Exploring and Anticipating Supramolecular Synthons: From Fundamental Science to Practical Applications

by

Bhupinder Kaur Sandhu

B.S., New Mexico Highlands University, 2013

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Chemistry
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2018

Abstract

Four different methods; molecular electrostatic potentials (MEPs), hydrogen-bond energies (HBE), hydrogen-bond propensities (HBP) and hydrogen-bond coordination (HBC) were used for mapping out the structural landscape of twelve pyrazole and twelve thiazole based molecules. In seven out of eight crystal structures obtained in pyrazoles, a combination of HBE and HBP predicted the experimentally observed synthons correctly. In all eight crystal structures obtained in thiazoles, the synthons were predicted correctly using all four methods.

A series of co-crystallizations between twelve pyrazole with twenty carboxylic acids (240 experiments), and twelve thiazole with twenty carboxylic acids (240 experiments) were carried out to build an experimental library that could be used for evaluating the ability of electrostatics, energies, propensities and molecular complementarity methods to rationalize the observed intermolecular interactions. The results suggested that a combination of electrostatics and molecular complementarity are essential for identifying the predominant molecular recognition events in the pyrazole based study, and methods such as MEPs, HBE, and HBP all predicted the observed synthons in co-crystals of the thiazole-based molecules.

In order to examine competition between hydrogen and halogen bonds, and to synthesize ternary co-crystals, four thiazole based molecules were co-crystallized with 15 hydrogen-bond donors and one halogen bond donor resulting in new co-crystals in 44 out of 60 experiments, and the crystal structures of two ternary co-crystals were obtained.

A series of eight unactivated and activated amide functionalized molecules were synthesized to establish a supramolecular halogen-bond hierarchy. The positive electrostatic potential on the halogen atoms was enhanced through an *sp*-hybridized carbon and electron-withdrawing fluoro group(s) next to amide group. Tetraflourinated and iodoethynyl based molecules were identified as the most effective halogen-bond donors and were therefore least successful for co-crystal synthesis.

In order to predict crystallizability of 83 drug-like molecules a molecule, logistic regression approach was employed using molecular descriptors such as molecular weight, rotatable bond, surface area, heteroatom, melting temperature, glass transition temperature, and molecular shape/volume. Four different models were developed, and the success rate was above 85% (using experimental DSC data for the crystallization classification).

Finally, the solid-form landscape of urea was explored using full interaction maps (FIMs), and data from the CSD to develop optimum protocols for synthesizing co-crystals of this compound. As a result, 49 of 60 attempted reactions produced new co-crystals. Moreover, the goal of reducing solubility and lowering hygroscopicity of the parent compound was achieved, which, in turn, offers new opportunities for a slow-release fertilizer with limited hygroscopicity thereby reducing many current problems of transport, handling, and storage of urea.

Exploring and Anticipating Supramolecular Synthons: From Fundamental Science to Practical Applications

by

Bhupinder Kaur Sandhu

B.S., New Mexico Highlands University, 2013

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Chemistry
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2018

Approved by:

Major Professor
Dr. Christer B Aakeröy

Copyright

Abstract

Four different methods; molecular electrostatic potentials (MEPs), hydrogen-bond energies (HBE), hydrogen-bond propensities (HBP) and hydrogen-bond coordination (HBC) were used for mapping out the structural landscape of twelve pyrazole and twelve thiazole based molecules. In seven out of eight crystal structures obtained in pyrazoles, a combination of HBE and HBP predicted the experimentally observed synthons correctly. In all eight crystal structures obtained in thiazoles, the synthons were predicted correctly using all four methods.

A series of co-crystallizations between twelve pyrazole with twenty carboxylic acids (240 experiments), and twelve thiazole with twenty carboxylic acids (240 experiments) were carried out to build an experimental library that could be used for evaluating the ability of electrostatics, energies, propensities and molecular complementarity methods to rationalize the observed intermolecular interactions. The results suggested that a combination of electrostatics and molecular complementarity are essential for identifying the predominant molecular recognition events in the pyrazole based study, and methods such as MEPs, HBE, and HBP all predicted the observed synthons in co-crystals of the thiazole-based molecules.

In order to examine competition between hydrogen and halogen bonds, and to synthesize ternary co-crystals, four thiazole based molecules were co-crystallized with 15 hydrogen-bond donors and one halogen bond donor resulting in new co-crystals in 44 out of 60 experiments, and the crystal structures of two ternary co-crystals were obtained.

A series of eight unactivated and activated amide functionalized molecules were synthesized to establish a supramolecular halogen-bond hierarchy. The positive electrostatic potential on the halogen atoms was enhanced through an *sp*-hybridized carbon and electron-withdrawing fluoro group(s) next to amide group. Tetraflourinated and iodoethynyl based molecules were identified as the most effective halogen-bond donors and were therefore least successful for co-crystal synthesis.

In order to predict crystallizability of 83 drug-like molecules a molecule, logistic regression approach was employed using molecular descriptors such as molecular weight, rotatable bond, surface area, heteroatom, melting temperature, glass transition temperature, and molecular shape/volume. Four different models were developed, and the success rate was above 85% (using experimental DSC data for the crystallization classification).

Finally, the solid-form landscape of urea was explored using full interaction maps (FIMs), and data from the CSD to develop optimum protocols for synthesizing co-crystals of this compound. As a result, 49 of 60 attempted reactions produced new co-crystals. Moreover, the goal of reducing solubility and lowering hygroscopicity of the parent compound was achieved, which, in turn, offers new opportunities for a slow-release fertilizer with limited hygroscopicity thereby reducing many current problems of transport, handling, and storage of urea.

Table of Contents

List of Figures

List of Tables

Glossary of terms

- **Hydrogen-bond propