  13. Guo, M., Sun, X., Chen, J., & Cai, T. Pharmaceutical cocrystals: A review of preparations, physicochemical properties and applications. *Acta Pharm. Sin. B* **11**, 2537–2564 (2021).
  14. Lee, E. H. A practical guide to pharmaceutical polymorph screening & selection. *Asian J. Pharm. Sci.* **9**, 163–175 (2014).
  15. Yousef, M. A. E. & Vangala, V. R. Pharmaceutical cocrystals: molecules, crystals, formulations, medicines. *Cryst. Growth Des.* **19**, 7420–7438 (2019).
  16. Food and Drug Administration. Guidance for industry regulatory classification of pharmaceutical co-crystals. *Food Drug Adm.* 1–5 (2013).
  17. European Medicines Agency. Reflection paper on the use of cocrystals and other solid state forms of active substances in medicinal products. *Eur. Med. Agency* **44**, 1–10 (2015).
  18. Food and Drug Administration. Regulatory Classification of Pharmaceutical Co-Crystals Guidance for Industry. *U.S. Dep. Heal. Hum. Serv. Food Drug Adm.* 1–5 (2016).
  19. Food and Drug Administration. Regulatory classification of pharmaceutical co-crystals, guidance for industry. *U.S. Dep. Heal. Hum. Serv.* 1–4 (2018).
  20. Cruz-Cabeza, A. J. Acid–base crystalline complexes and the pKa rule. *CrystEngComm* **14**, 6362 (2012).
  21. Cruz-Cabeza, A. J., Lusi, M., Wheatcroft, H. P., & Bond, A. D. The role of solvation

- in proton transfer reactions: implications for predicting salt/co-crystal formation using the  $\Delta p K_a$  rule. *Faraday Discuss.* **235**, 446–466 (2022).
22. Rajput, L., Banik, M., Yarava, J. R., Joseph, S., Pandey, M. K., Nishiyama, Y., & Desiraju, G. R. Exploring the salt-cocystal continuum with solid-state NMR using natural-abundance samples: Implications for crystal engineering. *IUCrJ* **4**, 466–475 (2017).
  23. Ito, S., White, F. J., Okunishi, E., Aoyama, Y., Yamano, A., Sato, H., Ferrara, J. D., Jasnowski, M., & Meyer, M. Structure determination of small molecule compounds by an electron diffractometer for 3D ED/MicroED. *CrystEngComm* **23**, 8622–8630 (2021).
  24. Saha, A., Nia, S. S., & Rodríguez, J. A. Electron Diffraction of 3D Molecular Crystals. *Chem. Rev.* **122**, 13883–13914 (2022).
  25. Miller, J. L. Electron diffraction sees hydrogen atoms. *Phys. Today* **70**, 16–19 (2017).
  26. McCusker, L. B. Electron diffraction and the hydrogen atom. *Science (80-. )*. **355**, 136–136 (2017).
  27. Hovmöller, S., Zou, X., & Weirich, T. E. Crystal structure determination from EM images and electron diffraction patterns. in *Advances in Imaging and Electron Physics* vol. 123 257–289 (2002).
  28. Van De Streek, J. & Neumann, M. A. Validation of molecular crystal structures from powder diffraction data with dispersion-corrected density functional theory (DFT-D). *Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater.* **70**, 1020–1032 (2014).
  29. Order-disorder phase transition induced by proton transfer in a co-crystal of 2,4-dichlorobenzoic acid and trimethylamine: N -oxide. *CrystEngComm* **19**, 3753–3759 (2017).
  30. Bhardwaj, R. M., Johnston, B. F., Oswald, I. D. H., & Florence, A. J. A complementary experimental and computational study of loxapine succinate and its monohydrate. *Acta Crystallogr. Sect. C Cryst. Struct. Commun.* **69**, 1273–1278 (2013).
  31. Rybarczyk-Pirek, A. J. Co-crystal/salt crystal structure disorder of trichloroacetic acid-N-methylurea complex with double system of homoand heteronuclear O-H O/N-H O hydrogen bonds: X-ray investigation, ab initio and DFT studies. *Struct. Chem.* **23**,
-

- 1739–1749 (2012).
32. Bernert, T., Ley, M. B., Ruiz-Fuertes, J., Fischer, M., Felderhoff, M., & Weidenthaler, C. Molecular structure of diethylaminoalane in the solid state: An X-ray powder diffraction, DFT calculation and Raman spectroscopy study. *Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater.* **72**, 232–240 (2016).
  33. Kumar, A. & Nanda, A. In-silico methods of cocrystal screening: A review on tools for rational design of pharmaceutical cocrystals. *J. Drug Deliv. Sci. Technol.* **63**, (2021).
  34. Etter, M. C., MacDonald, J. C., & Bernstein, J. Graph-set analysis of hydrogen-bond patterns in organic crystals. *Acta Crystallogr. Sect. B* **46**, 256–262 (1990).
  35. Desiraju, G. R. Supramolecular Synthons in Crystal Engineering—A New Organic Synthesis. *Angew. Chemie Int. Ed. English* **34**, 2311–2327 (1995).
  36. Abramov, Y. A. Computational Pharmaceutical Solid-State Chemistry. *Comput. Pharm. Solid State Chem.* 1–13 (2016).
  37. de Arenaza, I. M., Meaurio, E., & Sarasu, J.-R. Analysis of the Miscibility of Polymer Blends Through Molecular Dynamics Simulations. in *Polymerization* 29–46 (InTech, 2012).
  38. Walden, D. M., Bunday, Y., Jagarapu, A., Antontsev, V., Chakravarty, K., & Varshney, J. Molecular Simulation and Statistical Learning Methods toward Predicting Drug–Polymer Amorphous Solid Dispersion Miscibility, Stability, and Formulation Design. *Molecules* **26**, 182 (2021).
  39. Medarević, D., Djuriš, J., Barmapalexis, P., Kachrimanis, K., & Ibrić, S. Analytical and computational methods for the estimation of drug-polymer solubility and miscibility in solid dispersions development. *Pharmaceutics* vol. 11 at <https://doi.org/10.3390/pharmaceutics11080372> (2019).
  40. Chunsriviro, S., Santiso, E., & Trout, B. L. Binding affinity of a small molecule to an amorphous polymer in a solvent. Part 2: Preferential binding to local sites on a surface. *Langmuir* **27**, 12396–12404 (2011).
  41. Bansal, K., Baghel, U. S., & Thakral, S. Construction and Validation of Binary Phase Diagram for Amorphous Solid Dispersion Using Flory–Huggins Theory. *AAPS PharmSciTech* **17**, 318–327 (2016).
-

42. Gupta, J., Nunes, C., Vyas, S., & Jonnalagadda, S. Prediction of solubility parameters and miscibility of pharmaceutical compounds by molecular dynamics simulations. *J. Phys. Chem. B* **115**, 2014–2023 (2011).
43. Macháčková, M., Tokarský, J., & Čapková, P. A simple molecular modeling method for the characterization of polymeric drug carriers. *Eur. J. Pharm. Sci.* **48**, 316–322 (2013).
44. Maniruzzaman, M., Snowden, M. J., Bradely, M. S., & Douroumis, D. Studies of intermolecular interactions in solid dispersions using advanced surface chemical analysis. *RSC Adv.* **5**, 74212–74219 (2015).
45. Sarkar, N., Gonnella, N. C., Krawiec, M., Xin, D., & Aakeröy, C. B. Evaluating the Predictive Abilities of Protocols Based on Hydrogen-Bond Propensity, Molecular Complementarity, and Hydrogen-Bond Energy for Cocrystal Screening. *Cryst. Growth Des.* **20**, 7320–7327 (2020).
46. Akkermans, R. L. C., Spenley, N. A., & Robertson, S. H. Monte Carlo methods in Materials Studio. *Mol. Simul.* **39**, 1153–1164 (2013).
47. Yani, Y., Kanaujia, P., Chow, P. S., & Tan, R. B. H. Effect of API-Polymer Miscibility and Interaction on the Stabilization of Amorphous Solid Dispersion: A Molecular Simulation Study. *Ind. Eng. Chem. Res.* **56**, 12698–12707 (2017).
48. Pajula, K., Taskinen, M., Lehto, V. P., Ketolainen, J., & Korhonen, O. Predicting the formation and stability of amorphous small molecule binary mixtures from computationally determined flory-huggins interaction parameter and phase diagram. *Mol. Pharm.* **7**, 795–804 (2010).
49. Chakravarty, P., Lubach, J. W., Hau, J., & Nagapudi, K. A rational approach towards development of amorphous solid dispersions: Experimental and computational techniques. *Int. J. Pharm.* **519**, 44–57 (2017).
50. Ren, H., Zhang, Q., Chen, X., Zhao, W., Zhang, J., Zhang, H., Zeng, R., & Xu, S. A molecular simulation study of a series of cyclohexanone formaldehyde resins: Properties and applications in plastic printing. *Polymer (Guildf)*. **48**, 887–893 (2007).
51. Spackman, P. R., Thomas, S. P., & Jayatilaka, D. High throughput profiling of molecular shapes in crystals. *Sci. Rep.* **6**, 1–9 (2016).

52. Fábián, L. Cambridge Structural Database Analysis of Molecular Complementarity in Cocrystals. *Cryst. Growth Des.* **9**, 1436–1443 (2009).
53. Corpinot, M. K. & Bučar, D. K. A Practical Guide to the Design of Molecular Crystals. *Cryst. Growth Des.* **19**, 1426–1453 (2019).
54. Guidetti, M., Hilfiker, R., Kuentz, M., Bauer-Brandl, A., & Blatter, F. Exploring the Cocrystal Landscape of Posaconazole by Combining High-Throughput Screening Experimentation with Computational Chemistry. *Cryst. Growth Des.* **23**, 842–852 (2023).
55. Nauha, E. & Bernstein, J. ‘Predicting’ polymorphs of pharmaceuticals using hydrogen bond propensities: Probenecid and its two single-crystal-to-single-crystal phase transitions. *J. Pharm. Sci.* **104**, 2056–2061 (2015).
56. Stepanovs, D., Jure, M., Kuleshova, L. N., Hofmann, D. W. M., & Mishnev, A. Cocrystals of Pentoxifylline: In Silico and Experimental Screening. *Cryst. Growth Des.* **15**, 3652–3660 (2015).
57. Grecu, T., Hunter, C. A., Gardiner, E. J., & McCabe, J. F. Validation of a computational cocrystal prediction tool: Comparison of virtual and experimental cocrystal screening results. *Cryst. Growth Des.* **14**, 165–171 (2014).
58. Rajavel, A., Aditya Prasad, A., & Jeyakumar, T. Polymorphs of 4-isopropylbenzylidene thiophene-2-carbohydrazide: Crystal growth and density functional theory computations. *J. Mol. Struct.* **1130**, 138–149 (2017).
59. Grecu, T., Prohens, R., McCabe, J. F., Carrington, E. J., Wright, J. S., Brammer, L., & Hunter, C. A. Cocrystals of spironolactone and griseofulvin based on an in silico screening method. *CrystEngComm* **19**, 3592–3599 (2017).
60. Usha, C., Santhakumari, R., Meenakshi, R., Jayasree, R., & Bhuvaneshwari, M. Growth, density functional theory (DFT) and spectral studies on L-2-aminobutyric acid –biologically active material. *J. Mol. Struct.* **1149**, 655–661 (2017).
61. Khalaji, M., Potrzebowski, M. J., & Dudek, M. K. Virtual Cocrystal Screening Methods as Tools to Understand the Formation of Pharmaceutical Cocrystals - A Case Study of Linezolid, a Wide-Range Antibacterial Drug. *Cryst. Growth Des.* **21**, 2301–2314 (2021).

62. Motherwell, W. D. S., Ammon, H. L., Dunitz, J. D., Dzyabchenko, A., Erk, P., Gavezzotti, A., Hofmann, D. W. M., Leusen, F. J. J., Lommerse, J. P. M., Mooij, W. T. M., Price, S. L., Scheraga, H., Schweizer, B., Schmidt, M. U., Van Eijck, B. P., Verwer, P., & Williams, D. E. Crystal structure prediction of small organic molecules: A second blind test. *Acta Crystallogr. Sect. B Struct. Sci.* **58**, (2002).
63. Day, G. M., Cooper, T. G., Cruz-Cabeza, A. J., Hejczyk, K. E., Ammon, H. L., Boerrigter, S. X. M., Tan, J. S., Della Valle, R. G., Venuti, E., Jose, J., Gadre, S. R., Desiraju, G. R., Thakur, T. S., Van Eijck, B. P., Facelli, J. C., Bazterra, V. E., Ferraro, M. B., Hofmann, D. W. M., Neumann, M. A., *et al.* Significant progress in predicting the crystal structures of small organic molecules - A report on the fourth blind test. *Acta Crystallogr. Sect. B Struct. Sci.* **65**, (2009).
64. Bardwell, D. A., Adjiman, C. S., Arnautova, Y. A., Bartashevich, E., Boerrigter, S. X. M., Braun, D. E., Cruz-Cabeza, A. J., Day, G. M., Della Valle, R. G., Desiraju, G. R., Van Eijck, B. P., Facelli, J. C., Ferraro, M. B., Grillo, D., Habgood, M., Hofmann, D. W. M., Hofmann, F., Jose, K. V. J., Karamertzanis, P. G., *et al.* Towards crystal structure prediction of complex organic compounds - A report on the fifth blind test. *Acta Crystallogr. Sect. B Struct. Sci.* **67**, (2011).
65. Reilly, A. M., Cooper, R. I., Adjiman, C. S., Bhattacharya, S., Boese, A. D., Brandenburg, J. G., Bygrave, P. J., Bylsma, R., Campbell, J. E., Car, R., Case, D. H., Chadha, R., Cole, J. C., Cosburn, K., Cuppen, H. M., Curtis, F., Day, G. M., DiStasio, R. A., Dzyabchenko, A., *et al.* Report on the sixth blind test of organic crystal structure prediction methods. *Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater.* **72**, 439–459 (2016).
66. Nikhar, R. & Szalewicz, K. Reliable crystal structure predictions from first principles. *Nat. Commun.* **13**, 3095 (2022).
67. Day, G. M., Motherwell, W. D. S., & Jones, W. A strategy for predicting the crystal structures of flexible molecules: The polymorphism of phenobarbital. *Phys. Chem. Chem. Phys.* **9**, 1693–1704 (2007).
68. Tom, R., Rose, T., Bier, I., O'Brien, H., Vázquez-Mayagoitia, Á., & Marom, N. Genarris 2.0: A random structure generator for molecular crystals. *Comput. Phys. Commun.* **250**, 107170 (2020).
-

69. Mattei, A., Hong, R. S., Dietrich, H., Firaha, D., Helfferich, J., Liu, Y. M., Sasikumar, K., Abraham, N. S., Miglani Bhardwaj, R., Neumann, M. A., & Sheikh, A. Y. Efficient Crystal Structure Prediction for Structurally Related Molecules with Accurate and Transferable Tailor-Made Force Fields. *J. Chem. Theory Comput.* **18**, 5725–5738 (2022).
70. Hoja, J., Ko, H. Y., Neumann, M. A., Car, R., DiStasio, R. A., & Tkatchenko, A. Reliable and practical computational description of molecular crystal polymorphs. *Acta Hortic. Sin.* **45**, 1–10 (2019).
71. McDonagh, J. L., Palmer, D. S., Mourik, T. Van, & Mitchell, J. B. O. Are the sublimation thermodynamics of organic molecules predictable? *J. Chem. Inf. Model.* **56**, 2162–2179 (2016).
72. Sugden, I. J., Braun, D. E., Bowskill, D. H., Adjiman, C. S., & Pantelides, C. C. Efficient Screening of Coformers for Active Pharmaceutical Ingredient Cocrystallization. *Cryst. Growth Des.* **22**, 4513–4527 (2022).
73. Devogelaer, J., Meekes, H., Tinnemans, P., Vlieg, E., & Gelder, R. Co-crystal Prediction by Artificial Neural Networks. *Angew. Chemie Int. Ed.* **59**, 21711–21718 (2020).
74. Wang, D., Yang, Z., Zhu, B., Mei, X., & Luo, X. Machine-Learning-Guided Cocrystal Prediction Based on Large Data Base. *Cryst. Growth Des.* **20**, 6610–6621 (2020).
75. Musil, F., De, S., Yang, J., Campbell, J. E., Day, G. M., & Ceriotti, M. Machine learning for the structure-energy-property landscapes of molecular crystals. *Chem. Sci.* **9**, 1289–1300 (2018).
76. Heng, T., Yang, D., Wang, R., Zhang, L., Lu, Y., & Du, G. Progress in Research on Artificial Intelligence Applied to Polymorphism and Cocrystal Prediction. *ACS Omega* **6**, 15543–15550 (2021).
77. Kale, D. P., Zode, S. S., & Bansal, A. K. Challenges in Translational Development of Pharmaceutical Cocrystals. *J. Pharm. Sci.* **106**, 457–470 (2017).
78. Mswahili, M. E., Lee, M. J., Martin, G. L., Kim, J., Kim, P., Choi, G. J., & Jeong, Y. S. Cocrystal prediction using machine learning models and descriptors. *Appl. Sci.* **11**, 1–12 (2021).

79. Kostyrin, E. V., Ponkratov, V. V., & Salah Al-Shati, A. Development of machine learning model and analysis study of drug solubility in supercritical solvent for green technology development. *Arab. J. Chem.* **15**, 104346–104352 (2022).



## Chapter 2

# Experimental and Theoretical Methods

---

*This chapter covers experimental, characterisation and theoretical method applied during this work.*

---

### 2.1. Crystallization of Materials

#### 2.1.1. Agomelatine–phosphoric acid

##### (a) Patented Process (Patent No. EP2743255A1) for preparation of Agomelatine–phosphoric acid molecular complex for crystal structure determination

Agomelatine (50 g) and ethyl acetate (350 mL) were charged into a round bottom flask at 30°C, and the obtained mixture was stirred at 55°C until it ended with a clear solution. The solution was filtered through a micron filter and washed with ethyl acetate (100 mL). The phosphoric acid solution was also prepared by dissolving 22.4 g phosphoric acid in 150 mL of ethyl acetate. The prepared solution was added to the agomelatine solution was slowly for about 45 min at 30°C. The slurry containing precipitated product was stirred for 30 minutes at 30°C, cooled to 5°C, and maintained for 30 minutes. The precipitated product was collected by filtration, washed with ethyl acetate (50 mL) and dried at 35°C for 5 h to obtain the title compound<sup>1</sup>.

##### (b) Modified Process for preparation of agomelatine–phosphoric acid–molecular complex to understand phase transition behavior and transformation kinetics

Agomelatine (50 g) and 22.4 g phosphoric acid in ethyl acetate (350 mL) into a round bottom flask at 25°C and stirred to get a clear solution. After 30 min, precipitation will occur. Precipitated solid is filtered and dried at 25°C for 1 h.

##### (c) Crystallization attempts for Single crystal growth of agomelatine phosphoric acid molecular complex

- Agomelatine (1 g) and aqueous phosphoric acid (0.5 g) in 1:1 stoichiometric ratio was taken in conical flask and 10 mL of ethyl acetate was added and dissolved until it gets clear solution. The solution was kept on bench for crystallization. After one–week, microcrystalline material was precipitated.
- Diffusion crystallization was taken in the 50mL test tube and MTBE vapours are used as antisolvent. After two weeks, microcrystalline material was precipitated.

## (d) Screening experiments for API-PEG complexes

Table 2.1: List of Screening experiments for API-PEG complexes

API	PEG	Crystallization technique	Remarks
Dapsone	400	Antisolvent addition	<ul style="list-style-type: none"> <li>DAP-PEG co-crystal was formed.</li> <li>DAP-PEG co-crystal was formed.</li> <li>DAP-PEG co-crystal was formed. DAP-PEG400/4K material was observed slightly more crystalline than DAP-PEG 4K material.</li> </ul>
Dapsone	4K		
Dapsone	PEG400 + PEG4K		
Sulfanilamide	400		<ul style="list-style-type: none"> <li>SUL-PEG co-crystal was not formed.</li> <li>SUL-PEG co-crystal was not formed.</li> <li>Mixture of SUL-PEG (70%) and Sulfanilamide (30%) was formed.</li> </ul>
Sulfanilamide	4K		
Sulfanilamide	PEG400 + PEG4K		
Dapsone	4K	Fast solvent evaporation	SUL-PEG 4K was observed more crystalline compared to DAP-PEG 4K
Sulfanilamide	4K	Fast solvent evaporation	
Nicotinamide	4K	Antisolvent addition/DSC	Co-crystal not formed.
Lenalidomide	4K	Antisolvent addition/DSC	Co-crystal not formed.
Methyl paraben	4K	Antisolvent addition/DSC	Co-crystal formed.
Salicylic acid	4K	Antisolvent addition/DSC	Co-crystal not formed.
4-Aminosalicylic acid	4K	Antisolvent addition/DSC	Co-crystal formed.
Proline	4K	Antisolvent addition/DSC	Co-crystal not formed.

The above table (**Table 2.1**) is furnished with details on crystallization experiments performed to study (i) nobel order–disorder phenomena and symmetry in the long and local order, (ii) impact of process change on the crystallinity of API–PEG complex and (ii) API–PEG complexes screening. The first study on the order–disorder phenomenon was carried by crystallizing dapsone and sulfanilamide with higher molecular weight PEG (4K), lower molecular weight PEG (400) and mixture of both. To understand the impact of process, change on the crystallinity or molecular interactions in the dapsone–polyethylene glycol and sulfanilamide–polyethylene glycol complexes, fast solvent evaporation was performed. API–PEG complex screening was performed using anti–solvent addition and thermal analysis. (Differential Scanning Calorimetry). In anti–solvent addition crystallization approach, API (100 mg) was dissolved in methanol, ethanol, or acetone solvents (volume of solvent depends on solubility of the respective molecule) along with PEG at 50–60°C and kept for stirring until it becomes clear solution. As an anti–solvent, hexane was added to the mother solution drop–wise until it precipitates material or until it becomes hazy solution. The temperature of the solution is brought to 25°C and stirring is continued until the material precipitates out. The precipitated material was analyzed using DSC technique to confirm the formation of API–PEG complex. Physical solid mixtures of API and PEG were prepared in 1:1 stoichiometric ratio. These physical mixtures were analyzed on DSC by following the methodology given below.

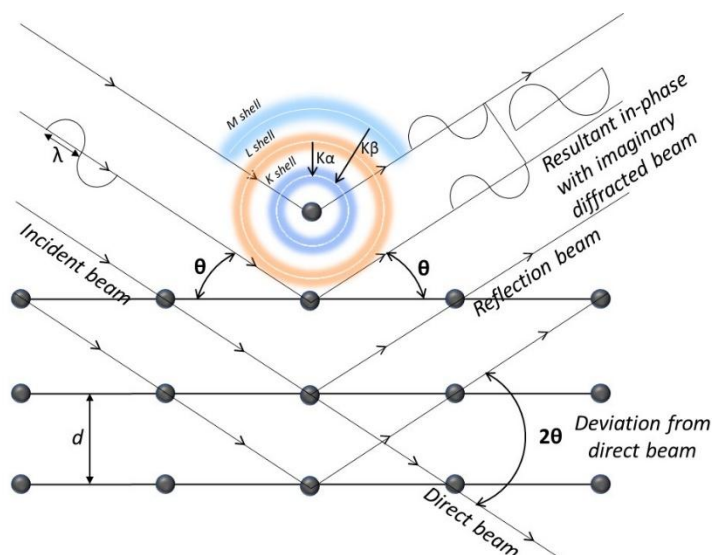
## 2.2. Characterization

### 2.2.1. Powder–XRD analysis

The phenomenon of diffraction is governed by Bragg’s Law. When X–rays incident on lattice plane of a crystal or crystalline material having atoms or molecules orderly arranged in 3D space and if the path difference between the waves is an integer multiple of wavelength ( $\lambda$ ), diffraction occurs (**Figure 2.1**).

$$n\lambda = 2d \sin\theta$$

Hence, X–ray diffraction is an unique tool for the phase identification. The phenomenon of polymorphism is exhibited by the material in which the material is having the same stoichiometry between the molecules but having different molecular arrangement. Hence powder X–ray diffraction is considered as promising and non–destructive tool for both qualitative and quantitative analysis. As qualitative analysis, it is used for the polymorphic form identification and detection of polymorphic impurities in the desired polymorphic forms. Also, it can quantify mixture of polymorphic forms.



**Figure 2.1:** Schematic illustration of Bragg's Law.

It is also a unique tool for the determination of crystallinity, since amorphous materials having molecular arrangement without order gives rise to the X-ray scattering background on contrary to the crystalline materials giving rise to the diffraction signals in the powder-XRD pattern. Crystallinity of material can be determined by determining the area under the whole pattern and area under the peaks in the powder-XRD pattern. Therefore, %crystallinity can be expressed as below equation.

$$\%Crystallinity = \frac{\text{Area under the peaks}}{\text{Total Area}} \times 100$$

Indexing of powder-XRD pattern can conclude the polymorphic purity of the material. Also, its crystal structure determination from either Single crystal-XRD or powder-XRD data can confirm purity as well as structural information in the solid state of the materials.

### **2.2.2. Powder-XRD data collection for polymorphic form identification**

XRPD data relevant to crystallization experiments were collected on Malvern PANalytical instrument, with the following instrumental configurations and parameters: PIXEL1D detector, reflection geometry, Cu K $\alpha$  radiation, scan range 3–40° 2 $\theta$ , step size of 0.013°, step time 49s.

### **2.2.3. Powder-XRD data collection for Indexing followed by Crystal structure determination**

For indexing of powder-XRD pattern and to solve crystal structure, powder-XRD data collection of AGL-P RT-form and AGL-P HT-form was done on Bruker D8 Advance instrument. Sample was packed in 1mm quartz capillary and mounted on the capillary sample

stage. Data was collected between 3–70° 2 $\theta$  range with 0.004° step size, 2s time per step for 10 h in Debye Scherrer geometry and focusing Göbel Mirror was used as a primary optics to generate a high-intensity K $\alpha$  beam.

### 2.2.4. Variable Temperature Powder–XRD studies

Variable temperature XRPD data for AGL–P were collected on Bruker D8 Advance Davinci instrument, attached with Anton Paar non-ambient stage. Data were collected at different temperature (°C) points at 30, 40, 50, 60 and the samples were cooled back to 30°C followed by data collection. The data was collected for 3–45° 2 $\theta$ , step size of 0.013°, step time 0.1s in reflection geometry using Cu K $\alpha$  radiation ( $\lambda = 1.5418 \text{ \AA}$ ), Lynxeye XE detector.

### 2.2.5. Crystal structure determination from SCXRD

Single crystal data for dapsone–polyethylene glycol and sulfanilamide–polyethylene glycol were collected on RIGAKU OXFORD/XtaLAB PRO with Cu–K $\alpha$  radiation ( $\lambda = 1.5418 \text{ \AA}$ ) at 293K. The data collection, integration, and data reduction were carried out by CrysAlisPro<sup>2</sup>. Structure solution were obtained by ShelXT<sup>3</sup> and refinement was carried by ShelXL<sup>4</sup> in Olex2<sup>5</sup>. All non-hydrogen atoms were refined with anisotropic displacement parameters. The aromatic and amino H–atoms were located from difference Fourier maps. In API–PEG systems, all non-hydrogen atoms of disordered PEG fragment were refined with isotropic thermal parameters and all H–atom positions were inserted in idealized positions and allowed to refine riding in the parent carbon atom. Molecular graphics were prepared using Mercury.<sup>6</sup>

### 2.2.6. Thermal analysis

Two thermal techniques were utilized to characterize the phase change of the materials as function of time and temperature. Differential Scanning Calorimetry (DSC) is a primary technique to measure thermal properties of materials to establish relation between temperature and physical properties of materials. It is potential technique to determine the enthalpy associated with material transitions as a function of temperature and time. It measures heat flow, which is radiated or absorbed excessively by the sample based on temperature difference between the sample and the reference material. A thermogram from DSC is a plot between temperature and heat Flow, which measures physical or chemical changes in the form of endothermic or exothermic events. An endothermic event indicates a solid–solid phase change or melting events and an exothermic event indicates recrystallization. Modulated DSC can measure both heat flow and heat capacity (reversing) simultaneously. It measures the glass transition temperature of amorphous materials. Thermogravimetric analysis (TGA) is another thermal technique measures the change in weight of materials as a function of temperature and

time. It is capable to measure identity anhydrous, solvates and hydrates and also measure stoichiometry between API and solvent molecules.

DSC analysis was performed using a Thermal Advantage (TA) discovery instrument. Around 5 mg of sample was sealed in aluminum pans and subjected to 10 °C/min ramp from 25 to 200 °C with continuous purging of dry nitrogen gas (50 mL/min). Thermogravimetric analysis (TGA) was performed on TA Q5000 series instrument by heating the samples at 10 °C/min ramp from 25 to 250 °C with a purging of dry nitrogen gas.

### 2.2.7. Hot Stage Microscopy

Thermomicroscopy is the combination of microscopy and thermal analysis to characterize the solid-state properties of pharmaceutical materials. Microscopy as a solo technique, it used to determine crystals or particles morphology and size of the materials. As a combined technique with thermal capability, it is used as a complimentary technique to routine thermal techniques (DSC and TGA). In this study, HSM imaging was done for AGL-P on Nikon's Eclipse polarized light microscope equipped with Mettler Toledo-FP28HT hot stage. Sample spread on a glass slide and images were taken at respective required temperatures. The microscopic images were recorded with a DP50 digital camera attached to Nikon BX-50 microscope (Nikon Optical GmbH, Vienna, A).

### 2.2.8. Spectroscopy

#### (a) IR Spectroscopy

In the qualitative prospective, Infrared spectroscopy analysis allows structural characterization of molecules giving information on functional groups of the molecules thereby supporting in the structural characterization. The IR absorption ranges from 400 to 4000  $\text{cm}^{-1}$ . Apart from classification of IR spectral range as functional group region and fingerprint region, effect of hydrogen bonding on it leading to the changes in the spectra is an interesting phenomenon in the present context. For example, hydrogen bonding shifts to characteristic bands from lower to higher frequency<sup>7,8</sup>. FT-IR spectrum was recorded for AGL Form-II, Phosphoric acid, AGL-P RT-form materials, DAP-PEG and sulfanilamide-PEG. About 5 mg of sample was triturated with 500 mg of dried KBr and a disc of 10-15mm diameter was prepared. IR Spectra were acquired over a range of 4000-450  $\text{cm}^{-1}$  at a resolution of 4  $\text{cm}^{-1}$  with 16 scans.

#### (b) $^1\text{H}$ and $^{13}\text{C}$ NMR

Multinuclear ( $^1\text{H}$  and  $^{13}\text{C}$ ) NMR spectroscopy is the most commonly used analytical-qualitative technique in the structural characterization of molecules and also as a quantitative technique to understand the stoichiometry between the individual components in the complexes

(cocrystals and salts). In this thesis work,  $^1\text{H}$  NMR and  $^{13}\text{C}$  NMR techniques were used to study the stoichiometric ratio between the API and coformer in DAP-PEG and sulphanilamide-PEG complexes. Spectra were recorded using Bruker Advance 500 MHz for  $^1\text{H}$  NMR and 126 MHz for  $^{13}\text{C}$  NMR using DMSO- $d_6$  as solvent. The  $^1\text{H}$  NMR chemical shifts values were referenced to the TMS (Tetra methyl Silane;  $\delta = 0.00$  ppm). All the experiments were carried out in the phase sensitive mode<sup>9</sup>.

### (c) Solid State NMR

Solid-state NMR is a well established technique in the pharmaceutical company for the polymorphic identification. In general, solid crystalline materials contain supramolecular building blocks of small molecules arranged in 3-dimensional array by means of intermolecular interactions. These intermolecular interactions between the molecules govern the Physico-chemical properties of the materials. By integrating solid state NMR with theoretical calculations of NMR chemical shifts (using DFT calculations) as a complementary has given birth to a new branch named as NMR crystallography. These combined experimental and experimental integration can probe and extract molecular level structural information of the materials such as nuclear proximities, torsional angles together giving molecular packing and spatial arrangement. These methods are becoming more prominent in the SDPD methodologies, where powder-XRD data may not be having adequate information due to peak overlapping problem. In such cases, NMR crystallography is getting recognition as a benchmark for the validation of determined crystal structures. Solid-state NMR was also applied in the study of the local order and impact of local defects on the scale of long-range order. In other words, crystallographic deduction of local order and global order in the material<sup>10,11</sup>. Solid-state NMR was used to study the protonation state of agomelatine.  $^{13}\text{C}$  CP/MAS NMR spectra of AGL form-II, AGL-P RT-form were measured by Bruker Advance III 500 MHz WB (wide bore) NMR spectrometer equipped with 4 mm probe and with 14 kHz spinning.  $^{13}\text{C}$  CP/MAS NMR spectra of DAP-PEG and sulphanilamide-PEG complexes were measured by Bruker 400 MHz NMR spectrometer equipped with 5 mm probe with 10 kHz spinning.

### 2.2.9. Theoretical Methods

#### 2.2.9.1. Geometric optimization of molecular structure using forcefield calculations

For geometric optimization of molecular structure, Forcite module in Material Studio was used. COMPASS III forcefield was used and charges were applied by forcefield assigned. Geometric



optimization was done using atom-based electrostatic and Van der Waals summation methods. COMPASS forcefield belongs to family of consistent forcefields parameterized to a range of functional groups. It covers broad range of covalent molecules, inorganic molecules and polymers. COMPASS was parameterized exclusively using quantum mechanical calculations (B3LYP/6-31G(d,p)) and Generalized Gradient Approximation (GGA) functional due to Perdew, Burke and Ernzerhof (PBE) from Density Functional Theory were used<sup>12</sup>. It was parameterized to predict various molecular properties such as molecular structures, vibrational frequencies, conformation energies, dipole moments, liquid structures, crystal structures, equations of state, and cohesive energy densities in isolation and condensed phases.

### **2.2.9.2. Polymorphic purity from powder-XRD: Determination of lattice parameters and their refinement by Pawley refinement**

Crystal structures were solved from powder-XRD data by using reflex module in Materials Studio 17.0. The lattice parameters were determined and refined through indexing powder-XRD patterns by X-cell<sup>13</sup>. Accurate lattice constants and various experimental setup parameters (zero-point shift, background, and profile parameters) and also the texture of the sample were determined by Pawley profile fitting<sup>14</sup>. All peak intensities were treated as independent parameters at this stage.

### **2.2.9.3. Crystal Structure Determination from powder-XRD**

#### **(a) Structure Solution by Monte Carlo Simulations**

In the structure solution step, geometric optimization of structural models were carried out by COMPASS forcefields. Molecular structures in the unit cell were defined by Z-matrix using in reflex powder solve module. Translational, rotational, and torsional angle movements were applied to the molecular models. By applying direct space methods and simulated annealing global optimization techniques, structure solution was obtained. Refined lattice parameters, peak shape and peak asymmetry values, which was obtained during Pawley refinement, were used for structure solution step.

#### **(b) Rietveld refinement**

In the Rietveld refinement step, all atoms were refined isotropically, and profile parameters (Pseudo-Voigt function), peak asymmetry (Berar-Baldinozzi), background (20<sup>th</sup> order of polynomial order), zero-point shift, preferred orientation (March-Dollase) were refined starting from default settings. During refinement, constraints and restraints were applied to the

bond lengths (0.1 Å) and torsional angles (10°). Similar structural models. Hydrogen atoms were constrained in relation to their corresponding bonded heavier atoms<sup>15</sup>.

### (c) Geometric optimization of crystal structure using VASP

For structural evaluation and optimization, plane wave–periodic DFT–D2 calculations were performed. Along with DFT, dispersion corrections were applied to correct the van der Waal’s interactions. Exchange and correlation effects were calculated in the GGA–PBE<sup>16</sup> level of theory using Vienna *Ab Initio* Simulation Package (VASP)<sup>17–19</sup> as implemented in the MedeA software<sup>20</sup>. The plane wave cut–off energy was set to 400 eV, the k–point spacing was set to 0.5 Å<sup>-1</sup> approximately. These energy minimizations were carried out keeping the cell dimensions fixed with atomic positions relaxed.

### (d) Geometric optimization using CASTEP

Geometric optimization of crystal structures were geometrically optimized using CASTEP module in Materials Studio 20.0<sup>21</sup>. Exchange–associative interactions were described by Generalized Gradient Approximation (GGA) and Perdew–Burke–Ernzerh (PBE) functional along with TS additive dispersion correction by keeping cell parameters constant and allowing only atomic positions to optimize<sup>16</sup>. The k–point meshes 1 × 1 × 1 were used for Brillouin zone sampling. Convergence criteria with cutoff energy of 571eV, maximum force of 0.05 eV/Å, max displacement of 0.002 Å.

### (e) Lattice energy calculation

Structural relaxation of crystal structures and corresponding components were geometrically optimized using CASTEP module in Materials Studio 20.0. Exchange–associative interactions were described by Generalized Gradient Approximation (GGA) and Perdew–Burke–Ernzerh (PBE) functional along with TS additive dispersion correction by keeping cell constant and allowing only atomic positions to optimize. The k–point meshes 1 × 1 × 1 were used for Brillouin zone sampling. Convergence criteria with cutoff energy of 571eV, maximum force of 0.05 eV/Å, max displacement of 0.002 Å. Lattice energy was calculated using total energy of the cell, E(bulk) and relaxed energies  $E_A$  and  $E_B$  of isolated molecules of A and B<sup>22</sup>.

$$E_{lattice} = \frac{E_{Bulk}}{Z} - E_A - E_B$$

#### 2.2.9.4. Comparison (overlay) of crystal structure CRYSTALCMP

The similarity between the crystal structures were calculated by CRYSTALCMP algorithm<sup>23</sup>. The similarity between the crystal structures was quantified in terms of RMSD

(root mean-squared deviation). If the crystal structure contains more than one component in its asymmetric unit, algorithm will consider only largest molecule among them by ignoring all hydrogen atoms. Clusters of 30 molecules from each crystal structure were compared with a distance penalty of 3 Å. The algorithm overlays the crystal structures of polymorphs and calculates packing similarity using the distance between molecular centre and rotation angles between related pairs by values. Finally, it returns the ( $PS_{ab}$ ) packing similarity index and RMSD based on crystal packing and molecular conformations<sup>23</sup>.

$$PS_{ab} = D_c + X \frac{A_d}{180}$$

$A_d$  = Average angle between molecular centres of related pairs

$D_c$  = Average distance between molecular centres of related pairs.

$X$  = weight

$$A_d = \frac{\sum A_i}{N}$$

$A_i$  = Angle between the molecules of related pair

$N$  = Number of related pairs.

$$A_i = \frac{1}{D} \sum_m^M \frac{A_m}{d_m^3}$$

$$D = \sum \frac{1}{d_m^3}$$

$M$  is the number of closest molecules

$A_m$  is the angle between molecules  $a_i$  and  $a_m$ .

$d_m$  is the distance molecules  $a_i$  and  $a_m$

### 2.2.9.5. Calculation of molecular and crystal structure descriptors using QSAR Module

- *Molecular Flexibility (No of rotatable bonds, No of torsion angles)*
- *Spatial descriptors (Molecular volume, Dipole moment)*
- *Crystal Structure similarity index using QSAR*

Descriptors are representation of molecules in the form of numerical or vectors to capture the structural characteristics and properties of molecules. In general, these molecular descriptors representing structural features of the molecules are used to construct a database. By the application of artificial intelligence and machine learning tools, viable models are constructed. By feeding these models with structural information, desired properties of the new molecules

---

can be predicted<sup>24</sup>. The above mentioned molecular flexibility, spatial descriptors and crystal structure similarity were calculated using QSAR modules in Materials Studio<sup>21</sup>.

### 2.2.9.6. Calculation of Thermodynamic descriptors

- *Calculation of Binding energy and Mixing energy using Blends Module*
- *Interaction energy*

To understand the affinity of drug molecules to form cocrystal with PEG as co-former, thermodynamics of mixtures were calculated by using Blends Module in Materials studio. By calculating binding energy, interaction parameter and mixing energies, the compatibility between the APIs and co-formers was estimated. Binding energy is calculated by applying extended Flory–Huggins’s theory by generating number of possible number of pair configurations using Boltzmann factor at certain temperature. Here algorithm packs molecules i and j as per their definition being ‘base’ or ‘screen’ molecule. A screen molecule is placed middle of the cluster and it is packed by surrounded by base molecules. Number of such configurations are generated randomly by using translation and rotation motions<sup>25–27</sup>.

$$\langle E_{ij} \rangle_T = \frac{\int dE E P_{ij}(E) e^{\frac{-E}{RT}}}{\int dE P_{ij}(E) e^{\frac{-E}{RT}}}$$

$P_{ij}$  is energy distribution of the binding energy between components i and j.  $\exp(-E_{ij}/RT)$  is Boltzmann factor at temperature T for pair coordination numbers. Binding energy is a measure of the energy of interaction between two components. Together with the coordination numbers, it enables generation of the mixing energy and the chi parameter. In mixing energy calculations, every molecule is given both base and screen role during calculation. For a lattice with coordination number Z, the mixing energy is

$$E_{mix} = \frac{1}{2}Z(E_{bs} + E_{sb} - E_{bb} - E_{ss})$$

Interaction energy ( $\chi$ ) was calculated by using

$$\chi = \frac{E_{mix}}{RT}$$

where  $E_{mix}$  is the mixing energy; the difference in free energy due to interaction between the mixed and the pure state. If interaction parameter is small, it implies that mixture favors the miscibility and it is positive, it implies that mixture disfavors miscibility.

Blends module combines improved Flory–Huggins’s theory and molecular simulations. By considering relative values of chi (interaction parameter) and binding energies, it was attempted to screen suitable drug molecules which are likely to form cocrystal with PEG. The Blends Module calculations were done using COMPASS II forcefield, charges by Gasteiger and atom–based summation. All calculations were performed by generating 10 million energy samples (one hundred thousand cluster samples) with 20 iterations per each cluster at 298K with 18.5Å Cutoff distance<sup>28</sup>. These thermodynamic mixing of binary mixtures was done between drug molecule and series of PEG chains differing in number of monomers (n = 2, 3, 4, 10, 15 and 20). General expression for the free energy of mixing for a binary system is

$$\frac{\Delta G}{RT} = \frac{\phi_b}{n_b} \ln \phi_b + \frac{\phi_s}{n_s} \ln \phi_s + \chi \phi_b \phi_s$$

$\Delta G$  is the free energy of mixing (per mole),  $\phi_i$  is the volume fraction of component,  $n_i$  is the degree of polymerization of component i,  $\chi$  is the interaction parameter, T is the absolute temperature, and R is the gas constant.

### 2.2.9.7. Calculation of Supramolecular descriptors

- *Hydrogen bond propensity*
- *Amorphous cell*

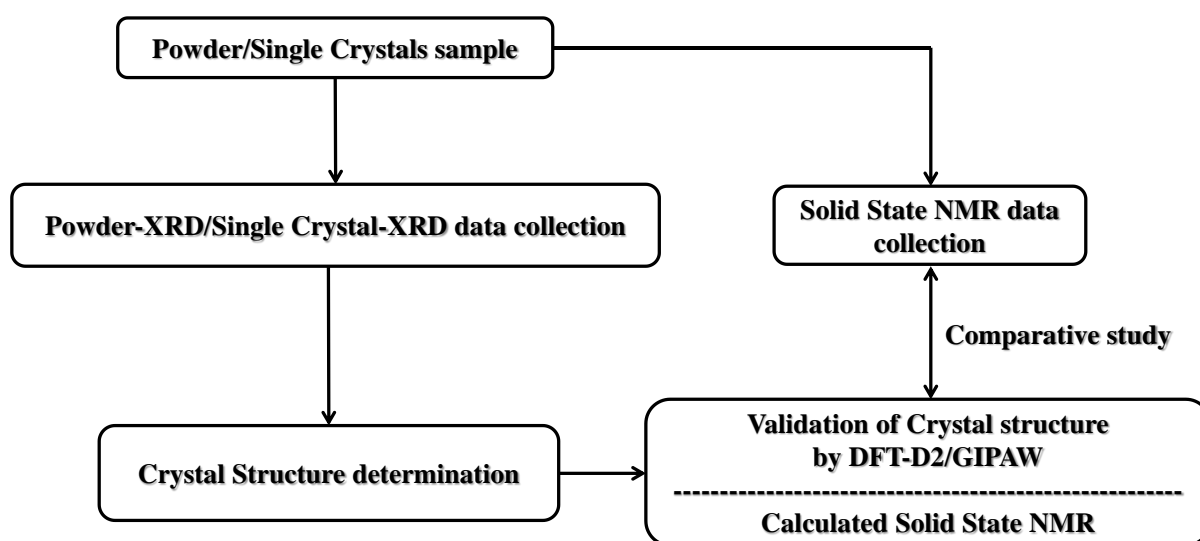
To understand drug–polymer affinity via hydrogen bonding interactions, a drug–polymer system was constructed using amorphous cell module from materials studio, by taking molecular cluster model as input generated from blends simulation calculations<sup>29</sup>.

Amorphous cell is typically used to perform structural characterization and property prediction for the modelled systems. For an isolated molecular system containing a small number of atoms, the atoms in the system would lie close to the periphery of the system and it can lead to significant surface effects. Typically, liquids and amorphous solids are most commonly studied using periodic boundary conditions and drug–polymer cluster generated from the Blends Module would be such ideal system to study. The Amorphous cell module in materials studio provides such a comprehensive set of tools to construct three–dimensional periodic structures of molecular and polymeric systems. Finally, the simulated drug–polymer system build in the amorphous cell is subjected to geometric optimization using molecular mechanics. The methodology adopted herein was validated using a known drug–polymer system i.e., DAP–PEG crystal structure.

### 2.2.9.8. Calculation of NMR chemical shift calculations using GIPAW method

Solid-State NMR (SS-NMR) spectroscopy can be used as complementary to the SDPD and SCXRD in the determination of structural properties of solid crystalline materials. SS-NMR combined with GIPAW calculations are proven to act as potential methodology for the improvement of crystal structures determined from powder-XRD or SC-XRD. In general, these methodologies can be applied for the following applications<sup>30</sup> (Figure 2.2).

1. Determination of number of molecules in the asymmetric unit
2. Detection and addressing disorders in the crystal structures
3. Enhancement and Validation of crystal structures determined from powder-XRD data

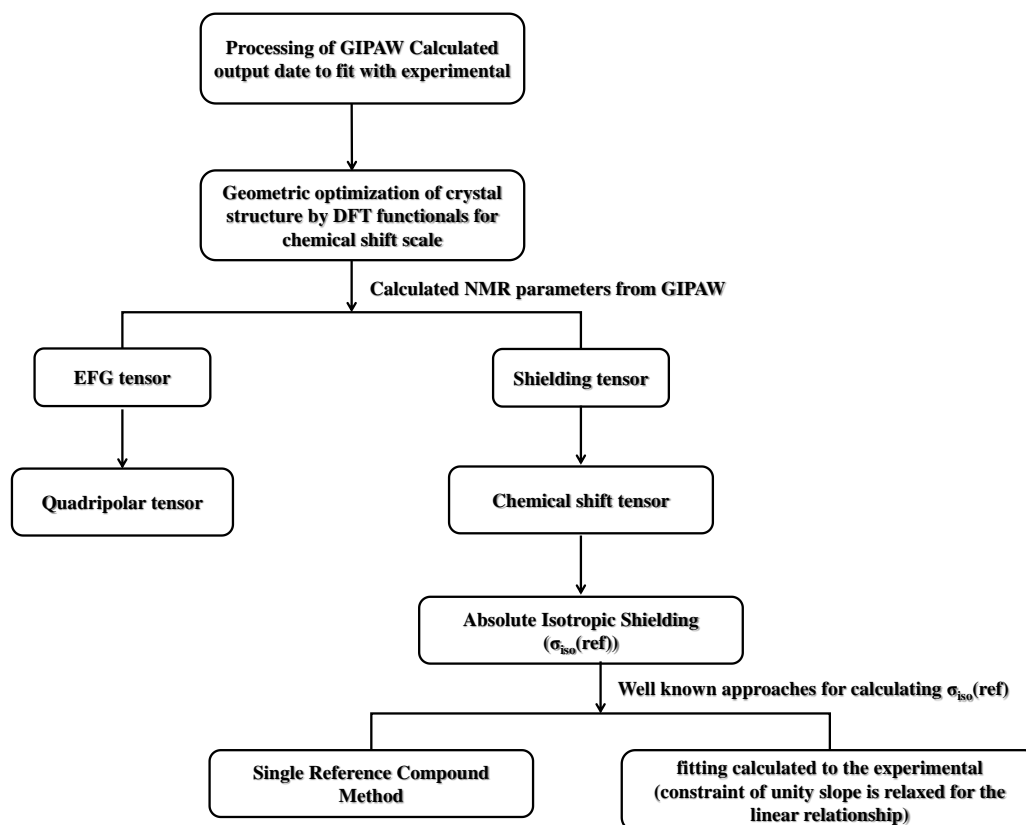


**Figure 2.2:** Flow chart showing the overview of validation of crystal structure by using powder/single crystal-XRD and SS-NMR as complementary techniques.

In this thesis work, GIPAW calculations are applied for the validation of crystal structures solved from the powder-XRD. Fundamentally SS-NMR gives information on dipolar couplings, J and quadrupolar couplings and chemical shifts<sup>31</sup>. GIPAW approach has evolved and matured enough to calculate key components of NMR interactions such as shielding and electric field gradient tensors. These theories are effectively implemented in many software packages such as CASTEP. These tools are proven to accurately calculate NMR parameters in periodic crystals. GIPAW calculations are carried by computing geometric optimization of crystal structures using DFT functionals for chemical shift scale. It calculates two tensors: (i) Shielding tensor and (ii) EFG tensor. In order to correlate, experimental and calculated shielding values, shielding tensors should be converted to chemical shift tensor and similarly EFG to quadrupolar tensor. As in the current context, since chemical shift tensors are of main interest, it will be discussed further. Chemical shift tensors need calculation of isotropic

shielding reference values. These calculations are mainly carried through well-known two approaches. The first method is single reference compound method (with known) and second method follows fitting calculated values to the experimental values (constraint of unity slope is relaxed for the linear relationship)<sup>32</sup>.

NMR chemical shifts for the crystal structure of API–Cofomer was calculated using CASTEP program using GIPAW approach<sup>32–34</sup>. Crystal structure was geometrically optimized using exchange–associative interactions were described by Generalized Gradient Approximation (GGA) and Perdew–Burke–Ernzerh (PBE) functional along with TS additive dispersion correction. Cell parameters were kept constant and only atomic positions were allowed to optimize. The k–point meshes  $1 \times 1 \times 1$  were used for Brillouin zone sampling. Convergence criteria with cutoff energy of 781 eV, maximum force of 0.05 eV/Å, max displacement of 0.002 Å. NMR chemical shielding have been calculated using the GIPAW method for the single crystal X–ray diffraction structure. Referencing of the  $^{13}\text{C}$  chemical shifts was achieved using the formula  $\delta_{\text{iso}}(\text{calc}) = \sigma_{\text{ref}} - \sigma_{\text{iso}}(\text{calc})$ . Here,  $\sigma_{\text{ref}}$  was calculated by deriving a correlation factor by drawing linearity between API NMR shift values (experimental) of selective reference atoms and theoretical chemical shifts of corresponding reference atoms and thereby calculated chemical shifts were scaled to reference chemical shifts (**Figure 2.3**).



**Figure 2.3:** Flow chart showing the overview of validation of crystal structure by using Powder–XRD and SS–NMR as complementary techniques.

### 2.2.9.9. Crystal morphology simulation from crystal structure

Crystallization steps in drug manufacturing process govern many fundamental properties of resulting crystalline materials such as polymorphism, crystal size and morphology (crystal habit and shape) etc., Among them, crystal morphology is one of the key properties in the pharma which can dictate the physicochemical properties of the final drug products such as appearance, stability, solubility, and dissolution properties<sup>35</sup>. Crystal shape is relative length to width of the crystal faces. Recent advances in the development of theoretical tools of morphology prediction have laid path to design experiments with rationality to modulate materials with desired morphological properties<sup>36</sup>. The fundamental key concept of predicting morphology involves calculating stable and unstable edges from most probable *hkl* planes<sup>35</sup>. The idea of employing theoretical methodologies for morphology prediction were seeded by Gibbs in 1875 stating that crystal surfaces are dominated by slow growing facets and the shape of a crystal depends on the relative growth rates of the facets in the crystal<sup>37</sup>.

**BFDH (Bravais–Friedel Donnay–Harkar):**



## Chapter 2: Experimental and Theoretical Methods

---

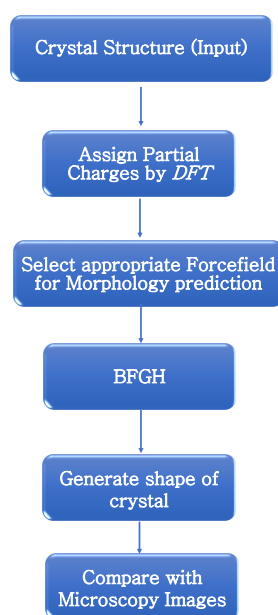
Bravias (1866) and Friedel (1907) proposed that the relative growth of facets in the crystals can be related to crystal structure and interactions within the crystal structure. As per this approach, the relative growth rate ( $R_{hkl}$ ) of  $h k l$  forms is inversely proportional to interplanar distance  $d_{hkl}$ <sup>36-38</sup>.

$$R_{hkl} \propto \frac{1}{d_{hkl}}$$

$R_{hkl}$  = Relative growth rate;  $d_{hkl}$  = interplanar distance.

Crystal structural information is used as input to extract symmetry and lattice planes information to generate list of possible growing faces.

### Methodology:



**Figure 2.4:** Flow chart showing the methodology for morphology prediction.

To compute morphology predictions, crystal structure information is the fundamental requirement. When a material exhibits polymorphism, it is advisable to select a crystal structure of a polymorphic form which is most likely to form. Partial charges on the atoms in the selected representative unit should be calculated by using DFT calculations or by the applying appropriate forcefield. Morphology predictions were computed using BFGH method implemented in WinXmorph by using optimized crystal structures as input<sup>39</sup>. Optimum and suitable morphology prediction method and compare the obtained crystal shape results with the experimental microscopy images for further evaluation<sup>35</sup> (**Figure 2.4**).

### 2.3. References

1. Kumar, V. S., Rao, B. V. K., Charyulu, K. S., Praveen, C., Reddy, Y. S., & Dey, A. Cocrystal of agomelatine with phosphoric acid, EP2743255A1. 1–18 (2014).
2. Agilent. CrysAlisPro Software System. *Technol. UK Ltd, Yarnton, Oxford, UK* **44**, (2014).
3. Sheldrick, G. M. SHELXT - Integrated space-group and crystal-structure determination. *Acta Crystallogr. Sect. A Found. Crystallogr.* **71**, 3–8 (2015).
4. Sheldrick, G. M. Crystal structure refinement with SHELXL. *Acta Crystallogr. Sect. C Struct. Chem.* **71**, 3–8 (2015).
5. Dolomanov, O. V., Bourhis, L. J., Gildea, R. J., Howard, J. A. K., & Puschmann, H. OLEX2: A complete structure solution, refinement and analysis program. *J. Appl. Crystallogr.* **42**, 339–341 (2009).
6. Macrae, C. F., Bruno, I. J., Chisholm, J. A., Edgington, P. R., McCabe, P., Pidcock, E., Rodriguez-Monge, L., Taylor, R., van de Streek, J., & Wood, P. A. Mercury CSD 2.0 – new features for the visualization and investigation of crystal structures. *J. Appl. Crystallogr.* **41**, 466–470 (2008).
7. Othman, N. IR Spectroscopy in Qualitative and Quantitative Analysis. *Infrared Spectrosc. - Perspect. Appl.* (2023).
8. Coates, J. Interpretation of Infrared Spectra, A Practical Approach. *Encycl. Anal. Chem.* (2006).
9. Portieri, A., Harris, R. K., Fletton, R. A., Lancaster, R. W., & Threlfall, T. L. Effects of polymorphic differences for sulfanilamide, as seen through <sup>13</sup>C and <sup>15</sup>N solid-state NMR, together with shielding calculations. *Magn. Reson. Chem.* **42**, 313–320 (2004).
10. Dib, E., Mineva, T., Gaveau, P., & Alonso, B. <sup>14</sup>N solid-state NMR: A sensitive probe of the local order in zeolites. *Phys. Chem. Chem. Phys.* **15**, 18349–18352 (2013).

11. Stebbins, J. F., McCarty, R. J., & Palke, A. C. Solid-state NMR and short-range order in crystalline oxides and silicates: a new tool in paramagnetic resonances. *Acta Crystallogr. Sect. C Struct. Chem.* **73**, 128–136 (2017).
12. Gupta, J., Nunes, C., Vyas, S., & Jonnalagadda, S. Prediction of solubility parameters and miscibility of pharmaceutical compounds by molecular dynamics simulations. *J. Phys. Chem. B* **115**, 2014–2023 (2011).
13. Neumann, M. A. X-Cell: A novel indexing algorithm for routine tasks and difficult cases. *J. Appl. Crystallogr.* **36**, 356–365 (2003).
14. Pawley, G. S. S. Unit-cell refinement from powder diffraction scans. *J. Appl. Crystallogr.* **14**, 357–361 (1981).
15. Rietveld, H. M. & IUCr. A profile refinement method for nuclear and magnetic structures. *J. Appl. Crystallogr.* **2**, 65–71 (1969).
16. Perdew, J. P., Burke, K., & Ernzerhof, M. Erratum: Generalized gradient approximation made simple. *Phys. Rev. Lett.* **78**, 1396 (1997).
17. Kresse, G. & Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **6**, 15–50 (1996).
18. Kresse, G. & Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B - Condens. Matter Mater. Phys.* **54**, 11169–11186 (1996).
19. Joubert, D. & Kresse, G. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B - Condens. Matter Mater. Phys.* **59**, 1758–1775 (1999).
20. MedeA®. *Mater. Des. Inc., San Diego, CA, USA* (2015).
21. BIOVIA. Dassault Systèmes, Biovia Materials Studio, 2017R2, San Diego: Dassault Systèmes, *Dassault Systèmes* (2017).
22. Basford, P. A., Connor, L. D., Whiteside, P., & Cruz-Cabeza, A. J. Does the Age of Pharmaceuticals Matter? Undetectable Hydrate Seeds Impact Hydration Behavior. *Cryst. Growth Des.* **21**, 1912–1916 (2021).
23. Rohlíček, J., Skořepová, E., Babor, M., & Čejka, J. CrystalCMP: An easy-to-use tool for fast comparison of molecular packing. *J. Appl. Crystallogr.* **49**, 2172–2183 (2016).
24. Kaneko, H. Molecular Descriptors, Structure Generation, and Inverse QSAR/QSPR Based on SELFIES. *ACS Omega* **8**, 21781–21786 (2023).
25. Pajula, K., Taskinen, M., Lehto, V. P., Ketolainen, J., & Korhonen, O. Predicting the formation and stability of amorphous small molecule binary mixtures from

- computationally determined flory-huggins interaction parameter and phase diagram. *Mol. Pharm.* **7**, 795–804 (2010).
26. Chakravarty, P., Lubach, J. W., Hau, J., & Nagapudi, K. A rational approach towards development of amorphous solid dispersions: Experimental and computational techniques. *Int. J. Pharm.* **519**, 44–57 (2017).
  27. Ren, H., Zhang, Q., Chen, X., Zhao, W., Zhang, J., Zhang, H., Zeng, R., & Xu, S. A molecular simulation study of a series of cyclohexanone formaldehyde resins: Properties and applications in plastic printing. *Polymer (Guildf)*. **48**, 887–893 (2007).
  28. de Arenaza, I. M., Meaurio, E., & Sarasu, J.-R. Analysis of the Miscibility of Polymer Blends Through Molecular Dynamics Simulations. in *Polymerization* 29–46 (InTech, 2012).
  29. Akkermans, R. L. C., Spenley, N. A., & Robertson, S. H. Monte Carlo methods in Materials Studio. *Mol. Simul.* **39**, 1153–1164 (2013).
  30. Harris, K. D. M. NMR Crystallography as a Vital Tool in Assisting Crystal Structure Determination from Powder XRD Data. *Crystals* **12**, 1277 (2022).
  31. Sturniolo, S., Green, T. F. G., Hanson, R. M., Zilka, M., Refson, K., Hodgkinson, P., Brown, S. P., & Yates, J. R. Visualization and processing of computed solid-state NMR parameters: MagresView and MagresPython. *Solid State Nucl. Magn. Reson.* **78**, 64–70 (2016).
  32. Charpentier, T. The PAW/GIPAW approach for computing NMR parameters: A new dimension added to NMR study of solids. *Solid State Nucl. Magn. Reson.* **40**, 1–20 (2011).
  33. Watts, A. E., Maruyoshi, K., Hughes, C. E., Brown, S. P., & Harris, K. D. M. Combining the Advantages of Powder X-ray Diffraction and NMR Crystallography in Structure Determination of the Pharmaceutical Material Cimetidine Hydrochloride. *Cryst. Growth Des.* **16**, 1798–1804 (2016).
  34. Dudenko, D. V., Yates, J. R., Harris, K. D. M., & Brown, S. P. An NMR crystallography DFT-D approach to analyse the role of intermolecular hydrogen bonding and  $\pi$ - $\pi$  interactions in driving cocrystallisation of indomethacin and nicotinamide. *CrystEngComm* **15**, 8797 (2013).
  35. Dandekar, P., Kuvadia, Z. B., & Doherty, M. F. Engineering Crystal Morphology. *Annu. Rev. Mater. Res.* **43**, 359–386 (2013).
  36. Moreno-Calvo, E., Calvet, T., Cuevas-Diarte, M. A., & Aquilano, D. Relationship between the Crystal Structure and Morphology of Carboxylic Acid Polymorphs.
-

- Predicted and Experimental Morphologies. *Cryst. Growth Des.* **10**, 4262–4271 (2010).
37. Docherty, R., Clydesdale, G., Roberts, K. J., & Bennema, P. Application of Bravais-Friedel-Donnay-Harker, attachment energy and Ising models to predicting and understanding the morphology of molecular crystals. *J. Phys. D. Appl. Phys.* **24**, 89–99 (1991).
38. Coombes, D. S., Catlow, C. R. A., Gale, J. D., Hardy, M. J., & Saunders, M. R. Theoretical and Experimental Investigations on the Morphology of Pharmaceutical Crystals. *J. Pharm. Sci.* **91**, 1652–1658 (2002).
39. Kaminsky, W. WinXMorph : a computer program to draw crystal morphology, growth sectors and cross sections with export files in VRML V2.0 utf8-virtual reality format. *J. Appl. Crystallogr.* **38**, 566–567 (2005).

## Chapter 3

# Agomelatine–Phosphoric Acid Molecular Complex

---

*This Chapter Covers Solid-State Phase Transition of Agomelatine–Phosphoric Acid Molecular Complexes along the Salt–Cocrystal Continuum: Ab Initio Powder X-ray Diffraction Structure Determination and DFT-D2 Analysis*

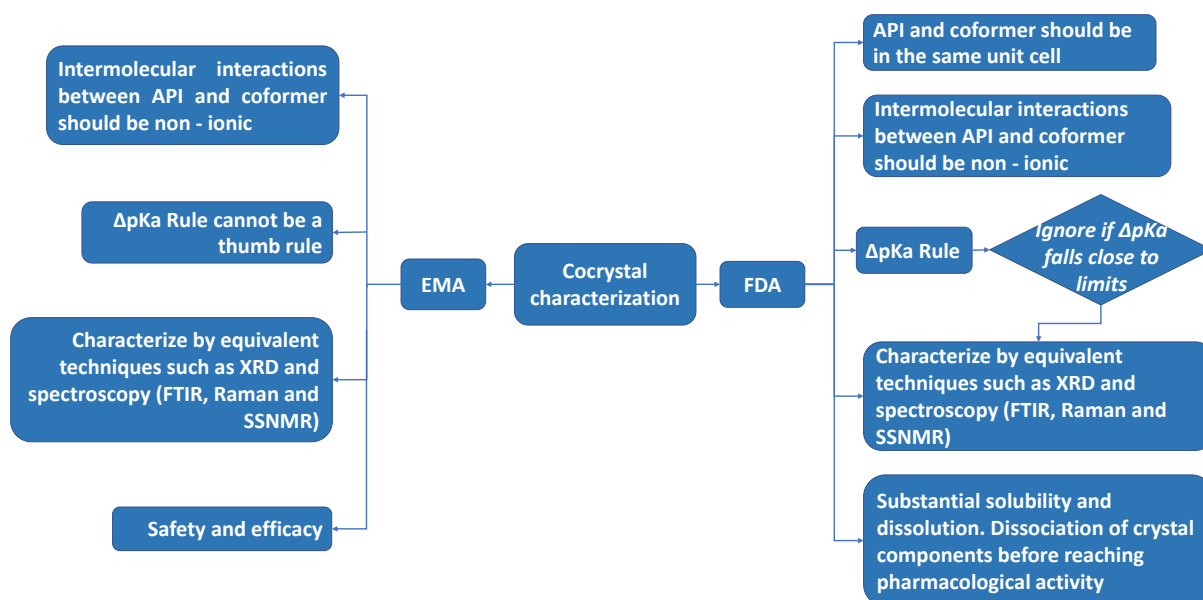
---

### 3.1. Introduction

The current work focuses on characterization of agomelatine phosphoric acid multi-component system using orthogonal experimental and theoretical techniques. Although traditionally salt form is the most preferable form<sup>1</sup>, in the last few decades, the concept of co-crystals is grabbing attention of both academic and industry interests. There are two requirements mainly which are being addressed effectively in this route:

1. Drug molecules without potential functional groups to form salts<sup>2</sup>
2. Intellectual property for generic industries

In 2013, FDA first released its guidance in 2013 and EMA has first released its reflection paper on cocrystals in 2014<sup>3</sup>. It has increased interest for academic and pharmaceutical industries on these entities to overcome both constraints from physical properties and IP reasons. As per FDA guidelines, intermolecular interactions between API and coformer should be non-ionic and they should present in the same unit cell. If  $\Delta pK_a$  rule does not distinguish the adduct between salt and cocrystal, characterization techniques should be used to confirm salt or cocrystal (**Figure 3.1**). As per EMA guidelines, intermolecular interactions between API and coformer should be non-ionic which should be characterized by the techniques such as XRD and spectroscopy techniques (FT-IR, Raman and SS-NMR) and must show equivalency between generic and innovator drugs with respect to safety and efficacy.



**Figure 3.1:** Outlines of FDA and EMA guidelines for cocrystals conformation.

Agomelatine is a BCS class II antidepressant drug marketed by Servier pharmaceutical company under the trade name of Valdoxan. The molecular formula of agomelatine is  $C_{15}H_{17}NO_2$  and its molecular weight is 243.3. The IUPAC name of the molecule is N-[2-(7-

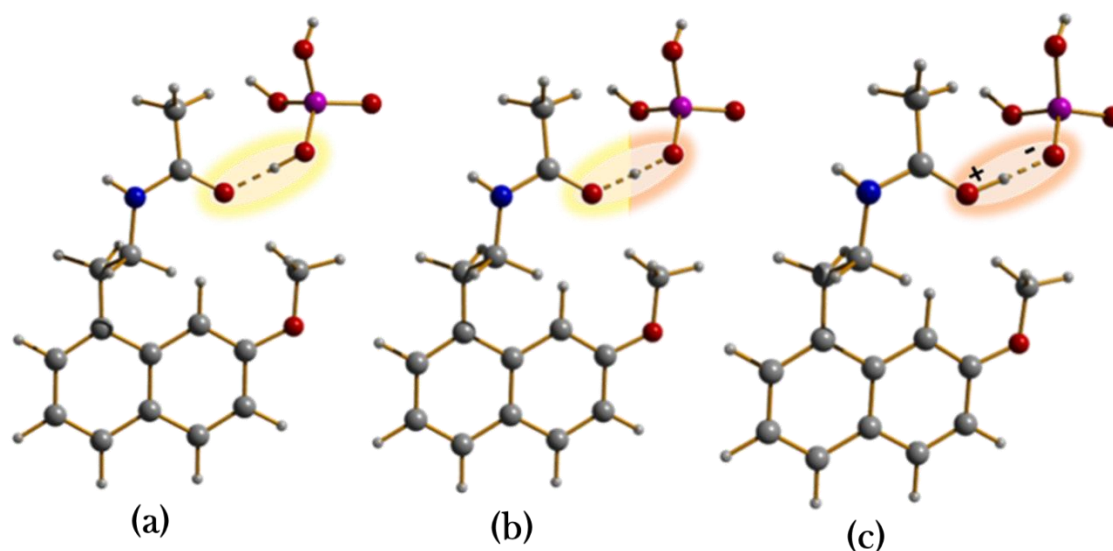
methoxy-1-naphthyl) ethyl] acetamide. Agomelatine phosphoric acid best fits with such challenges for which growing single crystal is a challenge and it poses lot of ambiguity to distinguish whether it is salt or cocrystal by just considering  $\Delta pK_a$  rule<sup>4</sup>. For agomelatine, 6 polymorphic forms of API, 13 cocrystals and 9 salts have been reported so far (**Table 3.1**). These numbers under discrete classes indicate that the acetamide functional group acts as potential hydrogen bonding functional group and molecule have adequate structural flexibility to exhibit polymorphism exhibiting wide landscape. A detailed literature review on reported agomelatine adducts with respect to their corresponding  $\Delta pK_a$  are tabulated in **Table 3.1**<sup>5–14</sup>.

The following observations are summarized from the **Table 3.1**:

- a) Abundance of cocrystals and salts is mostly equal.
- b) Success rate of  $\Delta pK_a$  rule is >90%.
- c) Crystal structural analysis in terms of a ratio of amide C-O and C-N bond distances also can be a useful indicator for deciding salt vs cocrystal.
- d) Agomelatine forms salts with inorganic acids (HI, HBr, H<sub>2</sub>SO<sub>4</sub> and H<sub>3</sub>PO<sub>4</sub>) and organic sulphated acids.
- e) Agomelatine benzenesulfonic acid and H<sub>3</sub>PO<sub>4</sub> entries are formed salt with negative  $\Delta pK_a$  values, which is contrary to  $\Delta pK_a$  rule.

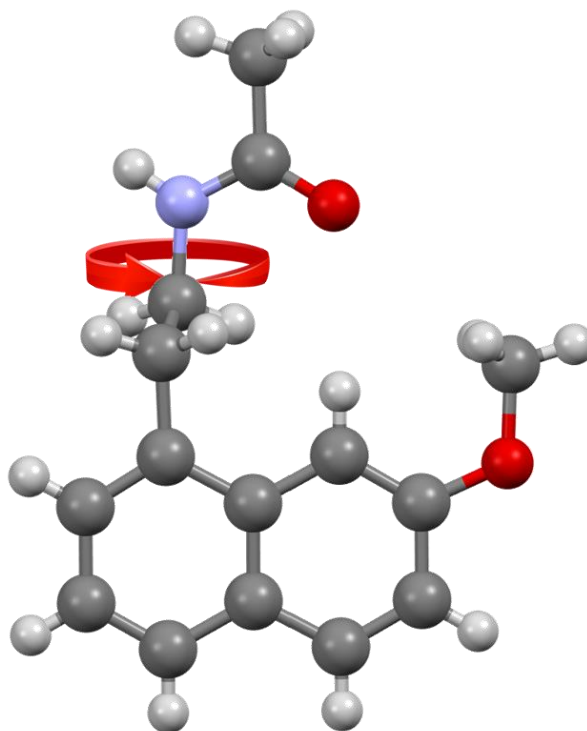
In this chapter, this anomaly of salt formation for the molecular complex of agomelatine and phosphoric acid is considered for probing. Process of agomelatine and phosphoric acid (AGL–P) process and its polymorphism was first patented by Dr. Reddy's Laboratories<sup>14</sup> and further solid-state characterization was published in an international conference<sup>15</sup>. Another conference paper by Eliška *et. al.* has reported the nature of AGL–P as salt, but crystal structure was not published<sup>13</sup>. In this chapter we aim to perform a detailed molecular and supramolecular level structural characterization on phase transition of AGL–P in view of 'salt-cocrystal continuum' (**Figure 3.2**) by combining complementary techniques and approaches (experimental and theoretical), namely DSC, VT-XRPD, hot stage optical microscopy (HSM), IR, SS–NMR, *ab initio* crystal structure determination from XRPD data (SDPD), and DFT-D2 calculations.





**Figure 3.2:** Schematic presentation of AGL–P in the form of (a) cocrystal (b) continuum and (c) salt of AGL–P via C=O...H–O–(P) interaction.

As discussed in **chapter 1**, growing single crystals and characterization of them becomes primary requirement in this scenario. Salt or cocrystal by means of proton transfer between the acid and base is a small thin line to differentiate, which needs advanced techniques and otherwise challenging to locate proton position.



**Figure 3.3:** Schematic presentation of AGL with torsion angle indicated. See details in Table 3.1.

**Table 3.1:** Crystal structure survey of agomelatine M-APIs (Multi-component Active Pharmaceutical Ingredients)

Sl. No.	Coformer	Reported (SXRDXRPD)	Calculated ( $\Delta pK_a$ )	Agomelatine	
				$d_{C-O}/d_{C-N}^{\#}$	Torsion angle (Figure 3.3)
1	Hydroquinone Form-I <sup>8</sup> *	Cocrystal SXRDXRPD	Cocrystal (-11.3)	0.95	95.3
				0.94	99.0
				0.93	93.1
				0.94	91.1
2	Hydroquinone Form - II <sup>8</sup>	Cocrystal XRPD	Cocrystal (-11.3)	0.97	86.6
3	Pyruvic acid <sup>8</sup>	Cocrystal SXRDXRPD	Cocrystal (-3.8)	0.94	84.4
4	Urea <sup>8</sup>	Cocrystal SXRDXRPD	Cocrystal (-1.6)	0.93	85.7
5	Oxalic acid i-1 <sup>8</sup>	Cocrystal SXRDXRPD	Cocrystal (-2.7)	0.96	89.8
6	Oxalic acid i-1 <sup>8</sup>			0.96	86.7
7	Acetic acid Form - I <sup>7</sup>	Cocrystal SXRDXRPD	Cocrystal (-6.2)	0.94	84.5
8	Acetic acid Form - II <sup>8</sup>	Cocrystal SXRDXRPD	Cocrystal (-6.2)	0.85	84.3
9	Ethylene glycol <sup>7</sup>	Cocrystal XRPD	Cocrystal (-16.6)	0.93	90.0
10	Glycolic acid <sup>12</sup>	Cocrystal SXRDXRPD	Cocrystal (-5.3)	0.93	80.3
11	Isonicotinamide <sup>12</sup>	Cocrystal SXRDXRPD	Cocrystal (-4.8)	0.92	88.7
12	Methyl 4-hydroxybenzoate <sup>12</sup>	Cocrystal SXRDXRPD	Cocrystal (-10)	0.92	88.9
13	Sulfuric acid <sup>10</sup>	Salt XRPD	Cocrystal/ Salt (1.4)	1.00	93.5

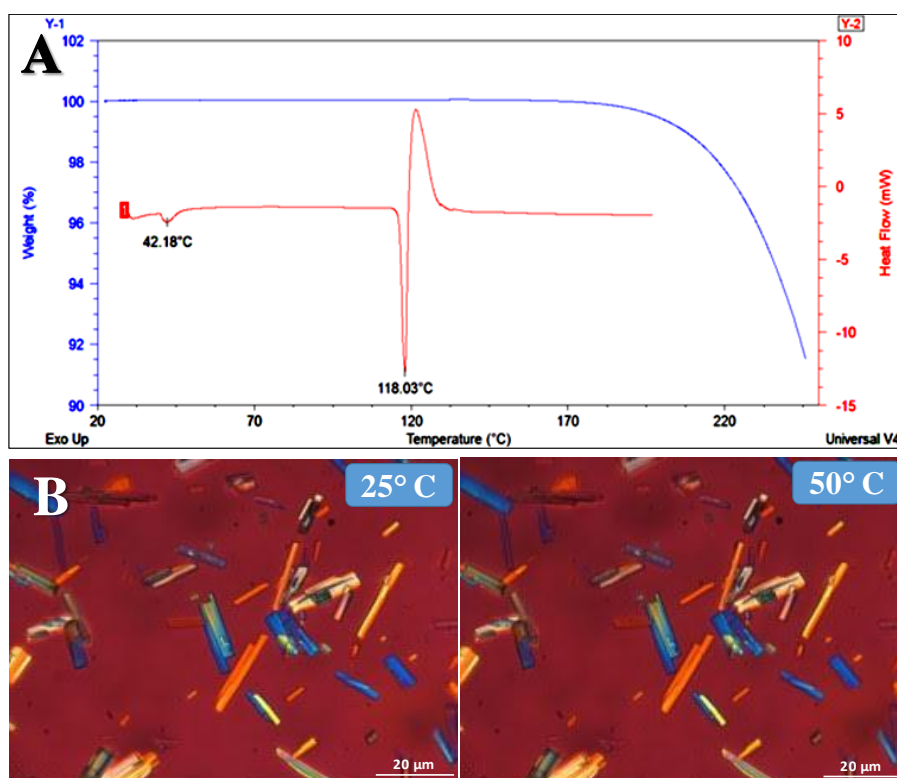
Sl. No.	Coformer	Reported (SXRDXRPD)	Calculated ( $\Delta pK_a$ )	Agomelatine		
				$d_{C-O}/d_{C-N}^{\#}$	Torsion angle (Figure 3.3)	
14	Methanesulfonic acid <sup>10</sup>	Salt SXRDX	Cocrystal (-0.3)	0.99	73.7	
15	Benzenesulfonic acid <sup>10</sup>	Salt SXRDX	Cocrystal (-2.2)	1.00	92.2	
16	Hydrogensulfate methanol solvate hemihydrate <sup>10</sup>	Salt SXRDX	Cocrystal/ Salt (1.4)	1.00	81.6	
					86.8	
17	mesylate monohydrate <sup>10</sup>	Salt SXRDX	Cocrystal/ Salt (-0.3)	Salt (0.98)	175.4	
18	HI (AgI) <sup>11</sup>	Salt SXRDX	Salt (8.4)	1.01	94.0	
19	HI (AgI3) <sup>11</sup>			0.97	81.4	
20	HI (AgI2) <sup>11*</sup>			0.96	88.6	
				0.96	90.2	
				0.97	91.8	
				0.96	94.0	
21	Phosphoric acid <sup>13</sup>	Salt SRXRDX	Cocrystal (-3.7)	Not available	Not available	
22	Resorcinol <sup>9</sup>	Cocrystal Not reported	Cocrystal (-10.7)	Not available	Not available	
23	AGL Form–I <sup>7</sup>	API SXRDX	NA	0.94	103.2	
				0.93	91.3	
				0.94	86.8	
				0.93	101.1	
24	AGL Form–II <sup>8</sup>		NA	NA	0.93	83.4
					0.94	96.7
25	AGL Form–III <sup>8</sup>		NA	NA	0.93	86.0
<b>Crystal Structures in This Study</b>						

Sl. No.	Coformer	Reported (SXR/DRP)	Calculated ( $\Delta pK_a$ )	Agomelatine	
				$d_{C-O}/d_{C-N}^{\#}$	Torsion angle (Figure 3.3)
26	AGL-P RT-Form (Rietveld-refined)	Cocrystal XRPD	Cocrystal (-3.7)	0.90	96.8
25	AGL-P RT-Form	Salt XRPD	Salt (-3.7)	0.99	93.7
27	AGL-P HT-Form ( Rietveld-refined)	Cocrystal XRPD	Cocrystal (-3.7)	0.90	79.7
28	AGL-P HT-Form	Cocrystal XRPD	Cocrystal (-3.7)	0.96	83.9

\*Z' = 4, # 1. As per isolated molecule: C-O/C-N criteria for cocrystal  $C-O/C-N \geq 0.82$  and for salt  $C-O/C-N \leq 1.14$ ; 2.  
As per CCDC: median ratio of the cocrystal (neutral amide) is 0.90 and salt (O-protonated amide) is 0.99

### 3.2. Polymorphism and thermal behaviour of AGL–P polymorphic forms

DSC thermogram of AGL–P is observed with two endothermic events (**Figure 3.4A**). First endothermic event at  $\sim 42^\circ\text{C}$  indicates polymorphic form conversion and  $\sim 118^\circ\text{C}$  indicates melting of the material. Observations from TGA analysis also support the same. Hot stage microscopy analysis of AGL–P indicates that with change in temperature (**Figure 3.4B**) has no impact on morphology, transparency and thermochromic behaviour<sup>16</sup>. However, slight changes in the crystal length were noticed. Therefore, observations from these complementary thermal techniques indicate that the endothermic event at  $\sim 42^\circ\text{C}$  suggests solid to solid-state structural phase transition (solid-state phase transition). This small endotherm is a clear indication of an enantiotropic solid-state phase transition. Such behaviour is not commonly observed in molecular crystals as they generally associated with relatively high energy barriers due to molecular arrangement (**Figure 3.6**)<sup>17</sup>.



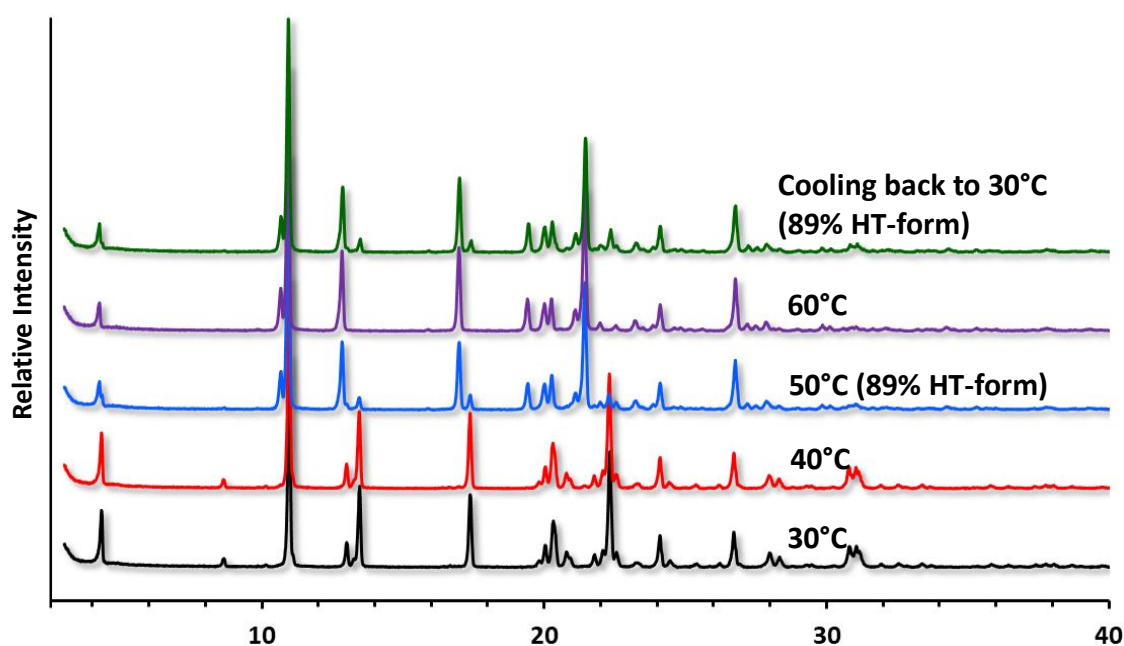
**Figure 3.4:** (A) Overlay of DSC and TGA thermograms of AGL-P with DSC curve showing two endothermic events at  $\sim 42^\circ\text{C}$  indicates phase transition and melting at  $\sim 118^\circ\text{C}$ . TGA indicates no weight loss. (B) HSM images of AGL-P at  $25^\circ\text{C}$  and  $50^\circ\text{C}$  representative of RT-form and HT-form, respectively.

Solid state phase transitions are mainly dictated by structural differences between initial form and polymorphic form after conversion. The mechanism of polymorphic form transformation is mainly known to happen in two ways.

1. Reconstructive transformation (nucleation and growth)
2. Diffusion less martensitic transformation (topotactic/epitactic)<sup>18,19</sup>

These mechanisms can be examined by means of HSM technique to understand phase transition mechanism<sup>20,21</sup> along with correlation between crystal packing and optical anisotropy. HSM study of nicotinamide-pimelic acid (1:1) polymorphic system<sup>22</sup> reported that the crystals became partially opaque upon phase transition. But, during phase transition in AGL–P crystals remain transparent and there is no change in birefringence which indicates retention of single crystal integrity. Reported co-crystal of 2,4-dichlorobenzoic acid and trimethylamine N- oxide (1:1)<sup>23</sup> shows such behaviour exhibiting order–disorder phase transition along with proton transfer<sup>23</sup>. This type of enantiotropic transition can be either reversible or irreversible in nature depending on the thermodynamic and kinetic interplay and activation energy barrier in between the polymorph interconversions<sup>24</sup>. To investigate further, thermal analysis (DSC) and in–situ variable temperature powder-XRD (VT-XRPD) analysis was performed.

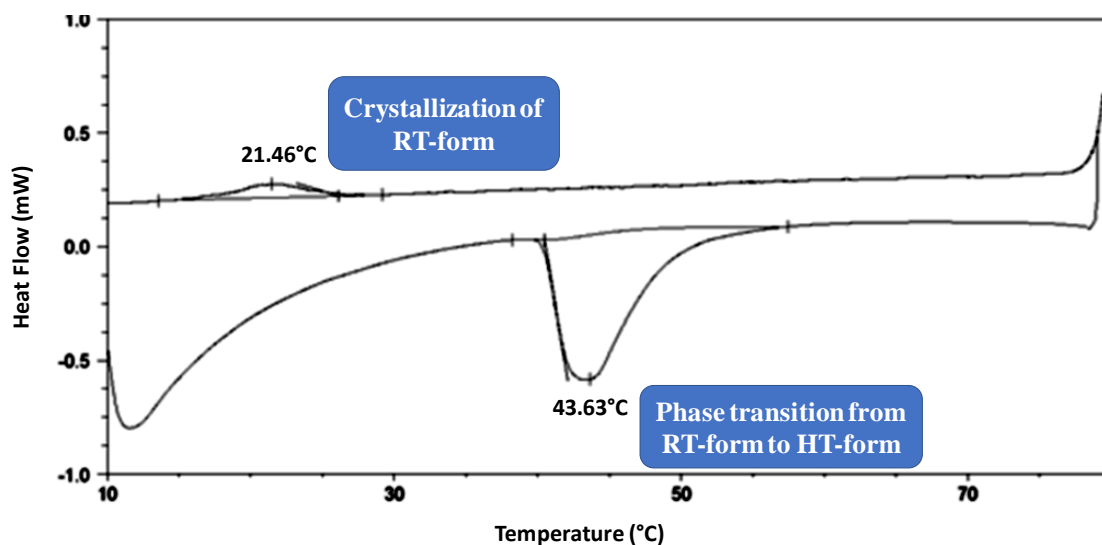
VT-XRPD studies were collected in the temperature range of 30°C–60°C at each 10°C increment with 10°C/minute ramp. No polymorphic form conversion was observed at 40°C. About 89% polymorphic form conversion from RT–form to HT–form was observed at 50°C (Figure 3.5 & 3.6; for corresponding Rietveld refinement profile fittings, refer Annexure 10 – 11).



**Figure 3.5:** Overlay of VT–XRPD data collected in the temperature range of 30°C to 60°C at each 10°C increment and at 30°C after cooling.

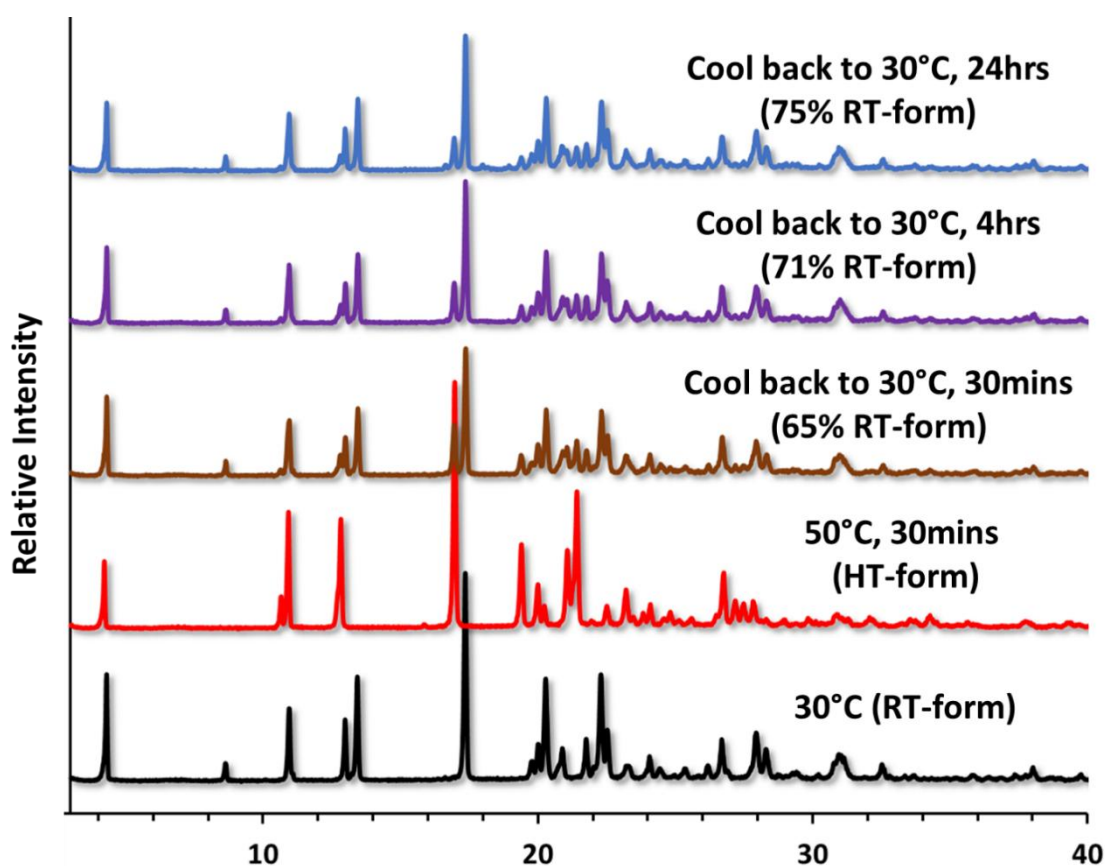
---

In DSC thermogram, endothermic event starts at  $\sim 40^{\circ}\text{C}$ , peak temperature at  $\sim 44^{\circ}\text{C}$  and it ends at  $\sim 60^{\circ}\text{C}$  (Figure 3.6).



**Figure 3.6:** DSC thermogram for heat–cool cycle of AGL–P with phase transition and recrystallization events are indicated.

The observations from VT-XRPD agree with DSC results (**Figure 3.6**) with a small difference between them due to the sample size differences used between these two techniques. However, if the sample was held at  $50^{\circ}\text{C}$  for 30 mins, complete polymorphic form conversion is observed. Upon cooling the sample from  $60^{\circ}\text{C}$  to  $30^{\circ}\text{C}$ , a small fraction of the sample conversion from RT-form to HT-form is observed. It indicates that AGL–P exhibits reversible enantiotropic phase transition behaviour. To understand further isothermal holding studies were conducted at different time intervals at 30 min, 4 h, and 24 h and 65%, 71% and 75% polymorphic form conversions were observed respectively (**Figure 3.7; for corresponding Rietveld refinement profile fittings, refer Annexure 12 – 14**). The trend followed by polymorphic form conversion from HT-form to RT-form resembles plateau like phase transition behaviour which indicates slow phase transition kinetics compared to transition from RT-form to HT-form. This behaviour can be readily understood from the DSC heat–cool thermograms, where a relatively large hysteresis indicating a large kinetic barrier for recrystallization event from HT–form to RT–form was observed at  $\sim 22^{\circ}\text{C}$ . This observation can be attributed to a diffusion less martensitic phase transformation due to a probably high degree of orientational relationship between the parent and the daughter phase<sup>18</sup>. However, factors such as material inhomogeneities and internal stresses in such a polycrystalline sample can lead to broad transformation temperatures<sup>25</sup>.



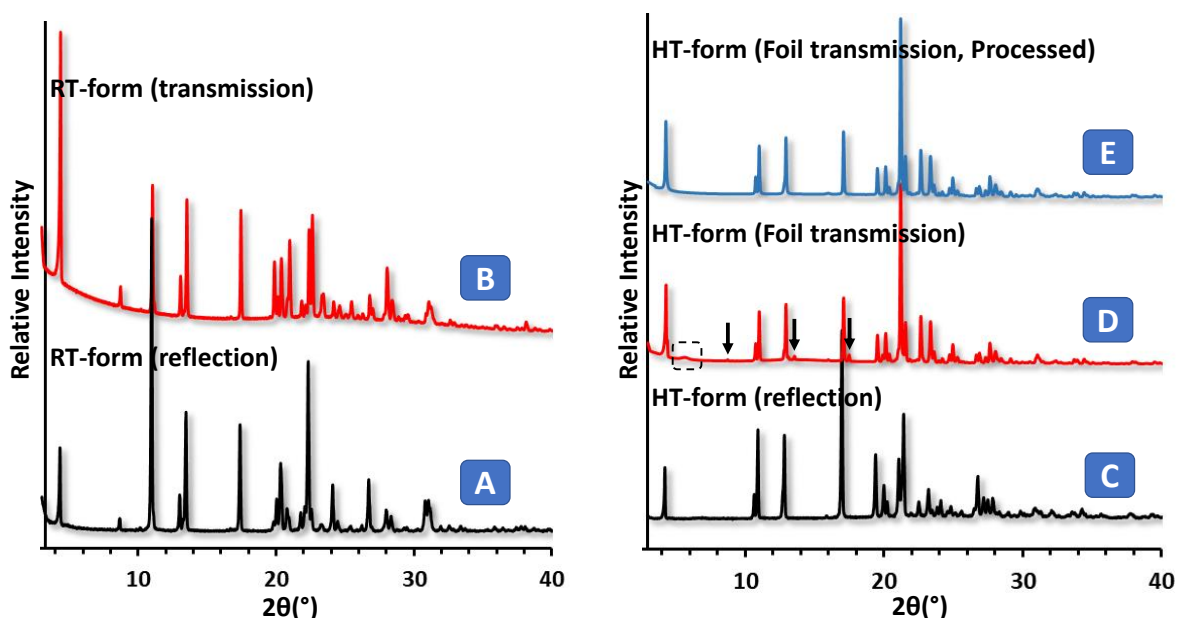
**Figure 3.7:** Overlay of VT–XRPD patterns of AGL–P RT–form and HT–form during heating and cooling monitoring.

### 3.3. Crystal structure determination: Salt vs co–crystal confirmation

Determination of crystal structures is often obligatory to understand whether the concerned multicomponent material is salt or cocrystal where proton position between acid and base is the critical event. Single crystal X–ray or neutron diffraction is the most definitive way to determine proton position<sup>26,27</sup>. Through the determination of crystal structure from powder–XRD may end up with less information comparatively from single crystal X–ray diffraction data due to the compression of three-dimensional data into one dimension. However, SDPD can be adequate for the characterization of cocrystal or salt<sup>28,29</sup>. Several attempts were made to grow single crystals of AGL–P, but all the trails were led to the polycrystalline materials. Therefore, crystal structures of both RT–form and HT–form of AGL–P were solved from powder–XRD. Preferred orientation effect in the powder–XRD data is another limiting factor during powder–XRD data collection. To overcome this issue, data collection was performed in transmission geometry for both RT–form (capillary) (**Figure 3.8 A - B**) and HT–form (foil) (**Figure 3.8 C - E**). Another challenge was data collection for HT–form due to unavailability



of nonambient transmission stage and the instability of HT–form. The behaviour of HT–form to transform into RT–form in slow manner have given scope for data collection of HT–form. Please note that peaks (marked with arrows) corresponding to the RT–form in HT–form data were trimmed. The broad beak peak indicated with dotted box corresponds to the background from foil used in foil transmission XRPD data collection.



**Figure 3.8:** Powder–XRD overlay (left) of RT–form collected in (A) transmission geometry and (B) reflection geometry. Powder-XRD overlay (right) of HT–form data collected in (C) reflection geometry, (D) transmission geometry and (E) processed data from transmission geometry.

**Table 3.2:** Lattice parameters of RT–form and RT–form.

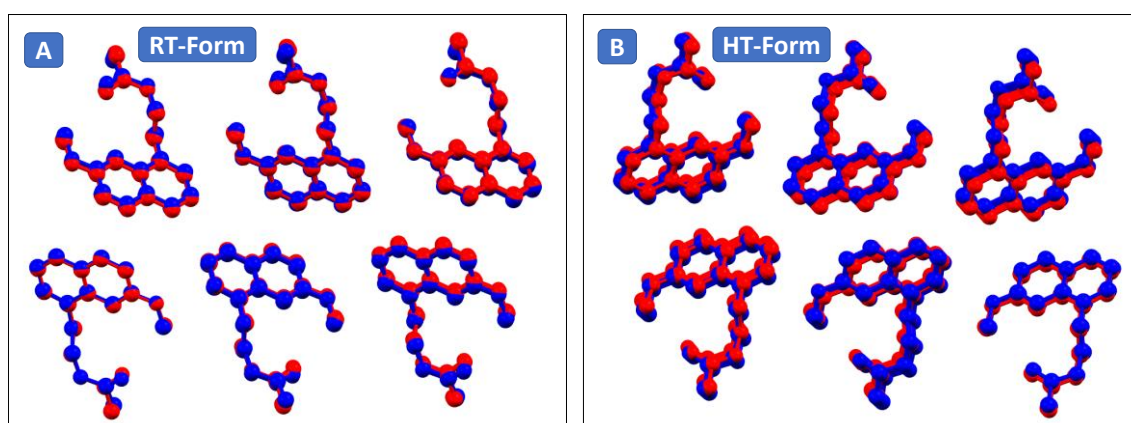
Lattice parameters	RT-form (Annexure 1)	HT-form (Annexure 2)
a, b, c (Å)	21.9914(6), 4.6552(2), 17.363(3)	22.304(4), 4.575(3), 17.710(7)
$\beta$ (°)	111.367(6)	111.476(9)
Volume (Å <sup>3</sup> )	1643.81(9)	1681.96(13)
Calculated density (g cm <sup>-3</sup> )	1.379	1.348
Z	4	4

The lattice parameters of RT and HT–forms are very close with same crystal system and space group (P2<sub>1</sub>/C). As per unit cell volume and space group, the most probable composition in the

asymmetric unit is one Agomelatine and one molecule of Phosphoric acid. Crystal structure of both forms were solved by direct space methods from powder-XRD data and refined using Rietveld refinement. The geometric optimization of crystal structure can improve the molecular bond lengths, bond angles and accuracy of supramolecular intermolecular interactions between the molecules within the crystal structure<sup>30</sup>. This may also help to establish the position of protons accurately which helps to form whether an adduct is a salt or cocrystal. As per Marcus A. Neumann, the accuracy of a crystal structure determined can be assessed by calculating RMSD (root mean-squared deviation) (i) between the experimental and optimized crystal structure. If  $\text{RMSD} < 0.25 \text{ \AA}^{30}$ , the determined crystal structure is accurate crystal structure. Along with RMSD, the similarity correlation between simulated patterns of refined and optimized crystal structure was used to finalize DFT–D2 optimized. The geometric optimization of crystal structure can be carried out in two ways:

- i. Fixed cell & relax atomic positions
- ii. Relax cell & relax atomic positions

Both options were attempted, and it was observed that fixed unit cell with atom position relaxed has given best results compared to the optimization procedure by relaxing unit cell and atom positions. RMSD between experimental and geometrically optimized crystal structures for RT and HT–forms were calculated as 0.2 and 0.51% respectively. Upon overlay of experimental and geometrically optimized crystal structures, it was observed that slight relatively higher displacement is observed for the HT–form which attributes to the higher RMSD value (**Figure 3.9 & Table 3.2**).



**Figure 3.9:** Overlay of RT and HT–forms crystal structures Rietveld–refined (red) and DFT–optimized (blue).

**Table 3.3:** Comparison between RT–form and HT–form from Rietveld refinement and DFT optimized with respect to crystal structure similarity and Amide group bond lengths and their ratio.

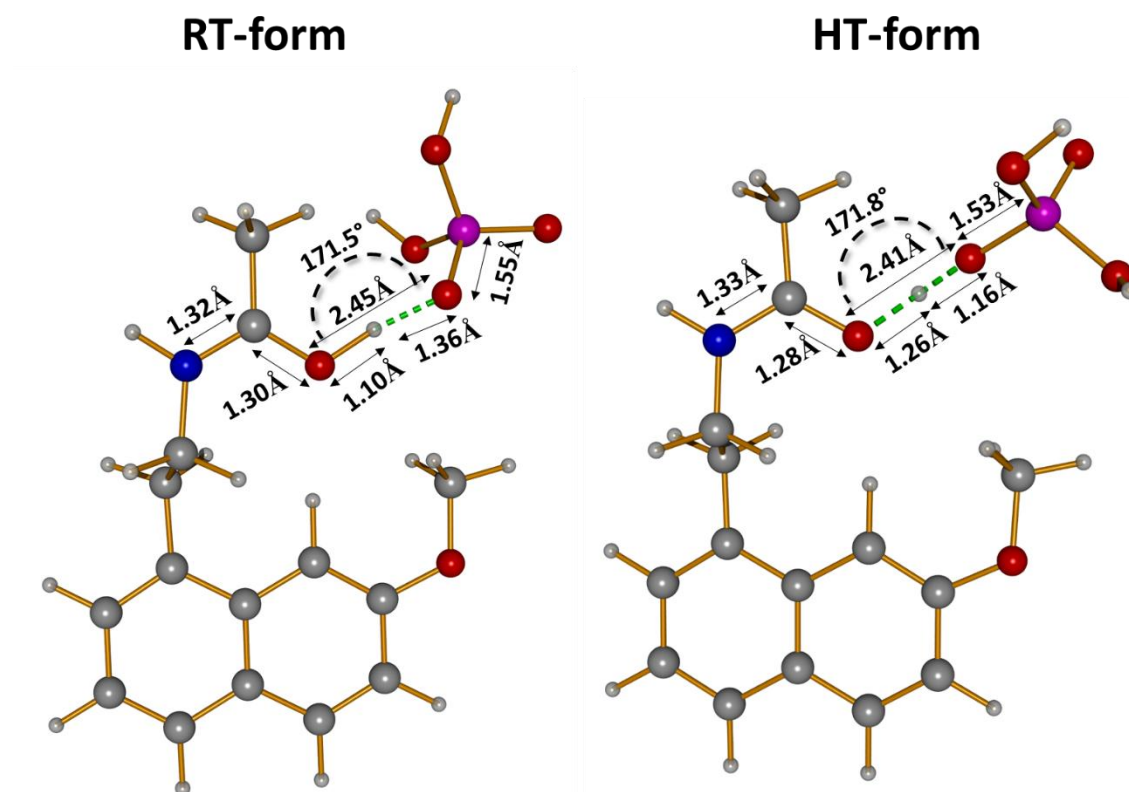
Crystal Structure Similarity			Amidic Group Bond Lengths & Their Ratio			
Crystal Structures		RMSD	Crystal Structure	C–O (Å)	C–N (Å)	C–O/C–N
RT–form (Rietveld-refined)	RT–form (DFT optimized)	0.20	RT–form (Rietveld-refined)	1.23	1.37	0.90
HT–form (Rietveld-refined)	HT–form (DFT optimized)	0.51	HT–form (Rietveld-refined)	1.23	1.37	0.90
RT–form (DFT optimized)	HT–form (DFT optimized)	1.35	RT–form	1.30	1.32	0.99
* Criteria for good structures RMSD < 0.25 Å ** Criteria for conformational polymorphs RMSD > 0.375 Å			HT–form	1.28	1.33	0.96
			AGL*	1.23	1.32	0.93
			* Average C–O/C–N of AGL (form I-III); as per CCDC median ratio of the neutral amide is 0.90 and O-protonated amide is 0.99 <sup>33</sup>			

The crystal structure similarity index between the RT and HT–form is higher than the RMSD criteria for conformational polymorphs (**Table 3.2**). This difference originates majorly from packing variance due to small changes in conformation<sup>31</sup>. Another difference contribution between RT–form and HT–form is slight difference due to resonance structure (O–protonated by the amide group functionality) and corresponding bond lengths. This kind of proton transfer can impact the C–O and C–N bond distances in the amide functional group. This phenomenon is not only observed in the AGL molecule<sup>10,11</sup>, but also in molecules with amide functional group<sup>5</sup>. Upon Rietveld refinement of RT and HT–forms, the C–O and C–N ratio was observed as 0.90 which is equivalent to the median value of a neutral amide group<sup>10</sup>. For AGL, C–O and C–N ratio was reported as 0.93. Both RT and HT–forms were subjected to DFT-D2 geometric optimization which has resulted in the proton transfer between hydroxyl group of phosphoric acid and amide group of agomelatine. It implies that the nature of agomelatine molecule in the RT–form is in ionised state i.e., O–protonated amide functional group.

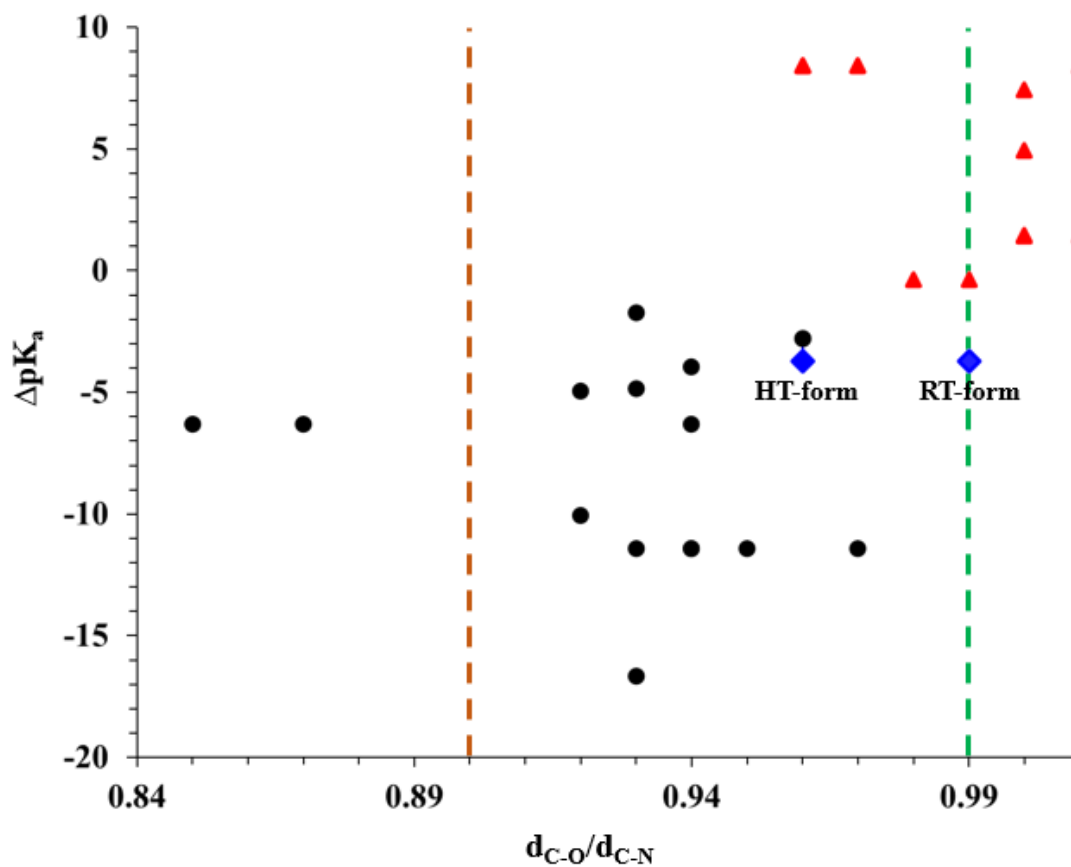
On the other hand, the resonance structure of amide functional group in HT–form has fallen between neutral amide and O–protonated amide functional group with C–O/C–N ratio as 0.96. This phenomenon of proton migration was also observed for adducts of pentachlorophenol–pyridine<sup>32</sup> and urea–phosphoric acid<sup>26,33</sup> characterized by in situ variable temperature neutron diffraction. Such structural changes due to proton migration also appeared in the phosphoric acid affecting the P–O bonds participated in the H–bonding interactions. Upon analysis of geometrically optimized crystal structures, decrease in P–O bond length is

observed from RT–form (1.55 Å) to HT–form (1.53 Å). Such phenomenon was also earlier reported in case of urea–phosphoric acid; decrease in P–O (1.50 Å to 1.46 Å) bond distance was observed upon increase in temperature and proton between phosphoric acid and urea was migrated and positioned between the O atoms<sup>26,33</sup>. But in the current study, proton position is close to phosphoric acid in the HT–form which indicates co–crystal nature.

This is a typical signature of resonance assisted Short Strong Hydrogen Bonds (SSHB)<sup>34</sup>. Theoretical studies suggest that such SSHB is the origin of the temperature–dependent proton migration phenomenon that can give rise to the high proton conductivity for phosphoric acid<sup>35</sup>. In general hydrogen bond with distance (O...O)  $\leq 2.55$  Å is considered to be a SSHB. This criteria can be used to classify salt vs co-crystal, considering salt having SSHB<sup>5</sup>. As per O...O distance, both RT and HT-forms have SSHB (**Figure 3.10**). Therefore, the ratio of C–O to C–N distances was considered for the verification of nature of crystal structures. Plot between C–O/C–N and  $\Delta pK_a$  was studied for all reported adducts with agomelatine and a clear distinction between salts and cocrystals can be observed (**Figure 3.11**). By including data points of RT–form and HT–form, it was observed that RT–form data falls in the pool of salts data points, whereas HT–form data point, interestingly falls in the intermediate area between salt and cocrystals. These observations indicate that RT–form is a salt and HT–form is a cocrystal or continuum.

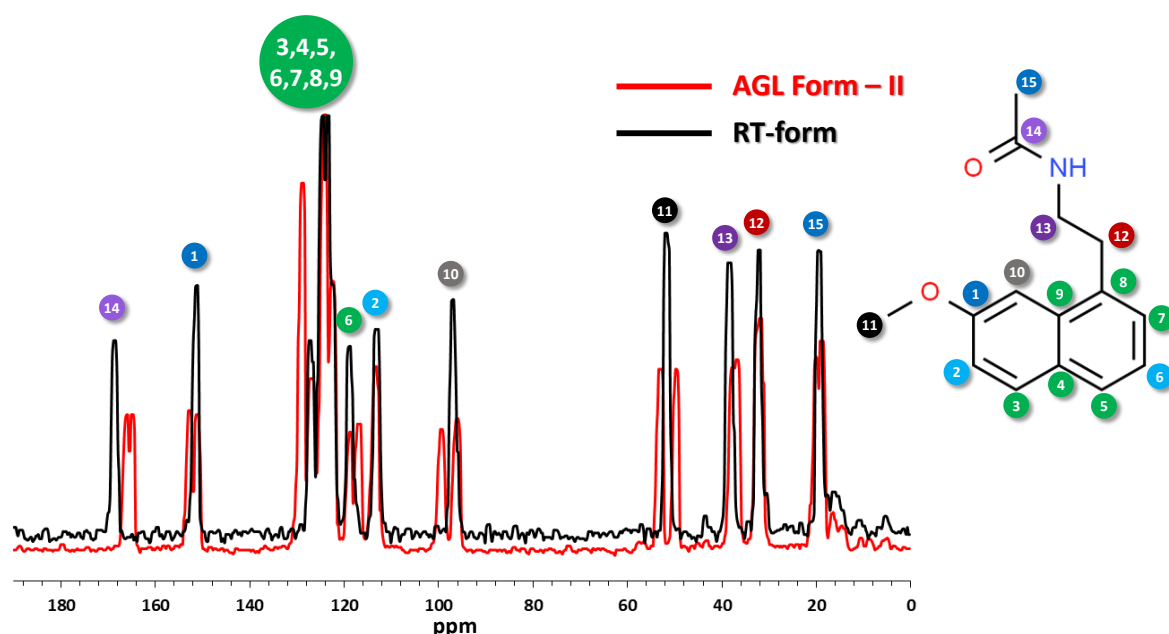


**Figure 3.10:** Representative molecular diagram showing proton shift between agomelatine and phosphoric molecules in RT and HT-forms and the associated bond distances and angle matrix.



**Figure 3.11:** Representative plot between  $\Delta pK_a$  and  $d_{C-O}/d_{C-N}$ . M-APIs of AGL reported as salts are indicated by red color triangles ( $\blacktriangle$ ) and M-APIs of AGL reports as cocrystal are indicated by black color circles ( $\bullet$ ). RT and HT-forms are marked by a blue diamond symbol ( $\blacklozenge$ ). For eye guidance, green and brown dotted lines represent the median for salt and cocrystal  $d_{C-O}/d_{C-N}$  ratio respectively.

IR and SS-NMR spectroscopic studies were carried out to derive supporting experimental evidence for ionization state of amide functional group of AGL in RT and HT - forms. In  $^{13}\text{C}$  SS-NMR, the amidic C atom of RT-form exhibits 3.3 ppm downfield shift compared to AGL form-II. In recent literature, SS-NMR  $^{13}\text{C}$  chemical shifts for AGL hydrogensulfate, mesylate, and besylate salts were reported with similar magnitude<sup>33</sup>. Due to the unavailability of variable temperature set up, the experimental chemical shift of HT-form is not obtainable. Therefore, to mitigate this gap, CASTEP module of Materials Studio Software was employed for the simulation of SS-NMR. Good agreement was achieved between experimental and simulated SS-NMR chemical shifts for AGL form-II and RT-form. Interestingly, the chemical shift value of HT-form falls between RT-form and AGL form-II which supports that HT-form is either co-crystal or continuum (**Figure 3.12 & Table 3.3**).

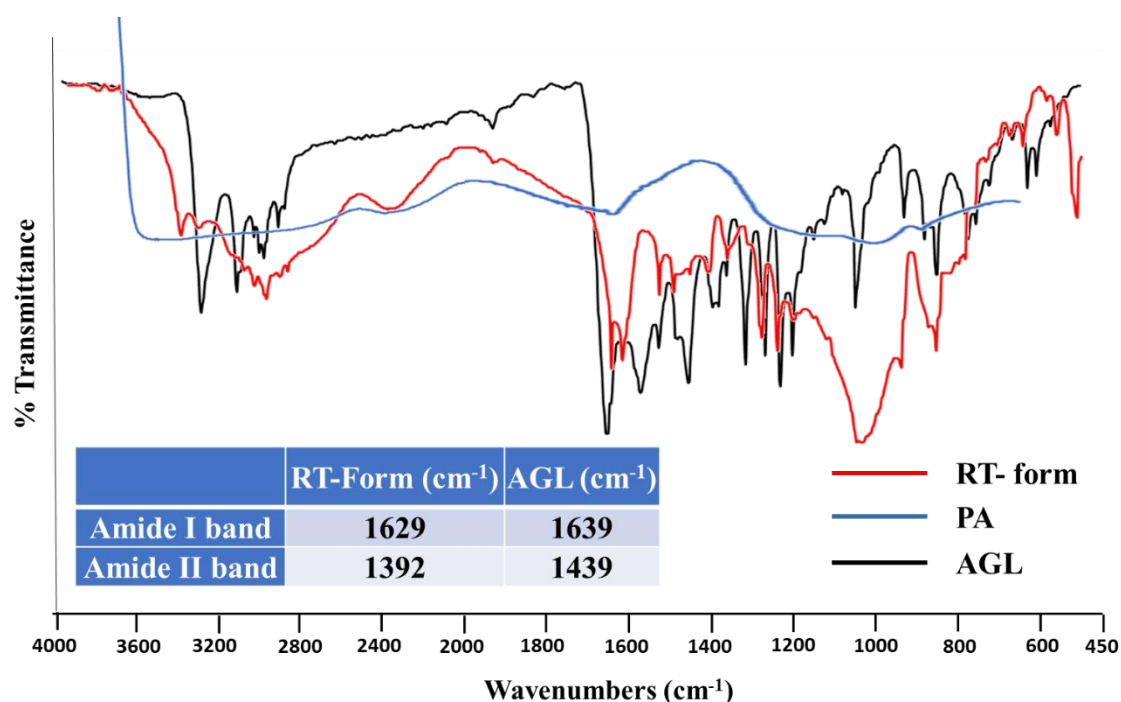


**Figure 3.12:** Overlay of solid-state  $^{13}\text{C}$  NMR spectrum of Agomelatine form-II and AGL-P RT-form.

**Table 3.4:** Chemical Shift of Amidic C in the Corresponding ss-NMR and Compared to AGL Form-II.

Crystal Form	Experimental		Calculated	
	Chemical shift $^{13}\text{C}$ [ppm]	$\Delta^{13}\text{C}$ (vs AGL) [ppm]	Chemical shift $^{13}\text{C}$ [ppm]	$\Delta^{13}\text{C}$ (vs AGL) [ppm]
AGL F-II	172.5	-	172.0	-
RT-form	175.7	-3.2	175.3	-3.3
HT-form	NA	NA	174.5	-2.5

The FT–IR spectra of AGL,  $\text{H}_3\text{PO}_4$ , and RT–form were compared and the data indicates that carbonyl frequency related to amide I band has red shifted by  $10\text{ cm}^{-1}$  in RT–form as compared to pure AGL. A similar observation is reported for dutasteride hydrochloride<sup>36</sup> and it is a clear attribute of amide O–protonation and subsequent decrease of the double bond character of the C=O bond. Further, the characteristic amide II band signature due to the change in resonance structure was observed for the RT–form<sup>37</sup> (Figure 3.13).



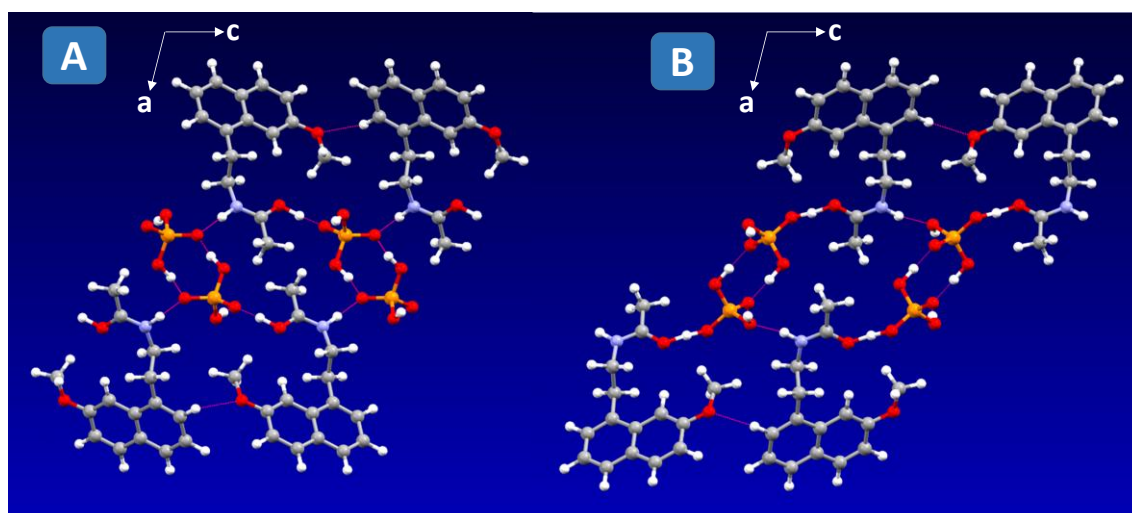
**Figure 3.13:** Overlay of experimental FT–IR spectrum of AGL, Phosphoric acid (PA), and RT–form. Wavenumbers for Amide–I and Amide–II band are listed in the table embedded in this figure.

#### 3.4. Crystal structure analysis: Mechanism of phase transition

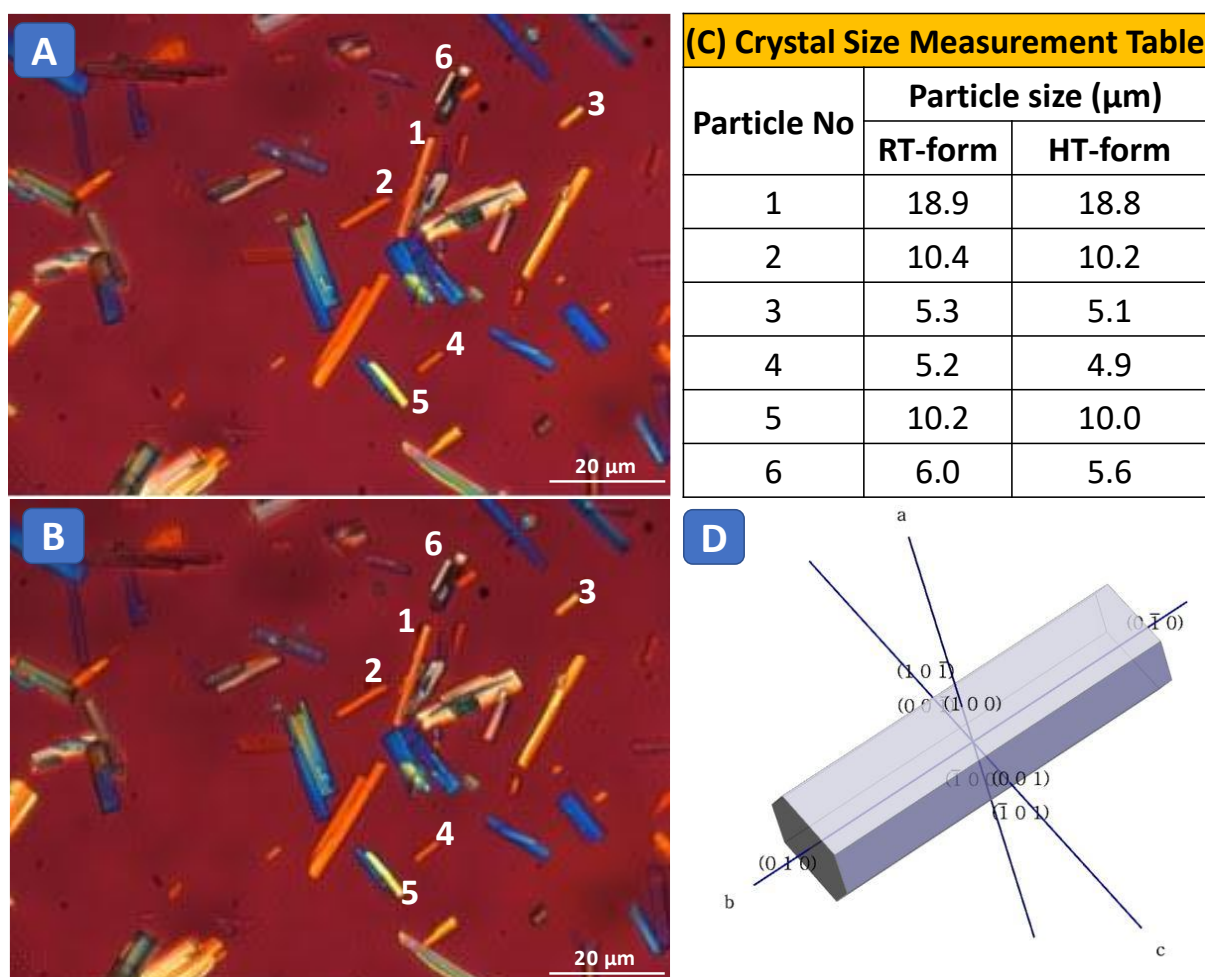
Anharmonic vibration of bonds and associated molecular rotations, motions, and translations of the individual molecules and their neighbours (cooperativity) are the triggering points of a

structural phase transformation. However, a successful solid–state phase transformation can occur in a concerted fashion between structures with a high degree of orientational relationship (structural closeness)<sup>18</sup>. Temporal and spatial averaged crystal structure obtained from X–ray diffraction studies are in general unable to yield molecular trajectories, which is essential for a molecular–level mechanistic understanding of phase transition<sup>38</sup>. Nevertheless, detailed crystal structure analysis to map the packing differences (precise molecular/atomic positions and their inter/intramolecular interactions) on both sides of the transitions can be an alternative way–out to elucidate the structural dynamics during the phase transition<sup>39–42</sup>. The building block of both the crystal forms can be considered as trilayer, which is facing head-to-head along the a–axis (**Figure 3.16**). The middle layer of this building block is polar, containing dimer of phosphoric acid interacting with the amide moiety of AGL. Strong hydrogen bonding interactions, N–H···O and O–H···O, are cementing the structural framework. The other two similar layers consist of hydrophobic naphthyl moiety of AGL and weak C–H···O hydrogen bonding interactions are observed (**Figure 3.14**). Crystal packing of both the forms can be visualized as stacking of trilayers along the c–axis in **Figure 3.16 A & B**. The length of the trilayer is 19.8 Å and 20.4 Å for RT–form and HT–form respectively. During the transition from RT–form to HT–form, there is an expansion of the breadth trilayer along the c–axis (RT–form: 8.8 Å to HT–form: 10 Å, phosphoric acid to phosphoric acid distance). This observation can be correlated with the increase in thermal expansion with decreasing interaction strength<sup>43</sup>. The HT–form contains relatively feeble SSHB O–H···O as compared to RT–form due to proton transfer. The N–H···O hydrogen bonds also follow a similar trend. Overall, anisotropic thermal expansion was witnessed during RT–form to HT–form transformation (along the a–axis and c–axis positive thermal expansion and negative thermal expansion along the b axis)<sup>44</sup>. This effect was also observed on crystal morphology, simulated crystal morphology and image analysis of HTM images confirmed the contraction along the b–axis (**Figure 3.16**). Such correspondence between molecular length scale (theory) and microscopic change in crystal size (experimental fact), provides a robust confirmation for our *ab initio* crystal structure elucidation protocol.



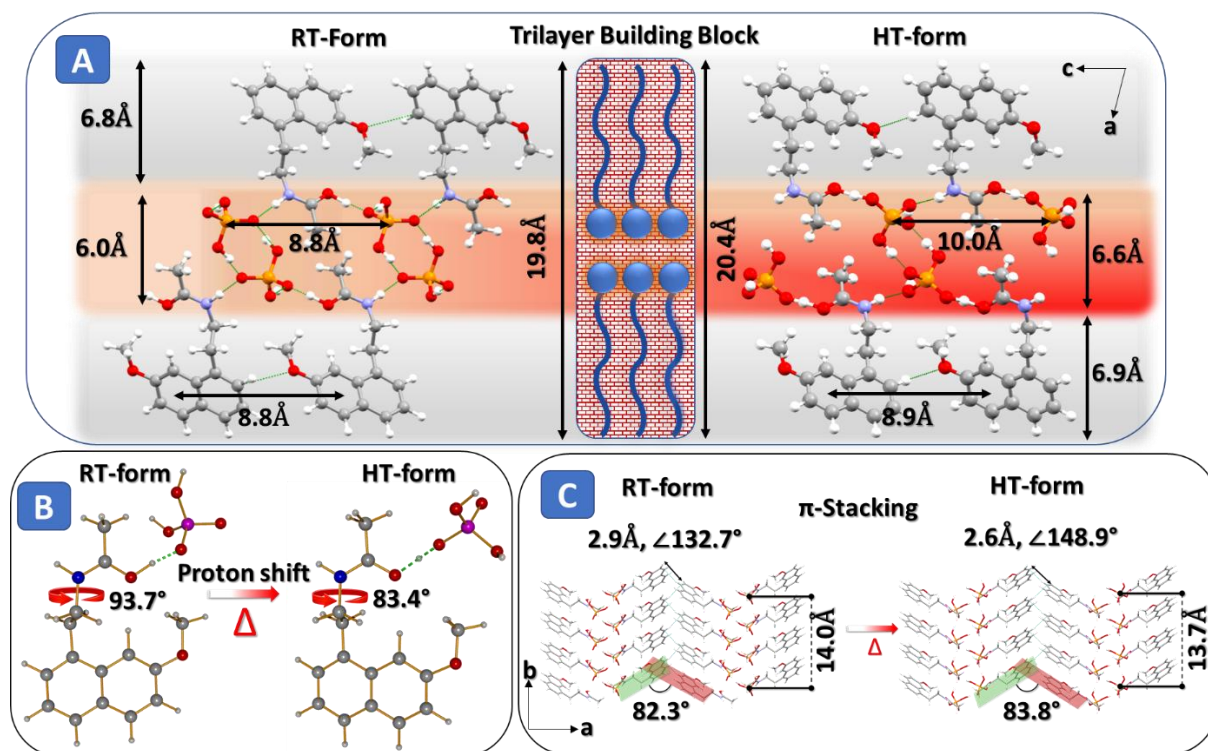


**Figure 3.14:** Crystal packing diagram showing possible hydrogen bonding interactions for RT-form (A) and HT-form (B).

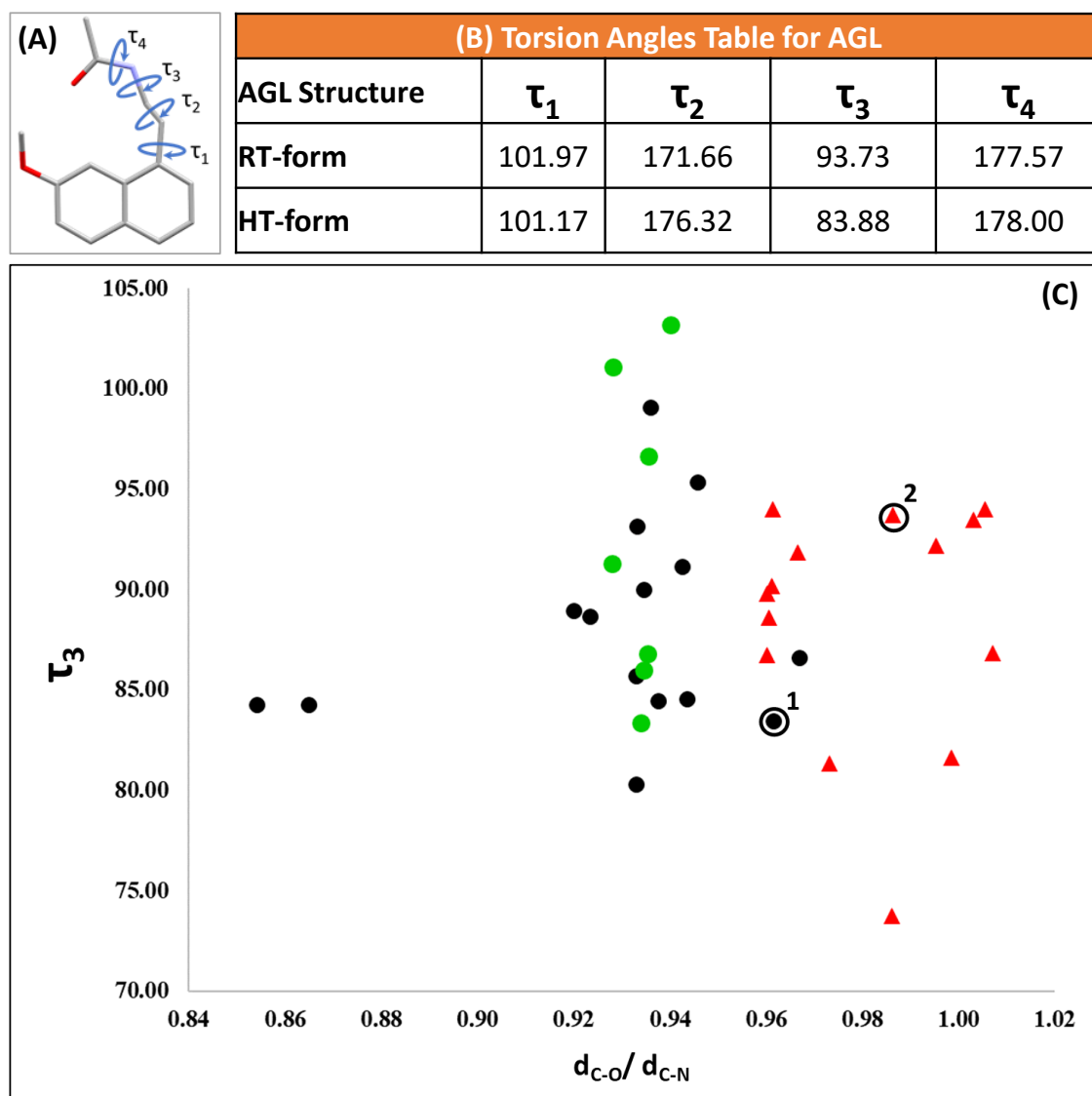


**Figure 3.15:** Image analysis of HSM pictures RT-form (A) and HT-form (B). (C) Table for crystal size measured along with long the major axis. (D) Simulated crystal morphology for RT and HT forms.

Upon comparison of predicted morphology with AGL–P crystal shape, the indexed miller facets were unambiguously assigned to the crystal morphology. According to the BFDH calculations, the growth rate and morphological importance for  $\{010\}$  is observed more compared to the  $\{100\}$  and  $\{001\}$  facets. Therefore, the major–axis is equivalent to b–axis. In general, the solid–state phase transition is a combined phenomenon, where molecular level conformational changes and supramolecular structural changes occur simultaneously in the crystal structure. Form this aspect, a conformational analysis was performed on all the four torsion angles of AGL. As shown in **Figure 3.17**, a maximum difference ( $\sim 10^\circ$ ) in molecular conformation in  $\tau_3$  (C–C–N–C) angle was noticed as compared to other torsion angles during the conversion from RT to HT–form. Further, to understand the correlation between conformational angles and degree of protonation of AGL, the ratio of C–O and C–N bond length ( $d_{C-O}/d_{C-N}$ ) is plotted against  $\tau_3$  (**Table 3.1** and **Figure 3.17**). This plot does not indicate any correlation between confirmation and protonation state. However, all these molecular level simultaneous changes are probably the triggering factors to overcome a local barrier to initiate the phase transition that finally creates a very minimal disruption in overall supramolecular structure (crystal packing). Expansion in length and breadth of trilayer for HT–form resulted in a slight increase in the naphthyl–naphthyl aromatic stacking angle between two consecutive trilayers (RT–form:  $\sim 82^\circ$  to HT–form:  $\sim 84^\circ$ ). This slight change in angle provided favorable aromatic–aromatic interaction (such as  $\pi \cdots \pi$  and C–H $\cdots\pi$ ) for HT–form and contraction of layers along the b–axis was noticed (**Figure 3.16**). In summary, it is more intricate than an order–disorder transition and this type of solid–state salt to cocrystal phase transition is rare<sup>45,46</sup>. The plausible underlying molecular mechanism of this martensitic–like phase transformation along the salt–cocrystal continuum could be the combination of the following concerted pathways: (i) Proton migration from O–protonated amide to phosphonyl group; (ii) Change the molecular structure of AGL (bond distance and torsion angle); (iii) Change in trilayer length and breadth and (iv) Change in aromatic stacking between the consecutive trilayers and so on. Such interesting structural changes were also earlier observed in the case of DL–Norleucine<sup>17</sup> and molecular dynamic simulation studies of this system revealed that the transformation occurs by concerted molecular displacements involving entire bilayers rather than on a molecule–by–molecule manner<sup>38</sup>.



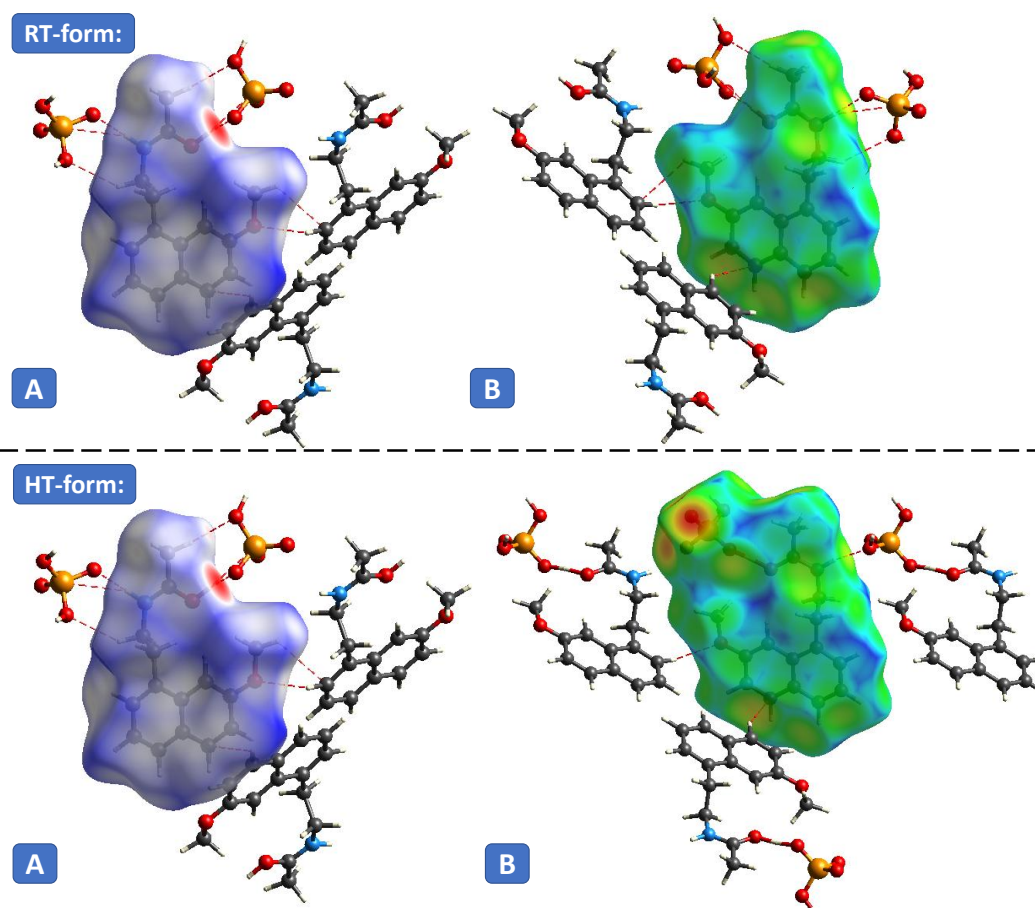
**Figure 3.16:** (A) Crystal structure showing trilayers building block of RT and HT forms with representative distance matrix. (B) Temperature-induced proton migration and change in torsion angle (C) Change in naphthyl–naphthyl stacking angle between consecutive trilayers and associated distances to support negative thermal expansion along the b-axis.



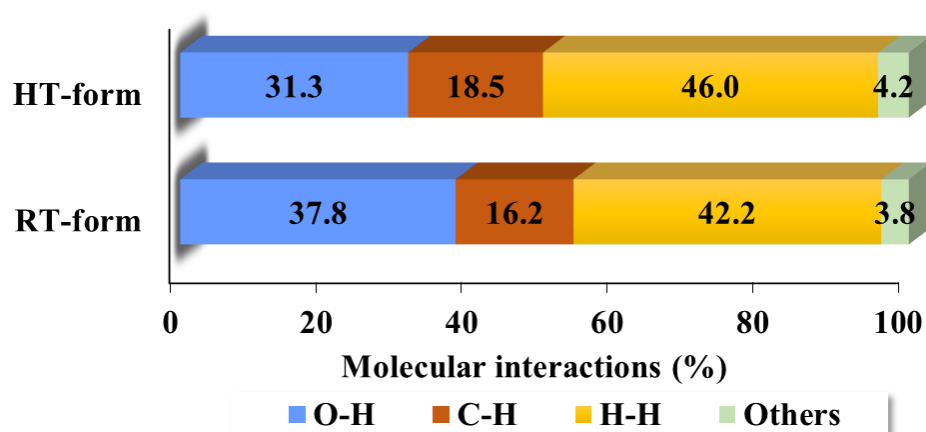
**Figure 3.17:** (A) Molecular structure of agomelatine showing four possible torsion angles. (B) All torsion angle values for RT and HT forms are tabulated. (C) Plot between  $\tau_3$  and  $d_{C-O}/d_{C-N}$  from the crystal structures reported with AGL M–APIs as one of its components. Circles (●) indicates co–crystal, red triangles (▲) indicate salt and green circles (●) indicates pure AGL polymorph. Data points corresponding to 1: HT–form and 2: RT–form are indicated.

Cognition of interaction preferences can guide to understand the thermodynamics and kinetics interplay in the polymorphic system. In some special cases, it can be observed that the less stable polymorph has better interactions while the more stable polymorph shows better packing<sup>47</sup>. From this context, quantification of interaction preferences was performed using the Hirshfeld surface analysis (Figure 3.18)<sup>48</sup> for RT and HT forms. It revealed that strong hydrogen bonding interactions are more pronounced for RT–form (37.8%) than HT–form (31.3%). Furthermore, the weak hydrogen bonding interactions and close packing are dominant

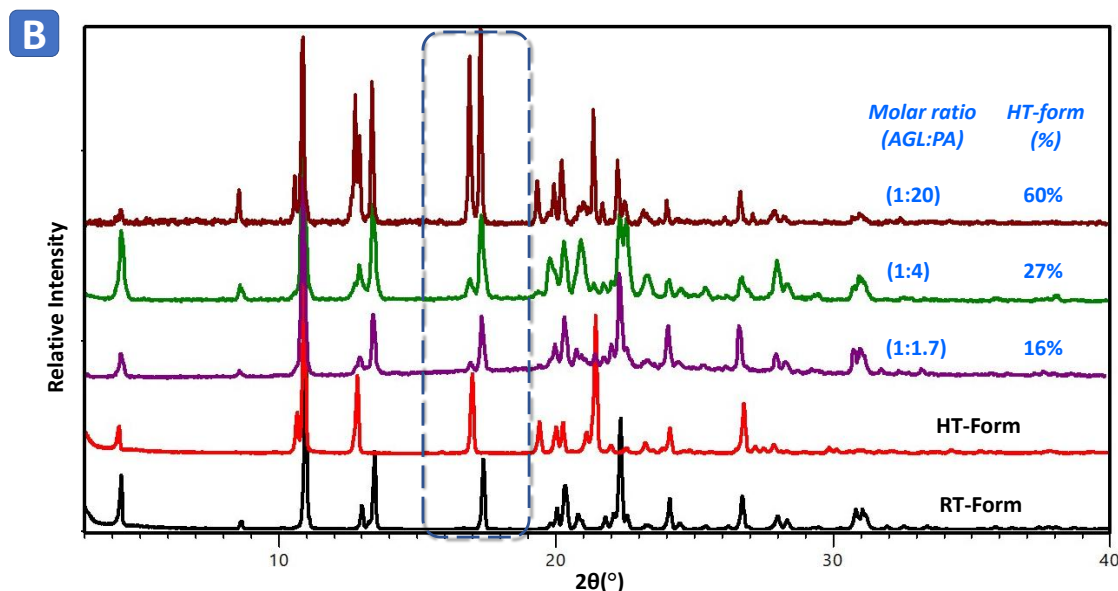
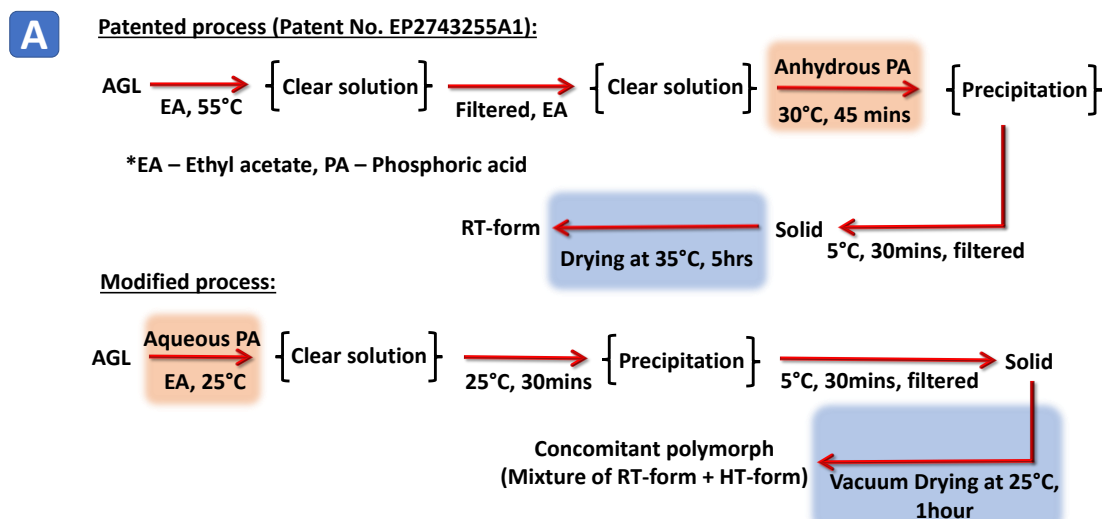
in HT–form (18.5 + 46.0%) as compared to RT–form (16.2 + 42.2%) (**Figure 3.19**). These data indicate that RT–form can be kinetically driven while HT–form may be controlled thermodynamically. However, as per lattice energy calculation, RT–form is relatively more stable by 4.5 kJ/mol than the HT–form, and this fact goes hand in hand with density rule (RT–form: 1.379 g cm<sup>-3</sup> and HT–form: 1.348 g cm<sup>-3</sup>)<sup>49</sup>. In fact, for this enantiotropic and reversible polymorphic system denoting thermodynamic or kinetic forms is not possible as they are kinetically switchable. Therefore, a more pertinent description would be: RT–form is more stable below transition temperature (42° C) and HT is more stable above transition temperature and the activation energy of RT–form↔HT–form kinetics is in the order of 30–40 kJ/mol<sup>15</sup>. The magnitude of activation energy for the AGL–P system is significantly less as compared to (1:1) nicotine–pimelic acid cocrystal<sup>50</sup> which was reported to exhibit a solid–state phase transition via nucleation and growth mechanism, unlike the current system. In summary, all this information refers to the point that the interplay of thermodynamic stability and kinetic activation barrier is very delicate for the AGL–P system. To verify this fact, slight tweaking of solution crystallization parameters such as supersaturation and solvent composition, which can influence the nucleation and growth of specific form(s) due to alternation of kinetic activation barrier, were conducted<sup>51,52</sup>. In all cases, RT and HT forms were obtained concomitantly (**Figure 3.20**). However, a correlation was observed between an increase in aqueous H<sub>3</sub>PO<sub>4</sub> concentration and HT–form enrichment, which can be attributed to the complex dissociation behaviour of phosphoric acid. These results indicate that there is a subtle interplay of thermodynamics and kinetics factors between the RT and HT forms. This system can pose a challenge to the current way of predicting the salt–cocrystal system mostly by the  $\Delta pK_a$  rule<sup>4</sup>, which is applicable only from the thermodynamic perspective. Kinetics interplay needs to amalgamate for enhancing the predictability of pharmaceutical salt–cocrystal systems<sup>52</sup>.



**Figure 3.18:** Comparison of Hirshfeld surface analysis and molecular interactions for RT and HT forms. (A) Surface is normalized with respect to internal and external interactions [dnorm with colour scale:  $-0.8$  (red) to  $1.4$  (blue)] (B) Surface with respect to external interactions only [de colour scale:  $0.5$  (red) to  $2.5$  (blue)].



**Figure 3.19:** Relative contributions (%) of intermolecular interactions to the Hirshfeld surface areas of the molecular constituents in RT and HT forms.



**Figure 3.20:** (A) Comparison of crystallization schematics for optimized patented process and modified current process. Differentiating steps were highlighted, namely  $\text{H}_3\text{PO}_4$  reagent grade (red) and drying step (blue). (B) Overlay of powder–XRD data of three experimental trials showing concomitant formation of RT and HT forms. However, a trend of increasing in HT–form content with  $\text{H}_3\text{PO}_4$  was noticed.

### 3.5. Conclusions

Crystal structures of the RT-form and HT-form were solved from powder–XRD data in combination with DFT–D2 analysis. This methodology, which can discriminate between the salt and cocrystal unambiguously, was also verified with the help of well–founded chemical crystallographic database analysis and spectroscopic studies. Detailed characterisation using a set of complementary techniques confirms that the nature of polymorphic form transformation

is a solid-state phase transition, exhibiting an enantiotropic and reversible relationship. Further, anisotropic thermal expansion and especially contraction of crystals along the b-axis were witnessed. This polymorphic transformation is more than an order–disorder transition where proton migration along the salt–cocrystal continuum and simultaneous changes in molecular structure are the triggering factors for martensitic-like (collective molecular layers displacement) transformation. Such unique solid-state salt to cocrystal phase transition is not well documented in the literature. This observation demands that phase transition stimuli such as temperature and humidity need to be considered while predicting and assigning any pharmaceutical multicomponent materials as a cocrystal or salt. Furthermore, these findings underscore the need for systematic evaluation of the dissociation behaviour of API and coformer in selective solvent systems for better process control to generate a salt or cocrystal form exclusively.

#### 3.6. References

1. Guillory, J. K. *Handbook of Pharmaceutical Salts: Properties, Selection, and Use. Journal of Medicinal Chemistry* vol. 46 (American Chemical Society, 2003).
2. Karimi-Jafari, M., Padrela, L., Walker, G. M., & Croker, D. M. Creating cocrystals: A review of pharmaceutical cocrystal preparation routes and applications. *Cryst. Growth Des.* **18**, 6370–6387 (2018).
3. Food and Drug Administration. Guidance for industry regulatory classification of pharmaceutical co-crystals. *Food Drug Adm.* 1–5 (2013).
4. Cruz-Cabeza, A. J. Acid–base crystalline complexes and the pKa rule. *CrystEngComm* **14**, 6362 (2012).
5. Perumalla, S. R. & Sun, C. C. Synthion preference in O-protonated amide crystals–dominance of short strong hydrogen bonds. *CrystEngComm* **15**, 8941–8946 (2013).
6. Tinant, B., Declercq, J.-P., Poupert, J. H., Yous, S., & Lesieur, D. N-[2-(7-Methoxy-1-naphthyl)ethyl]acetamide, a potent melatonin analog. *Acta Crystallogr. Sect. C Cryst. Struct. Commun.* **50**, 907–910 (1994).
7. Zheng, S. L., Chen, J. M., Zhang, W. X., & Lu, T. B. Structures of polymorphic agomelatine and its cocrystals with acetic acid and ethylene glycol. *Cryst. Growth Des.* **11**, 466–471 (2011).
8. Prohens, R., Barbas, R., Portell, A., Font-Bardia, M., Alcobé, X., & Puigjaner, C. Polymorphism of Cocrystals: The Promiscuous Behavior of Agomelatine. *Cryst. Growth Des.* **16**, 1063–1070 (2016).



9. Lee, M. J., Chun, N. H., Kim, H. C., Kim, M. J., Kim, P., Cho, M. Y., & Choi, G. J. Agomelatine co-crystals with resorcinol and hydroquinone: Preparation and characterization. *Korean J. Chem. Eng.* **35**, 984–993 (2018).
10. Skořepová, E., Bím, D., Hušák, M., Klimeš, J., Chatziadi, A., Ridvan, L., Boleslavská, T., Beránek, J., Šebek, P., & Rulišek, L. Increase in solubility of poorly-ionizable pharmaceuticals by salt formation: A case of agomelatine sulfonates. *Cryst. Growth Des.* **17**, 5283–5294 (2017).
11. Skořepová, E., Hušák, M., Ridvan, L., Tkadlecová, M., Havlíček, J., & Dušek, M. Iodine salts of the pharmaceutical compound agomelatine: The effect of the symmetric H-bond on amide protonation. *CrystEngComm* **18**, 4518–4529 (2016).
12. Yan, Y., Chen, J. M., Geng, N., & Lu, T. B. Improving the solubility of agomelatine via cocrystals. *Cryst. Growth Des.* **12**, 2226–2233 (2012).
13. Skořepová, E., Hušák, M., Ridvan, L., & Hanyková, T. Agomelatine phosphate: salt or co-crystal? *Acta Crystallogr. Sect. A Found. Adv.* **72**, s361–s361 (2016).
14. Kumar, V. S., Rao, B. V. K., Charyulu, K. S., Praveen, C., Reddy, Y. S., & Dey, A. Cocrystal of agomelatine with phosphoric acid, EP2743255A1. 1–18 (2014).
15. Voguri, R. S., Chappa, P., Chirla, R. R., & Dey, A. Molecular Complexes of Agomelatine-Phosphoric Acid: Crystal Structure Determination and Phase Transformation Kinetics By Non-Ambient. in *PPXRD-13 Presentations* (ICDD, 2015).
16. Jones, C. L., Skelton, J. M., Parker, S. C., Raithby, P. R., Walsh, A., Wilson, C. C., & Thomas, L. H. Living in the salt-cocrystal continuum: indecisive organic complexes with thermochromic behaviour. *CrystEngComm* **21**, 1626–1634 (2019).
17. Smets, M. M. H., Brugman, S. J. T., Van Eck, E. R. H., Van Den Ende, J. A., Meekes, H., & Cuppen, H. M. Understanding the solid-state phase transitions of dl-norleucine: An in situ DSC, microscopy, and solid-state NMR study. *Cryst. Growth Des.* **15**, 5157–5167 (2015).
18. Dunitz, J. D. Phase transitions in molecular crystals from a chemical viewpoint. *Pure Appl. Chem.* **63**, 177–185 (1991).
19. Van Den Ende, J. A., Smets, M. M. H., De Jong, D. T., Brugman, S. J. T., Ensing, B., Tinnemans, P. T., Meekes, H., & Cuppen, H. M. Do solid-to-solid polymorphic transitions in DL-norleucine proceed through nucleation? *Faraday Discuss.* **179**, 421–436 (2015).
20. Ge, C., Liu, J., Ye, X., Han, Q., Zhang, L., Cui, S., Guo, Q., Liu, G., Liu, Y., & Tao, X. Visualization of Single-Crystal-to-Single-Crystal Phase Transition of Luminescent

- Molecular Polymorphs. *J. Phys. Chem. C* **122**, 15744–15752 (2018).
21. Krishnan, B. P. & Sureshan, K. M. A spontaneous single-crystal-to-single-crystal polymorphic transition involving major packing changes. *J. Am. Chem. Soc.* **137**, 1692–1696 (2015).
  22. Aitipamula, S., Wong, A. B. H., Chow, P. S., & Tan, R. B. H. Polymorphism and phase transformations of a cocrystal of nicotinamide and pimelic acid. *CrystEngComm* **14**, 8193–8198 (2012).
  23. Order-disorder phase transition induced by proton transfer in a co-crystal of 2,4-dichlorobenzoic acid and trimethylamine: N -oxide. *CrystEngComm* **19**, 3753–3759 (2017).
  24. Kawakami, K. Reversibility of enantiotropically related polymorphic transformations from a practical viewpoint: Thermal analysis of kinetically reversible/irreversible polymorphic transformations. *J. Pharm. Sci.* **96**, 982–989 (2007).
  25. Bhattacharyya, A. & Lagoudas, D. C. A stochastic thermodynamic model for the gradual thermal transformation of SMA polycrystals. *Smart Mater. Struct.* **6**, 235–250 (1997).
  26. Parkin, A., Harte, S. M., Goeta, A. E., & Wilson, C. C. Imaging proton migration from X-rays and neutrons. *New J. Chem.* **28**, 718–721 (2004).
  27. Pindelska, E., Sokal, A., & Kolodziejski, W. Pharmaceutical cocrystals, salts and polymorphs: Advanced characterization techniques. *Adv. Drug Deliv. Rev.* **117**, 111–146 (2017).
  28. Tremayne, M. The impact of powder diffraction on the structural characterization of organic crystalline materials. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **362**, 2691–2707 (2004).
  29. Lapidus, S. H., Stephens, P. W., Arora, K. K., Shattock, T. R., & Zaworotko, M. J. A comparison of cocrystal structure solutions from powder and single crystal techniques. *Cryst. Growth Des.* **10**, 4630–4637 (2010).
  30. Van De Streek, J. & Neumann, M. A. Validation of molecular crystal structures from powder diffraction data with dispersion-corrected density functional theory (DFT-D). *Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater.* **70**, 1020–1032 (2014).
  31. Cruz-Cabeza, A. J. & Bernstein, J. Conformational polymorphism. *Chem. Rev.* **114**, 2170–2191 (2014).
  32. Steiner, T., Majerz, I., & Wilson, C. C. First O-H-N hydrogen bond with a centered proton obtained by thermally induced proton migration. *Angew. Chemie - Int. Ed.* **40**,
-

- 2651–2654 (2001).
33. Wilson, C. C., Shankland, K., & Shankland, N. Single-crystal neutron diffraction of urea-phosphoric acid: Evidence for H-atom migration in a short hydrogen bond between 150 K and 350 K. *Zeitschrift fur Krist.* **216**, 303–306 (2001).
  34. Morrison, C. A., Siddick, M. M., Camp, P. J., & Wilson, C. C. Toward understanding mobile proton behavior from first principles calculation: The short hydrogen bond in crystalline urea-phosphoric acid. *J. Am. Chem. Soc.* **127**, 4042–4048 (2005).
  35. Vilčiauskas, L., Tuckerman, M. E., Bester, G., Paddison, S. J., & Kreuer, K. D. The mechanism of proton conduction in phosphoric acid. *Nat. Chem.* **4**, 461–466 (2012).
  36. Nanubolu, J. B., Sridhar, B., & Ravikumar, K. Resonance-assisted amide protonation in dutasteride hydrochloride salt. *CrystEngComm* **14**, 2571–2578 (2012).
  37. Ji, Y., Yang, X., Ji, Z., Zhu, L., Ma, N., Chen, D., Jia, X., Tang, J., & Cao, Y. DFT-Calculated IR Spectrum Amide I, II, and III Band Contributions of N-Methylacetamide Fine Components. *ACS Omega* **15**, 8572–8578 (2020).
  38. Zahn, D. & Anwar, J. Collective displacements in a molecular crystal polymorphic transformation. *RSC Adv.* **3**, 12810–12815 (2013).
  39. Taniguchi, T., Sato, H., Hagiwara, Y., Asahi, T., & Koshima, H. Photo-triggered phase transition of a crystal. *Commun. Chem.* **2**, 1–10 (2019).
  40. Thakuria, R., Eddleston, M. D., Chow, E. H. H., Lloyd, G. O., Aldous, B. J., Krzyzaniak, J. F., Bond, A. D., & Jones, W. Use of In Situ Atomic Force Microscopy to Follow Phase Changes at Crystal Surfaces in Real Time. *Angew. Chemie Int. Ed.* **52**, 10541–10544 (2013).
  41. Banerjee, R., Bhatt, P. M., Kirchner, M. T., & Desiraju, G. R. Structural studies of the system Na(saccharinate)·n H<sub>2</sub>O: A model for crystallization. *Angew. Chemie - Int. Ed.* **44**, 2515–2520 (2005).
  42. Huang, Y.-G., Shiota, Y., Su, S.-Q., Wu, S.-Q., Yao, Z.-S., Li, G.-L., Kanegawa, S., Kang, S., Kamachi, T., Yoshizawa, K., Ariga, K., & Sato, O. Thermally Induced Intra-Carboxyl Proton Shuttle in a Molecular Rack-and-Pinion Cascade Achieving Macroscopic Crystal Deformation. *Angew. Chemie Int. Ed.* **55**, 14628–14632 (2016).
  43. Saha, B. K. Thermal Expansion in Organic Crystals. in *Journal of the Indian Institute of Science* vol. 97 177–191 (Springer International Publishing, 2017).
  44. Das, D., Jacobs, T., & Barbour, L. J. Exceptionally large positive and negative anisotropic thermal expansion of an organic crystalline material. *Nat. Mater.* **9**, 36–39 (2010).
-

45. Losev, E. A. & Boldyreva, E. Concomitant cocrystal and salt: No interconversion in the solid state. *Acta Crystallogr. Sect. C Struct. Chem.* **75**, 313–319 (2019).
46. Martins, D. M. S., Middlemiss, D. S., Pulham, C. R., Wilson, C. C., Weller, M. T., Henry, P. F., Shankland, N., Shankland, K., Marshall, W. G., Ibberson, R. M., Knight, K., Moggach, S., Brunelli, M., & Morrison, C. A. Temperature- and pressure-induced proton transfer in the 1:1 adduct formed between squaric acid and 4,4-bipyridine. *J. Am. Chem. Soc.* **131**, 3884–3893 (2009).
47. Dey, A. & Desiraju, G. R. Dimorphs of 4'-amino-4-hydroxy-2'-methylbiphenyl: Assessment of likelihood of polymorphism in flexible molecules. *CrystEngComm* **8**, 477–481 (2006).
48. Spackman, M. A. & McKinnon, J. J. Fingerprinting intermolecular interactions in molecular crystals. *CrystEngComm* **4**, 378–392 (2002).
49. Nelyubina, Y. V., Glukhov, I. V., Antipin, M. Y., & Lyssenko, K. A. 'Higher Density Does Not Mean Higher Stability' Mystery of Paracetamol Finally Unraveled. *Chem. Commun.* **46**, 3469–3471 (2010).
50. Lee, Y. J., Pahom, O., & Weeks, B. L. Kinetic Study for Comprehensive Understanding of Solid-State Polymorphic Transitions of Nicotinamide/Pimelic Acid Cocrystals. *Cryst. Growth Des.* **19**, 932–941 (2019).
51. Cao, Y., Du, S., Ke, X., Xu, S., Lan, Y., Zhang, T., Tang, W., Wang, J., & Gong, J. Interplay between Thermodynamics and Kinetics on Polymorphic Behavior of Vortioxetine Hydrobromide in Reactive Crystallization. *Org. Process Res. Dev.* **24**, 1233–1243 (2020).
52. Cruz-Cabeza, A. J., Lusi, M., Wheatcroft, H. P., & Bond, A. D. The role of solvation in proton transfer reactions: implications for predicting salt/co-crystal formation using the  $\Delta p K_a$  rule. *Faraday Discuss.* **235**, 446–466 (2022).

## Chapter 4

# Dapsone–Polyethylene glycol cocrystal and Virtual Screening of Drug–Polymer cocrystals

---

*This chapter covers the characterization of Dapsone–Polyethylene glycol cocrystal by combining experimental and theoretical methodologies and in-silico screening studies for Drug–polymer cocrystals.*

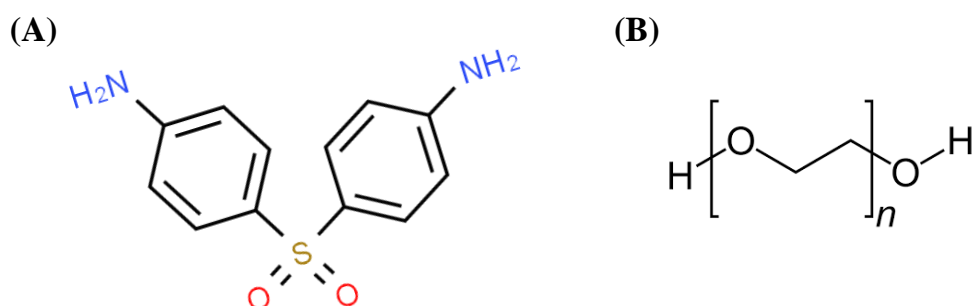
---

### 4.1. Introduction

The present work in this chapter has started with the characterization of dapsone–polyethylene glycol (DAP–PEG) cocystal (**Figure 4.1**), an interesting drug–polymer system where dapsone is drug and PEG is cofomer. Synthesis and demonstration of modulating its physiochemical properties through change in PEG molecular weights were published in the referenced paper<sup>1</sup>. Due to its peculiar crystal structure with a polymer chain as one of its components, it has posed challenges during its structural characterization. Hence, this has paved to incorporate computational methods such as DFT–D2 calculations for an effective crystal structure deduction<sup>2,3</sup>. However, with curiosity to discover more such drug–polymer complexes, an extended study aimed to discover more such polymeric cocystal systems with PEG as a cofomer. Cocystal screening for suitable cofomer is the most generally carried out studies by several conventional techniques such as mechanochemical grinding, solution evaporation, cooling crystallization, anti–solvent addition, and slurring. Also, high–throughput screening is an alternative approach that requires time and resources ending without any promise of the anticipated results. Hence, virtual screening procedures were explored to discover more such drug–PEG complexes to mitigate the expense of resources and time. This study was extended further to combine discrete potential computational methods effectively to calculate affinity between components (refers to drug molecules and polymers in the current context) to form co–crystals.

This chapter discusses primarily two topics:

- (i) Characterization of DAP–PEG cocystal, a drug–polymer cocystal complex characterized by combining experimental and theoretical methodologies.
- (ii) Virtual screening of drug–PEG cocystal complexes. Development of computational work methodologies to discover drug–polymer complexes.

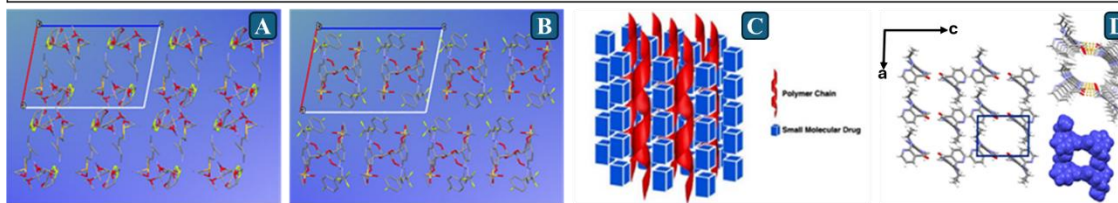


**Figure 4.1:** Molecular structures of (A) dapsone and (B) polyethylene glycol.

Speaking of DAP–PEG cocrystal characterization, dapsone (4–aminophenyl sulfone) is an antibiotic drug that belongs to the biopharmaceutical classification system class II (low solubility and high permeability)<sup>4</sup>. Three different polymorphic forms (Forms I, II and III)<sup>5,6</sup> and various cocrystals of this drug are known from the literature<sup>7–11</sup>. On the other hand, PEG is an FDA–approved and widely used polymer excipient/ligand in a variety of pharmaceutical and biopharmaceutical formulations such as oral, parenteral, topical, nasal and ocular applications<sup>12,13</sup> due to its superior properties such as congealing range, spreadability, consistency, penetrability and quick release of actives. Also, it does not support microbial growth<sup>14</sup>. Furthermore, PEG is known to form molecular complexes with small molecules, including urea, thiourea<sup>15</sup> and phenol derivatives<sup>16</sup> and few drug–PEG inclusion complexes/cocrystals were documented for drugs such as Mavacoxib,<sup>17</sup> Griseofulvin<sup>18</sup>, Diflunisal<sup>19–21</sup> and Nevirapine<sup>22</sup> showed polymer inclusion in the drug crystal channel almost no hydrogen bond interactions (**Table 4.1**).

**Table 4.1:** Crystal structure analysis of drug–polymer cocrystal/inclusion complexes available in the literature.

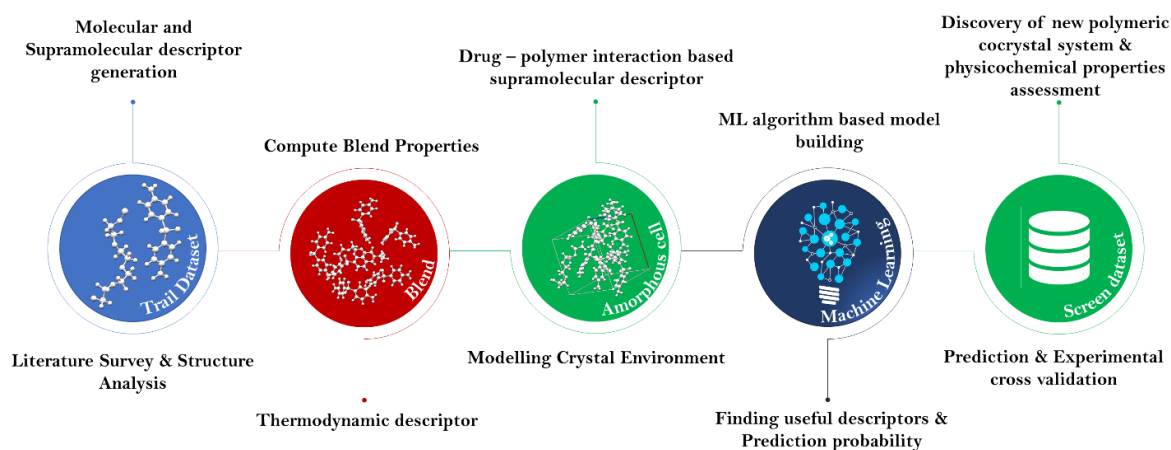
Drug	Polymer	Ref	Crystal Structure Analysis
Mavacoxib	PEG-600 & PEG-1000	Patent No: WO 2006/024930 A1	They are isostructural and were characterized by single crystal X-ray analysis. 3D coordinates are provided, but they incomplete/disordered structures. Therefore, difficult to comment on structural details (See Fig. A and Fig. B)
Griseofulvin	PEG-600	Chem Commun. 2014, 50, 6375	Cell indexed, no 3D coordinate available
Griseofulvin	PEG-1, 2, 4, 6, 20, 35, 100, 1000, 5000K	Int. J. Pharm. 2016, 508, 51	No 3D coordinates, but structure proposed as inclusion complex (See Fig. C)
Diflunisal	PTHF (n = 40), PCL (n = 88), P4HB (n = 47) and PBA (n = 60)	Cryst. Growth Des. 2016, 16, 1181	No 3D coordinates, but structure proposed as inclusion complex (See Fig. C)
Diflunisal	Poly( $\epsilon$ -caprolactone)	Cryst. Growth Des. 2017, 17, 355	No 3D coordinates, but structure proposed as inclusion complex (See Fig. C)
Nevirapine	Polycaprolactone	International journal of pharmaceutics. 2018 May 30; 543 (1-2): 121-9.	Structure proposed as inclusion complex (Fig. D). The polymer is disordered guest, not modelled.



It is well demonstrated in work published by Thiago Caon<sup>23</sup> that cocrystallization of dapsone with PEG (having a unique hydrophilic/hydrophobic nature)<sup>24</sup> may lead to a modulation in the permeability and work as a vehicle to deliver the drug<sup>25</sup>. As mentioned, this chapter was also extended to virtual cocrystal screening of drug–PEG cocrystals.

Thriving computing technologies worldwide have given scope to conduct extensive research in developing several virtual screening methodologies. These virtual methodologies already implemented by research people around are generally carried based on either knowledge–based (*molecular complementary and hydrogen bonding propensity*) or data–driven (*artificial intelligence and machine learning algorithms and Cambridge structural database*) or a

combination of both to carry *in-silico* coformer screening studies to predict the best coformer from a huge library of various molecules of interest in a short period. In the current work, screening studies were carried out contrary to usual, i.e., to identify the suitable API molecule to form cocrystal with PEG as coformer fixed. The accuracy of such methodologies depends on nature and strength of intermolecular interactions between drug and polymers. Predictability of such methodology entirely depends on the nature and strength of drug–polymer interactions, which also depends on miscibility or/and mobility of the molecules within the polymer matrix. Forcefield molecular simulations (amorphous cell and Flory Huggins) are proven effective in calculating dispersion and interaction parameters between the components in the solid dispersions, where various types of non-covalent molecular interactions need to be considered for further successful prediction. Therefore, our approach has been instrumental on miscibility using Blend modules in BIOVIA Materials Studio and further fueled by crystal engineering approach like drug–polymer molecular complementary and hydrogen bonding propensity<sup>26</sup>. However, such theoretical process and establishing complex correlations is time-consuming and computationally heavy. Machine learning algorithms are proven to establish the complex correlations between calculated structural, physical and chemical properties. In the present work, a tailor-made approach was taken which is combination of molecular simulation and machine learning algorithms-based predictive modelling that has been cross validated with respective to the outcomes of cocrystallization experiments. The overall framework is presented in **Figure 4.2**.



**Figure 4.2:** Schematic presentation of working framework for Virtual Screening of polymeric cocrystals.

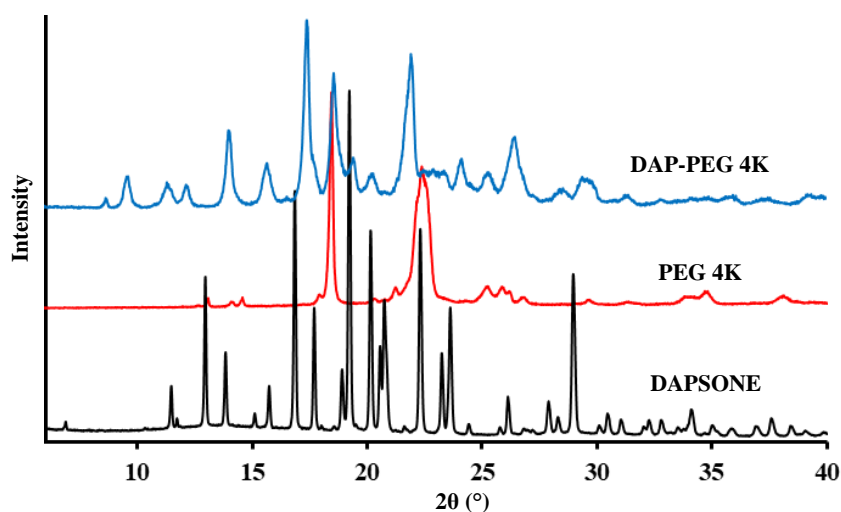
### 4.2. Characterization of DAP–PEG complex



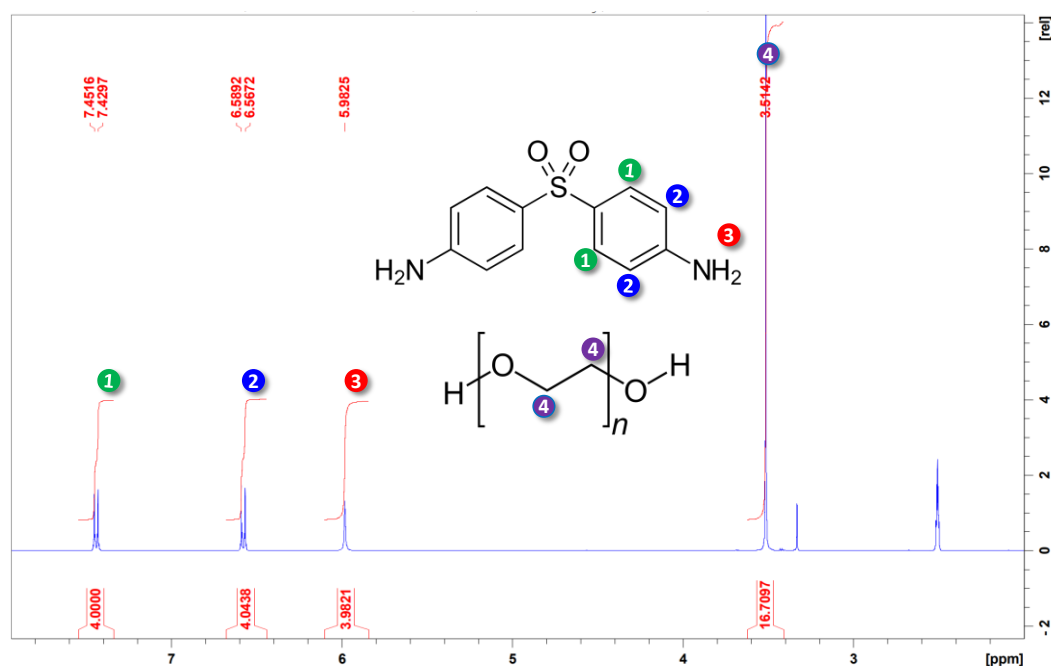
Molecular insight into intermolecular interactions, often drawn from the X-ray crystal structure analysis,<sup>27</sup> is a primary prerequisite for property-based pharmaceutical cocrystal design.<sup>28</sup> In this context, crystal structure determination is obligatory to correlate structure-property relationships. Multiple efforts towards growing single crystal of above cocrystals from different solvents using various methods (slow solvent evaporation, vapor diffusion, cooling crystallization, melt crystallization etc.) always lead to the formation of polycrystalline materials. However, the extent of better crystallinity i.e. large number of well-resolved peaks for DAP-PEG600 cocrystal gave us the confidence to solve the crystal structure from X-ray powder diffraction data.

### 4.2.1. Synthesis and Stoichiometry determination for DAP-PEG cocrystal

The crystallization process for single crystal growth as discussed in the thesis by Praveen Chappa from SRM University and the corresponding work was published<sup>29</sup>. The stoichiometry between dapsone and PEG in the cocrystal was determined by XRPD, DSC and <sup>1</sup>H NMR techniques. The XRPD and DSC data were reported in the above-referred thesis and research paper<sup>1</sup> (**Figure 4.3**). As a complementary, stoichiometry between DAP and PEG in the cocrystal was confirmed by <sup>1</sup>H NMR analysis and furnished in this thesis. In the <sup>1</sup>H NMR spectra of DAP-PEG complex, peaks corresponding to protons on the phenyl ring at ortho and meta positions were indicated with 1 and 2; Protons on amine groups were marked as 3 and protons in PEG monomer were indicated as 4. Upon normalizing integrated peak area corresponding to protons on dapsone at 1, 2 and 3 positions to 4 protons, area of the peak corresponding to protons on PEG (indicated as 4) was normalized to ~16 indicating stoichiometric ratio between dapsone and PEG in the asymmetric unit as 1:4 (**Figure 4.4**).



**Figure 4.3:** Powder-XRD overlay of (A) Dapsone, (B) PEG 4K and (C) DAP-PEG 4K.



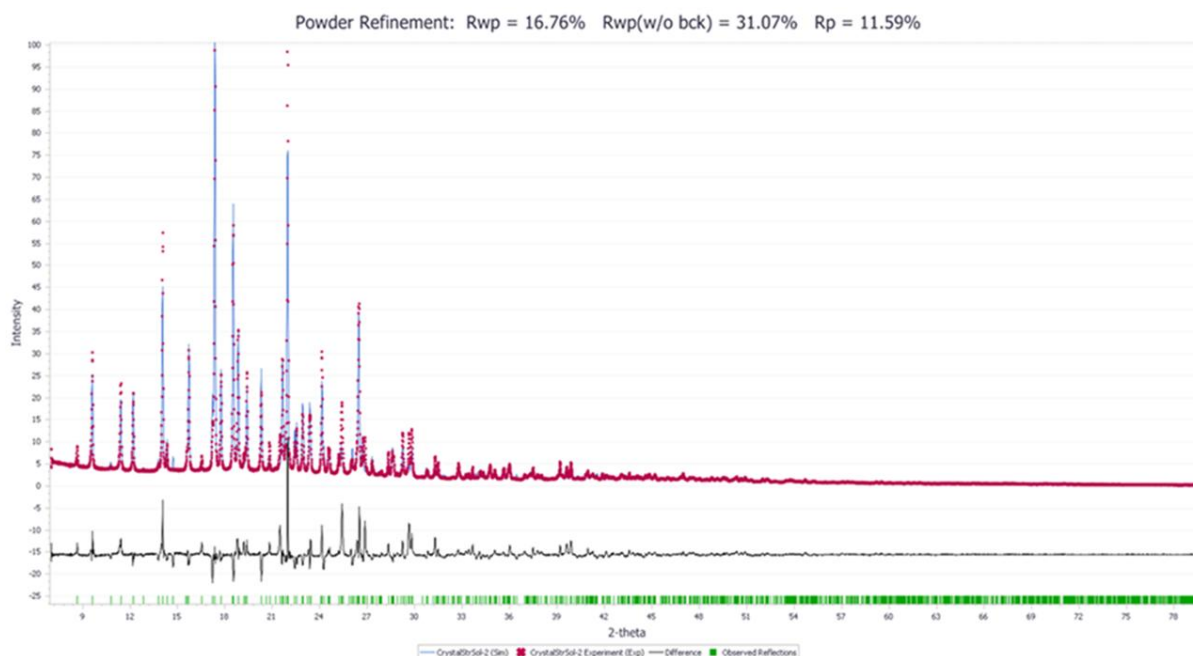
**Figure 4.4:**  $^1\text{H}$  NMR of DAP-PEG (1:4) using  $\text{CDCl}_3$  as solvent.  $^1\text{H}$  NMR 400 MHz ( $\text{CDCl}_3$ ):  $\delta = 7.4$  (d,  $J = 8.76$  Hz, 4H), 6.5 (d,  $J = 8.80$  Hz, 4H), 5.9 (s, 4H), 3.5 (s, 16H).

#### 4.2.2. Crystal structure analysis of DAP-PEG from Powder-XRD

Determination of crystal structure from XRPD was attempted using reflex module in Material Studio 17.1. A high-quality XRPD data was collected in transmission geometry for a period of 12 h. The polymorphic purity was confirmed, and Pawley profile fitting determined corresponding lattice parameters. The XRPD pattern of dapsone-PEG600 is found to be in an orthorhombic space group (Pbca) having chemically reasonable unit cell parameters:  $a = 20.514 \text{ \AA}$ ,  $b = 13.842 \text{ \AA}$ ,  $c = 15.567 \text{ \AA}$ . Considering the asymmetric unit with one molecule of dapsone and four monomers of PEG, the molecular structures in the asymmetric unit are modelled for the structure solution step. Structural solution was obtained by applying direct space methods and simulated annealing global optimization techniques. Final crystal structure was obtained by refinement of structure solution using Rietveld refinement with an Rwp (weighted profile R-factor) = 16.76% (**Figure 4.5**). R-factor indicates that the crystal solved from powder-XRD does not have enough accuracy and but it is reasonable value as one of its component is a flexible polymer PEG chain. For structural evaluation and optimization, plane wave-periodic DFT-D2 calculations were performed. Along with DFT, dispersion corrections were applied to correct van der Waal's interactions.

To understand the accuracy of crystal structure, crystal structure similarity was calculated as RMSD difference for crystal structures before and after optimization. As per the literature, the proposed assessment criteria of a correct structure are the crystal structure similarity

(experimental vs optimized crystal structure), with  $\text{RMSD} < 0.25 \text{ \AA}$ . RMSD between before and after geometrically optimized crystal structures was observed as  $0.57 \text{ \AA}$ . The crystal solved from Powder–XRD data does not appear highly accurate as RMSD values seems slightly higher than the acceptable limits.



**Figure 4.5:** Experimental, calculated and the difference profile of dapsone–PEG600 cocrystal obtained from Rietveld refinement.

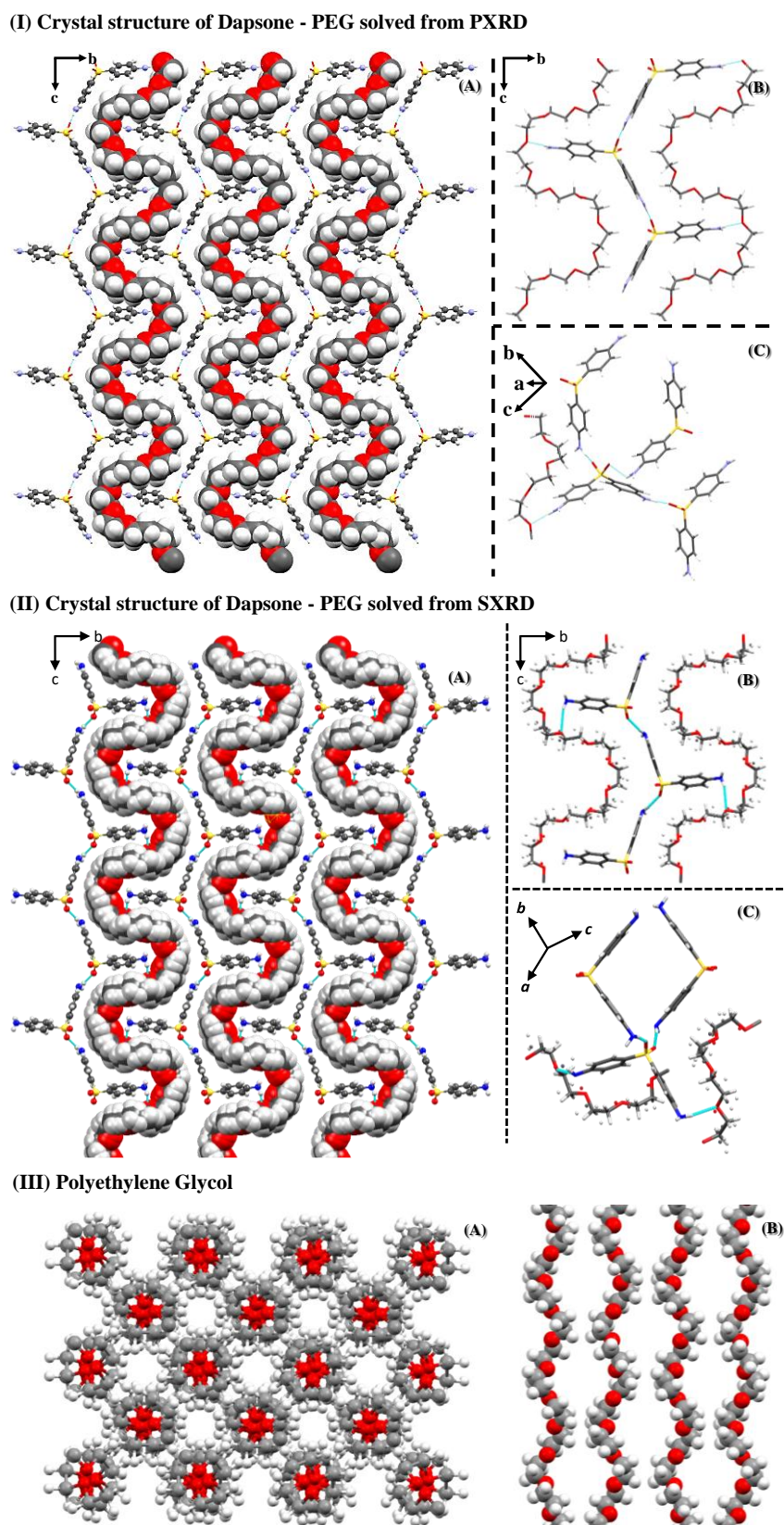
The crystal structure analysis reveals that PEO (polyethylene oxide) chains undergo adjustments of the conformation (pocket–like) to accommodate the aromatic rings. PEG chains adopt a coil–like pseudo helix structure along  $c$ –axis repeating for every four ethylene oxides monomers. This conformation differs completely from the  $7_2$  helix of pure PEG reported previously<sup>30</sup>. An infinite chain of dapsone molecules connected through  $\text{N–H}\cdots\text{O}=\text{S}$  interactions, run parallel to PEG chain. Likewise, it is well known that the oxygen group in the PEG chain is a strong hydrogen bond acceptor and can participate in hydrogen bonding. Therefore, alternative PEG chains are linked via  $\text{N–H}\cdots\text{O}$  interactions of dapsone molecules (Figure 4.6–I & Table 4.2).

**Table 4.2:** Intermolecular hydrogen–bonding interactions ( $\text{\AA}$ ,  $^\circ$ ).

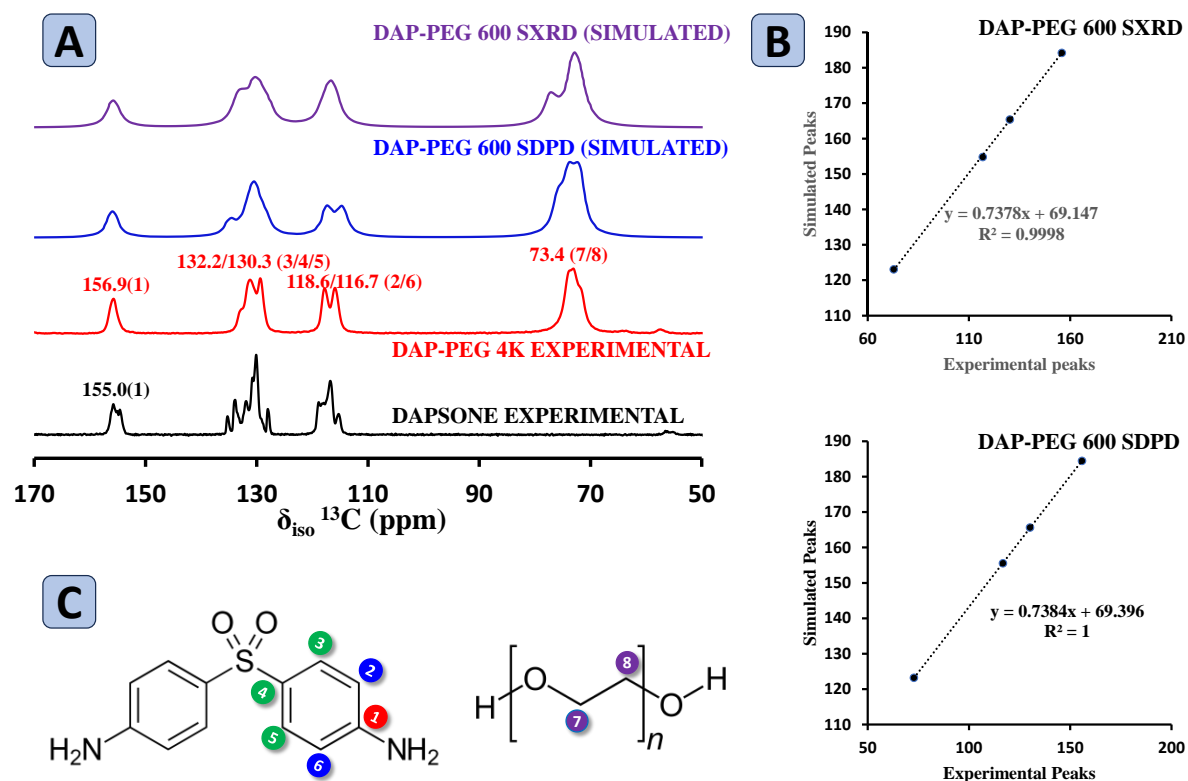
D–H $\cdots$ A	d(H $\cdots$ A) ( $\text{\AA}$ )	d(D $\cdots$ A) ( $\text{\AA}$ )	D–H $\cdots$ A ( $^\circ$ )
N4–H27 $\cdots$ O40 <sub>(-O-)</sub>	1.462	2.535	139.94
N4–H26 $\cdots$ O2 <sub>(-SO2-)</sub>	1.924	2.864	160.13
N5–H29 $\cdots$ O37 <sub>(-O-)</sub>	2.037	3.111	161.71
N5–H28 $\cdots$ O3 <sub>(-SO2-)</sub>	2.447	3.258	128.69

After multiple attempts, single crystals of DAP-PEG were successfully grown from methanol: acetone mixture 1:1(v/v). Crystal structure was solved from SCXRD which was discussed in detail in the next section. In this section, comparative studies between crystal structures solved from SDPD and SCXRD were also included to further comment on accuracy of the crystal structure from SDPD. From **Table 4.3**, we emphasize that the determined space group and lattice parameters from SXR D are similar to the values obtained from the indexing of XRPD data. From the comparison of **Figures 4.6-I & 4.6-II**, it can be observed that the crystal structure solved from SXR D is similar to the crystal structure solved from XRPD showing PEG chains with coil-like pseudo helix structure along the c-axis undergone adjustments conformation (pocket-like) to accommodate the aromatic rings. Also, the crystal structure similarity between the two optimized crystal structures was calculated as 0.17 Å, indicating that the crystal structure solved XRPD is accurate.

For further confirmation, <sup>13</sup>C SS-NMR experimental data of dapsone & dapsone-PEG 4K was compared with simulated <sup>13</sup>C SS-NMR from crystal structures solved from SDPD & SXR D. All peaks in the SS-NMR spectra of dapsone and DAP-PEG 4K experimental are assignable concerning corresponding NMR simulated spectra (**Figure 4.7**). It implies that SS-NMR spectra from crystal structures solved from XRPD and SCXRD are comparable with experimental spectrum of dapsone and DAP-PEG 4K. Therefore, from the above analysis, it can be concluded that the crystal structure solved from XRPD is reliable and accurate.



**Figure 4.6:** Overall close packing showing interdigitated chains of PEG and hydrogen bonded DAP molecules and Hydrogen bonding interactions in DAP–PEG crystal structures solved from (I) XRPD & (II) SXRD data and (III) Overall close packing showing interdigitated chains of PEG crystal structures.



**Figure 4.7:** (A) Overlay of experimental  $^{13}\text{C}$  SS-NMR spectra of dapsone and dapsone-PEG 4K experimental along with simulated  $^{13}\text{C}$  SS-NMR spectra of dapsone-PEG 600 (SDPD) and dapsone-PEG 600 (SXRD); (B) Plots between experimental  $^{13}\text{C}$  chemical shifts and simulated  $^{13}\text{C}$  chemical shifts for dapsone-PEG SDPD and dapsone-PEG SXRD. The  $R^2$  fit represents the correlation of chemical shift value of experimental and simulated; and (C) dapsone-PEG molecular structure with C-atoms assignment.

#### 4.2.3. Crystal structure analysis of DAP-PEG complex from SXRDR:

PEG600 molecular complex by vapor diffusion setup with acetone/diethyl ether solvent system<sup>1</sup>. The crystal structure was solved by single crystal XRD (SXRDR). As expected, the positional disorder was identified in PEG and modelled as a part of the SXRDR crystal structure. The detailed crystallography data is presented in **Table 4.3**.

**Table 4.3:** Comparison of crystallographic data obtained from SXRD and SDPD of dapsone–PEG cocrystal

Parameters	<i>SXRD</i>	<i>SDPD</i>
Empirical Formula	C <sub>12</sub> H <sub>12</sub> N <sub>2</sub> O <sub>2</sub> S, C <sub>8</sub> H <sub>16</sub> O <sub>4</sub>	C <sub>12</sub> H <sub>12</sub> N <sub>2</sub> O <sub>2</sub> S, C <sub>8</sub> H <sub>16</sub> O <sub>4</sub>
Crystal system	Orthorhombic	Orthorhombic
Space group	Pbca	Pbca
T/K	293	298
Radiation	CuK $\alpha$ (1.5418 Å)	CuK $\alpha$ (1.5418 Å)
a (Å)	20.0328(6)	20.5145(3)
b (Å)	13.6815(5)	13.8422(7)
c (Å)	15.4710(5)	15.5679(9)
$\alpha$ (°)	90	90
$\beta$ (°)	90	90
$\gamma$ (°)	90	90
V (Å <sup>3</sup> )	4240.3(2)	4420.7(5)
Z	8	8
Reflections collected	7853	
No of unique data	4216	1139
Observed reflections	3561	-
R <sub>1</sub> , I > 2 $\sigma$ (I)	0.062	11.59(Rp)
wR <sub>2</sub> , I > 2 $\sigma$ (I)	0.1776	16.76(Rwp)
Goodness of fit, S	1.046	-
CCDC number	1857934	1576320

This crystal structure has a packing coefficient<sup>31</sup> of 68.2 %, an intermediate value for organic crystals, implying a moderately close packing structure with strongly directional hydrogen bond interactions<sup>32</sup> (**Table 4.3**). The crystal structure analysis reveals that the PEG chains undergo adjustments of the conformation (pocket-like) to accommodate the aromatic rings. PEG chains adopt a coil-like pseudo helix structure along the c-axis, repeating every four ethylene oxides monomer units (**Figure 4.6–II**). The crystal structure of dapsone contains N–H...O=S hydrogen bonding between complementary functional groups such as –NH<sub>2</sub> and –SO<sub>2</sub> which are very good hydrogen bond donors and potential acceptors respectively.<sup>11</sup> A similar interactions motif is also observed in the crystal structure of the dapsone–PEG600 cocrystal (**Figure 4.6–II**). An infinite chain of dapsone molecules connected through N–H...O=S interactions runs parallel to the PEG chains. Likewise, it is well known that the oxygen groups in the PEG chain are strong hydrogen bond acceptors and can participate in hydrogen bonding

(Table 4.4). Therefore, alternative PEG chains are linked via the N–H···O interactions of dapsone molecules. As mentioned in the introduction, this study was further extended to a virtual screening of drug–PEG cocrystals.

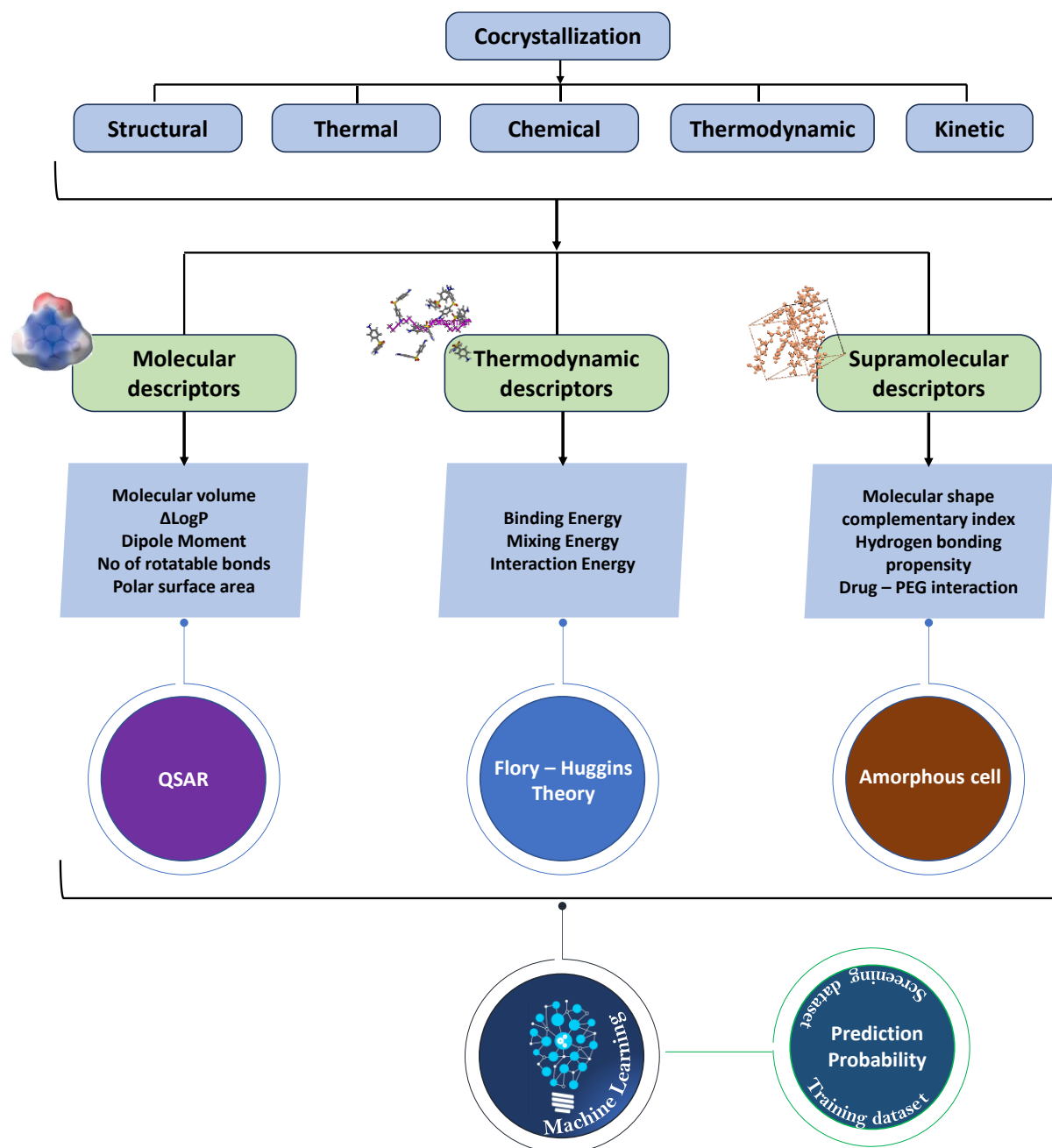
**Table 4.4:** Intermolecular hydrogen–bonding interactions (Å, °).

D–H···A	d(H···A) (Å)	d(D···A) (Å)	D–H···A (°)
N–H2A···O4A <sub>(-O-)</sub>	2.625	3.348	142.48
N–H2B···O6 <sub>(-SO2-)</sub>	2.155	2.995	165.52
N–H1B···O4A <sub>(-O-)</sub>	2.351	3.169	158.95
N–H1A···O5 <sub>(-SO2-)</sub>	2.133	2.987	171.45

### 4.3. Virtual Screening of drug–polymer co–crystals

Predictive models using machine learning (ML) algorithms have emerged as a promising tool for faster and accurate demonstration of new molecular materials and their properties<sup>33</sup>. The success of such methodologies largely depends on the careful mapping of appropriate descriptors, relevant to the challenge that needs to be solved. The present study is aimed at polymeric cocrystal prediction. Cocrystallization is composed of complex events that can be influenced by a wide variety of parameters, including structural, thermal, chemical, thermodynamic and kinetic in nature. Therefore, three–tier descriptors (molecular, supramolecular, and thermodynamic) coincided to solve the inherent complexities in crystallization. The molecular descriptors (Molecular volume,  $\Delta\log P$ , Number of rotatable bonds, Dipole moment and Polar surface area) selected in this study, are primarily as per literature on crystallizable cocrystal predictions<sup>34,35</sup>. The thermodynamic descriptors (binding energy, interaction parameter and mixing energy) are generated as a part of blend module calculation of drug–polymer miscibility. These parameters were commonly used to predict comparable drug–polymer pairs for successful ASD (amorphous solid dispersion) development<sup>36</sup>. The phenomenon of cocrystallization is related to the category of supramolecular chemistry. From this prospective, several indigenous supramolecular descriptors (molecular shape complementary index, hydrogen bonding propensity and drug–PEG interaction) are devised on the basis on the crystal engineering principles<sup>37</sup>. These descriptors are an amalgamation of molecular and supramolecular aspects of the co–crystallization process to ensure optimum crystal packing (close packing) and effective noncovalent interactions (hydrogen bonding), which can also impart some missing kinetic information during the crystallization process<sup>38</sup> in the prediction protocol (**Figure 4.8**).





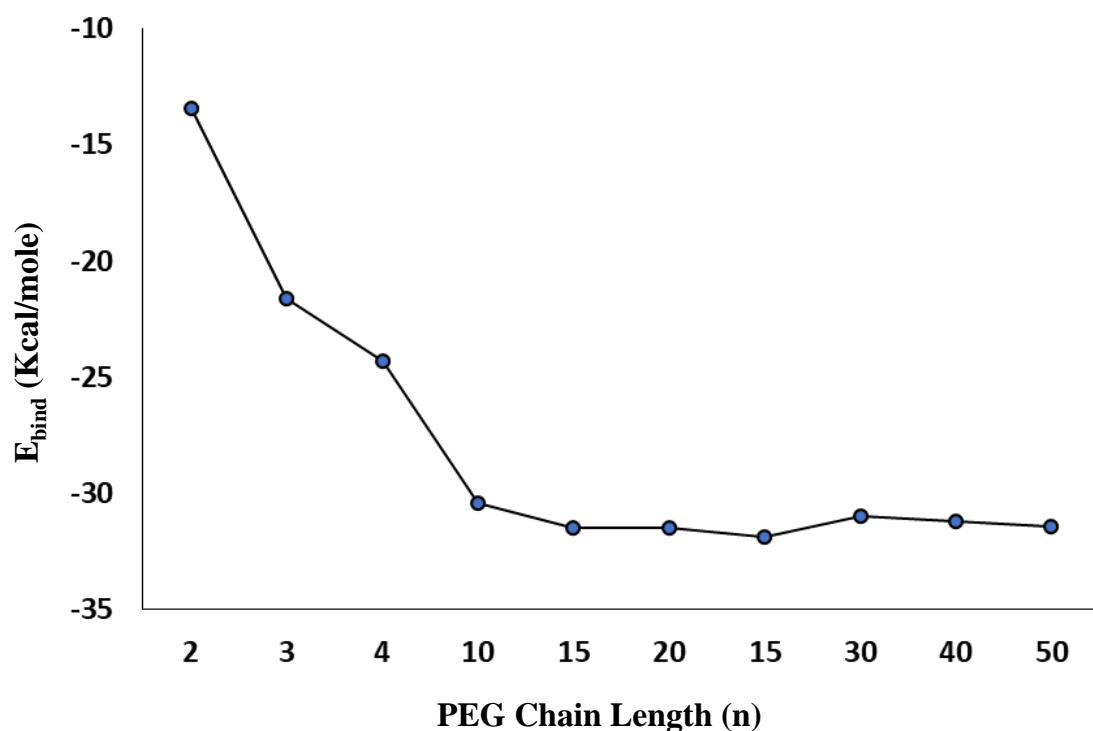
**Figure 4.8:** Flowchart showing protocol for virtual screening. \*Amorphous cell was employed to simulate Hydrogen bonding propensity and Drug-PEG interactions and mimic kinetic information.

(A) Overlay of experimental  $^{13}\text{C}$  SS-NMR spectra of DAP and DAP-PEG 4K experimental along with simulated  $^{13}\text{C}$  SS-NMR spectra of DAP-PEG 600 (SDPD) and dapsone-PEG 600 (SXR); (B) Plots between experimental  $^{13}\text{C}$  chemical shifts and simulated  $^{13}\text{C}$  chemical shifts for DAP-PEG SDPD and DAP-PEG SXR. The  $R^2$  fit represents the correlation of chemical shift value of experimental and simulated; and (C) DAP-PEG molecular structure with C-atoms assignment.

The molecular descriptors (molecular volume,  $\Delta\log P$ , number of rotatable bonds, dipole moment and polar surface area) and molecular shape complementary index under Supramolecular descriptor were calculated using QSAR module from Materials Studio. The thermodynamic descriptors (binding energy, interaction parameter and mixing energy) were computed using Blends module in Materials Studio.

In general, BIOVIA Materials Studio Blend predicts the miscibility behaviour of solvent and polymer systems and it requires only their molecular structures and a forcefield as inputs<sup>39</sup>. In pharmaceutical material development, Blend module is majorly used for suitable polymer screening for ASD<sup>40,41</sup> rather than cocrystal screening<sup>42</sup>. In this study, the Blend module is devised to predict drug–polymer cocrystal. The Flory–Huggins model is the simplest and best-known theory of the thermodynamics of mixing and phase separation in binary systems. The Blend module employs a modified Flory–Huggins model and molecular simulation techniques to improve predictability. This is accomplished by generating many pair configurations (by Monte Carlo Simulation) and calculating the binding energies, followed by temperature averaging the results using the Boltzmann factor and calculating the temperature–dependent interaction parameter. It is an off–lattice calculation, but takes care optimum arrangement with favorable coordination number. Three following parameters (thermodynamic descriptor) can be considered to characterize the optimized drug–polymer pairs: binding energy ( $E_{\text{bind}}$ ), mixing energy ( $E_{\text{mix}}$ ) and interaction parameters ( $\chi$ ).

Dapsone and PEG polymer cocrystal system was subjected to optimize the Blend methodology. Selecting a representative PEG chain length is desired to reduce the calculation's complexity. Therefore, Blend simulation was performed for dapsone and PEGs with varied chain lengths ( $n = 2, 3, 4, 10, 15, 20, 40$  and  $50$ ). It has been observed that there is more variation in  $E_{\text{mix}}$  and  $\chi$  values as compared to  $E_{\text{bind}}$  (**Table 4.5**). Each simulation was repeated thrice and an average of all three was reported. PEG chain lengths as a function of  $E_{\text{bind}}$  were plotted, from PEG  $n = 20$  onwards,  $E_{\text{bind}}$  values remained constant with further increase in PEG chain length (**Figure 4.9**). Hence, PEG polymer chain  $n = 20$  was considered for Blend simulation to compute the values of the thermodynamic descriptors for all entries in this study. Further, to understand drug–drug or drug–polymer association affinity via hydrogen bonding interactions, an artificial crystal environment is constructed with molecular cluster model as an input generated in Blend simulation, using the Amorphous Cell Module in Materials Studio.



**Figure 4.9:** Plot between calculated E<sub>bind</sub> (Kcal/Mole) between DAP–PEG verses change in PEG chain length (n).

**Table 4.5:** Comparison of E<sub>bind</sub>,  $\chi$  and E<sub>mix</sub> values calculated for dapson and PEG

PEG chain Length	$\chi$	$\chi$ Std Dev	E <sub>mix</sub>	E <sub>mix</sub> Std Dev	E <sub>bind</sub> Kcal/Mole	E <sub>bind</sub> Std Dev
2	25.6	<b>0.175</b>	15.2	<b>0.104</b>	-2.5	<b>0.003</b>
	25.3		15.0		-2.5	
	25.6		15.1		-2.5	
3	22.5	<b>0.359</b>	13.3	<b>0.213</b>	-3.1	<b>0.006</b>
	23.0		13.6		-3.1	
	23.2		13.7		-3.1	
4	15.3	<b>0.029</b>	9.1	<b>0.017</b>	-4.2	<b>0.008</b>
	15.3		9.0		-4.2	
	15.3		9.1		-4.2	
10	15.5	<b>3.948</b>	9.2	<b>2.338</b>	-10.2	<b>0.047</b>
	21.8		12.9		-10.1	
	14.6		8.6		-10.2	
15	4.60	<b>7.150</b>	2.7	<b>4.234</b>	-16.6	<b>0.144</b>
	-9.70		-5.7		-16.8	
	-1.70		-1.0		-16.9	
20	42.1	<b>11.216</b>	24.9	<b>6.642</b>	-17.2	<b>0.205</b>
	21.4		12.7		-17.4	
	24.2		14.3		-17.6	
25	144.0	<b>61.863</b>	85.2	<b>36.634</b>	-31.2	<b>0.742</b>
	249.2		147.6		-32.7	

	140.3		83.1		-31.8	
30	195.9	<b>21.046</b>	116.0	<b>12.463</b>	-30.1	<b>0.878</b>
	237.6		140.7		-30.9	
	211.4		125.2		-31.9	
40	325.6	<b>16.940</b>	192.8	<b>10.032</b>	-31.2	<b>1.182</b>
	335.7		198.8		-30.2	
	358.7		212.4		-32.6	
50	349.6	<b>35.077</b>	207.0	<b>20.772</b>	-30.4	<b>1.102</b>
	336.2		199.1		-32.6	
	402.6		238.4		-31.1	

Amorphous Cell tool from Materials Studio was employed to simulated and evaluate hydrogen bond propensity and drug–PEG interaction. The simulated crystal model is subjected to geometric optimization using forcefield calculation. It is to be noted that the periodic boundary conditions of the amorphous cell do not allow the molecules to be dispersed during optimization. This procedure to generate an artificial crystal environment is validated concerning dapsone and PEG polymer cocrystal system, having crystal structure reported. This native approach to a general artificial crystal environment would be beneficial to provide information on supramolecular descriptors, where no three–dimensional crystal structures are reported. In our study, more entries are in this category only.

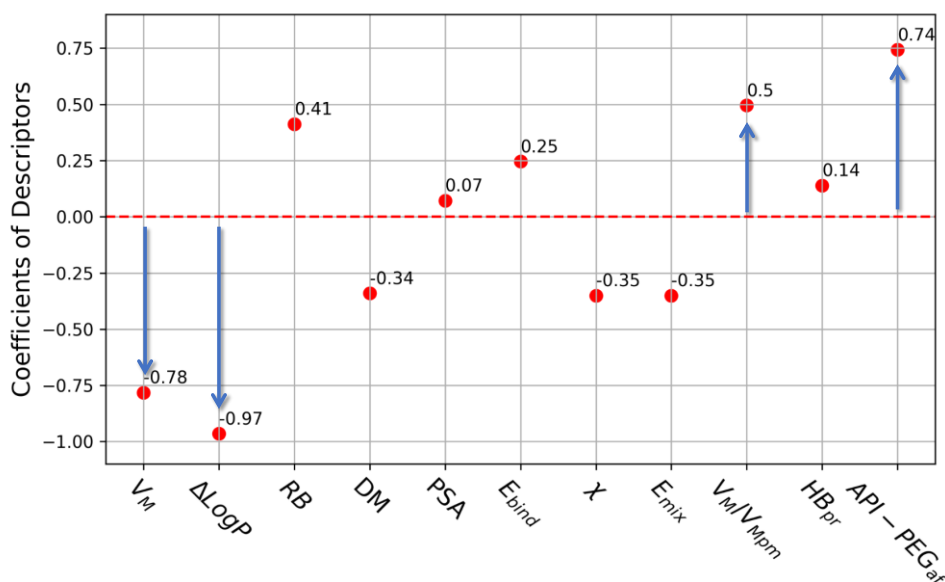
Finally, a machine learning algorithm was employed to establish a complex correlation between the three–tier descriptors using a training set. As per the prediction protocol, training dataset [1: Dapsone, 2: Griseofulvin, 3: Mavacoxib, 4: Efluniconazole 5: 4-Dichlorobenzene, 6: 4-Dibromobenzene, 7: 4-Bromochlorobenzene, 8: 4-Chloroiodobenzene, 9: 4-Bromofluorobenzene, 10: Resorcinol, 11: Hydroquinone, 12: 2-Methyl resorcinol (1:2), 13: 2-Methyl resorcinol (1:3.5), 14: p-Nitro phenol, 15: Urea, 16: Thiourea (2:1), 17: Thiourea (0.7:1) 18: Thiourea (1:4) and 19: Ammonia Borane]<sup>15,16,43–55</sup> is constructed with molecules that having a history to form supramolecular complex with PEG. Further, training dataset molecules can be divided into: inclusion complex (IC) and cocrystal (CC). The distinctive features were discussed elaboratively elsewhere<sup>48</sup>. Among them, strong drug–polymer interaction via hydrogen bonding is the main difference between IC and CC. In general, CC displaces strong drug–polymer, hydrogen bonding interaction whereas no or very weak interactions are formed for IC. It is noteworthy to mention that there is no negative control in the training dataset. Still, some IC entries can be considered as negative control from the perspective of descriptors like interaction and shape complementary index. Screening dataset [20: Sulfanilamide, 21: Pomalidomide, 22: Nicotinamide, 23: Lenalidomide, 24: Methylparaben, 25: Salicylic acid,

**26:** 4-Aminosalicylic acid, **27:** Proline] is selected with no history to form supramolecular complex with PEG and they have variation in their molecular functionality. All three-tier molecular, supramolecular and thermodynamic descriptors for both datasets are summarized in **Table 4.6**. The salient observations are listed below and these observations would provide us background to select a suitable algorithm for the AI-ML predictive model.

1. There is a clear trend observed for CC with prediction probability. For this class of crystal structures, prediction probability value can be considered as more than 0.5, representing a polymeric CC. Lower prediction probability numbers are indications of IC formation with PEG.
2. As  $\Delta\text{LogP}$  is an indicator of miscibility and  $\Delta\text{LogP}$  for PEG is -1.2. CC formation is observed for molecules with lower  $\Delta\text{LogP}$  and IC formation is observed for molecules with higher  $\Delta\text{LogP}$ .
3.  $\text{API-PEG}_{\text{af}}$  is clear indicator to understand drug-drug or drug-polymer association affinity via hydrogen bonding interactions. CC formation is observed for molecules exhibiting a high degree of hydrogen bonding interactions and indicated with 2. Molecules tending to show API-API interactions rather than API-PEG interactions are characterized by 0. Molecules exhibiting moderate hydrogen bonding interactions are indicated by 1.
4. A clear trend is observed for CC with Molecular shape complementary index ( $V_M/V_{M_{\text{pm}}}$ ). This class of crystal structures are properly closed packed and resulting in  $V_M/V_{M_{\text{pm}}} \sim 1$ . For IC, crystal structures are not packed efficiently due to mis-match of host-guest shape. Therefore,  $V_M/V_{M_{\text{pm}}}$  deviates from 1. A lower fraction number indicates IC formation with PEG being a host whereas larger number indicates a host-guest system where API is the host and PEG is the guest.
5. Obvious correlation is noticed with  $E_{\text{bind}}$  and CC or/and IC formation. When  $E_{\text{bind}}$  is higher than -15 kcal/mol, a clear sign of CC formation is noticed but  $E_{\text{bind}}$  is around or lower than -5 kcal/mol, a trend of IC formation. In this regard, only Mavacoxib-PEG is outlier an IC with higher  $E_{\text{bind}}$  (-28 kcal/mol). As per interaction examination, strong interaction among API molecules is the reason. Generally,  $\chi$  values are relatively little positive or negative and indicate compactable polymer-drug pair. Notably, no such correlation is found with  $\chi$  for CC or IC formation.

- Hydrogen bonding propensity is a clear indication of CC or IC formation. For CC preferable drug–polymer interactions were observed but for IC, API–API interaction was dominant in the expense of drug–polymer interactions. Therefore, the presence of strong API–API or no drug–polymer interaction can be used to confirm a negative entry.

Logistic regression is one of the most popular Machine Learning algorithms, which comes under the category of Supervised Learning technique. This methodology was employed to predict the probabilistic values (which lie between 0 and 1) to indicate the ability of the descriptors to predict the polymeric cocrystal formation as an independent variable. This learning technique was selected as it is suitable for small or big datasets and gives out binary classification. Identifying critical and suitable descriptors representing the cocrystallization phenomenon is an important step to predict the most probable drug–PEG combinations and give sensible inputs to ML algorithm. Analysis to select more significant and critical descriptors was done on the based on coefficient values, where if the coefficient value of a descriptor is either towards 1 or -1, it indicates a significant descriptor (**Figure 4.10**). It is observed that, out of 11 descriptors, 4 descriptors ( $\Delta\log P$ , API–PEG<sub>af</sub>,  $V_M/V_{Mpm}$  and  $V_m$ ) are exhibiting strong correlation with coefficient values more than 0.5.



**Figure 4.10:** Plot between descriptors and their coefficients. The coefficient values close to 1 or -1 indicate the corresponding better descriptors that fulfill the classification model.

Other than the four critical descriptors ( $\Delta\log P$ , API–PEG<sub>af</sub>,  $V_M/V_{Mpm}$  and  $V_m$ ), prediction probability and binding energy ( $E_{bind}$ ) were also considered due to their importance conceptually. Prediction probability was considered to leverage the advantage of machine

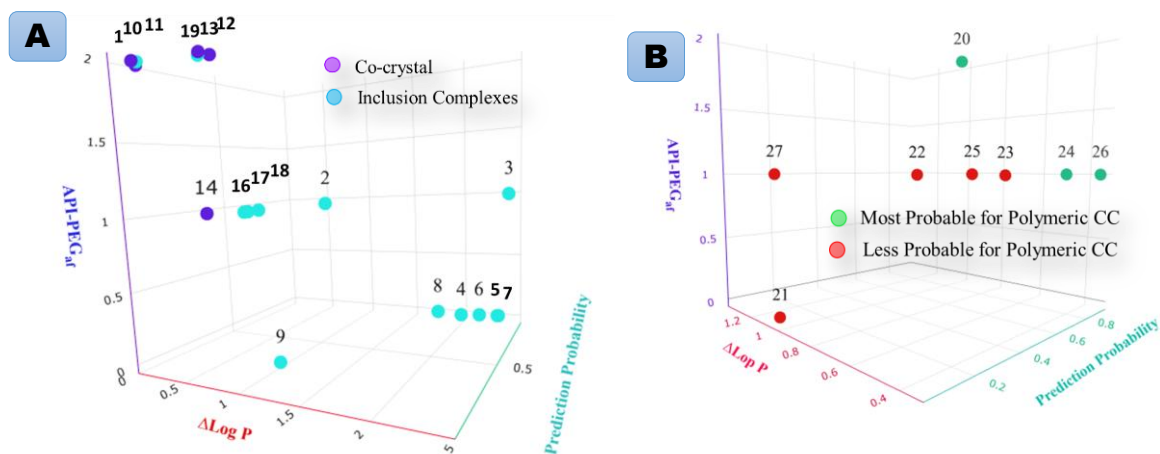
learning abilities into the prediction protocol. Binding energy is a well explored<sup>26</sup> and imminent choice for most scientists in the virtual screening of API–polymer solid dispersions. Hence, two 3D plots were strategically plotted for the training the dataset and screening dataset as below:

1. Plots–1:  $\Delta\log P$ , Prediction Probability and  $API-PEG_{af}$
2. Plots–2:  $\Delta\log P$ ,  $V_M/V_{Mpm}$  and  $E_{bind}$

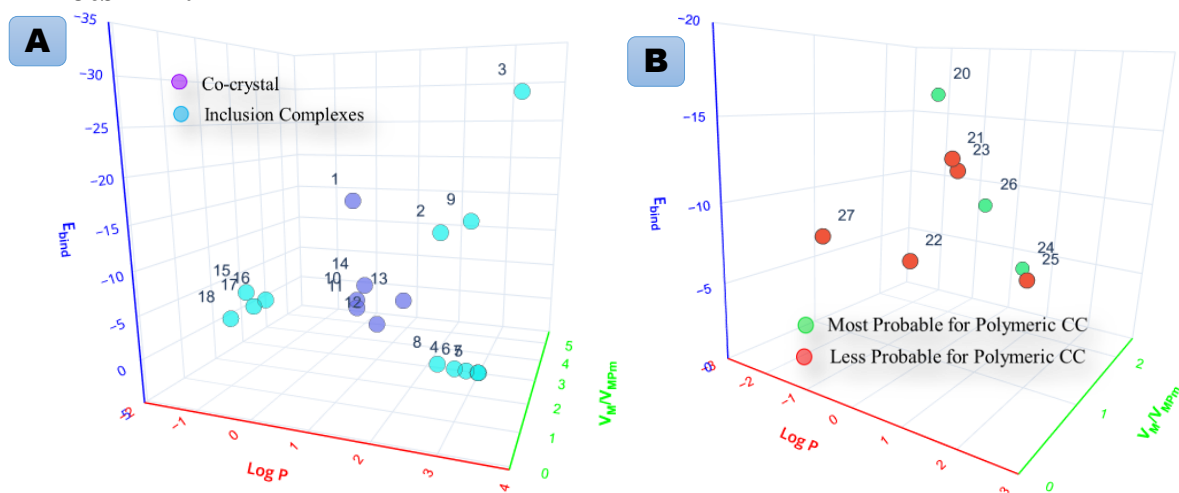
The two 3–dimensional plots are presented for training and screening datasets in **Figure 4.11**. Out of 19 entries in the training dataset, 17 entries were predicted as true positive (TP) which supports the accuracy of the predictive model. The False Positive (FP) entries are 1:2 urea–PEG and 1:4 thiourea–PEG complexes, in a true sense as per hydrogen bonding pattern, indicating scope to exist as CC whereas  $V_M/V_{Mpm}$  indicates to exist as IC. Following the literature nomenclature, these complexes were assigned as false negative. From the screening dataset, it can be inferred that a prediction probability value of more than 0.5 can represent a polymeric CC. If this criterion is applied to the test dataset then the most probable entries that can form polymeric CC are 4-aminosalicylic acid (0.92), methylparaben (0.89), sulfanilamide (0.83), salicylic acid (0.70) and lenalidomide (0.55). On the other hand, the rest of entries such as pomalidomide, nicotinamide and proline do not show affinity to form CC. To validate these predictions, crystallization experiments were performed either by anti–solvent addition, solvent evaporation, or melting crystallization.

The prediction ability of all test entries is defined and out of 8, one entry (Salicylic acid) is false positive and the other entries are true positive or negative. The result is very much encouraging, and this protocol shows the success rate of the prediction as ~90%. From this result, it can be inferred that a combination of various orthogonal descriptors can be a practical approach for virtual screening of cocrystal or ASD. Moreover, it can be put forward that such a predictive model's success largely depends on careful selection of indigenous supramolecular descriptors based on crystal engineering concepts. Due to the favourable synthesis route, the sulfanilamide–PEG system is considered for further study in **Chapter 5**.

Plots – I:



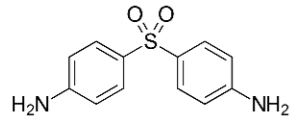
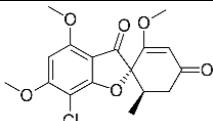
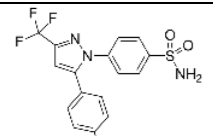



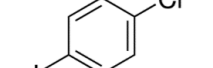
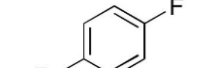
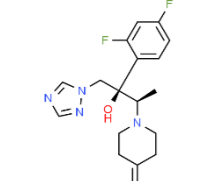
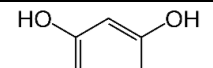
Plots – II:


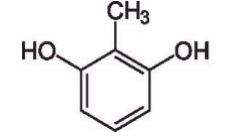
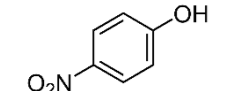
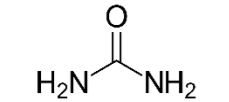
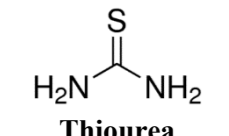
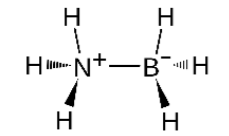
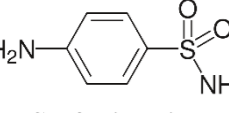
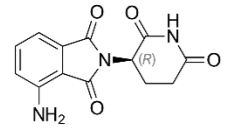


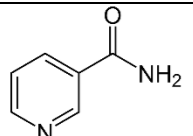
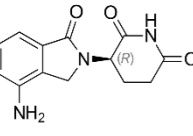
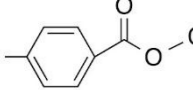
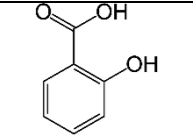
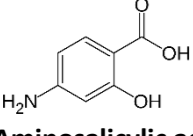
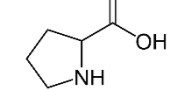
**Figure 4.11:** 3D plots–I ( $\Delta \text{log}P$ , Prediction Probability and API–PEG<sub>af</sub>) and 3D plots–II ( $\Delta \text{log}P$ ,  $V_M/V_{M_{\text{pm}}}$  and  $E_{\text{bind}}$ ) for (A) Training dataset and (B) Screening dataset.



Table 4.6: Master datasheet for ML modelling of various descriptors along with experimental and prediction result summary

S.No	Molecular structure	Molecular descriptors					Thermodynamic descriptors			Supramolecular descriptors				Classifications	Prediction probability	Literature evidence/experimental outcome	Predictive ability	Ref
		V <sub>M</sub> (Å <sup>3</sup> )	ΔLogP	RB	DM (D)	PSA (Å <sup>2</sup> )	E <sub>bind</sub>	χ	E <sub>mix</sub>	SR	V <sub>M</sub> /V <sub>Mpm</sub>	HB <sub>pr</sub>	API-PEG <sub>af</sub>					
Training Dataset																		
1	 <b>Dapsone</b>	211.2	0.07	2	3.7	86.2	-17.4	29.2	17.3	1:4	1.2	2	2	CC	0.92	As per literature it form CC with diffren greads of PEG and single crystal data also available.	TP	1
2	 <b>Griseofulvin</b>	293.9	0.97	3	3.7	71.1	-12.0	70.9	42.0	1:2	3.2	1	1	IC	0.40	As per literature reports it is cocrystal but crystal structure not reported.	TP	56
3	 <b>Mavacoxib</b>	290.4	2.43	3	1.4	78.0	-28.2	56.9	33.7	1:3	2.2	1	1	IC	0.11	As per literature reports, it is cocrystal but crystal structure not reported.	TP	17
4	 <b>4-Dichlorobenzene</b>	111.4	1.98	0	0	0	-3.4	122.5	72.5	3:10	0.1	0	0	IC	0.07	As per literature reports, it is cocrystal but crystal structure not reported.	TP	53
5	 <b>4-Dibromobenzene</b>	120.6	2.31	0	0	0	-3.4	124.3	73.6		0.1	0	0		0.04			
6	 <b>4-Bromochlorobenzene</b>	115.9	2.14	0	0.2	0	-3.4	124.6	73.8		0.1	0	0		0.05			
7	 <b>4-Chloriodobenzene</b>	122.5	2.30	0	0.4	0	-3.4	125.1	74.1		0.1	0	0		0.04			
8	 <b>4-Bromofluorobenzene</b>	106.7	1.74	0	0.4	0	-3.7	118.5	70.2		0.1	0	0		0.05			
9	 <b>Efinaconazole</b>	248.3	1.04	6	1.6	54.2	-11.0	64.4	38.1	1:1	5.6	0	0	CC	0.80	As per literature reports, it is cocrystal but crystal structure not reported.	TP	57
10	 <b>Hydroquinone</b>	99.9	0.16	2	0.9	40.5	-7.0	80.2	47.4	1:2	1.1	2	2	CC	0.97	# As per literature reports, it is cocrystal but crystal structure not reported.	TP	58,59

	<b>Resorcinol</b>																# Material can be synthesized easily but obtaining single crystal of XRD study was not successful		
11	 <b>Hydroquinone</b>	99.9	0.16	2	1.3	40.5	-6.1	92.3	54.7	1:4	1.1	1	2	CC	0.95	# As per literature reports, it is cocystal but crystal structure not reported. # Material can be synthesized easily but obtaining single crystal of XRD study was not successful	TP	50,60	
12	 <b>2-Methyl resorcinol</b>	116.3	0.68	2	1.1	40.5	-5.7	100.9	59.7	1:2	0.6	1	2	CC	0.86	# As per literature reports, it is cocystal but crystal structure not reported. # Material can be synthesized easily but obtaining single crystal of XRD study was not successful	TP	60,61	
13		116.3								1:3.5	1.9			CC	0.91				
14	 <b>p-Nitro Phenol</b>	114.7	0.41	1	2.4	63.4	-9.2	66.9	39.9	2:3	0.8	1	1	CC	0.80	# As per literature reports, it is cocystal but crystal structure not reported. # Material can be synthesized easily but obtaining single crystal of XRD study was not successful	TP	55,62	
15	 <b>Urea</b>	59.3	0.16	0	1.9	69.1	-4.1	107.8	63.8	2:1	2.2	2	2	IC	0.95	# As per literature reports, it is cocystal but crystal structure not reported. # Material can be synthesized easily but obtaining single crystal of XRD study was not successful.	FP	15,45,46	
16	 <b>Thiourea</b>	62.5	0.73	0	2.0	52.0	-5.2	96.1	56.9	2:1	1.4	2	1	CC	0.77	In literature, it was reported as a complex with layered crystal packing, where Thiourea is host and PEG is guest.	TP	15,63	
17										0.7:1	1.0			CC	0.75		TP		
18										1:4	0.3			IC	0.69		FP		
19	 <b>Ammonia Borane</b>	45.5	0.61	0	1.6	27.6	NA	NA	NA	1:5	0.2	2	2	CC	0.88	Literature reported single crystal structure.	TP	64	
<b>Screening Dataset</b>																			
20	 <b>Sulfanilamide</b>	139.5	0.95	1	3.9	86.2	-16.7	5.9	3.5	NA	1.1	2	2	NA	0.83	Attempted experimental trails on this material system has indicated ready formation of cocystal.	TP	Not Applicable	
21	 <b>Pomalidomide</b>	223.1	1.04	1	2.8	109.6	-12.8	60.3	36.5	NA	1.8	2	0	NA	0.05	Attempted experimental trails on this material system has resulted in unsuccessful results.	TN		

22	 <b>Nicotinamide</b>	109.0	0.81	1	2.4	56.0	-6.6	97.4	57.7	NA	0.9	2	1	NA	0.45	Attempted experimental trails on this material does not form cocystal.	TN
23	 <b>Lenalidomide</b>	221.5	0.5	1	2.5	92.5	-11.4	70.9	42.0	NA	1.8	2	1	NA	0.55	Attempted experimental trails on this material system has resulted in unsuccessful results.	TN
24	 <b>Methyl paraben</b>	136.3	0.47	3	1.4	46.5	-7.1	95.7	54.0	NA	1.1	1	1	NA	0.89	Attempted thermal experimental trails on this material system has indicated formation of cocystal.	TP
25	 <b>Salicylic acid</b>	116.7	0.77	3	2.2	57.5	-7.0	87.2	51.6	NA	0.9	2	1	NA	0.70	Attempted experimental trails on this material system has resulted in unsuccessful results.	FP
26	 <b>4-Aminosalicylic acid</b>	130.2	0.35	3	2.6	83.6	-6.9	96.9	57.4	NA	1.1	2	1	NA	0.92	Attempted thermal experimental trails on this material system has indicated formation of cocystal.	TP
27	 <b>Proline</b>	108.1	1.36	2	2.2	49.3	-6.9	90.9	53.9	NA	0.9	0	1	NA	0.20	Attempted experimental trails on this material system has resulted in unsuccessful results	TN

Not Applicable

$V_M (\text{\AA}^3)$  = Molecular Volume of API molecules.

$\Delta \log P = |\log P_{API}| - |\log P_{PEG}|$ ; Tabulated **Log P** (Partition coefficient) experimental data were taken from Chemicalize. **Log P** of PEG is -1.2.

**RB** = Number of Rotatable Bonds in molecular structures; Rotatable bonds were calculated using QSAR module from Materials Studio 20.0.

**DM** = Dipole Moment, calculated using

**PSA** = Polar Surface Area, calculated using

$E_{bind}$  = Binding energy;  $\chi$  = Interaction parameter;  $E_{mix}$  = Mixing energy; Thermodynamic descriptors were calculated using QSAR module from Materials Studio.

**SR** = Stoichiometric ratio of API/monomer of PEG; For screening molecules, API/monomer of PEG is 1.2.8 that is average of PEG monomer are derived from all the reported crystal structures.

$V_M/V_{Mpm}$  = Molecular shape complementary index, where  $V_M$  = Molecular Volume of API molecules and  $V_{Mpm}$  = Molecular volume of PEG monomer.

**HB<sub>pr</sub>** (Hydrogen bonding propensity scaled from 0 to 2) = No interactions between molecules → 0, Medium interactions between molecules → 1, Strong interactions between molecules → 2.

**API-PEG<sub>af</sub>** (Hydrogen bonding affinity between API and Polymer scaled from 0 to 2) = Affinity of molecules towards API-API → 0; API-PEG & API-API → 1; API-PEG → 2.

**Classification Criteria:** **Cocystal (CC)** has been defined with strong API-PEG hydrogen bonding interactions with layered packing with  $V_M/V_{Mpm}$  is near to 1 whereas **inclusion complex (IC)** has been defined with weak or no API-polymer interaction with  $V_M/V_{Mpm}$  near to zero (PEG host) or away from 1 (PEG is guest).

**Prediction Probability:** Towards "1" is cocystal and towards "0" is inclusion complex or no cocystal.

**Predictive Ability:** **TP:** true positive, **TN:** true negative, **FP:** false positive and **FN:** false negative

### 4.4. Summary Characterization of DAP–PEG complex

In summary, our work in this chapter covers the characterization of dapsone–PEG by combined experimental and computational studies followed by the development of virtual screening of drug–polymer cocrystals. In the initial stage, the crystal structure was solved from Powder–XRD data, and later it was solved from Single crystal–XRD. Upon solving the crystal structure, it was subjected to geometric optimization by DFT–D2 calculations.

Being motivated to discover more such drug–Polymer complexes with PEG as a coformer, this study was also extended further to develop potential computational methodologies together aimed at virtual screening of drug–PEG cocrystal complexes. These methodologies are incorporated with molecular descriptors, thermodynamic descriptors and supramolecular descriptors based on crystal engineering concepts and molecular simulation along with kinetic information associated with complex cocrystalization process combined with ML/AI studies. Experimental verification has confirmed that the success rate of the protocol is about 90% in discovering two drug–Polymer cocrystals. Among possible outcomes from screening suitable APIs to form drug–polymer cocrystals with PEG, Sulphanilamide–Polyethylene glycol was considered as an extended for further characterization in **Chapter 5**.

### 4.5. References

1. Chappa, P., Maruthapillai, A., Voguri, R., Dey, A., Ghosal, S., & Basha, M. A. Drug–Polymer Co-Crystals of Dapsone and Polyethylene Glycol: An Emerging Subset in Pharmaceutical Co-Crystals. *Cryst. Growth Des.* **18**, 7590–7598 (2018).
2. van de Streek, J. & Neumann, M. A. Validation of experimental molecular crystal structures with dispersion-corrected density functional theory calculations. *Acta Crystallogr. Sect. B Struct. Sci.* **66**, 544–558 (2010).
3. Van De Streek, J. & Neumann, M. A. Validation of molecular crystal structures from powder diffraction data with dispersion-corrected density functional theory (DFT-D). *Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater.* **70**, 1020–1032 (2014).
4. Lindenberg, M., Kopp, S., & Dressman, J. B. Classification of orally administered drugs on the World Health Organization Model list of Essential Medicines according to the biopharmaceutics classification system. *Eur. J. Pharm. Biopharm.* **58**, 265–278 (2004).
5. Maria Kuhnert-Brandstatter & IgoMoser. On the polymorphism of dapsone and ethambutol dihydrochloride. *Microchim. Acta* **71**, 125–136 (1979).

- Bertolasi, V., Ferretti, V., Gilli, P., & De Benedetti, P. G. Molecular structure and crystal packing of five 4-aminophenyl (4-substituted phenyl) sulfones. Correlations between structural distortions, spectroscopic parameters and electronic substituent effects. *J. Chem. Soc. Perkin Trans. 2* **2**, 213–219 (1993).
- Smith, G. & Wermuth, U. D. 4-Amino- N -(4,6-dimethylpyrimidin-2-yl)benzenesulfonamide–4-nitrobenzoic acid (1/1). *Acta Crystallogr. Sect. E Struct. Reports Online* **68**, 0669–0669 (2012).
- Smith, G. & Wermuth, U. D. 4-(4-Aminophenylsulfonyl)aniline–1,3,5- trinitrobenzene (1/2). *Acta Crystallogr. Sect. E Struct. Reports Online* **68**, 0494–0494 (2012).
- Martins, I., Martins, M., Fernandes, A., Andre, V., & Teresa Duarte, M. An insight into dapsone co-crystals: sulfones as participants in supramolecular interactions. *CrystEngComm* **15**, 8173–8179 (2013).
- Smith, G. & Wermuth, U. D. Proton-transfer and non-transfer compounds of the multi-purpose drug dapsone [4-(4-Aminophenylsulfonyl)aniline] with 3,5-dinitrosalicylic acid and 5-nitroisophthalic acid. *J. Chem. Crystallogr.* **43**, 664–670 (2013).
- Jiang, L., Huang, Y., Zhang, Q., He, H., Xu, Y., & Mei, X. Preparation and solid-state characterization of dapsone drug-drug co-crystals. *Cryst. Growth Des.* **14**, 4562–4573 (2014).
- Harris, J. M. & Chess, R. B. Effect of pegylation on pharmaceuticals. *Nat. Rev. Drug Discov.* **2**, 214–221 (2003).
- Kolate, A., Baradia, D., Patil, S., Vhora, I., Kore, G., & Misra, A. PEG - A versatile conjugating ligand for drugs and drug delivery systems. *J. Control. Release* **192**, 67–81 (2014).
- Ugriné Hunyadvári, E., Hadi, I. A., Kassem, M. A., Farouk, A. M., & Selmeczi, B. Formulation of Polyethylene Glycol ointment bases suitable for tropical and subtropical climates. II. *Acta Pharm. Hung.* **59**, 157–65 (1989).
- Bailey, F. E. & France, H. G. Molecular association complexes of polymers. Urea and thiourea complexes of high molecular weight poly(ethylene oxide). *J. Polym. Sci.* **49**, 397–406 (1961).
- Dosiere, M. Structure and crystallization of molecular complexes between

- poly(ethylene oxide) and hydroxybenzenes. *J. Macromol. Sci. - Phys.* **35**, 303–328 (1996).
17. Changquan Calvin Sun. Patent No. WO2006024930A1; Novel co-crystals between polyethylene glycols and 5-phenylpyrazolyl-1-benzenesulfonamides. (2005).
  18. Zhong, Z., Guo, C., Chen, L., Xu, J., & Huang, Y. Co-crystal formation between poly(ethylene glycol) and a small molecular drug griseofulvin. *Chem. Commun.* **50**, 6375–6378 (2014).
  19. Zhong, Z., Guo, C., Yang, X., Guo, B., Xu, J., & Huang, Y. Drug Molecule Diflunisal Forms Crystalline Inclusion Complexes with Multiple Types of Linear Polymers. *Cryst. Growth Des.* **16**, 1181–1186 (2016).
  20. Zhong, Z., Yang, X., Fu, X.-B., Yao, Y.-F., Guo, B.-H., Huang, Y., & Xu, J. Crystalline inclusion complexes formed between the drug diflunisal and block copolymers. *Chinese Chem. Lett.* **28**, 1268–1275 (2017).
  21. Zhong, Z., Yang, X., Guo, B., Xu, J., & Huang, Y. Dissolution Behavior of the Crystalline Inclusion Complex Formed by the Drug Diflunisal and Poly( $\epsilon$ -caprolactone). *Cryst. Growth Des.* **17**, 355–362 (2017).
  22. Yang, X., Yu, B., Zhong, Z., Guo, B., & Huang, Y. Nevirapine-polycaprolactone crystalline inclusion complex as a potential long-acting injectable solid form. *Int. J. Pharm.* **543**, 121–129 (2018).
  23. Da Costa Bernardo Port, B., Schneider-Rauber, G., Fretes Argenta, D., Arhangelskis, M., De Campos, C. E. M., João Bortoluzzi, A., & Caon, T. Effect of Vehicle Composition on the Preparation of Different Types of Dapsone Crystals for Topical Drug Delivery. *Mol. Pharm.* **19**, 2164–2174 (2022).
  24. Israelachvili, J. The different faces of poly(ethylene glycol). *Proc. Natl. Acad. Sci. U. S. A.* **94**, 8378–8379 (1997).
  25. Yan, Y., Chen, J. M., & Lu, T. B. Simultaneously enhancing the solubility and permeability of acyclovir by crystal engineering approach. *CrystEngComm* **15**, 6457–6460 (2013).
  26. Pajula, K., Taskinen, M., Lehto, V. P., Ketolainen, J., & Korhonen, O. Predicting the formation and stability of amorphous small molecule binary mixtures from
-

- computationally determined flory-huggins interaction parameter and phase diagram. *Mol. Pharm.* **7**, 795–804 (2010).
27. Aitipamula, S. & Vangala, V. R. X-Ray Crystallography and its Role in Understanding the Physicochemical Properties of Pharmaceutical Cocrystals. *J. Indian Inst. Sci.* **97**, 227–243 (2017).
  28. Berry, D. J. & Steed, J. W. Pharmaceutical cocrystals, salts and multicomponent systems; intermolecular interactions and property based design. *Adv. Drug Deliv. Rev.* **117**, 3–24 (2017).
  29. Chappa, P. Modulating Physicochemical Attributes of Pharmaceutical Co-Crystals: A Case Study Of Drug–Polymer/Small Molecular Systems. (SRM University, 2019).
  30. French, A. C., Thompson, A. L., & Davis, B. G. High-purity discrete PEG-oligomer crystals allow structural insight. *Angew. Chemie - Int. Ed.* **48**, 1248–1252 (2009).
  31. Spek, A. L. Structure validation in chemical crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **65**, 148–155 (2009).
  32. Kaplan, I. G. *Intermolecular Interactions. Intermolecular Interactions: Physical Picture, Computational Methods and Model Potentials* (Wiley, 2006). doi:10.1002/047086334X.
  33. Allen, A. E. A. & Tkatchenko, A. Machine learning of material properties: Predictive and interpretable multilinear models. *Sci. Adv.* **8**, 7185 (2022).
  34. Fábíán, L. Cambridge Structural Database Analysis of Molecular Complementarity in Cocrystals. *Cryst. Growth Des.* **9**, 1436–1443 (2009).
  35. Wicker, J. G. P. & Cooper, R. I. Will it crystallise? Predicting crystallinity of molecular materials. *CrystEngComm* **17**, 1927–1934 (2015).
  36. Saboo, S., Kestur, U. S., Flaherty, D. P., & Taylor, L. S. Congruent Release of Drug and Polymer from Amorphous Solid Dispersions: Insights into the Role of Drug-Polymer Hydrogen Bonding, Surface Crystallization, and Glass Transition. *Mol. Pharm.* **17**, 1261–1275 (2020).
  37. Nangia, A. K. & Desiraju, G. R. Heterosynthons, Solid Form Design and Enhanced Drug Bioavailability. *Angew. Chemie - Int. Ed.* **61**, e202207484 (2022).
-

38. Dey, A., Kirchner, M. T., Vangala, V. R., Desiraju, G. R., Mondal, R., & Howard, J. A. K. Crystal structure prediction of aminols: Advantages of a supramolecular synthon approach with experimental structures. *J. Am. Chem. Soc.* **127**, 10545–10559 (2005).
  39. Hlaváč, D. & Tokarský, J. Molecular modeling of chitosan/polyethylene oxide polymer blends. in *NANOCON 2013 - Conference Proceedings, 5th International Conference* 275–279 (2013).
  40. Yang, M., Wang, P., & Gogos, C. Prediction of acetaminophen's solubility in poly(ethylene oxide) at room temperature using the Flory-Huggins theory. *Drug Dev. Ind. Pharm.* **39**, 102–108 (2013).
  41. Chakravarty, P., Lubach, J. W., Hau, J., & Nagapudi, K. A rational approach towards development of amorphous solid dispersions: Experimental and computational techniques. *Int. J. Pharm.* **519**, 44–57 (2017).
  42. Zhou, J. H., Zhao, L., Shi, L. W., Xu, T. J., Li, Y. Y., Dai, J. F., & Zhao, Y. L. Analysis of the structures and interactions between CL-20 and its formers. *J. Mol. Struct.* **1207**, 127731 (2020).
  43. Gao, Y., Yao, S.-F., Ye, H.-M., Guo, B.-H., & Xu, J. Orientation of polymer chains in spherulites of poly(ethylene oxide)-urea inclusion compounds. *Polymer (Guildf)*. **130**, 209–217 (2017).
  44. Tadokoro, H., Yoshihara, T., Chatani, Y., & Murahashi, S. A preliminary report of structural studies on polyethylene oxide-urea complex. *J. Polym. Sci. Part B Polym. Lett.* **2**, 363–368 (1964).
  45. Ye, H.-M., Song, Y.-Y., Xu, J., Guo, B.-H., & Zhou, Q. Melting behavior of inclusion complex formed between polyethylene glycol oligomer and urea. *Polymer (Guildf)*. **54**, 3385–3391 (2013).
  46. Bogdanov, B. G., Michailov, M., Uzov, C. V., & Gavrilova, G. G. Complexation of poly(ethylene oxide) and urea. *J. Polym. Sci. Part B Polym. Phys.* **32**, 387–394 (1994).
  47. Ploszajski, A. R., Billing, M., Cockcroft, J. K., & Skipper, N. T. Crystalline structure of an ammonia borane–polyethylene oxide cocrystal: a material investigated for its hydrogen storage potential. *CrystEngComm* **20**, 4436–4440 (2018).
  48. Chappa, P., Maruthapillai, A., Voguri, R., Dey, A., Ghosal, S., & Basha, M. A. Drug–
-



- Polymer Co-Crystals of Dapsone and Polyethylene Glycol: An Emerging Subset in Pharmaceutical Co-Crystals. *Cryst. Growth Des.* **18**, 7590–7598 (2018).
49. Zhong, Z., Guo, C., Chen, L., Xu, J., & Huang, Y. Co-crystal formation between poly(ethylene glycol) and a small molecular drug griseofulvin. *Chem. Commun.* **50**, 6375–6378 (2014).
  50. Paternostre, L., Damman, P., & Dosièrè, M. Morphology and crystal structure of the poly(ethylene oxide)-hydroquinone molecular complex. *J. Polym. Sci. Part B Polym. Phys.* **37**, 1197–1208 (1999).
  51. Sun, C. C. Patent No. WO2006024930A1; Novel co-crystals between polyethylene glycols and 5-phenylpyrazolyl-1-benzenesulfonamides. (2005).
  52. Dosièrè, M. Lamellar structure of poly(ethylene oxide) molecular complexes. *Macromol. Symp.* **114**, 51–62 (1997).
  53. Point, J. J. & Damman, P. Poly(ethylene oxide)-p-dihalogenobenzene crystalline complex: FTIR and X-ray diffraction study of spherulitic fibers. *Macromolecules* **24**, 2019–2023 (1991).
  54. Paternostre, L., Damman, P., & Dosièrè, M. Polymorphism and Crystal Morphology of Poly(ethylene oxide)–2-Methyl Resorcinol Supramolecular Complexes. *Macromolecules* **32**, 153–161 (1999).
  55. Point, J. J. & Damman, P. Structure of a new crystalline complex: poly(ethylene oxide) with p-nitrophenol. 1. *Macromolecules* **25**, 1184–1188 (1992).
  56. Zhong, Z., Guo, C., Chen, L., Xu, J., & Huang, Y. Co-crystal formation between poly(ethylene glycol) and a small molecular drug griseofulvin. *Chem. Commun.* **50**, 6375–6378 (2014).
  57. Eun Ju Park, Hyun Ji, Eun Mi Kim, J. O. P. Patent No EP3928769A1; Co-crystalline efinaconazole, and method for producing same. (2020).
  58. Iannelli, P., Damman, P., Dosièrè, M., & Moulin, J.-F. Crystalline Structure of PEO–Resorcinol and PEO–Methylresorcinol Complexes. *Macromolecules* **32**, 2293–2300 (1999).
  59. Myasnikova, R. M., Titova, E. F., & Obolonkova, E. S. Study of 2:1 poly(ethylene
-

- oxide)—resorcinol molecular complex. *Polymer (Guildf)*. **21**, 403–407 (1980).
60. Paternostre, L., Damman, P., & Dosière, M. Crystalline structure of PEO-hydroquinone and PEO-2-methylresorcinol molecular complexes. *Macromol. Symp.* **114**, 205–210 (1997).
61. Paternostre, L., Damman, P., & Dosiere, M. Polymorphism and Crystal Morphology of Poly(ethylene oxide)–2-Methyl Resorcinol Supramolecular Complexes. *Macromolecules* **32**, 153–161 (1999).
62. Damman, P. & Point, J. J. Crystalline Structure of the Poly(ethylene oxide)-p-Nitrophenol Complex. 2. Fourier Transform Infrared Spectroscopy. *Macromolecules* **27**, 3919–3925 (1994).
63. Ye, H. M., Liu, K. S., & Zhou, Q. Inclusion complex formed between poly(ethylene oxide) and thiourea. *Colloids Surfaces A Physicochem. Eng. Asp.* **467**, 251–258 (2015).
64. Ploszajski, A. R., Billing, M., Cockcroft, J. K., & Skipper, N. T. Crystalline structure of an ammonia borane–polyethylene oxide cocrystal: a material investigated for its hydrogen storage potential. *CrystEngComm* **20**, 4436–4440 (2018).

# Chapter 5

## Sulfanilamide–Polyethylene glycol cocrystal

---

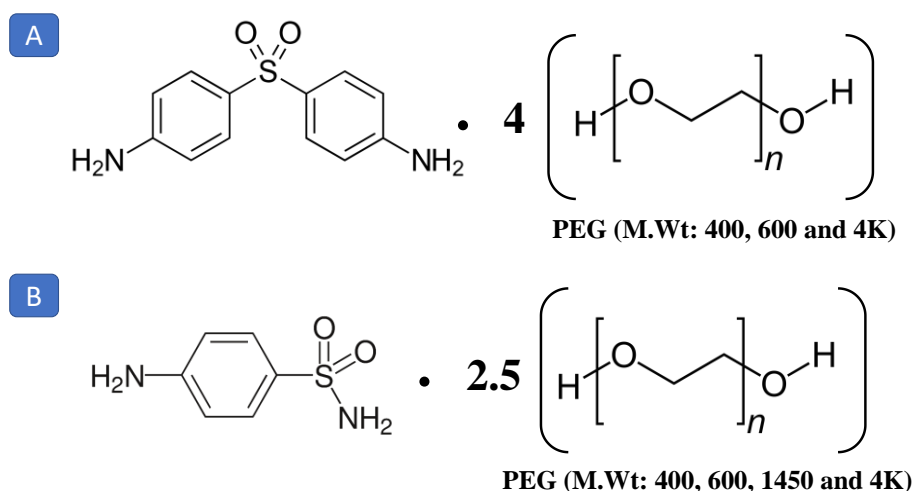
*This chapter covers characterisation of Sulfanilamide–PEG (Drug–Polymer cocrystal) and the establishment of Structure–Property Relationship.*

---

### 5.1. Introduction

The present chapter is a continuation of the previous chapter on ‘Characterization of DAP–PEG cocrystal and Virtual Screening of drug–polymer cocrystals’. The work on DAP–PEG was published, demonstrating the modulation of physico-chemical properties by means of associating dapsone with PEG of different molecular weights (M.wt: 600, 1450, 4K, 6K, 8K, 20K, 35K) forming a cocrystal. Inspired to explore and find more such drug - polymer systems, this work was extended towards developing novel computational work methodologies to discover new drug–polymer complexes. Our virtual screening of drug–polymer cocrystals, supports the formation of sulfanilamide and polyethylene glycol (SUL–PEG) cocrystals. A few other drug–polymer combinations were also discovered. Synthesis and physiochemical properties determination through change for SUL–PEG system were published in referenced thesis<sup>1</sup>. Sulfanilamide (4-Aminobenzenesulfonamide) is an antibiotic drug that belongs to biopharmaceutical classification system class III (low solubility and high permeability). Three different polymorphic forms (form- $\alpha$ , form- $\beta$ , and form- $\gamma$ ) and various cocrystals of this drug are known from the literature<sup>2</sup>. The importance of PEG as a cofomer to form cocrystals was already discussed in **Chapter 4**.

The present work in this chapter deals with detailed structural characterization of the SUL–PEG cocrystal system. As crystallinity is one of the essential physical properties of drug products, the impact of molecular weight (PEG polymer with different molecular weight grade) and process change on crystallinity were studied. The experimental results indicate that cocrystal with various PEG grades formed an isostructural crystal structure the same as DAP–PEG system, whereas the crystallinity remain unalter. In this chapter, the contradictory crystallinity behaviour<sup>3</sup> is understood using various techniques like X-ray diffraction; and SS-NMR in combination with computational simulation and molecular modelling. Throughout the studies, SUL–PEG 400 (1:2.5), SUL–PEG 600 (1:2.5), SUL–PEG 400/4K (1:2.5, solid solution), SUL–PEG 4K (1:2.5), DAP–PEG 400 (1:4), DAP–PEG 600 (1:4), DAP–PEG 400/4K (1:4, solid solution), DAP–PEG 4K (1:4) materials were used (**Figure 5.1**).



**Figure 5.1:** Molecular structures of (A) SUL–PEG 400, SUL–PEG 600, SUL–PEG 400/4K (1:1, solid solution), SUL–PEG 4K, SUL–PEG 1450 and (B) DAP–PEG 400 (1:4), DAP–PEG 600, DAP–PEG 400/4K (1:1, solid solution), DAP–PEG 4K.

## 5.2. Characterisation of SUL–PEG cocrystal:

The determination of stoichiometry between SUL and PEG was well-discussed elsewhere<sup>4</sup>. The powder-XRD overlay of SUL, PEG 4K and SUL-PEG 4K indicates the formation of cocrystal between SUL and PEG (**Figure 5.2**). Further, as a part of SUL–PEG structural characterization, stoichiometry between SUL and PEG was determined using two different approaches. In the first approach, the lattice parameters determination from Powder-XRD was attempted using the Reflex module in Material Studio 17.1<sup>5</sup>. A high-quality XRPD data was collected for SUL–PEG 600 sample in transmission geometry for 12 hours. The polymorphic purity was confirmed, and corresponding lattice parameters were determined and refined by Pawley profile fitting with  $R_{wp} = 4.41\%$  and  $R_p = 3.10\%$ . The lattice parameters of SUL - PEG 600 are found to be in an orthorhombic crystal system with *Pbca* space group having chemically reasonable unit cell parameters,  $a = 18.55\text{\AA}$ ,  $b = 10.82\text{\AA}$ ,  $c = 19.96\text{\AA}$  and  $V = 2803\text{\AA}^3$  with satisfactory refinement statistics (**Figure 5.3**). The obtained cell volume ( $2803\text{\AA}^3$ ) is reasonable enough to accommodate one molecule of SUL and 2.5 monomers of PEG polymer chain in its asymmetric unit with superior accuracy of 98% compared to 89.5% and 94.2% calculated for 2 and 3 monomers of PEG polymer chains respectively. Secondly, solution phase  $^1\text{H}$  and  $^{13}\text{C}$  NMR analysis was conducted for SUL–PEG4K sample. From  $^1\text{H}$  NMR spectra, signals integration of atom no. 1 (2H) and atom no. 5 & 6 (4H) is in the ratio (2.0/9.9) as 1:2.5 (**Figure 5.4**). Hence, from the above analysis, the stoichiometric composition in SUL–PEG cocrystal was reconfirmed.

An Interesting phenomenon was observed upon analysis of DAP–PEG and SUL–PEG cocrystals prepared by PEG 4K and analysed using the Powder-XRD technique. Upon comparison, a significant difference was observed in %crystallinity between them. The %crystallinity of SUL–PEG 4K was observed to be superior compared to the DAP–PEG 4K (Figure 5.5). Further experiments were planned towards understanding this interesting discrete difference between DAP–PEG and SUL–PEG crystals by studying the impact of different crystallization process and PEG M.wt grades on % crystallinity.

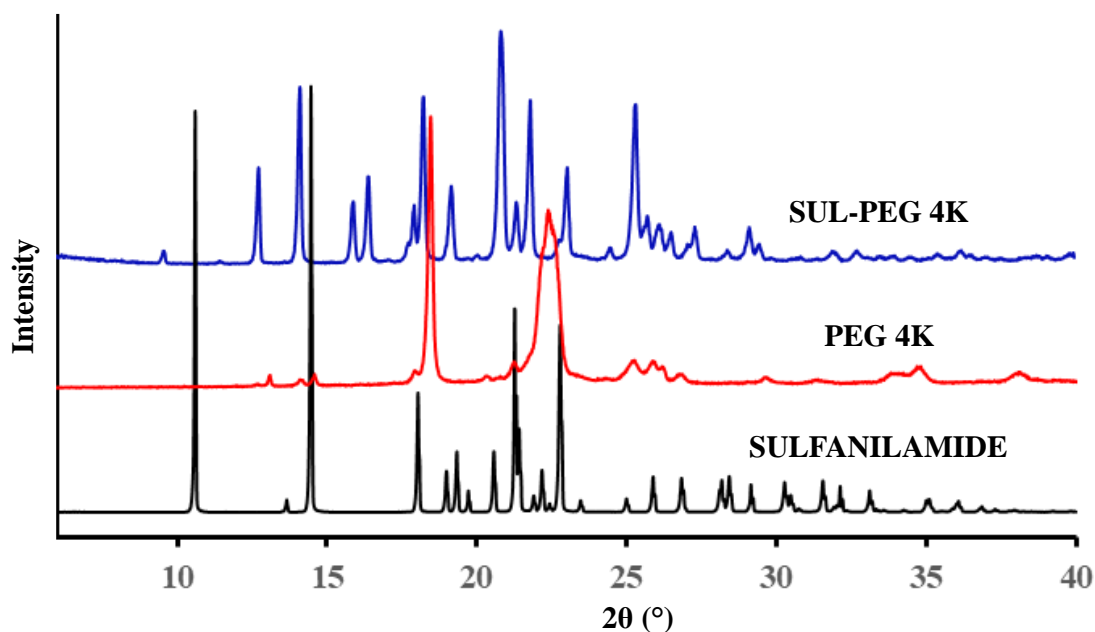


Figure 5.2: Powder-XRD overlay of (A) Sulfanilamide, (B) PEG 4K and (C) SUL-PEG 4K.

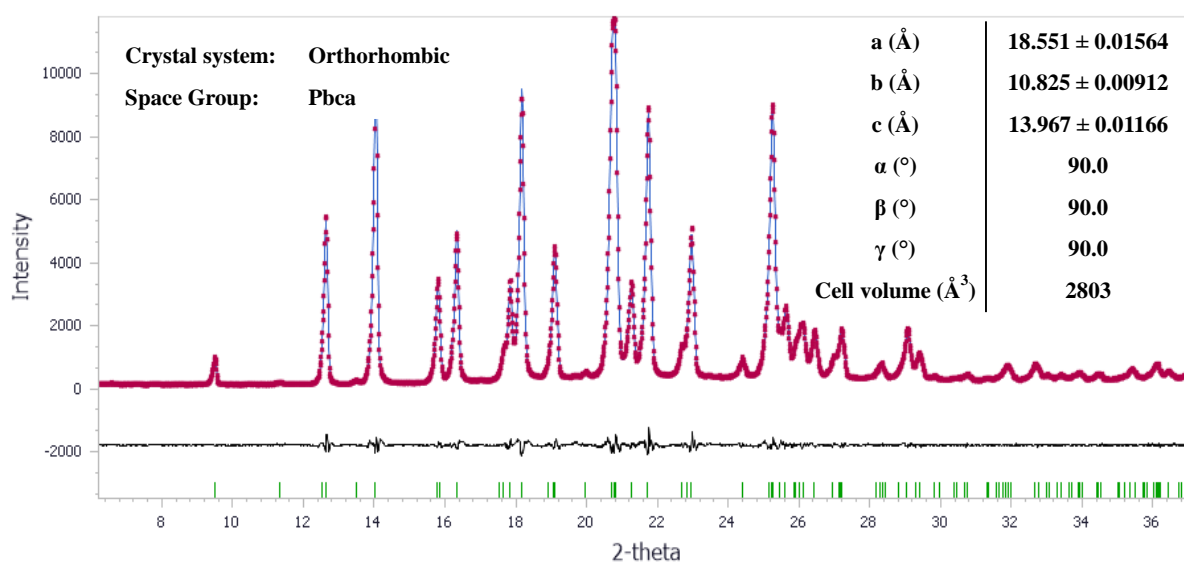
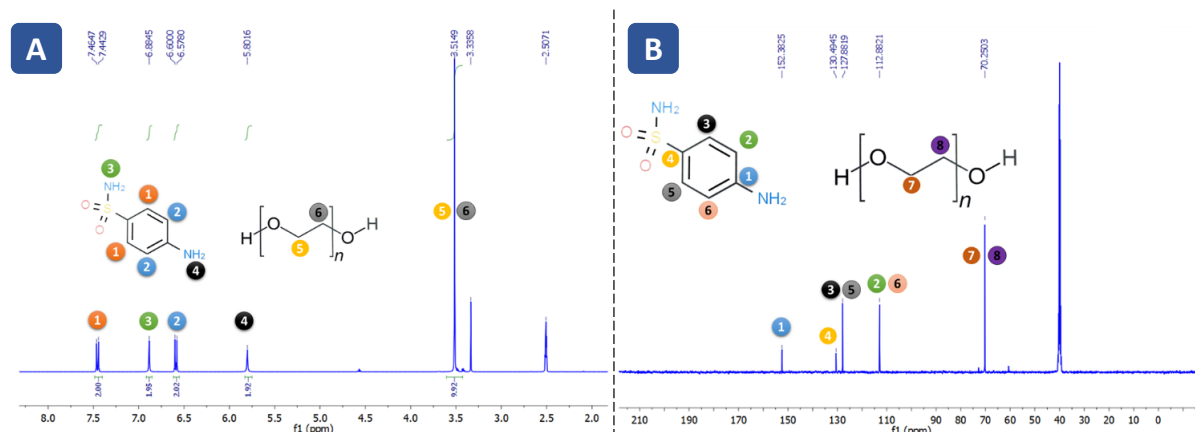
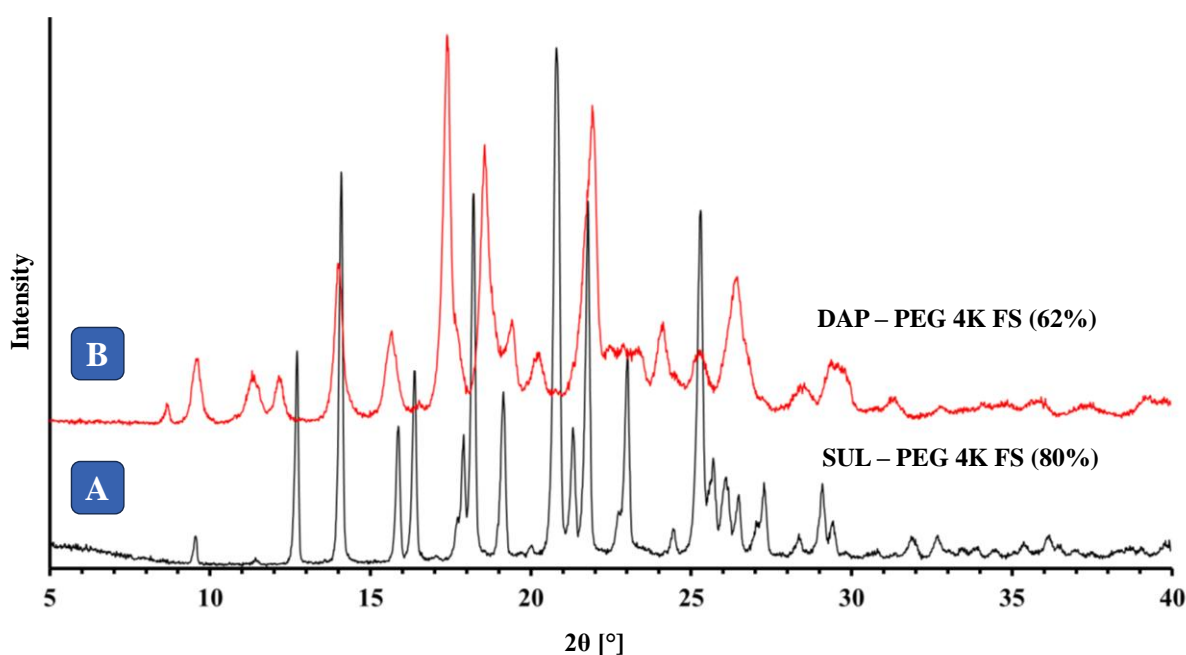


Figure 5.3: Pawley profile fitting of SUL–PEG 600 cocrystal with lattice parameters embedded in it. Red data points indicate experimental data, the blue pattern indicates the calculated pattern and the black pattern indicates the difference plot between experimental and calculated pattern. Green colour lines indicate  $h k l$  reflections.



**Figure 5.4:** NMR study of the SUL–PEG 4K cocrystal: (A)  $^1\text{H}$  NMR spectra in DMSO- $\text{D}_6$  and (B)  $^{13}\text{C}$  NMR spectra in DMSO- $\text{D}_6$ . Structural assignments for both spectra were indicated by labelling with assigned atom numbers. The signals integration in  $^1\text{H}$  NMR spectra for atom no. 1 (2H, SUL) and atom no. 5 & 6 (4H, PEG) is in the ratio of 2.0: 9.9 = 1:2.5, confirming the existence of both species with desired stoichiometry.



**Figure 5.5:** Powder-XRD overlay of (A) SUL–PEG 4K and (B) DAP–PEG 4K cocrystals synthesized using fast solvent (FS) evaporation process.

To evaluate the effect of different PEG molecular weight grades (M.wt = 600 and 4K) and the impact of different processes (slow solvent evaporation, fast solvent evaporation and anti-solvent addition) on the crystallinity of both SUL–PEG and DAP–PEG cocrystalline materials were studied. Surprisingly, no significant differences were observed on the crystallinity of SUL–PEG (600 & 4k) and vice versa in the case of DAP–PEG (600 & 4k). Property modulation of isostructural materials via forming solid solutions or organic alloys is a documented

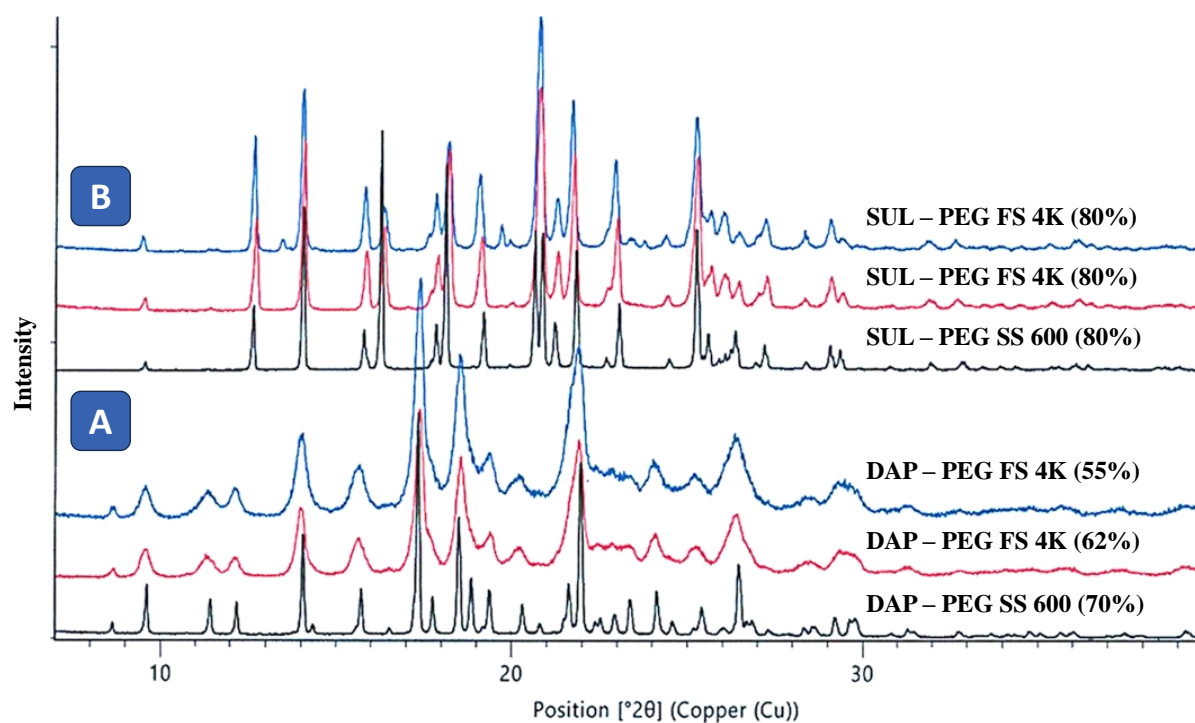
concept<sup>6</sup>, i.e., polymeric cocrystal of a given drug with a given polymer with various grades, forms isostructural cocrystals. The DAP–PEG polymeric cocrystal system demonstrates gradual crystallinity modulation but no such trend is observed for SUL–PEG polymeric cocrystal system (**Figure 5.6 –5.8, Table 5.1**).

From the above observations, it can be inferred that crystallinity can be an extrinsic property for DAP–PEG cocrystals. In contrast, for SUL–PEG cocrystals it is an intrinsic property concerning process change and PEG grades. This anomaly has triggered deep-diving research on their structural properties regarding the drug–polymer interactions and the folding (arrangement and conformational) of the PEG chain. This understanding can decipher the predicting and controlling crystallinity of molecular materials that have applications in a crystal engineering context, as well as process control and formulation in the pharmaceutical industry. In general, intrinsic properties for a material refers to its molecular level attributes<sup>7,8</sup>. Hence, characterisation was further developed by solving crystal structures of SUL - PEG cocrystal for a keen understanding in this direction.

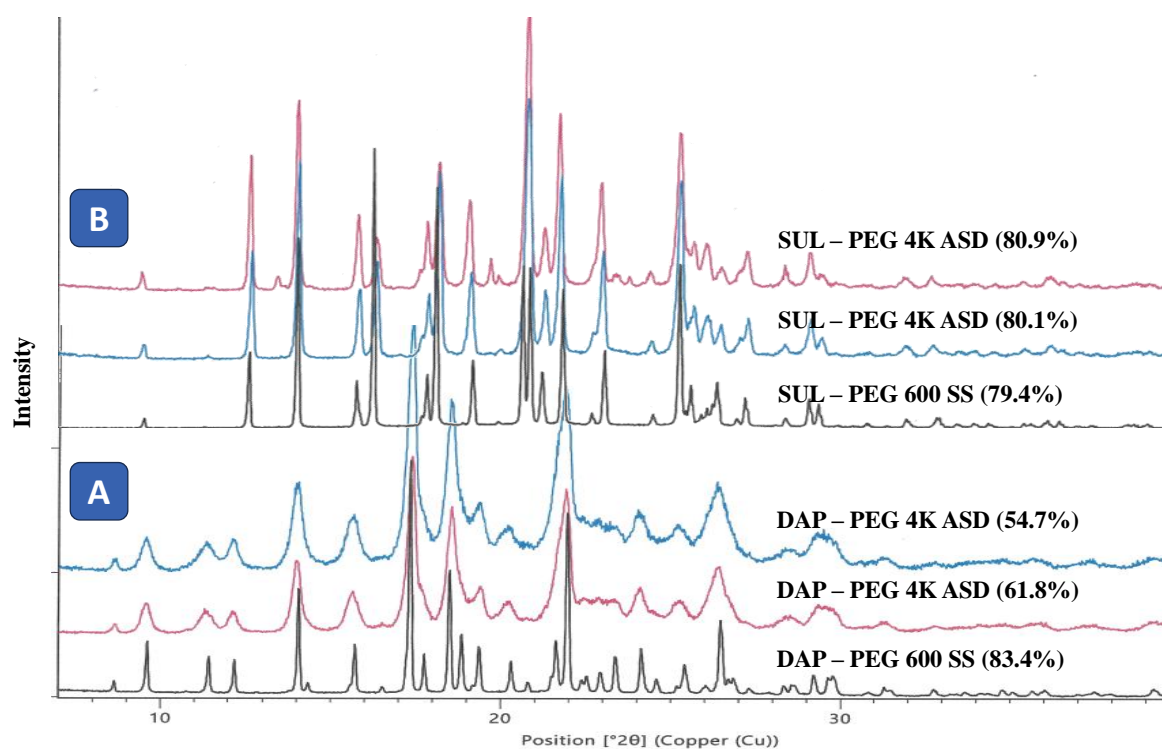
**Table 5.1:** Illustrations on the effects of different PEG molecular weights and processes on the crystallinity of DAP–PEG and SUL–PEG materials.

Drug - PEG cocrystal/ PEG M.wt	Preparation Method/%Crystallinity	
	FS*	ASD*
DAP–PEG 4K	58.5	58.3
SUL–PEG 4K	80.0	80.0
	<b>SS*</b>	
DAP–PEG 600	70.0	
SUL–PEG 600	80.0	
	<b>Solid-Solution (ASD)</b>	
DAP–PEG 4K	78	
DAP–PEG 400/4K (1:1)	68	
DAP–PEG 400	80	
SUL–PEG 4K	80	
SUL–PEG 400/4K (1:1)	80	
SUL–PEG 400	Cocrystal not formed	
*Average of %Crystallinity values from two preparations was considered FS–Fast solvent evaporation; ASD–Antisolvent addition; SS–Slow solvent		



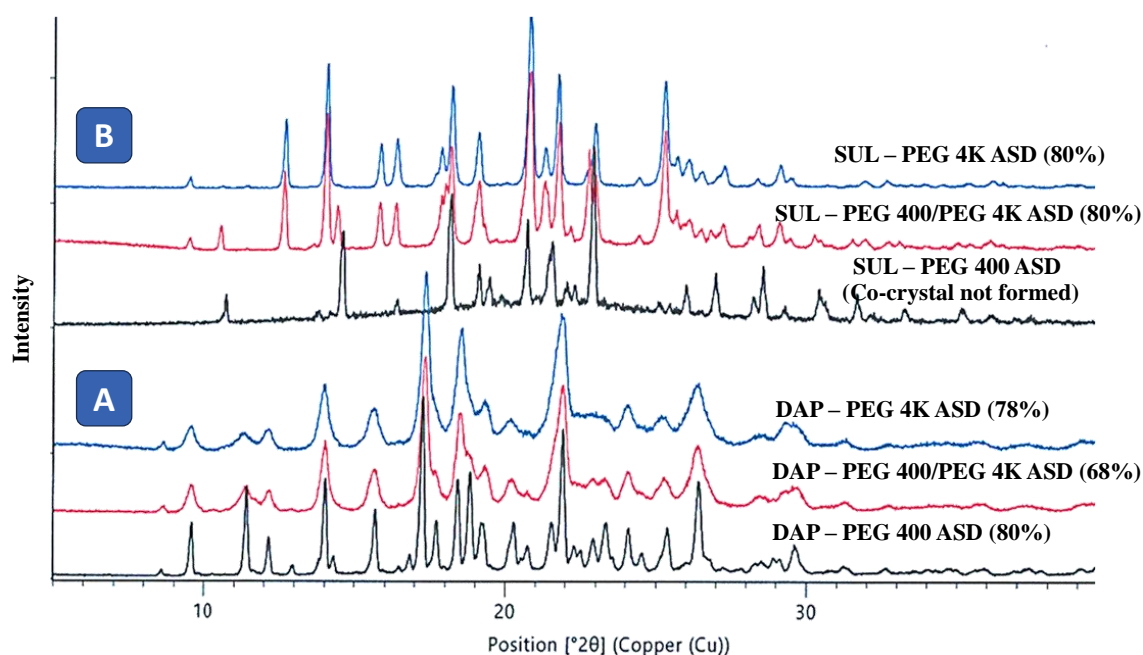


**Figure 5.6:** XRPD overlay of (A) DAP–PEG 600 and DAP–PEG 4K (2 Preparations) and (B) SUL–PEG 600 and SUL–PEG 4K (2 preparations). Here SS indicates that the sample was prepared by slow solvent evaporation method and FS indicates that samples were prepared by fast solvent evaporation method labelled with %crystallinity.



**Figure 5.7:** XRPD overlay of (A) DAP–PEG 600 and DAP–PEG 4K (2 preparations) and (B) SUL–PEG 600 and SUL–PEG 4K (2 preparations). Here SS indicates that the sample was

prepared by the slow solvent evaporation method and ASD indicates that samples were prepared by the anti-solvent addition method labelled with %crystallinity.



**Figure 5.8:** XRPD overlay of (A) DAP-PEG 400, DAP-PEG 400/4K (Solid Solution), DAP-PEG 4K and (B) SUL-PEG 600, SUL-PEG 400/4K (Solid Solution), SUL-PEG 4K. Here ASD indicates that samples were prepared by anti-solvent addition method labelled with %crystallinity. Since SUL and PEG 400 did not form cocrystals, hence %crystallinity was not mentioned.

### 5.3. Crystal structure analysis and insights into structure–property relationship:

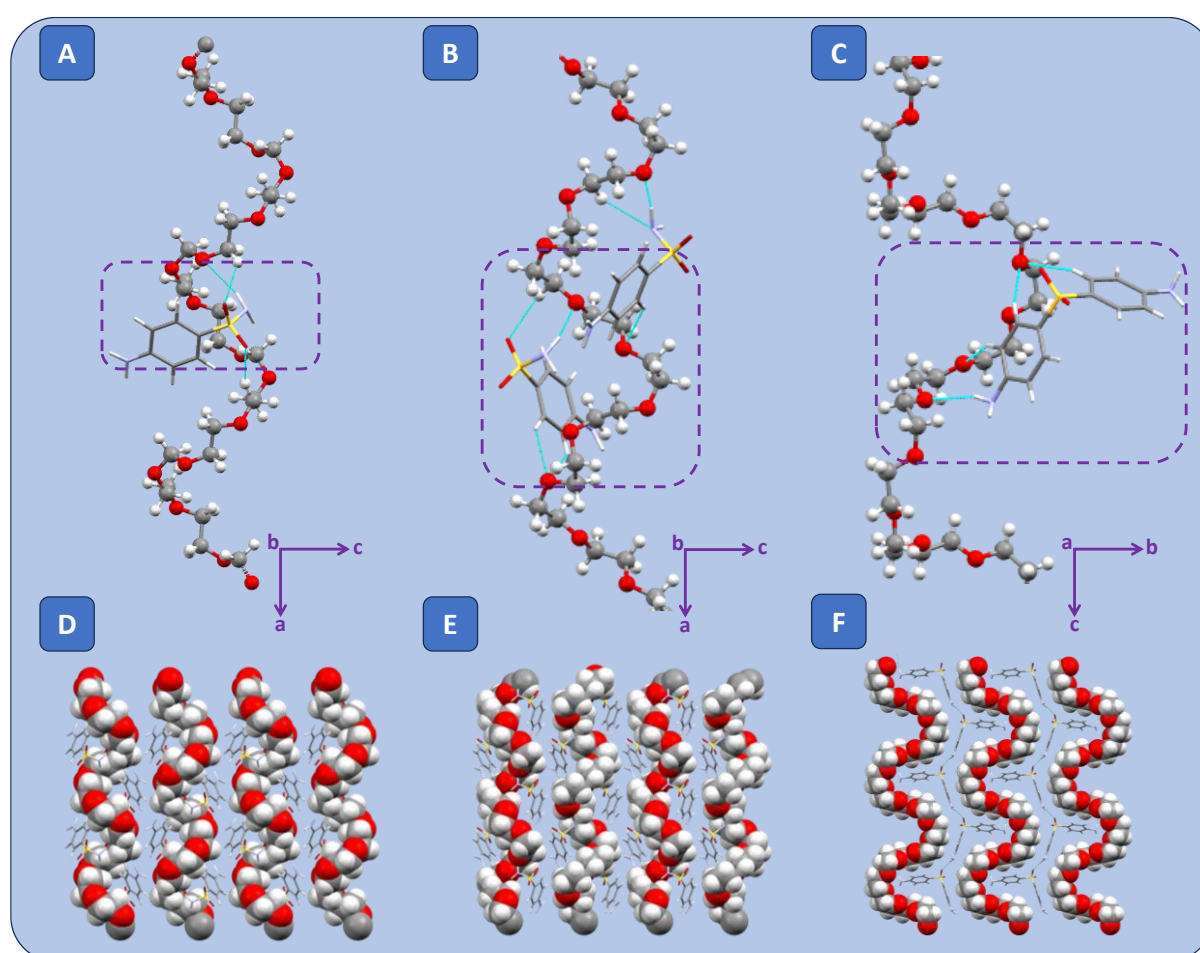
In general, 3-dimensional structural elucidation can be done either by X-ray or electron diffraction techniques. It can interpret molecular structure, crystal packing and intermolecular interactions<sup>9,10</sup>. In this study, the crystal structure was primarily solved by using single crystal XRD (SXRD) toolbox<sup>11</sup> (**Table 5.2**). Preliminary trials for solving the crystal structure of SUL-PEG 600 have resulted in a *Pbca* space group with an Rwp of 13.6%. The asymmetric unit (*Z'*) contains one molecule of SUL and undefinable PEO (Polyethylene Oxide) polymer chain monomers due to intractable disordered PEG chain. Consequently, another solution has popped up with similar unit cell dimensions in chiral space group *Pca2*<sub>1</sub>. Its asymmetric unit (*Z'*) contains two independent SUL molecules and a PEG chain with 5 monomer units (with relatively less disorder), i.e., ratio between SUL and PEG (monomer) becomes 2:5 (1:2.5) which is in agreement with other complementary experimental results (<sup>1</sup>H solution NMR). The possibility of two solutions in centrosymmetric and non-centrosymmetric space groups in SXRD has made to anticipate the existence of local and global order in the crystal structure. The possible leads on this interesting phenomenon will be discussed further<sup>12–14</sup>.

For crystal structure analysis, DFT-D2 optimized crystal structures were considered. The inferences on the hydrogen bonding interactions between the SUL and PEG from the two crystal structure solutions are furnished as follows: The hydrogen bonding interactions between SUL molecules were observed using (*Amine*)  $N-H \cdots O=S$  (*Sulfone*) hetero synthons which have formed an infinite chain of molecules along a–axis and b–axis. Hydrogen bonding interactions between SUL molecule and PEG chain were observed using (*Amine from SUL*)  $N-H \cdots O$  (*PEO*) and (*SUL*)  $C-H \cdots O$  (*PEO*) intermolecular interactions. Here, the PEO chain is running along a–axis and generated via 2–fold axis. It is well known that pure PEO adopts  $7_2$  helices and the PEO chain in SUL–PEG has adopted  $5_2$  helix (**Figure 5.9**). Considering the dapsone and Sufanilamide as molecular analog, upon comparing the crystal structures of SUL–PEG and DAP–PEG, the nature of homomolecular and heteromolecular intermolecular interactions observed in the DAP–PEG and SUL–PEG crystal structures are observed similar.

**Table 5.2:** Crystallographic and DFT-D2 optimization details of SUL–PEG cocrystal

	<b>Solution-I</b>	<b>Solution-II</b>
<b>a, b, c (Å)</b>	10.7497(2), 13.8725(3), 18.1229(3)	13.8725(3), 10.7497(2), 18.1229(3)
<b>Volume (Å<sup>3</sup>)</b>	2705.5(9)	2702.5(9)
<b>Calculated density (g cm<sup>-3</sup>)</b>	1.29	1.29
<b>Z, Z'</b>	8, 1	4, 2
<b>Chemical formula</b>	C <sub>5</sub> H <sub>10</sub> O <sub>2.5</sub> , C <sub>6</sub> H <sub>8</sub> N <sub>2</sub> O <sub>2</sub> S	C <sub>10</sub> H <sub>20</sub> O <sub>5</sub> , 2(C <sub>6</sub> H <sub>8</sub> N <sub>2</sub> O <sub>2</sub> S)
<b>Crystal system, space group</b>	Orthorhombic, <i>Pbca</i>	Orthorhombic, <i>Pca2</i> <sub>1</sub>
<b>Data Collection</b>		
<b>Radiation type</b>	CuK $\alpha$ (1.5418 Å)	
<b>Temperature (K)</b>	293	
<b>Data collection diffractometer</b>	RIGAKU OXFORD/XtaLAB PRO	
<b>Refinement</b>		
<b>Data reduction</b>	CrysAlisPro <sup>15</sup>	
<b>Structure solution</b>	ShelXT <sup>16</sup>	
<b>Structure refinement</b>	ShelXL <sup>17</sup>	
<b>Reflections collected</b>	8870	9020
<b>No of unique data</b>	2683	4685
<b>R1, I &gt; 2<math>\sigma</math> (I)</b>	0.1366	0.2930
<b>wR2, I &gt; 2<math>\sigma</math> (I)</b>	5.059	6.881
<b>Goodness of fit, S</b>	2.556	3.836
<b>Annexures</b>	Annexure 6 & 8	Annexure 7 & 9
<b>Crystal structure optimization</b>		

<b>Software suite</b>	Materials Studio, CASTEP <sup>5</sup>
<b>Energy minimization method</b>	DFT-D2 <sup>18</sup>
<b>Basis set</b>	GGE-PBE <sup>19</sup>
<b>Plane wave cut off energy (eV)</b>	780
<b>SCF convergence (eV)</b>	$1.0 \times 10^{-6}$
<b>Max no of steps</b>	1000
<b>Cell optimization parameters</b>	Unit cell fixed and relaxed atom positions



**Figure 5.9:** Asymmetric units (indicated with dotted boxes) of (A) SUL–PEG (*Pbca*), (B) SUL–PEG (*Pca2<sub>1</sub>*) and (C) DAP–PEG crystal structures showing hydrogen bonding interactions between Drug and PEG chain. Overall packing of interdigitated chains of PEG & drug molecules with hydrogen bonding interactions for (D) SUL–PEG (*Pbca*), (E) SUL–PEG (*Pca2<sub>1</sub>*) and (F) DAP–PEG.

The stoichiometric ratio for SUL–PEG was observed as 2:5 (1:2.5), whereas in DAP–PEG, the stoichiometric ratio was observed as 1:4. This significant difference in the stoichiometric ratio between SUL–PEG and DAP–PEG has probed an interesting behaviour difference observed concerning crystallinity against the molecular weight of PEG discussed in the preceding section. For each molecule of API in DAP–PEG, four monomers of PEG were observed in its asymmetric unit. Whereas for SUL–PEG, only 2.5 monomer ratio was observed for each molecule of SUL. The spatial distribution of the long chain in the DAP - PEG might have induced more flexibility in the PEG chain than the SUL–PEG. Hence, the crystallinity of SUL–PEG remained independent of the PEG grades and the synthesis procedure of the materials (**Figure 5.9**).

It is to be noted that just from SCXRD data, it can be complicated to distinguish between these (*Pbca* and *Pca2<sub>1</sub>*) two space groups due to equal intensities of Friedel Pairs  $I_{hkl}$  and  $I_{-h-k-l}$ <sup>20</sup>. Second Harmonic Generation (SHG) is a second-order non-linear optical technique forbidden for centrosymmetric systems. Non-centrosymmetric packing is a prerequisite for the observation of SHG signals<sup>21</sup>. Therefore, this measurement can clarify the ambiguity between the centrosymmetric and non-centrosymmetric space groups.

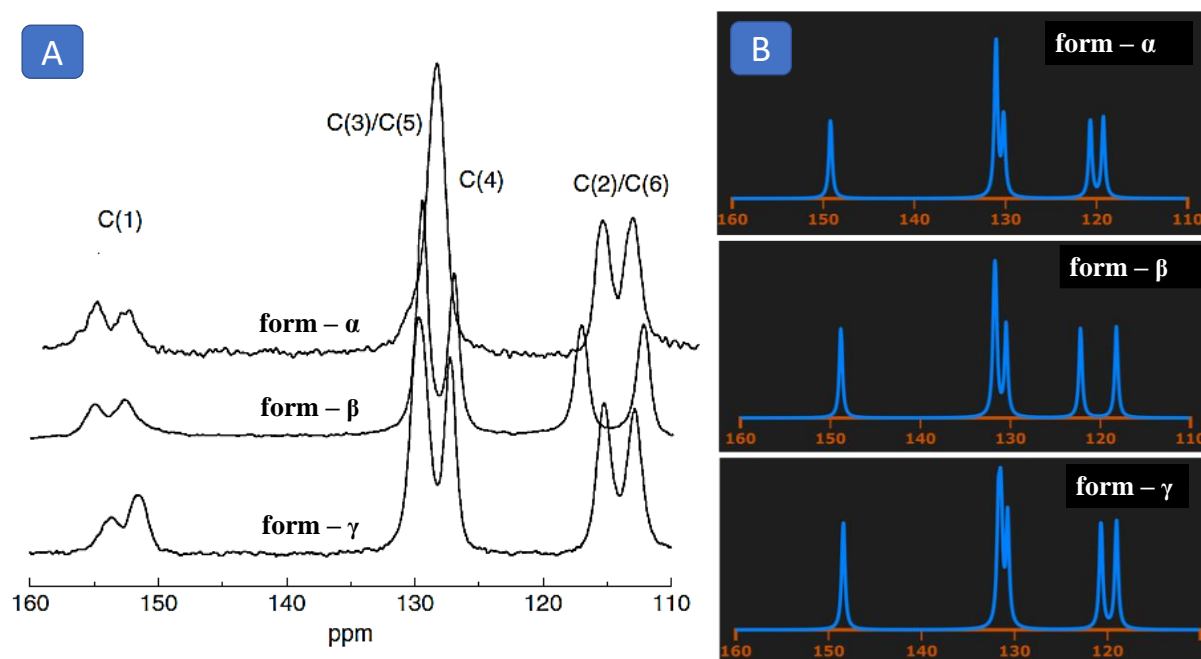
Powder SHG efficiency was measured using the Kurtz-Perry powder technique<sup>22</sup>. The powder sample is first densely packed in a capillary tubes and illuminated using Q-switched Nd:YAG laser (pulse width of 8 ns and a repetition rate of 10 Hz) having an energy of 4-5 mJ/pulse. The amplitude of the SHG output voltage was measured using a photomultiplier and a digitalizing oscilloscope assembly. The second harmonic response from the samples was confirmed by measuring the signal from the samples at 532 nm, the second harmonic of 1064 nm (fundamental wavelength). The SHG efficiencies reported are relative to the standard NLO material, urea and dihydrogen phosphate (KDP).

SHG data for SUL–PEG cocrystal is directional towards non-centrosymmetric arrangements (**Table 5.3**). Few reported findings from the literature on small molecular complexes has disclosed the scope on the presence of non-centrosymmetric regions within centrosymmetric crystal packing, which can be the origin of such SHG signals<sup>23</sup>. However, in recent years, the SHG effect from centrosymmetric crystals with vicinal faces has also been reported<sup>21</sup>. A strategic combined approach combining SS-NMR experimental and computational simulation brings forth further underpinning for a crystal structure model with local and global order<sup>12-14,24,25</sup>.

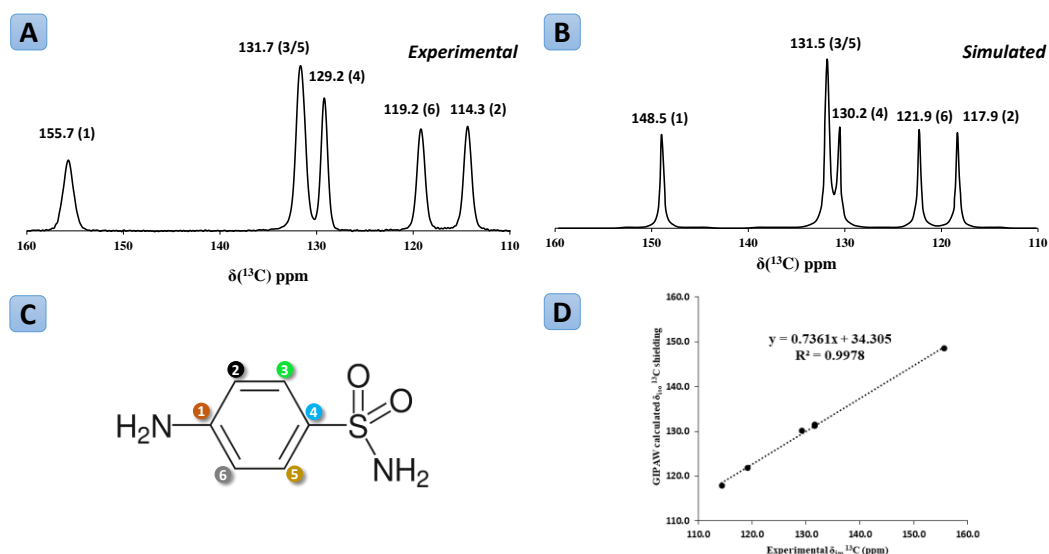
**Table 5.3:** Summary of second harmonic generator (SHG) results

S.No.	Sample	SHG signal (mV)
1.	SUL–PEG 1450	200
2.	Urea	200

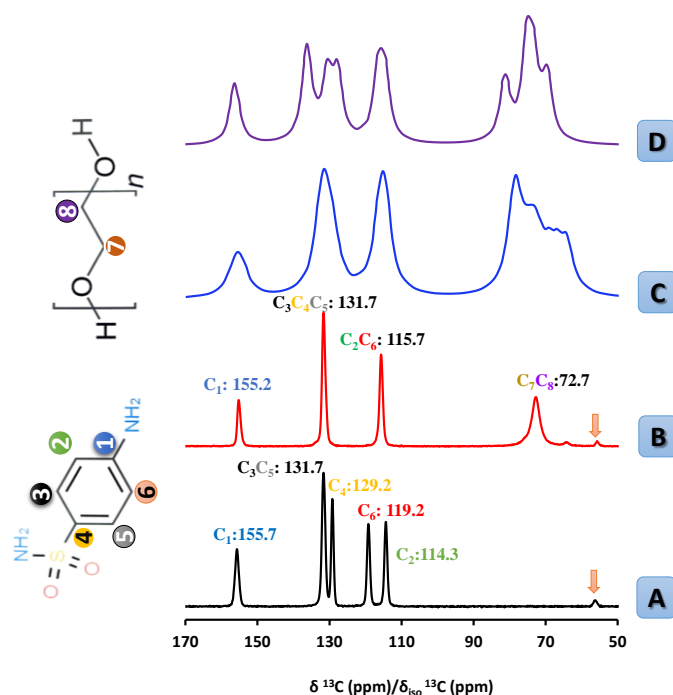
X-ray diffraction studies can evaluate the study of crystal structures at long range or global order. However, the local order cannot be accessible by X-ray diffraction studies. Typical techniques such as Electron diffraction or SS  $^{13}\text{C}$  NMR can allow direct study of local order<sup>26</sup>. For interpretation and evaluation of the splitting of peaks in the SS  $^{13}\text{C}$  NMR spectra, analysis of SS  $^{13}\text{C}$  NMR spectra for form- $\alpha$ , form- $\beta$  and form- $\gamma$  of SUL were taken as the starting point to draw benchmarking points for further evaluation. Also, SS  $^{13}\text{C}$  NMR spectra were collected for SUL–PEG and as a control SUL (form- $\beta$ ) sample was also analysed. For accurate assignment of C atoms to the corresponding peaks in the experimental spectra of SUL, SS  $^{13}\text{C}$  NMR spectra of form- $\alpha$ , form- $\beta$  and form- $\gamma$  were calculated (**Figure 5.10**). To confirm the correctness of simulated patterns from three sulfanilamide polymorphic forms, form- $\beta$  was considered as representative and compared its simulated spectrum with experimental spectra. Experimental and simulated form- $\beta$  patterns are compared and agree with each other (**Figure 5.11**). Comparing the SS  $^{13}\text{C}$  NMR spectra of three polymorphic forms, the main difference is the extent of peaks splitting (i.e C2/C6 & C3/C5). As per reference, the splitting in heights occurs due to N-atoms relaxation time/mobility and shielding asymmetry/anisotropy<sup>27</sup>. The peak splitting in the experimental and simulated patterns are in agreement with each other.



**Figure 5.10:** (A) Reported and (B) Simulated SS  $^{13}\text{C}$  NMR patterns of form- $\alpha$ , form- $\beta$  and form- $\gamma$  SUL crystal structures<sup>2</sup>.



**Figure 5.11:** (A) Experimental and (B) Simulated  $^{13}\text{C}$  SS-NMR patterns of form- $\beta$  SUL crystal structures, (C) sulfanilamide molecular structure with C-atoms assignment and (D) Plots between experimental  $^{13}\text{C}$  chemical shifts and simulated  $^{13}\text{C}$  chemical shifts for form- $\beta$  SUL. The  $R^2$  fit represents the correlation of chemical shift value of experimental and simulated.



**Figure 5.12:** (A) SUL ( $\beta$ -polymorph) experimental  $^{13}\text{C}$  SS-NMR spectra with peak assignment and read arrow indicate a spinning side band (B) SUL-PEG experimental  $^{13}\text{C}$  SS-NMR spectra with peak assignment and read arrow indicate a spinning side band; (C) SUL-PEG simulated  $^{13}\text{C}$  SS-NMR spectra of local crystal structure that solved in  $Pca2_1$  space group. (D) SUL-PEG simulated  $^{13}\text{C}$  SS-NMR spectra of average crystal structure involved in the  $Pbca$  space group.

For SUL-PEG, peaks corresponding to SUL in  $^{13}\text{C}$  SS NMR experimental spectrum were observed at 155.2, 131.7 and 115.7 ppm and a relatively broad peak at 72.7 ppm can be assigned to PEG<sup>27</sup>. Among all C atoms, the more deshielded C-atom is  $\text{C}_1$ , which is directly connected to  $-\text{NH}_2$  group and upfield change (0.5 ppm) is observed for SUL-PEG compared to SUL<sup>28</sup> (Figure 5.12). This chemical shift can be attributed to hydrogen-bonding interactions<sup>29</sup> around the  $-\text{NH}_2$  group. Other distinct differences are noticed for the overlap of phenylene doublet of  $\text{C}_2$ - $\text{C}_6$  and  $\text{C}_3$ / $\text{C}_5$ - $\text{C}_4$  in the SUL-PEG spectrum, In contrast, SUL shows a precise splitting with 4.9 ppm and 2.5 ppm respectively.

As discussed, shielding in asymmetry/anisotropy i.e., symmetry in the hydrogen bonding interactions is one of the significant parameters influencing splitting in the spectra. The extent of signal splitting, will increase with an increase in symmetry in hydrogen bonding interactions. To understand further, intermolecular interactions between the SUL molecules were measured and summarized (Table 5.4). A comparison of symmetry in the intermolecular interactions with splitting peaks of corresponding splitting of peaks of the corresponding C atoms has led to the following inferences.



Higher asymmetry was observed in intermolecular hydrogen bonds SUL (form-β) which explains higher resonance splitting for corresponding ‘C’ atoms and similarity relatively lesser asymmetric intermolecular interactions in SUL (form-α), which explains relatively lower splitting compared to SUL (form-β). Hence, higher symmetrical intermolecular interactions in SUL-PEG (*Pca2<sub>1</sub>* & *Pbca*) explains lesser extent of splitting resonance for the corresponding ‘C’ atoms. Similar analysis was also performed for DAP and DAP - PEG (*Pbca*) to strengthen this hypothesis (Refer Chapter 4). In DAP, the two amino groups have participated in other intermolecular interactions, reflecting the greater extent of resonance splitting in peaks. Similarly, in DAP-PEG, although the intermolecular hydrogen bonds exhibited by the amine groups in dapson are similar, the strength of their interactions is significantly different. Hence, moderate resonance of peaks was observed. The above analysis indicates that nature of intermolecular hydrogen bonds between the SUL molecules is reflected in the <sup>13</sup>C SS-NMR spectra of materials. Hence, Further analysis of <sup>13</sup>C SS-NMR will be continued by taking SUL molecules as basis.

**Table 5.4:** Composition of hydrogen bonding interaction and <sup>13</sup>C SS-NMR spectra (Geometry Optimized Structure).

Crystal Structure	Hydrogen Bonding Interactions	Hydrogen Bonding Matrices	Comments
SUL ( <i>β-P2<sub>1</sub>/n</i> )	(C <sub>1</sub> )-N-H···O	d: 2.035Å and θ: 162°	The asymmetry of hydrogen bonding interactions for C <sub>1</sub> -NH <sub>2</sub> group can explain the greater extent of splitting resonance for C <sub>2</sub> and C <sub>6</sub> .
	(S)-N-H···O	d: 2.019Å and θ: 156°	
	(S)-N-H···O	d: 2.019Å and θ: 165°	
	(C <sub>2</sub> )-H···O	d: 2.542Å and θ: 137°	
SUL ( <i>α-Pbca</i> )	(C <sub>1</sub> )-N-H···O	d: 2.061Å and θ: 168°	Lesser asymmetry of hydrogen bonding interactions for -NH <sub>2</sub> group can explain the lower extent of splitting resonance for C <sub>2</sub> and C <sub>6</sub> .
	(C <sub>1</sub> )-N-H···N	d: 2.331Å and θ: 160°	
	(S)-N-H···O	d: 2.655Å and θ: 125°	
	(S)-N-H···O	d: 1.961Å and θ: 174°	
SUL-PEG ( <i>Pca2<sub>1</sub></i> )	(C <sub>2</sub> )-H···O	d: 2.575Å and θ: 121°	The symmetry of hydrogen bonding interactions for -NH <sub>2</sub> group can explain the resonance overlap for C <sub>2</sub> and C <sub>6</sub> .
	(C <sub>3</sub> )-H···O	d: 2.503Å and θ: 124°	
	(S <sub>2</sub> )-O···H	d: 2.558Å and θ: 124°	
	(S)-N-H···O	d: 1.959Å and θ: 166°	
	(C <sub>5</sub> )-H···O	d: 2.437Å and θ: 127°	
	(S)-N-H···O	d: 1.892Å and θ: 172°	
	(S)-N···H-C(PEG)	d: 2.721Å and θ: 135°	
	(S)-N-H···O	d: 2.159Å and θ: 153°	

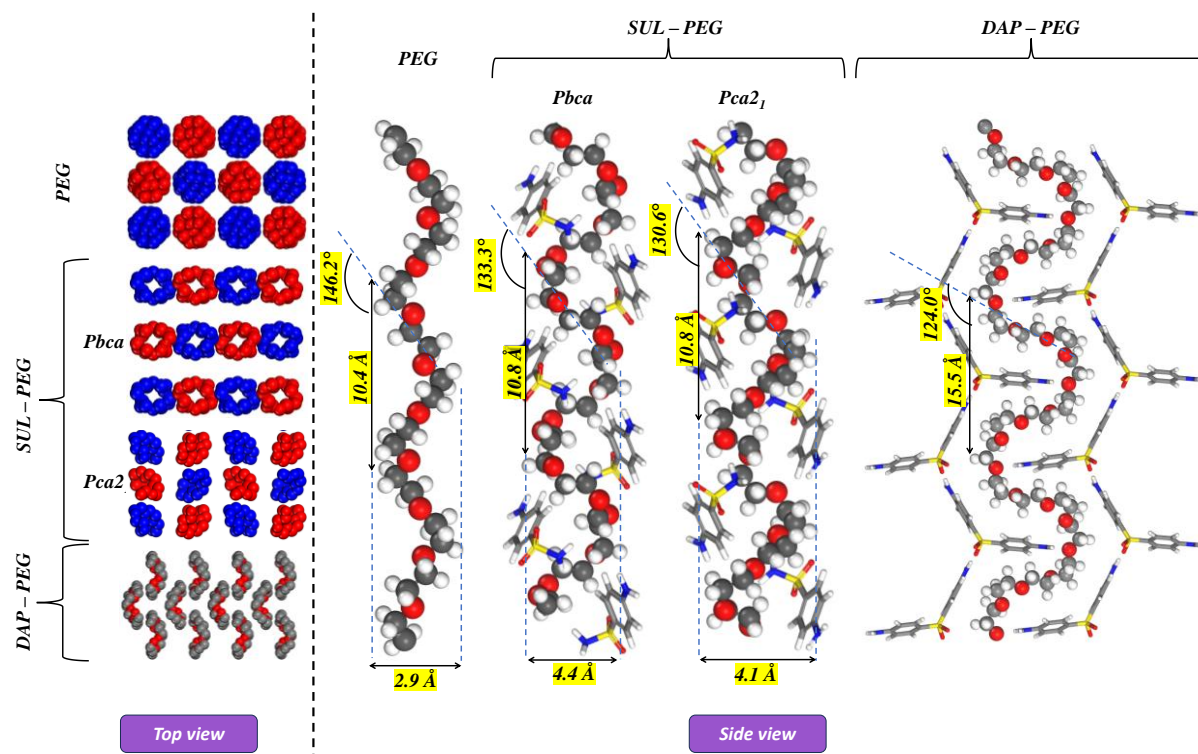
	(C1)-N-H...O=S	d: 1.918Å and $\theta$ : 167°	
	(C1)-N-H...O=S	d: 1.882Å and $\theta$ : 168°	
SUL-PEG ( <i>Pbca</i> )	(C <sub>1</sub> )-N-H...O	d: 2.672Å and $\theta$ : 137°	Symmetry of hydrogen bonding interactions for the –NH <sub>2</sub> group can explain the resonance overlap for C <sub>2</sub> and C <sub>6</sub> .
	O...H-C (PEG)	d: 2.466Å and $\theta$ : 130°	
	O...H-C (PEG)	d: 2.543Å and $\theta$ : 153°	
	(S)-N-H...O (PEG)	d: 1.923Å and $\theta$ : 169°	
DAP ( <i>Anhydrous</i> )	(C <sub>1</sub> )-N-H...H (N)	d: 2.268Å and $\theta$ : 161°	Both the –NH <sub>2</sub> groups do not display similar hydrogen bonding. Asymmetry of hydrogen bonding interactions for the –NH <sub>2</sub> group can explain the more splitting in all C-atoms.
	(C <sub>1</sub> )-N-H...O (S)	d: 1.927Å and $\theta$ : 159°	
	(C <sub>12</sub> )-N-H...O (S)	d: 2.259Å and $\theta$ : 165°	
	(C <sub>12</sub> )-N-H	No H-Bonding	
DAP-PEG ( <i>Pbca</i> )	(C <sub>1</sub> )-N-H...O (PEG)	d: 2.351Å and $\theta$ : 158°	Both the –NH <sub>2</sub> groups display similar hydrogen bonding interactions but the strength of interactions is different. Moderately asymmetry of hydrogen bonding interactions for –NH <sub>2</sub> group can explain the lesser extent of splitting in all C-atoms.
	(C <sub>10</sub> )-N-H...O (PEG)	d: 2.625Å and $\theta$ : 142°	
	C <sub>10</sub> -H...O (PEG)	d: 2.598Å and $\theta$ : 148°	
	C <sub>11</sub> -H...O (PEG)	d: 2.543Å and $\theta$ : 154°	
	C <sub>16</sub> -H...O (PEG)	d: 2.665Å and $\theta$ : 160°	
	(C <sub>1</sub> )-N-H...O=S	d: 2.133Å and $\theta$ : 171°	
	(C <sub>11</sub> )-N-H...O=S	d: 2.640Å and $\theta$ : 137°	

As per the crystal structures evaluation, it was observed that intermolecular interactions between the SUL molecules are dictated by sulfone–amine interactions chain motifs. As per statistical analysis, Achiral molecules containing Amide chain shows 11.1% probability for non-centrosymmetric space group and 74.8% for centrosymmetric. But if  $Z' > 1$ , the frequency of the centrosymmetric space group will decrease to 2.7% and the frequency of non-centrosymmetric will increase to 7.2%. i.e., the frequency of crystal structures with amide chains having high  $Z'$  structures is very low and hence these results indicate that such motifs can perturb general space group statistics<sup>30</sup>.

As per the reference article<sup>30</sup>, it was stated that Amide also forms inversion symmetry-mediated dimers whereas chain motifs can be created through glide planes, screw axes or translational symmetry. Analysis of the symmetry operator used to propagate the chain motif showed that the glide plane was the most popular propagator (31.5%) followed closely by translation (29.3%), then the 2<sub>1</sub>-screw axis (23.6%). Similarly, structural analogue was also observed in the SUL-PEG crystal structure, where continuous chain motifs were observed between SUL

molecules via Sulfone and Amine functional groups propagated through glide planes and a  $2_1$ -screw axis. Thus, the propensity or competition of the chain motif influenced by the  $Pbca$  is higher than the  $Pca2_1$  space group. Hence, it is appropriate to consider the sulfone–amine interactions between the SUL molecules which are dictating the scope for the existence of crystal symmetry in  $Pbca$  and  $Pca2_1$  space groups which may be indicating the presence of the local and global order in the materials<sup>30</sup>.

From the above analysis, the following hypothesis can be put forward.  $Pbca$  crystal model with higher symmetry may be taken as descriptive of the order in the long-range, or other terms, it may describe an average structure. The  $Pca2_1$  model presents symmetry breaking and it is probably a descriptive of the local situation of the order; that is, the symmetry, locally, is broken. In such a situation, symmetry breaking owing to the presence of disorder is very commonplace for polymers. SHG signal of SUL–PEG system can be a clear support of this hypothesis however further elaboration is required for complete understanding this noble phenomena.



**Figure 5.13:** Figure showing top view and side view of PEG chains in (A) PEG (*As such*), (B) SUL–PEG ( $Pbca$ ), (C) SUL–PEG ( $Pca2_1$ ) and (D) DAP–PEG. In top view, right-handed helices were indicated in blue and left-handed helices were marked in red. PEG chain helix parameters such as helix pitch, helix pitch angle and helix diameter were indicated. The diameter of the Pseudo helix PEG chain was not calculated for DAP - PEG.

Apart from sulfone and amine interactions between SUL molecules, conformations and arrangements of PEG chains in the crystal structures of SUL–PEG can also provide substantial information to establish structure–property relationship related to peculiar phenomena of crystals existing with local and global order. Comparing the helix shape parameters and arrangement of PEG chains in ‘As such PEG’, SUL–PEG (*Pbca* and *Pca2<sub>1</sub>*) and DAP–PEG are measured and compared, it can be observed that the helical shape and arrangement of PEG chains parameters in SUL–PEG are similar to ‘As such PEG’ with minimal conformational and helical shape deformations. From reference paper<sup>31</sup>, it can be inferred that the molecular weight of PEG does not significantly impact the crystallinity in solid PEG materials (PEG with M.Wt above 1000 Da). Also, it is worth noting that ‘As such PEG’ takes the shape of helical parameters to attain minimum energy structure. Similarly, PEG in SUL–PEG also mimics the conformation of the ‘As Such PEG’ chain to attain structure with minimum energy (**Table 5.4**). The arrangement and shape of ‘As such PEG’ is retained in the SUL–PEG crystal structure having helically shaped chains arranged parallelly with minimal differences in the helical conformations. **Table 5.5** shows that the pitch angle values of PEG chain in PEG (As Such) and SUL–PEG are closer compared to the measurement values of DAP–PEG. Hence, the impact of PEG molecular weight was not observed on the crystallinity of SUL–PEG (**Figure 5.13**). But in DAP–PEG crystal structure, PEG exhibits a pseudo helical structure away from minimum structure energy. In the case of DAP–PEG, the pseudo helical conformation of PEG has two vectors perpendicular to the chain axis.

The difference in the pitch and pitch angle parameters for PEG (As such) and DAP–PEG indicates the same (**Table 5**). Such PEG arrangements can lead to the up and down disorders that comes under limit ordered disorder. This phenomenon is evident from the powder–XRD data of SUL–PEG and DAP–PEG materials that were prepared using PEG of the same molecular weight and process of crystallinity. From powder–XRD patterns it is evident that the DAP–PEG pattern exhibits broad peaks, whereas SUL–PEG shows sharp peaks.

**Table 5.5:** Comparison table on Helix parameters of PEG (As such), SUL–PEG (*Pbca*), SUL–PEG (*Pca2<sub>1</sub>*) and DAP–PEG.

Helix Parameter	PEG ( <i>As such</i> )	SUL–PEG ( <i>Pbca</i> )	SUL–PEG ( <i>Pca2<sub>1</sub></i> )	DAP–PEG
Pitch (Å)	10.4	10.8	10.8	15.5
Pitch Angle (°)	146.2	133.3	130.6	124.0
Pitch Diameter (Å)	2.9	4.4	4.1	NA

Based on the literature review on systems exhibiting local and global order such as polymer crystallography, statistical analysis on relevant symmetries and type of intermolecular interactions in crystal structures, many structural aspects of PEG chains and corresponding arrangements in SUL–PEG can reveal clues in understanding the existence of SUL–PEG with local and global order in the crystal lattice. In polymer systems, the molecular conformations and intermolecular interactions are determinants of structural parameters in the crystal structure. The following are the fundamental structural aspects of crystalline polymer materials<sup>26</sup>.

- A crystallographic polymer must be stereoregular.
- The chain must be parallel to each other.
- The symmetry of the polymer chains must be maintained in the crystal lattice such that local symmetry becomes crystallographic symmetry.

Interestingly, the PEG chains in the SUL–PEG exhibit all above-mentioned properties in the crystal structure (**Figure 5.12**). Also, the conformational and structural arrangement of PEG in the asymmetric unit of the crystal structures are determinants of the chiral nature of the structural arrangement<sup>26</sup>. Hence to examine the chirality, it is also appropriate to structurally evaluate the PEG polymer chains<sup>32,33</sup>. Therefore, the concepts relevant to the polymer crystallography were adopted for the structural deduction of SUL–PEG. In SUL–PEG, PEG helical chains are running along *c*-axis in opposite directions. This kind of arrangement refers to chiral crystallization in polymers<sup>26</sup>. It is also well-known phenomenon in crystalline polymers, where chemical repeats are achiral and the crystallization state is chiral<sup>26</sup>. These achiral polymer structures resemble as packing of enantiomorphs helical chains<sup>25,34–38</sup>.

In SUL–PEG, the circumstances of solving crystal structures in *Pbca* and *Pca2<sub>1</sub>* can also be explained as symmetry-breaking phenomenon which will adhere to the study of local arrangement in polymer chains<sup>26</sup>. In general, such interesting phenomenon in polymer arises which above postulated rules of ideal crystal structures weaken. Several such polymers structures have been reported in the literature where deviations in the ideal crystal structures due to symmetry breaking or disorders in polymer crystals exhibited<sup>26</sup>. As per this article<sup>32</sup>, we know another interesting phenomenon on “Polymorphs with fractional composition (PFC)” as a common property in polymorphs. Most competing PFCs are neither enantiopure or racemic, but with mixtures of left-handed and right-handed with different compositions. SUL–PEG being a cocrystal formed by a polymer and small molecule, there might be a change that PFC existence is leading to two different domains in the same crystal structure.

Phenomenon was also reported for the syndiotactic polypropylene Form–I, which is a structural analogue example to the current scenario<sup>39</sup>. The ideal ordered structure of these polymer chains in the two-fold helix are packed in orthorhombic crystal system with a space group of *Ibca*. The space group defined local order corresponds to lattice having low symmetry with local 2-fold rotation symmetry axis of helical chain (*P2<sub>1</sub>/a*). A similar phenomenon was already reported by G. DesiRaju in quinone-hydroquinone Complexes<sup>23</sup>. In the current work, observing such structural properties in the drug–polymer system (SUL–PEG)<sup>25,36,38</sup>.

As per the article<sup>32</sup>, oligomers exhibiting strong heterogeneous interactions will show a tendency in resolution of racemic mixtures during crystallization and such materials will crystallize better in enantiopure than forming racemic mixtures. In SUL–PEG, the asymmetric unit of both *Pbca* and *Pca2<sub>1</sub>* crystal structures contains homomolecular and heteromolecular interactions. In *Pca2<sub>1</sub>*, heteromolecular interactions dominate compared to *Pbca* (Table 5.4). As per ref<sup>32</sup>, If hetero interactions dominate during crystallization of molecules, it can promote the resolution of the chirality of the system. Hence, there might be a scope for the formation of two different symmetries in the lattice when crystallized.

### 5.4. Summary

**Chapter 5** is an extension for the virtual screening of drug–polymer cocrystals from **chapter 4**, where SUL–PEG cocrystal is one of the screened outcomes from the developed computational screening methodology for drug–polymer cocrystals. The material was successfully synthesized and characterized by combined experimental and computational tools. During this process, two exciting phenomena were observed:

1. Structural characterization from SXRD, where crystal structure was solved in two different symmetries indicating scope for the existence of multi–domain single crystal having local and global order.
2. A possible hypothesis was build by combining experimental (SC–XRD and <sup>13</sup>C SS–NMR) and computational studies (DFT-D2 optimizations and computing NMR simulations) approaches.
3. Keeping the possibilities in view, there might be a scope for the existence of local order and global order in the SUL–PEG crystal lattice for which this work needs to be further extended.
4. Secondly, while establishing structure–property relationship (%Crystallinity), a phenomenon was observed during comparative studies between SUL–PEG and DAP–PEG concerning %crystallinity.

5. With the alternation in Molecular weight of PEG, the modulation in physical properties (crystallinity) was observed in DAP–PEG cocrystals and it is contradictory to the behaviour of SUL–PEG cocrystals where crystallinity of SUL–PEG cocrystals with different PEG grades does not change.
6. The possible grounds for the above exciting observations on structure–property relationship were stoichiometry between Drug and PEG monomers in the asymmetric unit and helical properties of the PEG in the crystal structure.

### 5.5. References

1. Chappa, P., Maruthapillai, A., Voguri, R., Dey, A., Ghosal, S., & Basha, M. A. Drug–Polymer Co-Crystals of Dapsone and Polyethylene Glycol: An Emerging Subset in Pharmaceutical Co-Crystals. *Cryst. Growth Des.* **18**, 7590–7598 (2018).
2. Portieri, A., Harris, R. K., Fletton, R. A., Lancaster, R. W., & Threlfall, T. L. Effects of polymorphic differences for sulfanilamide, as seen through  $^{13}\text{C}$  and  $^{15}\text{N}$  solid-state NMR, together with shielding calculations. *Magn. Reson. Chem.* **42**, 313–320 (2004).
3. Chappa, P., Maruthapillai, A., Voguri, R., Dey, A., Ghosal, S., & Basha, M. A. Drug–Polymer Co-Crystals of Dapsone and Polyethylene Glycol: An Emerging Subset in Pharmaceutical Co-Crystals. *Cryst. Growth Des.* **18**, 7590–7598 (2018).
4. Chappa, P. Modulating Physicochemical Attributes of Pharmaceutical Co-Crystals: A Case Study Of Drug–Polymer/Small Molecular Systems. (SRM University, 2019).
5. BIOVIA. Dassault Systèmes, Biovia Materials Studio, 2017R2, San Diego: Dassault Systèmes, . *Dassault Systèmes* (2017).
6. Dabros, M., Emery, P. R., & Thalladi, V. R. A supramolecular approach to organic alloys: Cocrystals and three- and four-component solid solutions of 1,4-diazabicyclo[2.2.2]octane and 4-X-phenols (X = Cl, CH<sub>3</sub>, Br). *Angew. Chemie - Int. Ed.* **46**, (2007).
7. Dib, E., Mineva, T., Gaveau, P., & Alonso, B.  $^{14}\text{N}$  solid-state NMR: A sensitive probe of the local order in zeolites. *Phys. Chem. Chem. Phys.* **15**, 18349–18352 (2013).
8. Shi, R. & Tanaka, H. Impact of local symmetry breaking on the physical properties of tetrahedral liquids. *Proc. Natl. Acad. Sci.* **115**, 1980–1985 (2018).
9. Ito, S., White, F. J., Okunishi, E., Aoyama, Y., Yamano, A., Sato, H., Ferrara, J. D., Jasnowski, M., & Meyer, M. Structure determination of small molecule compounds by an electron diffractometer for 3D ED/MicroED. *CrystEngComm* **23**, 8622–8630 (2021).
10. Saha, A., Nia, S. S., & Rodríguez, J. A. Electron Diffraction of 3D Molecular Crystals. *Chem. Rev.* **122**, 13883–13914 (2022).

11. Newman, J. A., Iuzzolino, L., Tan, M., Orth, P., Bruhn, J., & Lee, A. Y. From Powders to Single Crystals: A Crystallographer's Toolbox for Small-Molecule Structure Determination. *Mol. Pharm.* **19**, 2133–2141 (2022).
12. Jo, S., Lee, S., Park, S., Tandon, A., Nguyen, T. B. N., Vu, T. H. N., Raza, M. T., & Park, S. H. Multi-Domains in a Single Lattice Formed by DNA Self-Assembly. *ACS Omega* **7**, 26514–26522 (2022).
13. Zhang, X., Zhao, F., Wu, Y., Yang, J., Han, G. W., Zhao, S., Ishchenko, A., Ye, L., Lin, X., Ding, K., Dharmarajan, V., Griffin, P. R., Gati, C., Nelson, G., Hunter, M. S., Hanson, M. A., Cherezov, V., Stevens, R. C., Tan, W., *et al.* Crystal structure of a multi-domain human smoothed receptor in complex with a super stabilizing ligand. *Nat. Commun.* **2017 81** **8**, 1–10 (2017).
14. Li, Q., Kartikowati, C. W., Horie, S., Ogi, T., Iwaki, T., & Okuyama, K. Correlation between particle size/domain structure and magnetic properties of highly crystalline Fe<sub>3</sub>O<sub>4</sub> nanoparticles. *Sci. Reports* **2017 71** **7**, 1–7 (2017).
15. Agilent. CrysAlisPro Software System. *Technol. UK Ltd, Yarnton, Oxford, UK* **44**, (2014).
16. Sheldrick, G. M. SHELXT - Integrated space-group and crystal-structure determination. *Acta Crystallogr. Sect. A Found. Crystallogr.* **71**, 3–8 (2015).
17. Sheldrick, G. M. Crystal structure refinement with SHELXL. *Acta Crystallogr. Sect. C Struct. Chem.* **71**, 3–8 (2015).
18. Dudenko, D. V., Yates, J. R., Harris, K. D. M., & Brown, S. P. An NMR crystallography DFT-D approach to analyse the role of intermolecular hydrogen bonding and  $\pi$ – $\pi$  interactions in driving cocrystallisation of indomethacin and nicotinamide. *CrystEngComm* **15**, 8797 (2013).
19. Perdew, J. P., Burke, K., & Ernzerhof, M. Erratum: Generalized gradient approximation made simple. *Phys. Rev. Lett.* **78**, 1396 (1997).
20. McCabe, E. E., Jones, I. P., Zhang, D., Hyatt, N. C., & Greaves, C. Crystal structure and electrical characterisation of Bi<sub>2</sub>NbO<sub>5</sub>F: An Aurivillius oxide fluoride. *J. Mater. Chem.* **17**, 1193–1200 (2007).
21. Nalla, V., Medishetty, R., Wang, Y., Bai, Z., Sun, H., Wei, J., & Vittal, J. J. Second harmonic generation from the 'centrosymmetric' crystals. *IUCrJ* **2**, 317–321 (2015).
22. Kurtz, S. K. & Perry, T. T. A Powder Technique for the Evaluation of Nonlinear Optical Materials. *J. Appl. Phys.* **39**, 3798–3813 (1968).
23. Desiraju, G. R., Curtin, D. Y., & Paul, I. C. Structural Studies of 1:1 Quinone-



- Hydroquinone Complexes. *Mol. Cryst. Liq. Cryst.* **52**, 259–266 (1979).
24. Walls, B., Mazilkin, A. A., Mukhamedov, B. O., Ionov, A., Smirnova, I. A., Ponomareva, A. V., Fleischer, K., Kozlovskaya, N. A., Shulyatev, D. A., Abrikosov, I. A., Shvets, I. V., & Bozhko, S. I. Nanodomain structure of single crystalline nickel oxide. *Sci. Reports 2021 111* **11**, 1–10 (2021).
  25. di Gregorio, M. C., Shimon, L. J. W., Brumfeld, V., Houben, L., Lahav, M., & van der Boom, M. E. Emergence of chirality and structural complexity in single crystals at the molecular and morphological levels. *Nat. Commun.* **11**, 380 (2020).
  26. De Rosa, C. Chain Conformation, Crystal Structures, and Structural Disorder in Stereoregular Polymers. in 71–155 (2003). doi:10.1002/0471471895.ch2.
  27. Portieri, A., Harris, R. K., Fletton, R. A., Lancaster, R. W., & Threlfall, T. L. Effects of polymorphic differences for sulfanilamide, as seen through <sup>13</sup>C and <sup>15</sup>N solid-state NMR, together with shielding calculations. *Magn. Reson. Chem.* **42**, 313–320 (2004).
  28. Maddileti, D., Jayabun, S. K., & Nangia, A. Soluble Cocrystals of the Xanthine Oxidase Inhibitor Febuxostat. *Cryst. Growth Des.* **13**, 3188–3196 (2013).
  29. Maddileti, D., Jayabun, S. K., & Nangia, A. Soluble Cocrystals of the Xanthine Oxidase Inhibitor Febuxostat. *Cryst. Growth Des.* **13**, 3188–3196 (2013).
  30. Dey, A. & Pidcock, E. The relevance of chirality in space group analysis: A database study of common hydrogen-bonding motifs and their symmetry preferences. *CrystEngComm* **10**, 1258–1264 (2008).
  31. Johansson, P., Paberit, R., Rilby, E., Göhl, J., Swenson, J., Refaa, Z., & Jansson, H. Cycling stability of poly(ethylene glycol) of six molecular weights: Influence of thermal conditions for energy applications. *ACS Appl. Energy Mater.* **3**, (2020).
  32. Carpenter, J. E. & Grünwald, M. Heterogeneous Interactions Promote Crystallization and Spontaneous Resolution of Chiral Molecules. *J. Am. Chem. Soc.* **142**, 10755–10768 (2020).
  33. Carpenter, J. E. & Grünwald, M. Pre-Nucleation Clusters Predict Crystal Structures in Models of Chiral Molecules. *J. Am. Chem. Soc.* **143**, 21580–21593 (2021).
  34. Yoshimoto, Y., Suzuki, Y., Imai, T., Kimura, S., & Uji, H. Supramolecular chiral emergence in water even after compensating for helix chirality in vesicular helix-peptide-aromatic frameworks. *Polym. J. 2023 558* **55**, 877–883 (2023).
  35. Huang, S., Yu, H., & Li, Q. Supramolecular Chirality Transfer toward Chiral Aggregation: Asymmetric Hierarchical Self-Assembly. *Adv. Sci.* **8**, 2002132 (2021).
  36. Sung, B., De La Cotte, A., & Grelet, E. Chirality-controlled crystallization via screw
-

- dislocations. *Nat. Commun.* 2018 91 **9**, 1–7 (2018).
37. Maeda, K. & Yashima, E. Helical Polyacetylenes Induced via Noncovalent Chiral Interactions and Their Applications as Chiral Materials. *Top. Curr. Chem.* 2017 3754 **375**, 1–31 (2017).
38. Makiguchi, W., Kobayashi, S., Furusho, Y., & Yashima, E. Chiral amplification in double-stranded helical polymers through chiral and achiral amidinium–carboxylate salt bridges. *Polym. J.* 2012 4411 **44**, 1071–1076 (2012).
39. Nguyen, N. Q., Chen, T.-F., & Lo, C.-T. Confined crystallization and chain conformational change in electrospun poly(ethylene oxide) nanofibers. *Polym. J.* **53**, 895–905 (2021).

### **Future perspective**

This thesis aims to combine experimental techniques with computational studies to address the challenging structural characterization of cocrystals where coformer is either a small molecule or a polymer (polyethylene glycol). Also, it includes in-silico screening studies combined with AI/ML studies to discover drug–polymer cocrystals. During this work, multiple challenges in the structural characterization and establishing structure–property relationships were addressed for agomelatine–phosphoric Acid (1:1) molecular complexes and dapsone–polyethylene glycol cocrystal. Established In–silico screening protocol has discovered sulphanilamide–Polyethylene cocrystal. This In–silico screening protocol can be used further to find new drug–polymer cocrystal where polymer can be more sophisticated for targeted and controlled drug delivery purposes. Moreover, sulphanilamide–polyethylene cocrystal’s structural characterization combined with computational studies has allowed the revealing of interesting structure–property phenomena and scope for the existence of multi-domain single crystal with local order and global order in the sulphanilamide–polyethylene crystal lattice for which this work needs to be further extended. The interesting outcomes in the thesis have paved the way for further extension of the present work for drug–polymer systems in particular of interest. The MM/QM calculations pose many limitations to dig deeper into the further exploration of the prediction of structural and property aspects of drug–polymer systems. Coarse grain molecular dynamics is an emerging wing in the area of molecular modeling. As an extension to this work, further research can be carried out by employing Coarse grain molecular dynamics to deduce structure-property relationships for drug–polymer systems. For sulphanilamide–polyethylene cocrystal, the existence of local order and global order in its crystal lattice can be further characterized by advanced techniques such as micro Electron Diffraction (microED) techniques as help us to underpin the local crystallography that would provide a better understanding of such multi-domain single crystal systems to bring in practical real-world applications.

## Annexures

### Annexure 1:

**Table 1:** Fractional atomic coordinates for agomelatine–phosphoric acid (RT–form)

Atom	x	y	z
C1	0.84510(6)	0.38787(52)	0.12855(15)
C2	0.90369(7)	0.54513(51)	0.17217(15)
H3	0.9228	0.68542	0.13626
C4	0.93449(5)	0.51805(37)	0.25696(15)
H5	0.9795	0.63883	0.29033
C6	0.90911(5)	0.33424(33)	0.30329(15)
C7	0.94090(6)	0.30583(42)	0.39104(15)
H8	0.98581	0.42867	0.42332
C9	0.91566(7)	0.12617(57)	0.43459(16)
H10	0.94026	0.10379	0.50204
C11	0.85756(7)	-0.03133(56)	0.39155(16)
H12	0.83825	-0.17464	0.42675
C13	0.82458(6)	-0.01090(41)	0.30605(16)
C14	0.85007(5)	0.17536(33)	0.25942(15)
C15	0.81890(5)	0.20653(38)	0.17121(15)
H16	0.77357	0.09041	0.13741
C17	0.75972(25)	0.28665(173)	-0.00295(16)
H18	0.7474	0.35544	-0.06763
H19	0.7192	0.34082	0.01705
H20	0.7666	0.05143	0.00143
C21	0.76198(6)	-0.17997(41)	0.26299(17)
H22	0.75747	-0.34747	0.30538
H23	0.7636	-0.29113	0.20783
C24	0.69975(6)	0.01018(47)	0.23452(25)
H25	0.69248	0.10105	0.28893
H26	0.70559	0.19045	0.19706
C27	0.61604(8)	-0.13252(70)	0.09901(20)

C28	0.56793(26)	-0.36856(127)	0.05395(26)
H29	0.56588	-0.3832	-0.01022
H30	0.51795	-0.33566	0.05246
H31	0.58692	-0.57538	0.08441
O32	0.81900(8)	0.43309(52)	0.04486(15)
O33	0.63194(18)	0.05665(97)	0.06094(23)
N34	0.64013(6)	-0.14336(44)	0.18392(20)
H35	0.62288	-0.25983	0.20744
P36	0.57454(6)	3.88670(53)	0.85155(12)
O37	0.59195(14)	4.11756(60)	0.91924(16)
O38	0.57716(26)	4.01379(76)	0.77518(13)
O39	0.62250(12)	3.63782(58)	0.87963(28)
O40	0.50670(11)	3.75256(71)	0.83335(30)
H41	0.48625	3.67073	0.79608
H42	0.60257	4.083	0.97159
H43	0.61362	3.46925	0.89388
H42	0.60257	4.083	0.97159
H43	0.61362	3.46925	0.89388

**Table 2:** Bond lengths for agomelatine–phosphoric acid (RT–form)

Atom	Atom	Length (Å)	Bond type
C1	C2	1.424(2)	S
C1	C15	1.385(3)	S
C1	O32	1.360(3)	S
C2	H3	1.091	S
C2	C4	1.372(3)	S
C4	H5	1.093	S
C4	C6	1.426(3)	S
C6	C7	1.420(3)	S
C6	C14	1.435(2)	S
C7	H8	1.093	S
C7	C9	1.378(4)	S

C9	H10	1.092	S
C9	C11	1.415(3)	S
C11	H12	1.094	S
C11	C13	1.385(3)	S
C13	C14	1.438(3)	S
C13	C21	1.513(2)	S
C14	C15	1.427(3)	S
C15	H16	1.089	S
C17	H18	1.098	S
C17	H19	1.102	S
C17	H20	1.104	S
C17	O32	1.425(6)	S
C21	H22	1.102	S
C21	H23	1.101	S
C21	C24	1.545(2)	S
C24	H25	1.101	S
C24	H26	1.099	S
C24	N34	1.458(2)	S
C27	C28	1.520(6)	S
C27	O33	1.229(6)	S
C27	N34	1.365(5)	S
C28	H29	1.1	S
C28	H30	1.1	S
C28	H31	1.101	S
N34	H35	0.85	S
P36	O37	1.530(4)	S
P36	O38	1.473(4)	S
P36	O39	1.517(3)	S
P36	O40	1.535(3)	S
O37	H42	0.863	S
O39	H43	0.867	S
O40	H41	0.738	S

**Table 3:** Bond angles for agomelatine–phosphoric acid (RT–form)

Atom	Atom	Atom	Bond angle (°)
C2	C1	C15	120.4(2)
C2	C1	O32	114.7(2)
C15	C1	O32	125.0(2)
C1	C2	H3	118.1
C1	C2	C4	120.1(2)
H3	C2	C4	121.8
C2	C4	H5	119.9
C2	C4	C6	121.4(2)
H5	C4	C6	118.7
C4	C6	C7	121.5(2)
C4	C6	C14	118.6(2)
C7	C6	C14	119.9(2)
C6	C7	H8	118.8
C6	C7	C9	120.6(2)
H8	C7	C9	120.7
C7	C9	H10	120.5
C7	C9	C11	119.7(2)
H10	C9	C11	119.7
C9	C11	H12	119
C9	C11	C13	122.0(2)
H12	C11	C13	119
C11	C13	C14	119.2(2)
C11	C13	C21	120.0(2)
C14	C13	C21	120.8(2)
C6	C14	C13	118.7(2)
C6	C14	C15	119.0(2)
C13	C14	C15	122.3(2)
C1	C15	C14	120.5(2)
C1	C15	H16	119.8
C14	C15	H16	119.7
H18	C17	H19	110

H18	C17	H20	109.5
H18	C17	O32	105.7
H19	C17	H20	108.1
H19	C17	O32	112.2
H20	C17	O32	111.3
C13	C21	H22	109.5
C13	C21	H23	110.5
C13	C21	C24	112.9(2)
H22	C21	H23	106.8
H22	C21	C24	109
H23	C21	C24	108
C21	C24	H25	110.2
C21	C24	H26	109.8
C21	C24	N34	113.3(2)
H25	C24	H26	107.6
H25	C24	N34	108.3
H26	C24	N34	107.4
C28	C27	O33	121.7(3)
C28	C27	N34	115.9(3)
O33	C27	N34	122.3(3)
C27	C28	H29	108.6
C27	C28	H30	113.9
C27	C28	H31	108.8
H29	C28	H30	109.2
H29	C28	H31	107.1
H30	C28	H31	109
C1	O32	C17	118.0(3)
C24	N34	C27	121.6(2)
C24	N34	H35	119.1
C27	N34	H35	118.8
O37	P36	O38	109.0(2)
O37	P36	O39	110.5(2)
O37	P36	O40	112.5(2)



O38	P36	O39	109.6(2)
O38	P36	O40	110.2(2)
O39	P36	O40	104.9(2)
P36	O37	H42	124.5
P36	O39	H43	125.3
P36	O40	H41	127.6

*Annexure 2:*

**Table 1:** Fractional atomic coordinates for agomelatine–phosphoric acid (HT–form)

Atom	x	y	z
C1	0.34886(17)	0.65653(100)	0.23553(23)
C2	0.40530(17)	0.48325(92)	0.25247(27)
H3	0.42277	0.36071	0.3092
C4	0.43613(11)	0.47397(60)	0.19819(32)
H5	0.47946	0.34113	0.21128
C6	0.41291(8)	0.63527(54)	0.12422(40)
C7	0.44474(13)	0.62606(84)	0.06796(35)
H8	0.48793	0.49146	0.08205
C9	0.42154(21)	0.78448(116)	-0.00308(33)
H10	0.44615	0.77781	-0.04631
C11	0.36565(21)	0.95778(105)	-0.02014(27)
H12	0.34791	1.08388	-0.07658
C13	0.33272(14)	0.97389(67)	0.03261(24)
C14	0.35603(9)	0.81027(51)	0.10704(23)
C15	0.32478(10)	0.81697(69)	0.16441(22)
H16	0.28105	0.94566	0.15146
C17	0.26136(40)	0.78258(296)	0.27636(46)
H18	0.26702	1.00918	0.25939
H19	0.24645	0.78084	0.32896
H20	0.22356	0.67351	0.22542
C21	0.27245(15)	1.15843(61)	0.01095(28)
H22	0.27333	1.33091	-0.03205
H23	0.2706	1.26941	0.06525
C24	0.20997(14)	0.97829(63)	-0.02684(28)
H25	0.20727	0.8839	-0.08514
H26	0.20986	0.79636	0.01364
C27	0.13045(17)	1.20847(112)	0.02098(34)
C28	0.07060(17)	1.39850(115)	-0.00120(52)
H29	0.06109	1.44532	0.05432

H30	0.02733	1.29949	-0.04653
H31	0.08025	1.60893	-0.02463
O32	0.32251(24)	0.64745(145)	0.29366(25)
O33	0.15780(26)	1.11457(198)	0.09016(33)
N34	0.15164(15)	1.14824(57)	-0.04052(27)
H35	0.13318	1.23091	-0.08622
P2	0.91872(11)	0.75705(70)	0.26308(23)
O6	0.87557(20)	0.88617(74)	0.30421(32)
O8	0.93682(31)	1.01791(70)	0.20736(37)
O7	0.88660(10)	0.50226(84)	0.21034(30)
O5	0.97583(21)	0.63465(92)	0.31947(28)
H8P	0.9651	0.98268	0.18605
H7P	0.89288	0.32573	0.20094
H	0.86517	0.75292	0.33062

**Table 2:** Bond lengths for agomelatine–phosphoric acid (HT–form)

Atom	Atom	Length (Å)	Bond type
C1	C2	1.424(6)	D
C1	C15	1.385(5)	S
C1	O32	1.360(7)	S
C2	H3	1.091	S
C2	C4	1.372(7)	S
C4	H5	1.093	S
C4	C6	1.426(7)	D
C6	C7	1.420(9)	S
C6	C14	1.435(3)	S
C7	H8	1.093	S
C7	C9	1.378(7)	D
C9	H10	1.092	S
C9	C11	1.414(7)	S
C11	H12	1.094	S
C11	C13	1.386(7)	D
C13	C14	1.438(5)	S

C13	C21	1.513(4)	S
C14	C15	1.427(6)	D
C15	H16	1.089	S
C17	H18	1.098	S
C17	H19	1.098	S
C17	H20	1.103	S
C17	O32	1.420(1)	S
C21	H22	1.102	S
C21	H23	1.101	S
C21	C24	1.545(4)	S
C24	H25	1.101	S
C24	H26	1.099	S
C24	N34	1.458(4)	S
C27	C28	1.519(6)	S
C27	O33	1.229(8)	D
C27	N34	1.365(8)	S
C28	H29	1.1	S
C28	H30	1.1	S
C28	H31	1.101	S
N34	H35	0.85	S
P2	O6	1.523(7)	S
P2	O8	1.689(7)	S
P2	O7	1.502(5)	S
P2	O5	1.415(5)	D
O6	H	0.85	S
O8	H8P	0.86	S
O7	H7P	0.849	S
C1	C15	1.385(5)	S

**Table 3:** Bond angles for agomelatine–phosphoric acid (HT–form)

Atom	Atom	Atom	Bond angle (°)
C2	C1	C15	120.4(2)
C2	C1	O32	114.7(2)
C2	C1	C15	120.4(4)
C2	C1	O32	114.7(4)
C15	C1	O32	125.0(4)
C1	C2	H3	118.1
C1	C2	C4	120.1(4)
H3	C2	C4	121.9
C2	C4	H5	120
C2	C4	C6	121.4(3)
H5	C4	C6	118.6
C4	C6	C7	121.5(3)
C4	C6	C14	118.6(3)
C7	C6	C14	119.9(3)
C6	C7	H8	118.8
C6	C7	C9	120.5(4)
H8	C7	C9	120.7
C7	C9	H10	120.5
C7	C9	C11	119.8(5)
H10	C9	C11	119.7
C9	C11	H12	119.1
C9	C11	C13	122.0(4)
H12	C11	C13	118.9
C11	C13	C14	119.2(3)
C11	C13	C21	120.0(3)
C14	C13	C21	120.8(3)
C6	C14	C13	118.7(3)
C6	C14	C15	119.0(3)
C13	C14	C15	122.3(3)
C1	C15	C14	120.5(3)
C1	C15	H16	119.8

C14	C15	H16	119.6
H18	C17	H19	109.8
H18	C17	H20	109.4
H18	C17	O32	105.7
H19	C17	H20	108.2
H19	C17	O32	112.2
H20	C17	O32	111.4
C13	C21	H22	109.6
C13	C21	H23	110.5
C13	C21	C24	112.9(3)
H22	C21	H23	106.8
H22	C21	C24	108.9
H23	C21	C24	107.9
C21	C24	H25	110.3
C21	C24	H26	109.8
C21	C24	N34	113.3(3)
H25	C24	H26	107.6
H25	C24	N34	108.3
H26	C24	N34	107.3
C28	C27	O33	121.8(5)
C28	C27	N34	115.9(4)
O33	C27	N34	122.3(5)
C27	C28	H29	108.6
C27	C28	H30	113.9
C27	C28	H31	108.9
H29	C28	H30	109.1
H29	C28	H31	107.2
H30	C28	H31	109
C1	O32	C17	118.0(6)
C24	N34	C27	121.6(4)
C24	N34	H35	119.1
C27	N34	H35	118.8
O6	P2	O8	109.5(3)

O6	P2	O7	111.0(3)
O6	P2	O5	112.3(3)
O8	P2	O7	110.3(3)
O8	P2	O5	110.2(3)
O7	P2	O5	103.5(3)
P2	O6	H	109.3
P2	O8	H8P	119.6
P2	O7	H7P	141.9

*Annexure 3:*

**Table 1:** Fractional atomic coordinates for dapsone-polyethylene glycol (SDPD)

Atom	x	y	z
S1	0.61732(22)	1.39611(37)	0.28747(26)
O2	0.65533(22)	1.36346(43)	0.36833(29)
O3	0.55163(22)	1.34566(41)	0.30709(28)
N4	0.73581(47)	1.19260(43)	-0.01865(46)
N5	0.57461(44)	1.82685(40)	0.31879(45)
C6	0.65243(26)	1.33465(33)	0.19602(30)
C7	0.61263(31)	1.30225(79)	0.12880(45)
C8	0.64011(38)	1.25525(80)	0.05802(48)
C9	0.70768(39)	1.24065(37)	0.05372(40)
C10	0.74748(33)	1.27376(84)	0.12077(56)
C11	0.71995(27)	1.32076(82)	0.19151(52)
C12	0.60462(23)	1.52487(38)	0.29751(25)
C13	0.65190(32)	1.58327(42)	0.33655(64)
C14	0.64196(37)	1.68288(43)	0.34370(67)
C15	0.58478(36)	1.72498(39)	0.31151(37)
C16	0.53775(40)	1.66651(42)	0.27190(63)
C17	0.54773(35)	1.56691(41)	0.26478(60)
C18	0.34758	0.93404	-0.28828
C19	0.32281	0.98837	-0.20989
O20	0.37585	1.03391	-0.1687
C21	0.38059	1.00018	-0.08332
C22	0.42417	1.06799	-0.03191
O23	0.41755	1.16289	-0.06401
C24	0.43741	1.22964	-0.0006
O25	0.3295	1.49682	0.14767
C26	0.31306	1.4968	0.06025
C27	0.37148	1.46693	0.00565
O28	0.36727	1.36731	-0.01305
C29	0.42961	1.33194	-0.03519
H30	0.70409	1.1661	-0.07139



H31	0.78936	1.1818	-0.02208
H32	0.52466	1.85705	0.31331
H33	0.61667	1.87573	0.32992
H34	0.55769	1.31418	0.13172
H35	0.60757	1.22866	0.00357
H36	0.80247	1.26239	0.11771
H37	0.75265	1.3478	0.24561
H38	0.69861	1.54959	0.36258
H39	0.6804	1.73015	0.37578
H40	0.49119	1.70015	0.24537
H41	0.50937	1.51991	0.23221
H42	0.32029	0.86304	-0.29742
H43	0.40227	0.92143	-0.28178
H44	0.2987	0.93547	-0.16363
H45	0.28579	1.04508	-0.23086
H46	0.40248	0.92448	-0.08311
H47	0.32997	0.99803	-0.05318
H48	0.47709	1.04388	-0.03815
H49	0.40912	1.06622	0.03856
H50	0.49069	1.2163	0.01649
H51	0.40596	1.2203	0.0592
H52	0.29673	1.57237	0.04065
H53	0.27153	1.44366	0.04868
H54	0.41842	1.48196	0.04249
H55	0.37148	1.50995	-0.05679
H56	0.435	1.33124	-0.10807
H57	0.46865	1.3805	-0.00599

**Table 2:** Bond lengths for dapsons-polyethylene glycol (SDPD)

Atom	Atom	Length (Å)	Bond type
S1	O2	1.548(6)	D
S1	O3	1.548(7)	D
S1	C6	1.808(6)	S
S1	C12	1.808(7)	S
N4	C9	1.43(1)	S
N4	H30	1.11	S
N4	H31	1.11	S
N5	C15	1.430(8)	S
N5	H32	1.11	S
N5	H33	1.11	S
C6	C7	1.401(9)	D
C6	C11	1.400(8)	S
C7	C8	1.40(1)	S
C7	H34	1.14	S
C8	C9	1.40(1)	D
C8	H35	1.14	S
C9	C10	1.40(1)	S
C10	C11	1.40(1)	D
C10	H36	1.14	S
C11	H37	1.139	S
C12	C13	1.401(9)	S
C12	C17	1.400(9)	D
C13	C14	1.398(8)	D
C13	H38	1.14	S
C14	C15	1.40(1)	S
C14	H39	1.14	S
C15	C16	1.40(1)	D
C16	C17	1.398(8)	S
C16	H40	1.14	S
C17	H41	1.14	S
C18	C19	1.52	S

C18	H42	1.14	S
C18	H43	1.139	S
C19	O20	1.411	S
C19	H44	1.139	S
C19	H45	1.14	S
O20	C21	1.412	S
C21	C22	1.523	S
C21	H46	1.14	S
C21	H47	1.139	S
C22	O23	1.41	S
C22	H48	1.139	S
C22	H49	1.139	S
O23	C24	1.412	S
C24	C29	1.523	S
C24	H50	1.14	S
C24	H51	1.14	S
O25	C26	1.4	S
C26	C27	1.526	S
C26	H52	1.14	S
C26	H53	1.139	S
C27	O28	1.41	S
C27	H54	1.139	S
C27	H55	1.139	S
O28	C29	1.412	S
C29	H56	1.139	S
C29	H57	1.14	S

**Table 3:** Bond angles for dapsone-polyethylene glycol (SDPD)

Atom	Atom	Atom	Bond angle (°)
O2	S1	O3	98.4(4)
O2	S1	C6	107.6(3)
O2	S1	C12	106.8(3)
O3	S1	C6	106.8(3)

O3	S1	C12	107.6(3)
C6	S1	C12	126.1(3)
C9	N4	H30	119.99
C9	N4	H31	120
H30	N4	H31	120.01
C15	N5	H32	120
C15	N5	H33	120
H32	N5	H33	120
S1	C6	C7	120.4(5)
S1	C6	C11	119.9(5)
C7	C6	C11	119.7(6)
C6	C7	C8	120.2(7)
C6	C7	H34	120
C8	C7	H34	119.83
C7	C8	C9	120.2(8)
C7	C8	H35	120.01
C9	C8	H35	119.79
N4	C9	C8	120.2(7)
N4	C9	C10	120.2(7)
C8	C9	C10	119.5(7)
C9	C10	C11	120.2(8)
C9	C10	H36	120
C11	C10	H36	119.79
C6	C11	C10	120.2(7)
C6	C11	H37	120
C10	C11	H37	119.8
S1	C12	C13	120.4(4)
S1	C12	C17	119.9(4)
C13	C12	C17	119.7(5)
C12	C13	C14	120.2(6)
C12	C13	H38	120
C14	C13	H38	119.84
C13	C14	C15	120.2(7)

C13	C14	H39	120
C15	C14	H39	119.78
N5	C15	C14	120.2(6)
N5	C15	C16	120.2(6)
C14	C15	C16	119.5(6)
C15	C16	C17	120.2(7)
C15	C16	H40	120
C17	C16	H40	119.78
C12	C17	C16	120.2(6)
C12	C17	H41	120
C16	C17	H41	119.79
C19	C18	H42	111.25
C19	C18	H43	109.47
C19	C18	O25	98.15
H42	C18	H43	111.25
H42	C18	O25	111.24
H43	C18	O25	114.79
C18	C19	O20	109.13
C18	C19	H44	109.53
C18	C19	H45	109.47
O20	C19	H44	109.53
O20	C19	H45	109.64
H44	C19	H45	109.52
C19	O20	C21	109.45
O20	C21	C22	109.34
O20	C21	H46	109.49
O20	C21	H47	109.47
C22	C21	H46	109.49
C22	C21	H47	109.54
H46	C21	H47	109.49
C21	C22	O23	109.32
C21	C22	H48	109.5
C21	C22	H49	109.48

O23	C22	H48	109.5
O23	C22	H49	109.54
H48	C22	H49	109.49
C22	O23	C24	109.48
O23	C24	C29	109.32
O23	C24	H50	109.5
O23	C24	H51	109.47
C29	C24	H50	109.5
C29	C24	H51	109.55
H50	C24	H51	109.49
C26	O25	C18	137.63
O25	C26	C27	110.59
O25	C26	H52	109.29
O25	C26	H53	109.47
C27	C26	H52	109.28
C27	C26	H53	108.91
H52	C26	H53	109.28
C26	C27	O28	109.35
C26	C27	H54	109.49
C26	C27	H55	109.47
O28	C27	H54	109.49
O28	C27	H55	109.53
H54	C27	H55	109.49
C27	O28	C29	109.48
C24	C29	O28	109.34
C24	C29	H56	109.49
C24	C29	H57	109.47
O28	C29	H56	109.49
O28	C29	H57	109.54
H56	C29	H57	109.49

*Annexure 4:*

**Table 1:** Fractional atomic coordinates for dapsone-polyethylene glycol (SDPD; DFT-D2 Optimized)

Atom	x	y	z
S1	0.61324	1.38898	0.31386
O2	0.65839	1.37171	0.38426
O3	0.54928	1.34402	0.3182
N4	0.74323	1.25948	-0.00778
N5	0.57633	1.81266	0.2764
C6	0.65127	1.35012	0.21829
C7	0.61369	1.32502	0.14627
C8	0.64419	1.29437	0.07147
C9	0.71311	1.28783	0.0668
C10	0.7502	1.31469	0.1397
C11	0.7196	1.34602	0.21406
C12	0.60186	1.51448	0.30408
C13	0.65341	1.57783	0.32317
C14	0.64536	1.67654	0.31349
C15	0.58546	1.71541	0.28394
C16	0.534	1.65024	0.26492
C17	0.54219	1.55146	0.2746
C18	0.34492	0.93518	-0.31829
C19	0.33492	0.93121	-0.22209
O20	0.37503	1.00091	-0.1809
C21	0.36953	0.99463	-0.08985
C22	0.41958	1.0583	-0.04675
O23	0.40357	1.15771	-0.05645
C24	0.45076	1.21666	-0.01483
O25	0.31742	1.47797	0.14821
C26	0.30713	1.48014	0.05743
C27	0.36571	1.4507	0.00383
O28	0.37841	1.35054	0.01511
C29	0.43591	1.32222	-0.02951

H30	0.71609	1.22186	-0.05123
H31	0.79152	1.24223	-0.00528
H32	0.53924	1.83609	0.23842
H33	0.6148	1.85824	0.28273
H34	0.56066	1.32821	0.14927
H35	0.61514	1.27387	0.01573
H36	0.80321	1.31068	0.13671
H37	0.74862	1.36796	0.26929
H38	0.69951	1.549	0.34618
H39	0.68524	1.72525	0.32915
H40	0.48741	1.6791	0.24327
H41	0.5016	1.50319	0.26112
H42	0.32065	0.87185	-0.34733
H43	0.39749	0.93031	-0.33335
H44	0.34756	0.85796	-0.19954
H45	0.28298	0.944	-0.20675
H46	0.37806	0.91947	-0.06874
H47	0.31987	1.01514	-0.06913
H48	0.46826	1.04406	-0.07436
H49	0.42109	1.03918	0.02212
H50	0.49963	1.20166	-0.04158
H51	0.4518	1.20065	0.05462
H52	0.29136	1.55294	0.03761
H53	0.26658	1.43038	0.04503
H54	0.40932	1.49333	0.02182
H55	0.35522	1.46703	-0.06444
H56	0.43057	1.33687	-0.09889
H57	0.47787	1.36508	-0.00581



**Table 2:** Bond lengths for dapsone-polyethylene glycol (SDPD; DFT-D2 Optimized)

Atom	Atom	Length (Å)	Bond type
S1	O2	1.455	D
S1	O3	1.454	D
S1	C6	1.764	S
S1	C12	1.759	S
N4	C9	1.373	S
N4	H30	1.019	S
N4	H31	1.02	S
N5	C15	1.364	S
N5	H32	1.017	S
N5	H33	1.015	S
C6	C7	1.404	D
C6	C11	1.404	S
C7	C8	1.388	S
C7	H34	1.09	S
C8	C9	1.419	D
C8	H35	1.09	S
C9	C10	1.416	S
C10	C11	1.386	D
C10	H36	1.09	S
C11	H37	1.089	S
C12	C13	1.405	S
C12	C17	1.404	D
C13	C14	1.385	D
C13	H38	1.087	S
C14	C15	1.418	S
C14	H39	1.088	S
C15	C16	1.42	D
C16	C17	1.386	S
C16	H40	1.089	S
C17	H41	1.088	S
C18	C19	1.513	S

C18	H42	1.105	S
C18	H43	1.106	S
C19	O20	1.421	S
C19	H44	1.104	S
C19	H45	1.106	S
O20	C21	1.425	S
C21	C22	1.51	S
C21	H46	1.105	S
C21	H47	1.106	S
C22	O23	1.423	S
C22	H48	1.105	S
C22	H49	1.105	S
O23	C24	1.422	S
C24	C29	1.51	S
C24	H50	1.105	S
C24	H51	1.104	S
O25	C26	1.429	S
C26	C27	1.519	S
C26	H52	1.102	S
C26	H53	1.097	S
C27	O28	1.422	S
C27	H54	1.108	S
C27	H55	1.108	S
O28	C29	1.424	S
C29	H56	1.104	S
C29	H57	1.108	S

**Table 3:** Bond angles for dapsone-polyethylene glycol (SDPD; DFT-D2 Optimized)

Atom	Atom	Atom	Bond angle (°)
O2	S1	C6	107.68
O2	S1	C12	108.18
O3	S1	C6	107.93
O3	S1	C12	107.88

C6	S1	C12	106.67
O2	S1	O3	117.99
C19	O20	C21	111.21
C22	O23	C24	110.45
C18	O25	C26	113.71
C27	O28	C29	111.1
O25	C18	C19	109.81
O20	C19	C18	110.1
O20	C21	C22	110.64
O23	C22	C21	111.12
O23	C24	C29	110.42
O25	C26	C27	114.88
O28	C27	C26	109.81
O28	C29	C24	111.08
O25	C18	H43	111
O25	C18	H42	110
H42	C18	H43	108
C19	C18	H42	108
C19	C18	H43	110
C18	C19	H45	110
O20	C19	H44	110
O20	C19	H45	111
C18	C19	H44	108
H44	C19	H45	108
O20	C21	H46	110
O20	C21	H47	110
H46	C21	H47	108
C22	C21	H47	110
C22	C21	H46	108
C21	C22	H48	110
O23	C22	H49	110
O23	C22	H48	110
C21	C22	H49	108

H48	C22	H49	108
H50	C24	H51	108
O23	C24	H50	110
O23	C24	H51	110
C29	C24	H50	108
C29	C24	H51	110
O25	C26	H53	106
C27	C26	H52	109
O25	C26	H52	110
H52	C26	H53	108
C27	C26	H53	110
C26	C27	H54	111
O28	C27	H55	111
O28	C27	H54	110
C26	C27	H55	109
H54	C27	H55	107
O28	C29	H57	109
C24	C29	H56	110
C24	C29	H57	108
H56	C29	H57	108
O28	C29	H56	110
H30	N4	H31	116
C9	N4	H30	117
C9	N4	H31	118
C15	N5	H32	118
C15	N5	H33	120
H32	N5	H33	116
S1	C6	C7	120.4
S1	C6	C11	119.56
C7	C6	C11	120.03
C6	C7	C8	119.86
C7	C8	C9	120.77
N4	C9	C10	120.74

C8	C9	C10	118.54
N4	C9	C8	120.68
C9	C10	C11	120.53
C6	C11	C10	120.24
C13	C12	C17	119.84
S1	C12	C17	120.26
S1	C12	C13	119.86
C12	C13	C14	120.19
C13	C14	C15	120.87
N5	C15	C14	121.42
N5	C15	C16	120.46
C14	C15	C16	118.09
C15	C16	C17	120.93
C12	C17	C16	120.07
C6	C7	H34	120
C8	C7	H34	120
C7	C8	H35	120
C9	C8	H35	119
C9	C10	H36	119
C11	C10	H36	120
C6	C11	H37	120
C10	C11	H37	120
C12	C13	H38	120
C14	C13	H38	120
C13	C14	H39	120
C15	C14	H39	119
C15	C16	H40	119
C17	C16	H40	120
C12	C17	H41	120
C16	C17	H41	119

**Annexure 5:**

**Table 1:** Fractional atomic coordinates for dapsone-polyethylene glycol (SCXRD)

Atom	x	y	z	U <sub>iso</sub> /U <sub>equiv</sub>	adp type
S1	0.61362(3)	0.38531(4)	0.18388(4)	0.02455(19)	Uani
O5	0.65940(9)	0.36845(13)	0.11321(11)	0.0311(4)	Uani
O6	0.54825(9)	0.34064(13)	0.17951(12)	0.0321(4)	Uani
N2	0.58008(12)	0.81222(15)	0.22435(16)	0.0335(5)	Uani
H2A	0.570951	0.836979	0.174691	0.04	Uiso
H2B	0.547799	0.823647	0.259535	0.04	Uiso
O2A	0.87503(18)	0.4986(3)	0.3193(3)	0.0331(9)	Uiso
N1	0.74625(12)	0.26224(18)	0.50702(15)	0.0371(5)	Uani
H1A	0.722427	0.219611	0.533983	0.044	Uiso
H1B	0.784887	0.237792	0.495843	0.044	Uiso
O1A	0.81842(19)	0.4768(3)	0.1466(3)	0.0369(9)	Uiso
O3A	0.9086(2)	0.3420(3)	0.4455(2)	0.0390(9)	Uiso
C9	0.71542(13)	0.28710(18)	0.43123(16)	0.0278(5)	Uani
C18	0.58826(13)	0.71332(17)	0.21573(16)	0.0264(5)	Uani
C12	0.65209(12)	0.34551(17)	0.27888(16)	0.0243(5)	Uani
C15	0.60293(12)	0.51139(18)	0.19462(16)	0.0243(5)	Uani
C13	0.61388(12)	0.31980(17)	0.35105(17)	0.0266(5)	Uani
H13	0.567529	0.321886	0.348272	0.032	Uiso
C11	0.72177(12)	0.34229(17)	0.28300(16)	0.0261(5)	Uani
H11	0.747098	0.360154	0.235167	0.031	Uiso
C14	0.64506(13)	0.29145(18)	0.42604(17)	0.0283(5)	Uani
H14	0.61952	0.274891	0.474057	0.034	Uiso
C10	0.75287(13)	0.31268(18)	0.35786(17)	0.0287(5)	Uani
H10	0.799217	0.30957	0.359912	0.034	Uiso
C16	0.65515(12)	0.57420(19)	0.17286(17)	0.0278(5)	Uani
H16	0.694967	0.549141	0.151308	0.033	Uiso
C20	0.54324(13)	0.54951(18)	0.22662(17)	0.0287(5)	Uani
H20	0.508348	0.507722	0.241008	0.034	Uiso
C19	0.53594(13)	0.64838(18)	0.23686(18)	0.0302(6)	Uani
H19	0.495923	0.673133	0.258057	0.036	Uiso

C17	0.64744(13)	0.67355(19)	0.18345(17)	0.0297(6)	Uani
H17	0.682419	0.715048	0.168785	0.036	Uiso
O4A	0.8836(3)	0.1477(4)	0.5161(3)	0.0572(12)	Uiso
C3A	0.8700(3)	0.5039(4)	0.4128(4)	0.0365(11)	Uiso
H3AA	0.82636	0.481125	0.430848	0.044	Uiso
H3AB	0.87479	0.571337	0.431176	0.044	Uiso
C2A	0.8329(2)	0.5664(4)	0.2794(4)	0.0326(10)	Uiso
H2AA	0.842001	0.631368	0.30165	0.039	Uiso
H2AB	0.786741	0.550561	0.292008	0.039	Uiso
C4A	0.9221(3)	0.4439(4)	0.4543(3)	0.0373(11)	Uiso
H4AA	0.96496	0.458524	0.428303	0.045	Uiso
H4AB	0.924589	0.460356	0.515191	0.045	Uiso
C1A	0.8440(2)	0.5647(4)	0.1841(4)	0.0352(10)	Uiso
H1AA	0.82215	0.620641	0.157989	0.042	Uiso
H1AB	0.891441	0.569574	0.172254	0.042	Uiso
C5A	0.9564(3)	0.2855(5)	0.4908(4)	0.0484(13)	Uiso
H5AA	0.953554	0.299235	0.552209	0.058	Uiso
H5AB	1.000988	0.302395	0.471322	0.058	Uiso
C6A	0.9432(3)	0.1762(5)	0.4745(4)	0.0494(13)	Uiso
H6AA	0.939501	0.164328	0.412854	0.059	Uiso
H6AB	0.980233	0.137805	0.496427	0.059	Uiso
C7A	0.8702(4)	0.0453(5)	0.5035(5)	0.0658(17)	Uiso
H7AA	0.907979	0.006647	0.522876	0.079	Uiso
H7AB	0.86282	0.031912	0.44269	0.079	Uiso
C1B	0.8432(4)	0.5768(5)	0.2301(6)	0.0358(15)	Uiso
H1BA	0.814535	0.630866	0.213867	0.043	Uiso
H1BB	0.88918	0.59811	0.224174	0.043	Uiso
C8B	0.8454(5)	-0.0297(7)	0.5845(6)	0.055(2)	Uiso
H8BA	0.823344	-0.091808	0.574643	0.066	Uiso
H8BB	0.893103	-0.039004	0.577106	0.066	Uiso
C2B	0.8304(4)	0.5504(6)	0.3202(6)	0.0395(17)	Uiso
H2BA	0.832757	0.608239	0.356313	0.047	Uiso
H2BB	0.786022	0.522724	0.325571	0.047	Uiso

C6B	0.9129(6)	0.1373(8)	0.4684(7)	0.059(2)	Uiso
H6BA	0.897155	0.131408	0.409336	0.071	Uiso
H6BB	0.942258	0.082611	0.480431	0.071	Uiso
O3B	0.9035(3)	0.3046(5)	0.4506(4)	0.0436(15)	Uiso
O2B	0.8789(3)	0.4809(4)	0.3477(4)	0.0346(13)	Uiso
C5B	0.9513(4)	0.2342(7)	0.4796(5)	0.0398(17)	Uiso
H5BA	0.963381	0.245083	0.539524	0.048	Uiso
H5BB	0.991265	0.235473	0.444266	0.048	Uiso
C4B	0.9297(4)	0.4013(7)	0.4655(5)	0.0361(16)	Uiso
H4BA	0.939937	0.409207	0.526396	0.043	Uiso
H4BB	0.97071	0.409762	0.433026	0.043	Uiso
O1B	0.8312(3)	0.4945(5)	0.1721(4)	0.0422(15)	Uiso
O4B	0.8574(4)	0.1359(5)	0.5267(5)	0.0534(17)	Uiso
C3B	0.8819(4)	0.4743(7)	0.4394(6)	0.0418(18)	Uiso
H3BA	0.838186	0.457351	0.461887	0.05	Uiso
H3BB	0.894525	0.537145	0.463339	0.05	Uiso
C8A	0.8099(3)	0.0198(5)	0.5542(4)	0.0497(14)	Uiso
H8AA	0.774255	0.064388	0.538406	0.06	Uiso
H8AB	0.795791	-0.045559	0.538179	0.06	Uiso
C7B	0.8196(5)	0.0473(8)	0.5197(7)	0.058(2)	Uiso
H7BA	0.823242	0.021886	0.461312	0.069	Uiso
H7BB	0.772906	0.060798	0.531109	0.069	Uiso



**Table 2:** Bond lengths for dapsone-polyethylene glycol (SXR D)

Atom	Atom	Length (Å)	Bond type
S1	O5	1.4455(19)	D
S1	O6	1.4467(18)	D
S1	C12	1.747(3)	S
S1	C15	1.746(3)	S
N2	H2A	0.8594	S
N2	H2B	0.8597	S
N2	C18	1.369(3)	S
O2A	C3A	1.452(6)	S
O2A	C2A	1.398(6)	S
N1	H1A	0.8613	S
N1	H1B	0.8608	S
N1	C9	1.368(3)	S
O1A	C1A	1.430(6)	S
O1A	C8A	1.440(7)	S
O3A	C4A	1.425(7)	S
O3A	C5A	1.417(7)	S
C9	C14	1.413(4)	S
C9	C10	1.405(4)	D
C18	C19	1.412(3)	D
C18	C17	1.397(4)	S
C12	C13	1.399(3)	S
C12	C11	1.398(3)	D
C15	C16	1.395(3)	S
C15	C20	1.395(3)	D
C13	H13	0.93	S
C13	C14	1.374(4)	D
C11	H11	0.93	S
C11	C10	1.376(4)	S
C14	H14	0.93	S
C10	H10	0.93	S
C16	H16	0.93	S

C16	C17	1.378(4)	D
C20	H20	0.93	S
C20	C19	1.370(4)	S
C19	H19	0.93	S
C17	H17	0.93	S
O4A	C6A	1.413(8)	S
O4A	C7A	1.439(8)	S
C3A	H3AA	0.97	S
C3A	H3AB	0.97	S
C3A	C4A	1.474(8)	S
C2A	H2AA	0.97	S
C2A	H2AB	0.97	S
C2A	C1A	1.491(7)	S
C4A	H4AA	0.97	S
C4A	H4AB	0.97	S
C1A	H1AA	0.97	S
C1A	H1AB	0.97	S
C5A	H5AA	0.97	S
C5A	H5AB	0.97	S
C5A	C6A	1.540(8)	S
C6A	H6AA	0.97	S
C6A	H6AB	0.97	S
C7A	H7AA	0.97	S
C7A	H7AB	0.97	S
C7A	C8A	1.483(9)	S
C1B	H1BA	0.97	S
C1B	H1BB	0.97	S
C1B	C2B	1.463(11)	S
C1B	O1B	1.460(10)	S
C8B	H8BA	0.97	S
C8B	H8BB	0.97	S
C8B	O1B	1.467(11)	S
C8B	C7B	1.544(14)	S

C2B	H2BA	0.97	S
C2B	H2BB	0.97	S
C2B	O2B	1.424(10)	S
C6B	H6BA	0.97	S
C6B	H6BB	0.97	S
C6B	C5B	1.542(13)	S
C6B	O4B	1.432(13)	S
O3B	C5B	1.430(10)	S
O3B	C4B	1.441(10)	S
O2B	C3B	1.423(11)	S
C5B	H5BA	0.97	S
C5B	H5BB	0.97	S
C4B	H4BA	0.97	S
C4B	H4BB	0.97	S
C4B	C3B	1.441(12)	S
O4B	C7B	1.433(13)	S
C3B	H3BA	0.97	S
C3B	H3BB	0.97	S
C8A	H8AA	0.97	S
C8A	H8AB	0.97	S
C7B	H7BA	0.97	S
C7B	H7BB	0.97	S

**Table 3:** Bond angles for dapsone-polyethylene glycol (SXRDR)

Atom	Atom	Atom	Bond angle (°)
O5	S1	O6	118.13(11)
O5	S1	C12	107.87(11)
O5	S1	C15	107.90(11)
O6	S1	C12	107.88(11)
O6	S1	C15	108.11(11)
C15	S1	C12	106.38(11)
H2A	N2	H2B	109.5
C18	N2	H2A	109.1
C18	N2	H2B	109.4
C2A	O2A	C3A	111.4(4)
H1A	N1	H1B	109.4
C9	N1	H1A	109.5
C9	N1	H1B	109.3
C1A	O1A	C8A	114.8(4)
C5A	O3A	C4A	111.0(5)
N1	C9	C14	120.6(2)
N1	C9	C10	120.9(2)
C10	C9	C14	118.4(2)
N2	C18	C19	120.7(2)
N2	C18	C17	121.4(2)
C17	C18	C19	117.9(2)
C13	C12	S1	120.58(19)
C11	C12	S1	119.26(19)
C11	C12	C13	120.1(2)
C16	C15	S1	119.58(19)
C16	C15	C20	119.9(2)
C20	C15	S1	120.54(19)
C12	C13	H13	120.1
C14	C13	C12	119.8(2)
C14	C13	H13	120.1
C12	C11	H11	120

C10	C11	C12	120.0(2)
C10	C11	H11	120
C9	C14	H14	119.6
C13	C14	C9	120.9(2)
C13	C14	H14	119.6
C9	C10	H10	119.6
C11	C10	C9	120.8(2)
C11	C10	H10	119.6
C15	C16	H16	120.2
C17	C16	C15	119.7(2)
C17	C16	H16	120.2
C15	C20	H20	119.9
C19	C20	C15	120.1(2)
C19	C20	H20	119.9
C18	C19	H19	119.5
C20	C19	C18	121.0(2)
C20	C19	H19	119.5
C18	C17	H17	119.3
C16	C17	C18	121.5(2)
C16	C17	H17	119.3
C6A	O4A	C7A	111.4(6)
O2A	C3A	H3AA	109.4
O2A	C3A	H3AB	109.4
O2A	C3A	C4A	110.9(4)
H3AA	C3A	H3AB	108
C4A	C3A	H3AA	109.4
C4A	C3A	H3AB	109.4
O2A	C2A	H2AA	109.7
O2A	C2A	H2AB	109.7
O2A	C2A	C1A	109.6(4)
H2AA	C2A	H2AB	108.2
C1A	C2A	H2AA	109.7
C1A	C2A	H2AB	109.7

O3A	C4A	C3A	111.7(4)
O3A	C4A	H4AA	109.3
O3A	C4A	H4AB	109.3
C3A	C4A	H4AA	109.3
C3A	C4A	H4AB	109.3
H4AA	C4A	H4AB	107.9
O1A	C1A	C2A	111.2(4)
O1A	C1A	H1AA	109.4
O1A	C1A	H1AB	109.4
C2A	C1A	H1AA	109.4
C2A	C1A	H1AB	109.4
H1AA	C1A	H1AB	108
O3A	C5A	H5AA	109.8
O3A	C5A	H5AB	109.8
O3A	C5A	C6A	109.4(5)
H5AA	C5A	H5AB	108.2
C6A	C5A	H5AA	109.8
C6A	C5A	H5AB	109.8
O4A	C6A	C5A	109.8(5)
O4A	C6A	H6AA	109.7
O4A	C6A	H6AB	109.7
C5A	C6A	H6AA	109.7
C5A	C6A	H6AB	109.7
H6AA	C6A	H6AB	108.2
O4A	C7A	H7AA	110.1
O4A	C7A	H7AB	110.1
O4A	C7A	C8A	108.0(6)
H7AA	C7A	H7AB	108.4
C8A	C7A	H7AA	110.1
C8A	C7A	H7AB	110.1
H1BA	C1B	H1BB	108
C2B	C1B	H1BA	109.3
C2B	C1B	H1BB	109.3

O1B	C1B	H1BA	109.3
O1B	C1B	H1BB	109.3
O1B	C1B	C2B	111.5(6)
H8BA	C8B	H8BB	108.4
O1B	C8B	H8BA	110.1
O1B	C8B	H8BB	110.1
O1B	C8B	C7B	108.1(8)
C7B	C8B	H8BA	110.1
C7B	C8B	H8BB	110.1
C1B	C2B	H2BA	109.8
C1B	C2B	H2BB	109.8
H2BA	C2B	H2BB	108.3
O2B	C2B	C1B	109.3(7)
O2B	C2B	H2BA	109.8
O2B	C2B	H2BB	109.8
H6BA	C6B	H6BB	108.3
C5B	C6B	H6BA	109.8
C5B	C6B	H6BB	109.8
O4B	C6B	H6BA	109.8
O4B	C6B	H6BB	109.8
O4B	C6B	C5B	109.2(8)
C5B	O3B	C4B	108.9(6)
C3B	O2B	C2B	111.7(7)
C6B	C5B	H5BA	111.3
C6B	C5B	H5BB	111.3
O3B	C5B	C6B	102.1(7)
O3B	C5B	H5BA	111.3
O3B	C5B	H5BB	111.3
H5BA	C5B	H5BB	109.2
O3B	C4B	H4BA	109.6
O3B	C4B	H4BB	109.6
H4BA	C4B	H4BB	108.1
C3B	C4B	O3B	110.4(7)

C3B	C4B	H4BA	109.6
C3B	C4B	H4BB	109.6
C6B	O4B	C7B	112.0(8)
O2B	C3B	C4B	110.6(7)
O2B	C3B	H3BA	109.5
O2B	C3B	H3BB	109.5
C4B	C3B	H3BA	109.5
C4B	C3B	H3BB	109.5
H3BA	C3B	H3BB	108.1
O1A	C8A	C7A	114.9(5)
O1A	C8A	H8AA	108.6
O1A	C8A	H8AB	108.6
C7A	C8A	H8AA	108.6
C7A	C8A	H8AB	108.6
H8AA	C8A	H8AB	107.5
C8B	C7B	H7BA	109.5
C8B	C7B	H7BB	109.5
O4B	C7B	C8B	110.6(8)
O4B	C7B	H7BA	109.5
O4B	C7B	H7BB	109.5
H7BA	C7B	H7BB	108.1



**Annexure 6:****Table 1:** Fractional atomic coordinates for sulfanilamide–polyethylene glycol (SCXRD; Solution – I, Pbc<sub>a</sub>)

Atom	x	y	z	U <sub>iso</sub> /U <sub>equiv</sub>	adp type
S1	0.61362(3)	0.38531(4)	0.18388(4)	0.02455(19)	Uani
O5	0.65940(9)	0.36845(13)	0.11321(11)	0.0311(4)	Uani
O6	0.54825(9)	0.34064(13)	0.17951(12)	0.0321(4)	Uani
N2	0.58008(12)	0.81222(15)	0.22435(16)	0.0335(5)	Uani
H2A	0.570951	0.836979	0.174691	0.04	Uiso
H2B	0.547799	0.823647	0.259535	0.04	Uiso
O2A	0.87503(18)	0.4986(3)	0.3193(3)	0.0331(9)	Uiso
N1	0.74625(12)	0.26224(18)	0.50702(15)	0.0371(5)	Uani
H1A	0.722427	0.219611	0.533983	0.044	Uiso
H1B	0.784887	0.237792	0.495843	0.044	Uiso
O1A	0.81842(19)	0.4768(3)	0.1466(3)	0.0369(9)	Uiso
O3A	0.9086(2)	0.3420(3)	0.4455(2)	0.0390(9)	Uiso
C9	0.71542(13)	0.28710(18)	0.43123(16)	0.0278(5)	Uani
C18	0.58826(13)	0.71332(17)	0.21573(16)	0.0264(5)	Uani
C12	0.65209(12)	0.34551(17)	0.27888(16)	0.0243(5)	Uani
C15	0.60293(12)	0.51139(18)	0.19462(16)	0.0243(5)	Uani
C13	0.61388(12)	0.31980(17)	0.35105(17)	0.0266(5)	Uani
H13	0.567529	0.321886	0.348272	0.032	Uiso
C11	0.72177(12)	0.34229(17)	0.28300(16)	0.0261(5)	Uani
H11	0.747098	0.360154	0.235167	0.031	Uiso
C14	0.64506(13)	0.29145(18)	0.42604(17)	0.0283(5)	Uani
H14	0.61952	0.274891	0.474057	0.034	Uiso
C10	0.75287(13)	0.31268(18)	0.35786(17)	0.0287(5)	Uani
H10	0.799217	0.30957	0.359912	0.034	Uiso
C16	0.65515(12)	0.57420(19)	0.17286(17)	0.0278(5)	Uani
H16	0.694967	0.549141	0.151308	0.033	Uiso
C20	0.54324(13)	0.54951(18)	0.22662(17)	0.0287(5)	Uani
H20	0.508348	0.507722	0.241008	0.034	Uiso
C19	0.53594(13)	0.64838(18)	0.23686(18)	0.0302(6)	Uani

H19	0.495923	0.673133	0.258057	0.036	Uiso
C17	0.64744(13)	0.67355(19)	0.18345(17)	0.0297(6)	Uani
H17	0.682419	0.715048	0.168785	0.036	Uiso
O4A	0.8836(3)	0.1477(4)	0.5161(3)	0.0572(12)	Uiso
C3A	0.8700(3)	0.5039(4)	0.4128(4)	0.0365(11)	Uiso
H3AA	0.82636	0.481125	0.430848	0.044	Uiso
H3AB	0.87479	0.571337	0.431176	0.044	Uiso
C2A	0.8329(2)	0.5664(4)	0.2794(4)	0.0326(10)	Uiso
H2AA	0.842001	0.631368	0.30165	0.039	Uiso
H2AB	0.786741	0.550561	0.292008	0.039	Uiso
C4A	0.9221(3)	0.4439(4)	0.4543(3)	0.0373(11)	Uiso
H4AA	0.96496	0.458524	0.428303	0.045	Uiso
H4AB	0.924589	0.460356	0.515191	0.045	Uiso
C1A	0.8440(2)	0.5647(4)	0.1841(4)	0.0352(10)	Uiso
H1AA	0.82215	0.620641	0.157989	0.042	Uiso
H1AB	0.891441	0.569574	0.172254	0.042	Uiso
C5A	0.9564(3)	0.2855(5)	0.4908(4)	0.0484(13)	Uiso
H5AA	0.953554	0.299235	0.552209	0.058	Uiso
H5AB	1.000988	0.302395	0.471322	0.058	Uiso
C6A	0.9432(3)	0.1762(5)	0.4745(4)	0.0494(13)	Uiso
H6AA	0.939501	0.164328	0.412854	0.059	Uiso
H6AB	0.980233	0.137805	0.496427	0.059	Uiso
C7A	0.8702(4)	0.0453(5)	0.5035(5)	0.0658(17)	Uiso
H7AA	0.907979	0.006647	0.522876	0.079	Uiso
H7AB	0.86282	0.031912	0.44269	0.079	Uiso
C1B	0.8432(4)	0.5768(5)	0.2301(6)	0.0358(15)	Uiso
H1BA	0.814535	0.630866	0.213867	0.043	Uiso
H1BB	0.88918	0.59811	0.224174	0.043	Uiso
C8B	0.8454(5)	-0.0297(7)	0.5845(6)	0.055(2)	Uiso
H8BA	0.823344	-0.091808	0.574643	0.066	Uiso
H8BB	0.893103	-0.039004	0.577106	0.066	Uiso
C2B	0.8304(4)	0.5504(6)	0.3202(6)	0.0395(17)	Uiso
H2BA	0.832757	0.608239	0.356313	0.047	Uiso

H2BB	0.786022	0.522724	0.325571	0.047	Uiso
C6B	0.9129(6)	0.1373(8)	0.4684(7)	0.059(2)	Uiso
H6BA	0.897155	0.131408	0.409336	0.071	Uiso
H6BB	0.942258	0.082611	0.480431	0.071	Uiso
O3B	0.9035(3)	0.3046(5)	0.4506(4)	0.0436(15)	Uiso
O2B	0.8789(3)	0.4809(4)	0.3477(4)	0.0346(13)	Uiso
C5B	0.9513(4)	0.2342(7)	0.4796(5)	0.0398(17)	Uiso
H5BA	0.963381	0.245083	0.539524	0.048	Uiso
H5BB	0.991265	0.235473	0.444266	0.048	Uiso
C4B	0.9297(4)	0.4013(7)	0.4655(5)	0.0361(16)	Uiso
H4BA	0.939937	0.409207	0.526396	0.043	Uiso
H4BB	0.97071	0.409762	0.433026	0.043	Uiso
O1B	0.8312(3)	0.4945(5)	0.1721(4)	0.0422(15)	Uiso
O4B	0.8574(4)	0.1359(5)	0.5267(5)	0.0534(17)	Uiso
C3B	0.8819(4)	0.4743(7)	0.4394(6)	0.0418(18)	Uiso
H3BA	0.838186	0.457351	0.461887	0.05	Uiso
H3BB	0.894525	0.537145	0.463339	0.05	Uiso
C8A	0.8099(3)	0.0198(5)	0.5542(4)	0.0497(14)	Uiso
H8AA	0.774255	0.064388	0.538406	0.06	Uiso
H8AB	0.795791	-0.045559	0.538179	0.06	Uiso
C7B	0.8196(5)	0.0473(8)	0.5197(7)	0.058(2)	Uiso
H7BA	0.823242	0.021886	0.461312	0.069	Uiso
H7BB	0.772906	0.060798	0.531109	0.069	Uiso

**Table 2:** Bond lengths for sulfanilamide–polyethylene glycol (SXRd; Solution – I, Pbc<sub>a</sub>)

Atom	Atom	Length (Å)	Bond type
C1B	O1B	1.460(10)	S
C8B	H8BA	0.97	S
C8B	H8BB	0.97	S
C8B	O1B	1.467(11)	S
C8B	C7B	1.544(14)	S
C2B	H2BA	0.97	S
C2B	H2BB	0.97	S
C2B	O2B	1.424(10)	S
C6B	H6BA	0.97	S
C6B	H6BB	0.97	S
C6B	C5B	1.542(13)	S
C6B	O4B	1.432(13)	S
O3B	C5B	1.430(10)	S
O3B	C4B	1.441(10)	S
O2B	C3B	1.423(11)	S
C5B	H5BA	0.97	S
C5B	H5BB	0.97	S
C4B	H4BA	0.97	S
C4B	H4BB	0.97	S
C4B	C3B	1.441(12)	S
O4B	C7B	1.433(13)	S
C3B	H3BA	0.97	S
C3B	H3BB	0.97	S
C8A	H8AA	0.97	S
C8A	H8AB	0.97	S
C7B	H7BA	0.97	S
C7B	H7BB	0.97	S

**Table 3:** Bond angles for sulfanilamide–polyethylene glycol (SXR D; Solution – I, P bca)

Atom	Atom	Atom	Bond angle (°)
O5	S1	O6	118.13(11)
O5	S1	C12	107.87(11)
O5	S1	C15	107.90(11)
O6	S1	C12	107.88(11)
O6	S1	C15	108.11(11)
C15	S1	C12	106.38(11)
H2A	N2	H2B	109.5
C18	N2	H2A	109.1
C18	N2	H2B	109.4
C2A	O2A	C3A	111.4(4)
H1A	N1	H1B	109.4
C9	N1	H1A	109.5
C9	N1	H1B	109.3
C1A	O1A	C8A	114.8(4)
C5A	O3A	C4A	111.0(5)
N1	C9	C14	120.6(2)
N1	C9	C10	120.9(2)
C10	C9	C14	118.4(2)
N2	C18	C19	120.7(2)
N2	C18	C17	121.4(2)
C17	C18	C19	117.9(2)
C13	C12	S1	120.58(19)
C11	C12	S1	119.26(19)
C11	C12	C13	120.1(2)
C16	C15	S1	119.58(19)
C16	C15	C20	119.9(2)
C20	C15	S1	120.54(19)
C12	C13	H13	120.1
C14	C13	C12	119.8(2)
C14	C13	H13	120.1
C12	C11	H11	120

C10	C11	C12	120.0(2)
C10	C11	H11	120
C9	C14	H14	119.6
C13	C14	C9	120.9(2)
C13	C14	H14	119.6
C9	C10	H10	119.6
C11	C10	C9	120.8(2)
C11	C10	H10	119.6
C15	C16	H16	120.2
C17	C16	C15	119.7(2)
C17	C16	H16	120.2
C15	C20	H20	119.9
C19	C20	C15	120.1(2)
C19	C20	H20	119.9
C18	C19	H19	119.5
C20	C19	C18	121.0(2)
C20	C19	H19	119.5
C18	C17	H17	119.3
C16	C17	C18	121.5(2)
C16	C17	H17	119.3
C6A	O4A	C7A	111.4(6)
O2A	C3A	H3AA	109.4
O2A	C3A	H3AB	109.4
O2A	C3A	C4A	110.9(4)
H3AA	C3A	H3AB	108
C4A	C3A	H3AA	109.4
C4A	C3A	H3AB	109.4
O2A	C2A	H2AA	109.7
O2A	C2A	H2AB	109.7
O2A	C2A	C1A	109.6(4)
H2AA	C2A	H2AB	108.2
C1A	C2A	H2AA	109.7
C1A	C2A	H2AB	109.7

O3A	C4A	C3A	111.7(4)
O3A	C4A	H4AA	109.3
O3A	C4A	H4AB	109.3
C3A	C4A	H4AA	109.3
C3A	C4A	H4AB	109.3
H4AA	C4A	H4AB	107.9
O1A	C1A	C2A	111.2(4)
O1A	C1A	H1AA	109.4
O1A	C1A	H1AB	109.4
C2A	C1A	H1AA	109.4
C2A	C1A	H1AB	109.4
H1AA	C1A	H1AB	108
O3A	C5A	H5AA	109.8
O3A	C5A	H5AB	109.8
O3A	C5A	C6A	109.4(5)
H5AA	C5A	H5AB	108.2
C6A	C5A	H5AA	109.8
C6A	C5A	H5AB	109.8
O4A	C6A	C5A	109.8(5)
O4A	C6A	H6AA	109.7
O4A	C6A	H6AB	109.7
C5A	C6A	H6AA	109.7
C5A	C6A	H6AB	109.7
H6AA	C6A	H6AB	108.2
O4A	C7A	H7AA	110.1
O4A	C7A	H7AB	110.1
O4A	C7A	C8A	108.0(6)
H7AA	C7A	H7AB	108.4
C8A	C7A	H7AA	110.1
C8A	C7A	H7AB	110.1
H1BA	C1B	H1BB	108
C2B	C1B	H1BA	109.3
C2B	C1B	H1BB	109.3

O1B	C1B	H1BA	109.3
O1B	C1B	H1BB	109.3
O1B	C1B	C2B	111.5(6)
H8BA	C8B	H8BB	108.4
O1B	C8B	H8BA	110.1
O1B	C8B	H8BB	110.1
O1B	C8B	C7B	108.1(8)
C7B	C8B	H8BA	110.1
C7B	C8B	H8BB	110.1
C1B	C2B	H2BA	109.8
C1B	C2B	H2BB	109.8
H2BA	C2B	H2BB	108.3
O2B	C2B	C1B	109.3(7)
O2B	C2B	H2BA	109.8
O2B	C2B	H2BB	109.8
H6BA	C6B	H6BB	108.3
C5B	C6B	H6BA	109.8
C5B	C6B	H6BB	109.8
O4B	C6B	H6BA	109.8
O4B	C6B	H6BB	109.8
O4B	C6B	C5B	109.2(8)
C5B	O3B	C4B	108.9(6)
C3B	O2B	C2B	111.7(7)
C6B	C5B	H5BA	111.3
C6B	C5B	H5BB	111.3
O3B	C5B	C6B	102.1(7)
O3B	C5B	H5BA	111.3
O3B	C5B	H5BB	111.3
H5BA	C5B	H5BB	109.2
O3B	C4B	H4BA	109.6
O3B	C4B	H4BB	109.6
H4BA	C4B	H4BB	108.1
C3B	C4B	O3B	110.4(7)



C3B	C4B	H4BA	109.6
C3B	C4B	H4BB	109.6
C6B	O4B	C7B	112.0(8)
O2B	C3B	C4B	110.6(7)
O2B	C3B	H3BA	109.5
O2B	C3B	H3BB	109.5
C4B	C3B	H3BA	109.5
C4B	C3B	H3BB	109.5
H3BA	C3B	H3BB	108.1
O1A	C8A	C7A	114.9(5)
O1A	C8A	H8AA	108.6
O1A	C8A	H8AB	108.6
C7A	C8A	H8AA	108.6
C7A	C8A	H8AB	108.6
H8AA	C8A	H8AB	107.5
C8B	C7B	H7BA	109.5
C8B	C7B	H7BB	109.5
O4B	C7B	C8B	110.6(8)
O4B	C7B	H7BA	109.5
O4B	C7B	H7BB	109.5
H7BA	C7B	H7BB	108.1

*Annexure 7:*

**Table 1:** Fractional atomic coordinates for sulfanilamide–polyethylene glycol (SXR D;  
Solution – II, Pca2<sub>1</sub>)

Atom	x	y	z	U <sub>iso</sub> /U <sub>equiv</sub>	adp type
S1	0.6819	0.8016	0.3619	0.055	Uiso
S2	0.8185	0.3028	0.6284	0.053	Uiso
O3	0.679	0.9024	0.309	0.066	Uiso
N4	0.8049	0.3584	0.548	0.062	Uiso
O5	0.6019	0.7195	0.3691	0.084	Uiso
O6	0.8234	0.4043	0.6791	0.078	Uiso
C7	0.547	0.0696	0.6829	0.057	Uiso
C8	0.6255	0.0256	0.641	0.055	Uiso
N9	1.0243	0.499	0.2822	0.107	Uiso
O10	0.9019	0.2216	0.6273	0.078	Uiso
C11	0.7933	0.593	0.367	0.069	Uiso
C12	0.8521	0.7649	0.2971	0.032	Uiso
C13	0.5631	0.191	0.7117	0.042	Uiso
C14	0.7906	0.7068	0.3411	0.054	Uiso
C15	0.7213	0.2185	0.6511	0.042	Uiso
N16	0.6945	0.8631	0.4406	0.068	Uiso
C17	0.7054	0.0969	0.6221	0.053	Uiso
C18	0.9399	0.6992	0.2804	0.061	Uiso
C19	0.641	0.2698	0.6937	0.081	Uiso
N20	0.4713	0.0017	0.7029	0.065	Uiso
C21	0.8792	0.5258	0.3515	0.047	Uiso
C22	0.9426	0.5752	0.3039	0.059	Uiso
O23	0.6305	0.50109	0.53817	0.05	Uiso
C24	0.59729	0.61103	0.57591	0.05	Uiso
O25	0.76246	0.70371	0.5668	0.05	Uiso
C26	0.88398	0.90411	0.45176	0.05	Uiso
O27	0.88462	0.9055	0.53856	0.05	Uiso
O28	0.90174	0.16177	0.41076	0.05	Uiso
C29	0.8081	0.1835	0.433	0.05	Uiso

C30	0.9253	0.0234	0.393	0.05	Uiso
C31	0.72798	0.18723	0.36929	0.05	Uiso
C32	0.69918	0.66324	0.61648	0.05	Uiso
C33	0.6326	0.318	0.4329	0.05	Uiso
C34	0.5992	0.416	0.467	0.05	Uiso
O35	0.6403	0.1932	0.424	0.05	Uiso
C36	0.829	0.821	0.585	0.043	Uiso
C37	0.889	0.812	0.55	0.039	Uiso
H38	0.54366	1.08803	0.17484	0	Uiso
H39	0.57695	0.96527	0.24766	0	Uiso
H40	0.86632	0.35911	0.50875	0	Uiso
H41	0.73396	0.3966	0.53099	0	Uiso
H42	0.62609	-0.07452	0.62027	0	Uiso
H43	1.06942	0.52673	0.23437	0	Uiso
H44	1.0416	0.41305	0.31345	0	Uiso
H45	0.72945	0.55461	0.39939	0	Uiso
H46	0.83401	0.8613	0.27483	0	Uiso
H47	0.51038	0.23248	0.75321	0	Uiso
H48	0.76773	0.87724	0.4638	0	Uiso
H49	0.63011	0.89212	0.47263	0	Uiso
H50	0.76129	0.05857	0.58198	0	Uiso
H51	1.00086	0.74865	0.25009	0	Uiso
H52	0.64046	0.3711	0.7123	0	Uiso
H53	0.89285	0.4323	0.37929	0	Uiso
H54	0.54095	0.58599	0.61923	0	Uiso
H55	0.56637	0.67941	0.53406	0	Uiso
H56	0.80654	0.88371	0.43453	0	Uiso
H57	0.92922	0.82274	0.43105	0	Uiso
H58	0.80611	0.276	0.46372	0	Uiso
H59	0.78445	0.10711	0.4727	0	Uiso
H60	0.89693	0.00291	0.33523	0	Uiso
H61	1.00738	0.02096	0.39037	0	Uiso
H62	0.72755	0.09787	0.33542	0	Uiso

H63	0.7399	0.27266	0.33316	0	Uiso
H64	0.73306	0.58446	0.64966	0	Uiso
H65	0.68092	0.74336	0.65525	0	Uiso
H66	0.71134	0.34621	0.42626	0	Uiso
H67	0.6085	0.31995	0.37277	0	Uiso
H68	0.58702	0.48659	0.42099	0	Uiso
H69	0.52475	0.37441	0.47702	0	Uiso
H70	0.78873	0.91096	0.57237	0	Uiso
H71	0.85401	0.82042	0.64492	0	Uiso
H72	0.95757	0.78667	0.58125	0	Uiso
H73	0.88043	0.75665	0.49675	0	Uiso

**Table 2:** Bond lengths for sulfanilamide–polyethylene glycol (SXR D; Solution – II, Pca2<sub>1</sub>)

Atom	Atom	Length (Å)	Bond type
S1	O3	1.447	D
S1	O5	1.424	D
S1	C14	1.859	S
S1	N16	1.582	S
S2	N4	1.586	S
S2	O6	1.428	D
S2	O10	1.449	D
S2	C15	1.676	S
N4	H40	1.11	S
N4	H41	1.11	S
C7	C8	1.409	D
C7	C13	1.423	S
C7	N20	1.329	S
C8	C17	1.39	S
C8	H42	1.14	S
N9	C22	1.453	S
N9	H43	1.11	S
N9	H44	1.11	S
C11	C14	1.311	S

C11	C21	1.422	D
C11	H45	1.14	S
C12	C14	1.324	D
C12	C18	1.44	S
C12	H46	1.14	S
C13	C19	1.411	D
C13	H47	1.14	S
C15	C17	1.426	D
C15	C19	1.463	S
N16	H48	1.11	S
N16	H49	1.11	S
C17	H50	1.14	S
C18	C22	1.4	D
C18	H51	1.14	S
C19	H52	1.14	S
N20	H38	1.11	D
N20	H39	1.11	D
C21	C22	1.342	S
C21	H53	1.14	S
O23	C24	1.441	S
O23	C34	1.64	S
C24	C32	1.689	S
C24	H54	1.14	S
C24	H55	1.14	S
O25	C32	1.331	S
O25	C36	1.597	S
C26	O27	1.573	S
C26	H56	1.14	S
C26	H57	1.14	S
C26	C30	1.763	S
O27	C37	1.028	S
O28	C29	1.38	S
O28	C30	1.557	S

C29	C31	1.603	S
C29	H58	1.14	S
C29	H59	1.14	S
C30	H60	1.14	S
C30	H61	1.14	S
C30	C26	1.763	S
C31	O35	1.571	S
C31	H62	1.14	S
C31	H63	1.14	S
C32	H64	1.14	S
C32	H65	1.14	S
C33	C34	1.306	S
C33	O35	1.355	S
C33	H66	1.14	S
C33	H67	1.14	S
C34	H68	1.14	S
C34	H69	1.14	S
C36	C37	1.051	S
C36	H70	1.14	S
C36	H71	1.14	S
C37	H72	1.14	S
C37	H73	1.14	S
H38	N20	1.11	D
H39	N20	1.11	D

**Table 3:** Bond angles for sulfanilamide–polyethylene glycol (SXR; Solution – Ii, Pca<sub>21</sub>)

Atom	Atom	Atom	Bond angle (°)
N16	S1	C14	108.81
O3	S1	C14	107.38
O3	S1	O5	120.2
O3	S1	N16	106.7
O5	S1	N16	105.22
O5	S1	C14	108.13

O6	S2	O10	115.53
O10	S2	C15	108.66
N4	S2	C15	109.48
O6	S2	C15	107.07
O10	S2	N4	108.01
O6	S2	N4	107.99
C24	O23	C34	138.26
C32	O25	C36	119.98
C36	O27	C37	46.07
C26	O27	C37	101.11
C26	O27	C36	124.6
C29	O28	C30	114.82
C31	O35	C33	100.17
O23	C24	C32	102.21
O28	C29	C31	116.52
O35	C31	C29	94.78
O25	C32	C24	111.48
O35	C33	C34	151.94
O23	C34	C33	136.72
O25	C36	O27	132.69
O27	C36	C37	44.79
O25	C36	C37	105.1
O27	C37	C36	89.14
O23	C24	H55	109
C32	C24	H54	111
O23	C24	H54	111
C32	C24	H55	113
H54	C24	H55	111
H56	C26	H57	106
O27	C26	H56	106
O27	C26	H57	109
C31	C29	H58	108
C31	C29	H59	106

O28	C29	H59	109
H58	C29	H59	108
O28	C29	H58	108
O28	C30	H60	108
H60	C30	H61	108
O28	C30	H61	104
C29	C31	H63	109
H62	C31	H63	112
C29	C31	H62	112
O35	C31	H62	112
O35	C31	H63	116
C24	C32	H65	109
C24	C32	H64	109
H64	C32	H65	109
O25	C32	H64	109
O25	C32	H65	108
O35	C33	H67	86
O35	C33	H66	100
C34	C33	H67	109
H66	C33	H67	100
C34	C33	H66	100
O23	C34	H69	109
C33	C34	H68	104
O23	C34	H68	104
C33	C34	H69	95
H68	C34	H69	104
O27	C36	H71	113
H70	C36	H71	110
C37	C36	H70	110
C37	C36	H71	109
O25	C36	H70	110
O25	C36	H71	112
O27	C36	H70	67



O27	C37	H73	109
H72	C37	H73	113
O27	C37	H72	113
C36	C37	H72	113
C36	C37	H73	118
H43	N9	H44	120
C22	N9	H43	120
C22	N9	H44	120
C14	C11	C21	115.31
C14	C12	C18	116.11
S1	C14	C11	117.53
S1	C14	C12	112.72
C11	C14	C12	129.59
H48	N16	H49	120
S1	N16	H48	120
S1	N16	H49	120
C12	C18	C22	115.19
C11	C21	C22	118.39
N9	C22	C21	117.54
C18	C22	C21	123.71
N9	C22	C18	118.39
C14	C11	H45	120
C21	C11	H45	125
C18	C12	H46	124
C14	C12	H46	120
C12	C18	H51	120
C22	C18	H51	125
C22	C21	H53	122
C11	C21	H53	120
S2	N4	H40	120
S2	N4	H41	120
H40	N4	H41	120
N20	C7	C8	124.97

N20	C7	C13	121.84
C8	C7	C13	112.59
C7	C8	C17	124.31
C7	C13	C19	125.86
S2	C15	C19	122.57
S2	C15	C17	121.98
C17	C15	C19	114.97
C8	C17	C15	122.53
C13	C19	C15	118.6
C7	N20	H38	120
C7	N20	H39	120
H38	N20	H39	120
C7	C8	H42	120
C17	C8	H42	116
C7	C13	H47	120
C19	C13	H47	114
C8	C17	H50	120
C15	C17	H50	117
C13	C19	H52	120
C15	C19	H52	121

**Annexure 8:**

**Table 1:** Fractional atomic coordinates for sulfanilamide–polyethylene glycol (SXR);  
Solution – I, *Pbca*, DFT-D2 Optimized)

Atom	x	y	z
S1	0.47204	0.89859	0.6283
O2	0.36789	0.911	0.67713
O3	0.56557	0.97359	0.62236
N4	0.41112	0.89624	0.544
H5	0.47614	0.89654	0.50307
H6	0.34599	0.84341	0.53777
C7	0.6591	0.76492	0.62255
H8	0.70752	0.81564	0.58722
C9	0.71498	0.67833	0.63996
H10	0.80681	0.66071	0.61836
N11	0.71181	0.52789	0.70563
H12	0.67658	0.48456	0.74576
H13	0.7945	0.50961	0.68241
C14	0.54217	0.789	0.65156
C15	0.65592	0.61207	0.68824
C16	0.5371	0.63719	0.71677
H17	0.48853	0.58644	0.75236
C18	0.48094	0.72374	0.69868
H19	0.38856	0.74058	0.71952
O20	0.26064	0.59784	0.5968
C21	0.65833	0.84174	0.41022
C22	0.36106	0.61656	0.5471
O23	0.45698	0.56545	0.43061
C24	0.48484	0.6567	0.39863
C25	0.34976	0.55728	0.47599
C26	0.62426	0.67095	0.413
O27	0.68253	0.75034	0.37746
H28	0.44612	0.5948	0.5758
H29	0.36789	0.69438	0.53587

H30	0.42964	0.715	0.42369
H31	0.46453	0.65578	0.339
H32	0.26418	0.57311	0.44501
H33	0.34524	0.48094	0.49084
H34	0.67227	0.60646	0.39266
H35	0.63746	0.67568	0.47353
H36	0.63342	0.82933	0.46898
H37	0.58189	0.87837	0.38121

**Table 2:** Bond lengths for sulfanilamide–polyethylene glycol (SXR D; Solution – I, *Pbca*, DFT-D2 Optimized)

Atom	Atom	Length (Å)	Bond type
S1	O2	1.437	D
S1	O3	1.451	D
S1	N4	1.662	S
S1	C14	1.749	S
N4	H5	1.019	S
N4	H6	1.02	S
C7	H8	1.084	S
C7	C9	1.38	D
C7	C14	1.403	S
C9	H10	1.09	S
C9	C15	1.419	S
N11	H12	1.017	S
N11	H13	1.016	S
N11	C15	1.35	S
C14	C18	1.408	D
C15	C16	1.421	D
C16	H17	1.088	S
C16	C18	1.383	S
C18	H19	1.088	S
O20	C22	1.43	S
O20	C21	1.389	D

C21	O27	1.424	S
C21	H36	1.111	S
C21	H37	1.1	S
C21	O20	1.389	D
C22	C25	1.534	S
C22	H28	1.094	S
C22	H29	1.101	S
O23	C24	1.424	S
O23	C25	1.42	S
C24	C26	1.534	S
C24	H30	1.101	S
C24	H31	1.103	S
C25	H32	1.1	S
C25	H33	1.094	S
C26	O27	1.421	S
C26	H34	1.097	S
C26	H35	1.108	S

**Table 3:** Bond angles for sulfanilamide–polyethylene glycol (SXR D; Solution – I, *Pbca*, DFT-D2 Optimized)

Atom	Atom	Atom	Bond angle (°)
N4	S1	C14	111.98
O2	S1	C14	106.95
O2	S1	O3	119.97
O2	S1	N4	105.15
O3	S1	N4	102.64
O3	S1	C14	110.04
C21	O20	C22	115.51
C24	O23	C25	118.5
C21	O27	C26	114.87
O20	C21	O27	110.77
O20	C22	C25	111.85
O23	C24	C26	104.53

O23	C25	C22	112.3
O27	C26	C24	116.97
O27	C21	H36	108
O20	C21	H37	106
O20	C21	H36	112
O27	C21	H37	110
H36	C21	H37	110
H28	C22	H29	108
O20	C22	H28	106
O20	C22	H29	110
C25	C22	H28	109
C25	C22	H29	112
O23	C24	H31	110
H30	C24	H31	108
O23	C24	H30	112
C26	C24	H31	111
C26	C24	H30	111
C22	C25	H32	113
O23	C25	H33	105
O23	C25	H32	112
C22	C25	H33	108
H32	C25	H33	106
C24	C26	H35	108
O27	C26	H34	106
O27	C26	H35	110
H34	C26	H35	109
C24	C26	H34	107
S1	N4	H5	113
S1	N4	H6	113
H5	N4	H6	113
C9	C7	C14	120.78
C7	C9	C15	120.68
H12	N11	H13	118

C15	N11	H12	121
C15	N11	H13	121
S1	C14	C7	120.2
S1	C14	C18	120.25
C7	C14	C18	119.54
N11	C15	C9	120.33
N11	C15	C16	121.79
C9	C15	C16	117.88
C15	C16	C18	121.25
C14	C18	C16	119.86
C9	C7	H8	119
C14	C7	H8	120
C7	C9	H10	121
C15	C9	H10	119
C15	C16	H17	119
C18	C16	H17	120
C14	C18	H19	120
C16	C18	H19	120

**Annexure 9:**

**Table 1:** Fractional atomic coordinates for sulfanilamide–polyethylene glycol (SXRDX; Solution – I, *Pca2<sub>1</sub>*, DFT-D2 Optimized)

Atom	x	y	z
S1	0.57486	0.79856	0.34275
S2	0.92832	0.30878	0.62778
O3	0.56278	0.87193	0.27642
N4	0.90403	0.36779	0.54602
O5	0.50043	0.7104	0.36373
O6	0.93107	0.41587	0.6763
C7	0.66711	0.06944	0.68336
C8	0.74839	0.01795	0.64749
N9	0.94191	0.52546	0.29773
O10	1.01399	0.23418	0.61725

C11	0.69875	0.60319	0.36456
C12	0.75846	0.77281	0.28936
C13	0.67377	0.19369	0.7092
C14	0.68477	0.71988	0.33206
C15	0.83414	0.21226	0.65759
N16	0.58704	0.89079	0.41495
C17	0.8305	0.08771	0.6347
C18	0.84386	0.70861	0.27838
C19	0.75603	0.26347	0.69656
N20	0.58641	-0.00027	0.69549
C21	0.78417	0.53904	0.35419
C22	0.85841	0.58905	0.30952
O23	0.65368	0.38033	0.47764
C24	0.61257	0.47001	0.52659
O25	0.76265	0.57409	0.56137
C26	0.84392	0.82244	0.48522
O27	0.88833	0.82766	0.55603
O28	0.87855	0.04794	0.4647
C29	0.7826	0.09532	0.46135
C30	0.88741	-0.07627	0.43692
C31	0.75604	0.13859	0.38373
C32	0.6789	0.51044	0.5875
C33	0.60864	0.25829	0.36739
C34	0.57708	0.32035	0.43864
O35	0.65369	0.13991	0.37573
C36	0.79948	0.65335	0.61845
C37	0.89375	0.71333	0.595
H38	0.42564	1.07876	0.16564
H39	0.47212	0.95957	0.21923
H40	0.90238	0.30429	0.50431
H41	0.84895	0.43244	0.54526
H42	0.74645	-0.07803	0.62862
H43	0.98857	0.55349	0.25765



H44	0.95221	0.43904	0.31923
H45	0.64212	0.56179	0.39754
H46	0.74772	0.86321	0.26362
H47	0.61359	0.2344	0.73918
H48	0.60522	0.98017	0.39913
H49	0.52421	0.88952	0.44495
H50	0.89034	0.04515	0.60529
H51	0.90004	0.74974	0.24437
H52	0.7604	0.35829	0.71736
H53	0.79371	0.44893	0.38062
H54	0.54943	0.42993	0.55459
H55	0.58628	0.55136	0.4954
H56	0.76487	0.83104	0.48969
H57	0.85781	0.73248	0.45793
H58	0.77784	0.17241	0.50076
H59	0.7304	0.02477	0.47969
H60	0.85707	-0.08413	0.38099
H61	0.96521	-0.09189	0.43265
H62	0.78505	0.07379	0.34238
H63	0.7865	0.23062	0.37164
H64	0.70042	0.42952	0.62083
H65	0.63658	0.57282	0.62317
H66	0.65476	0.32068	0.33479
H67	0.54339	0.24051	0.3347
H68	0.52193	0.3898	0.42325
H69	0.54303	0.25112	0.47526
H70	0.74471	0.72391	0.63201
H71	0.81304	0.59819	0.66869
H72	0.93571	0.73509	0.64469
H73	0.934	0.64413	0.56242

**Table 2:** Bond lengths for sulfanilamide–polyethylene glycol (SXR D; Solution – I,  $Pca2_1$ ,

DFT-D2 Optimized)

Atom	Atom	Length (Å)	Bond type
S1	O3	1.448	D
S1	O5	1.452	D
S1	C14	1.754	S
S1	N16	1.65	S
S2	N4	1.647	S
S2	O6	1.449	D
S2	O10	1.446	D
S2	C15	1.754	S
N4	H40	1.019	S
N4	H41	1.033	S
C7	C8	1.414	D
C7	C13	1.418	S
C7	N20	1.365	S
C8	C17	1.383	S
C8	H42	1.087	S
N9	C22	1.362	S
N9	H43	1.019	S
N9	H44	1.017	S
C11	C14	1.399	S
C11	C21	1.384	D
C11	H45	1.083	S
C12	C14	1.403	D
C12	C18	1.385	S
C12	H46	1.088	S
C13	C19	1.385	D
C13	H47	1.088	S
C15	C17	1.403	D
C15	C19	1.406	S
N16	H48	1.034	S
N16	H49	1.027	S

C17	H50	1.087	S
C18	C22	1.418	D
C18	H51	1.087	S
C19	H52	1.088	S
N20	H38	1.016	D
N20	H39	1.018	D
C21	C22	1.416	S
C21	H53	1.089	S
O23	C24	1.429	S
O23	C34	1.43	S
C24	C32	1.501	S
C24	H54	1.1	S
C24	H55	1.103	S
O25	C32	1.429	S
O25	C36	1.434	S
C26	O27	1.424	S
C26	H56	1.103	S
C26	H57	1.103	S
C26	C30	1.522	S
O27	C37	1.42	S
O28	C29	1.427	S
O28	C30	1.432	S
C29	C31	1.527	S
C29	H58	1.096	S
C29	H59	1.1	S
C30	H60	1.101	S
C30	H61	1.095	S
C30	C26	1.522	S
C31	O35	1.427	S
C31	H62	1.099	S
C31	H63	1.098	S
C32	H64	1.1	S
C32	H65	1.101	S

C33	C34	1.518	S
C33	O35	1.426	S
C33	H66	1.099	S
C33	H67	1.099	S
C34	H68	1.105	S
C34	H69	1.103	S
C36	C37	1.519	S
C36	H70	1.101	S
C36	H71	1.103	S
C37	H72	1.097	S
C37	H73	1.102	S
H38	N20	1.016	D
H39	N20	1.018	D

**Table 3:** Bond angles for sulfanilamide–polyethylene glycol (SXR D; Solution – I,  $Pca2_1$ , DFT-D2 Optimized)

Atom	Atom	Atom	Bond angle (°)
N16	S1	C14	106.75
O3	S1	C14	105.76
O3	S1	O5	119.39
O3	S1	N16	110.06
O5	S1	N16	104.91
O5	S1	C14	109.41
O6	S2	O10	119.93
O10	S2	C15	108.95
N4	S2	C15	110.65
O6	S2	C15	107.62
O10	S2	N4	105.25
O6	S2	N4	104.2
C24	O23	C34	108.33
C32	O25	C36	109.57
C26	O27	C37	115.91
C29	O28	C30	113.45

C31	O35	C33	117.11
O23	C24	C32	114.03
O27	C26	C30	108.57
O28	C29	C31	111.91
O28	C30	C26	115.52
O35	C31	C29	109.68
O25	C32	C24	113.16
O35	C33	C34	115.37
O23	C34	C33	113.85
O25	C36	C37	110.92
O27	C37	C36	117.46
C32	C24	H55	110
H54	C24	H55	106
O23	C24	H54	110
O23	C24	H55	110
C32	C24	H54	105
H56	C26	H57	106
O27	C26	H56	111
O27	C26	H57	111
C30	C26	H56	112
C30	C26	H57	107
C31	C29	H58	111
C31	C29	H59	109
O28	C29	H59	111
H58	C29	H59	107
O28	C29	H58	107
O28	C30	H60	111
C26	C30	H60	109
H60	C30	H61	107
O28	C30	H61	105
C26	C30	H61	109
C29	C31	H62	110
O35	C31	H63	111

H62	C31	H63	107
C29	C31	H63	111
O35	C31	H62	108
C24	C32	H64	110
C24	C32	H65	106
H64	C32	H65	108
O25	C32	H65	110
O25	C32	H64	110
H66	C33	H67	107
C34	C33	H66	111
O35	C33	H66	110
O35	C33	H67	105
C34	C33	H67	107
O23	C34	H68	110
H68	C34	H69	108
O23	C34	H69	109
C33	C34	H68	106
C33	C34	H69	110
H70	C36	H71	108
C37	C36	H70	111
C37	C36	H71	108
O25	C36	H71	110
O25	C36	H70	109
O27	C37	H73	110
C36	C37	H73	107
O27	C37	H72	105
H72	C37	H73	108
C36	C37	H72	109
H43	N9	H44	117
C22	N9	H43	120
C22	N9	H44	121
C14	C11	C21	120.55
C14	C12	C18	120.02

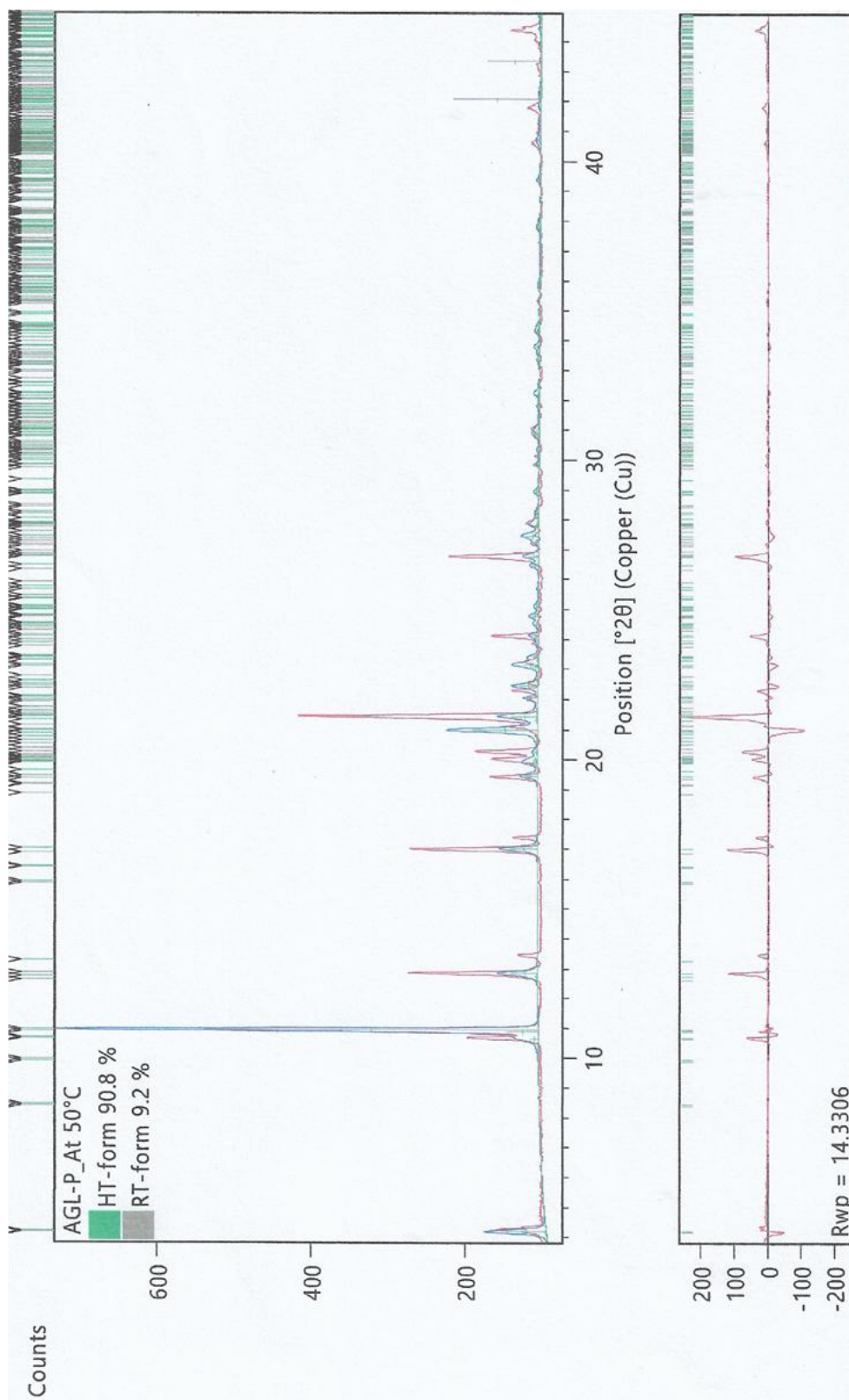
S1	C14	C11	120.41
S1	C14	C12	119.92
C11	C14	C12	119.66
H48	N16	H49	111
S1	N16	H48	111
S1	N16	H49	109
C12	C18	C22	121.07
C11	C21	C22	120.75
N9	C22	C21	121.18
C18	C22	C21	117.91
N9	C22	C18	120.9
C14	C11	H45	120
C21	C11	H45	119
C18	C12	H46	120
C14	C12	H46	120
C12	C18	H51	119
C22	C18	H51	119
C22	C21	H53	120
C11	C21	H53	119
S2	N4	H40	114
S2	N4	H41	115
H40	N4	H41	115
N20	C7	C8	120.87
N20	C7	C13	121.15
C8	C7	C13	117.93
C7	C8	C17	121.42
C7	C13	C19	120.61
S2	C15	C19	119.83
S2	C15	C17	120.03
C17	C15	C19	119.65
C8	C17	C15	119.84
C13	C19	C15	120.41
C7	N20	H38	120

C7	N20	H39	119
H38	N20	H39	117
C7	C8	H42	120
C17	C8	H42	119
C7	C13	H47	120
C19	C13	H47	120
C8	C17	H50	119
C15	C17	H50	121
C13	C19	H52	120
C15	C19	H52	120



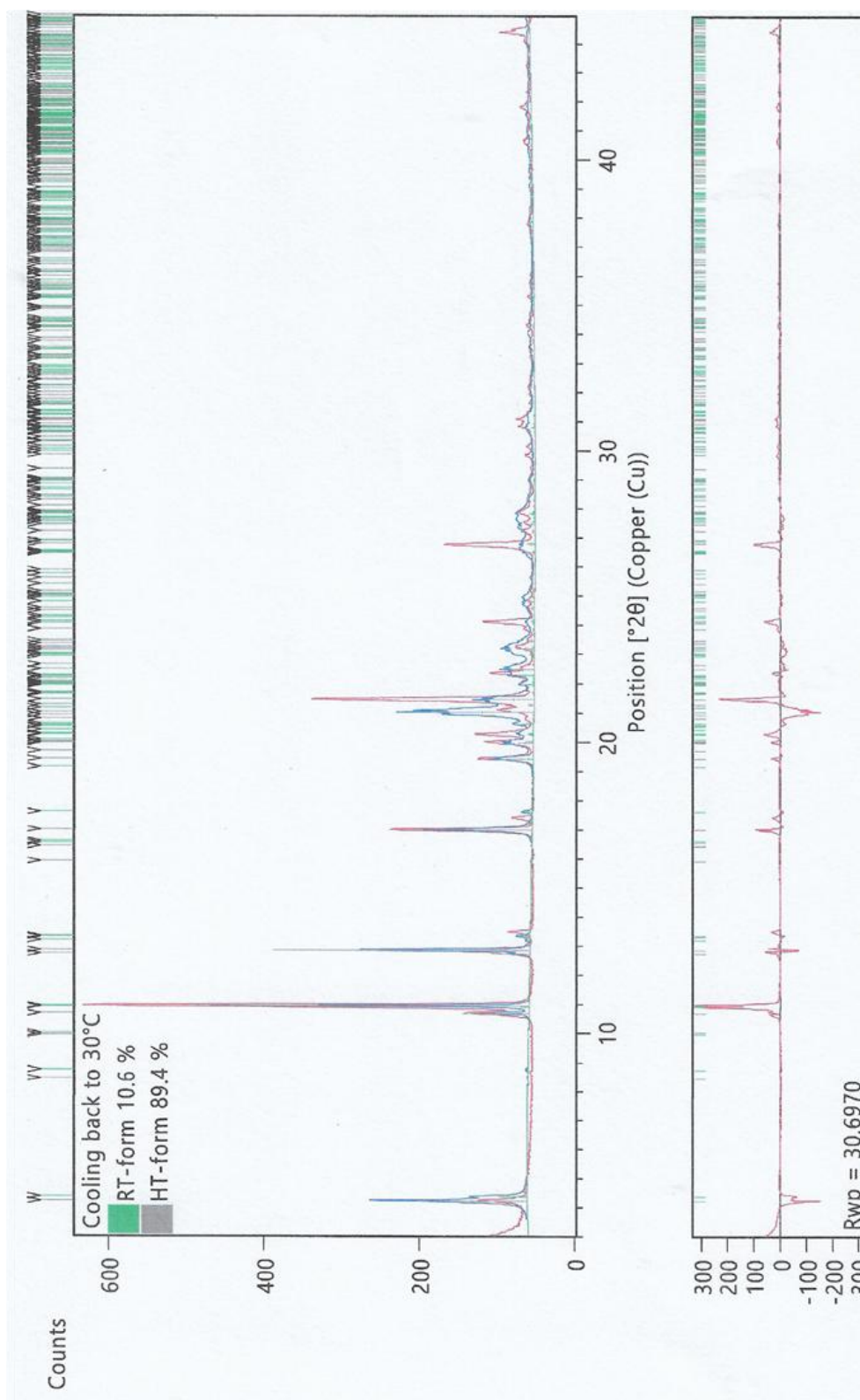
**Annexure 10:**

Experimental, calculated and the difference profile of AGL-P co-crystal VT-PXRD data at 50°C obtained from Rietveld refinement.



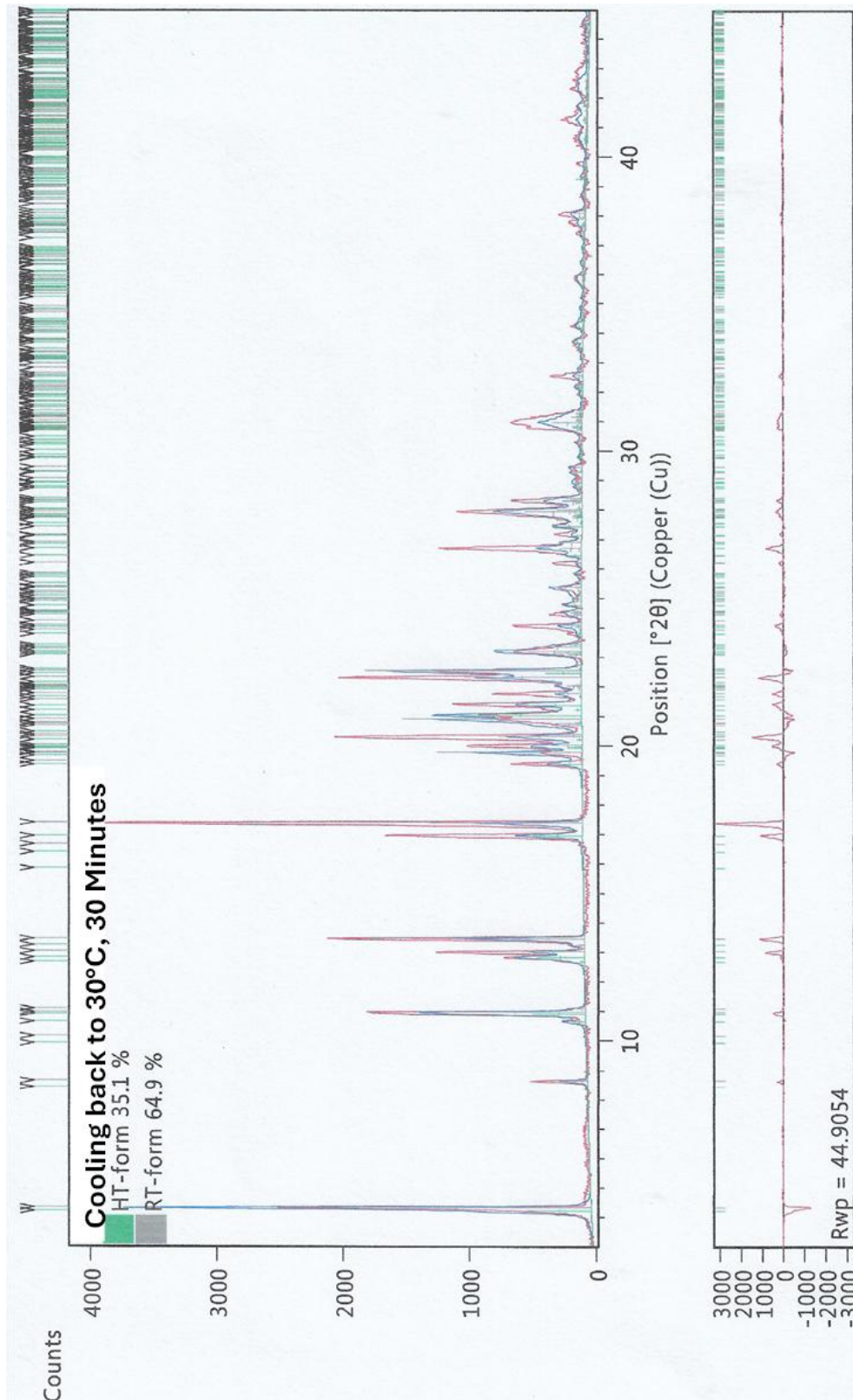
**Annexure 12:**

Experimental, calculated and the difference profile of AGL-P co-crystal VT-PXRD data at 50°C obtained from Rietveld refinement.



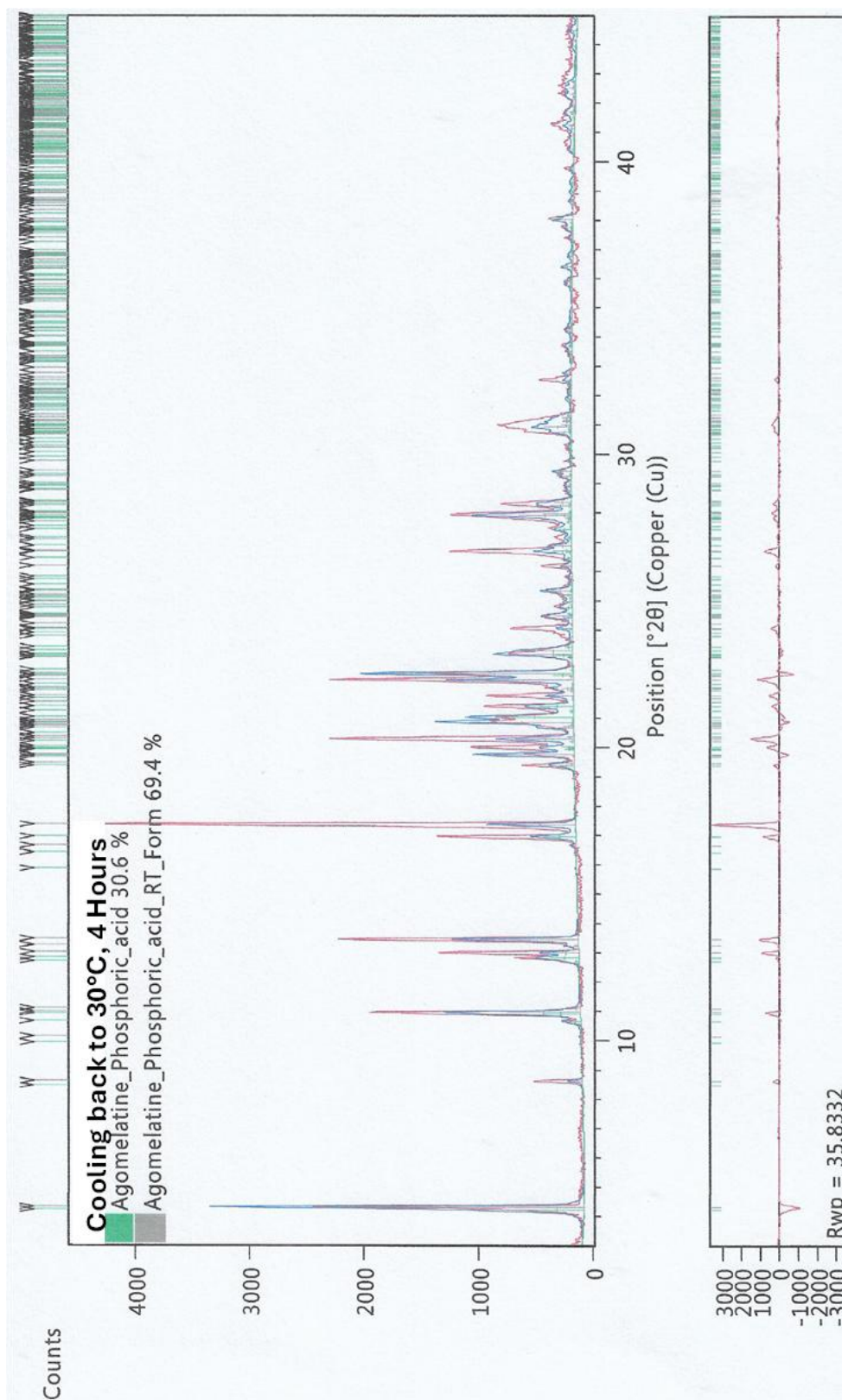
**Annexure 12:**

Experimental, calculated and the difference profile of AGL-P co-crystal VT-PXRD data by cooling back to 30°C after 30 minutes obtained from Rietveld refinement.



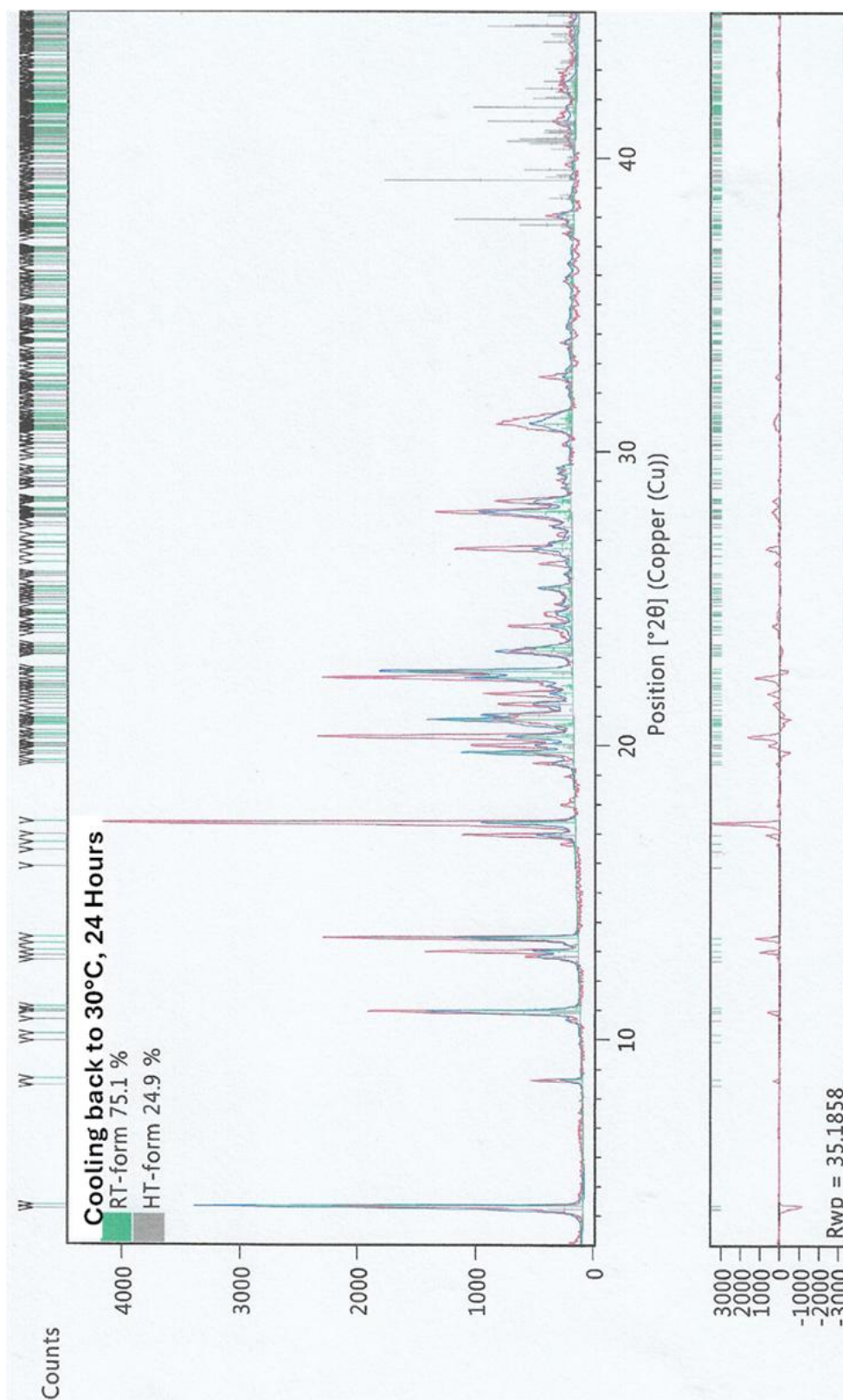
**Annexure 13:**

Experimental, calculated and the difference profile of AGL-P co-crystal VT-PXRD data by cooling back to 30°C after 4 hours obtained from Rietveld refinement.



**Annexure 14:**

Experimental, calculated and the difference profile of AGL-P co-crystal VT-PXRD data by cooling back to 30°C after 24 hours obtained from Rietveld refinement.



### List of Publications

- **Voguri, R. S.**; Ranga, S.; Dey, A.; Ghosal, S. Solid-State Phase Transition of Agomelatine–Phosphoric Acid Molecular Complexes along the Salt–Cocrystal Continuum: Ab Initio Powder X-Ray Diffraction Structure Determination and DFT-D2 Analysis. *Cryst. Growth Des.* **2020**, *20* (12), 7647–7657.
- Chappa, P.; Maruthapillai, A.; **Voguri, R. S.**; Dey, A.; Ghosal, S.; Basha, M. A. Drug–Polymer Co-Crystals of Dapsone and Polyethylene Glycol: An Emerging Subset in Pharmaceutical Co-Crystals. *Cryst. Growth Des.* **2018**, *18* (12), 7590–7598.
- Ganesan, T.; Muthudoss, P.; **Voguri, R. S.**; Ghosal, S.; Ann, E. Y. C.; Kwok, J.; Shahnawaz, S. S.; Omar, M. F.; Allada, R.; See, H. H. A New Febuxostat-Telmisartan Drug-Drug Cocrystal for Gout-Hypertension Combination Therapy. *J. Pharm. Sci.* **2022**, *111* (12).
- **Voguri, R. S.**; Chappa, P.; Dey, Chakravarty, M.; Ghosal, S.; Dey, A.; Polymeric Cocrystal of Sulfanilamide and Polyethylene glycol: From Virtual Screening to Structure-Property-Process Correspondence (**Manuscript under preparation**)
- **Voguri, R. S.**; Dey, Chakravarty, M.; Ghosal, S.; Dey, A.; Crystal structure of L–Proline and Salicylic Acid molecular complex (**Manuscript under preparation**)

### Biography of the candidate



Mr. *Raja Sekhar Voguri* was born in Guntur, Andhra Pradesh. He has completed his Bachelor of Science in Chemical Technology from Loyola Academy, Hyderabad, Telangana; Master of Science in Chemistry (Organic chemistry specialization) from ANU, Guntur, A.P and Master of Research in Nano Chemistry & Materials Chemistry from University of Birmingham, UK. He worked as a Lecturer for 2<sup>nd</sup> year M.Sc Analytical Chemistry Coarse, Bapatla Engineering College, Andhra Pradesh from 2011 to 2012. He worked as a Scientist in Dr. Reddy's from 2012 to 2018. He is currently working as Associate Scientific Manager-II at Biocon Limited. He has joined the Department of Chemistry, BITS-Pilani, Hyderabad campus for his doctoral research under the supervision of Prof. Subhas Ghosal, Prof. Manab Chakravarty, and Dr. Archan Dey in 2016. He has published 7 scientific publications in well-renowned international journals and presented papers at various national and international conferences.

### Biography of the supervisor



*Prof. Manab Chakravarty* was born in Hooghly, West Bengal. He completed his M. Sc. from Visva-Bharati University, Shantiniketan with the specialization in inorganic chemistry. By securing CSIR junior/senior research fellowship, he registered for Ph.D. degree in organophosphonate chemistry from School of Chemistry, University of Hyderabad under the supervision of Prof. K. C. Kumara Swamy. After completing Ph. D. in 2007, he joined at Chembiotek Pvt. Ltd as a research associate for few months and then moved to department of chemistry, University of New Mexico, Albuquerque, New Mexico, USA for his postdoctoral studies with Prof. Robert T. Paine on synthesis of phosphorus-based molecules that act as ligands for radioactive elements. Prof. Manab Chakravarty joined BITS Pilani, Hyderabad campus in 2010 as an assistant professor and promoted as associate professor in 2015 then promoted to full professor in 2021. He was appointed as Head of the department from 2016-2020. He has 23 years of research experience and 13 years of teaching experience. His research interests are developing fluorophores with AIE properties with applications as piezofluorochromic, acidchromic, solvatochromic and thermochromic properties with a group of two Ph.D students and a post-doctoral fellow. He has completed three research projects as PI sponsored by DST and CSIR. Currently, he has two major projects from DST and LSRB.



### Biography of the co-supervisor



*Dr. Subhas Ghosal* was born in 1977 at Bolpur, West Bengal, India. He had completed his Master's degree in Physical Chemistry from Visva Bharati University, Santiniketan, India in 2001. Following Master's degree, he had pursued his PhD degree in theoretical Chemistry from the University of Hyderabad, India in 2007. Further, he has worked as a Royal Society of Chemistry (Landon) Post-doctoral Fellow during 2007 - 08 at University of Durham, UK and later worked as a Postdoctoral Research Associate at University of Connecticut, USA from 2009 - 12. He worked as an Assistant Professor in Chemistry at the Birla Institute of Technology and Science Hyderabad from 2012 to 2018. Currently, he is working as an Associate Professor at National Institute of Technology, Durgapur India. He has around twenty years of research experience and more than ten years of teaching experience. His research interests are understanding the correlation between structure and reactivity of experimental systems through theoretical calculations using Density Functional Theory (DFT) and Molecular Dynamics (MD) Simulations. His research interest also expands to theoretical investigation of hydrocarbon molecules having astrochemical Interest and Low-temperature Chemical reaction dynamics.

### Biography of the co-supervisor



*Dr Archan Dey* was born in Bhastara (Hooghly, West Bengal, 1978). He completed B.Sc. (Hons) and M.Sc. in chemistry, specialization in organic chemistry from the University of Burdwan (Bejoy Narayan Mahavidyalaya, Itachuna, 1999) and Visva-Bharat University (Santiniketan, 2001) respectively. Awarded Ph.D. in chemical science (Solid-State Supramolecular Structural Chemistry; Crystal Engineering) in 2007 from the University of Hyderabad. Following a post-doctoral stint in the Laboratory of Nanostructured Fluorinated

Materials, Milan, Italy, he moved to the Laboratory of Materials and Interface Chemistry, TU/e (Netherlands) in 2008 as a Research Associate. I had worked in the research areas of Biomineralization and Biomimetic Materials using Advanced Electron Microscopy techniques. He had successfully, handled two projects: *Hierarchical Materials for Load-Bearing Bone Replacement* (European Project) and *Biomimetic Magnetic Materials for MRI Agents* (NOW Project). During these Ph.D. and Postdoctoral periods, He had published several research articles and book chapters (~25) that includes high-impact journals, such as Nature Materials and Journal of America Chemical Society). He joined Dr. Reddy's Lab as senior scientist, 2012. Currently, he leads the Advanced Characterization Technology Lab with focus on developing generics drug products by means of reverse engineering, pharmaceuticals material science concepts and cutting-edge technologies. His research interest is understanding the crystallization (self-assembly & co-self-assembly processes, eventually, utilize this knowledge to develop materials or protocols for real-world applications. In the DRL tenure also, he has been continuing on the scientific research track by patents (2), articles (3), and symposium communications/invited talks (3). He has been co-supervising two Ph.D. students. He was selected for the prestigious Ramanujan Fellowship Award 2013 (India). He is a member of the International Centre for Diffraction Data (ICDD).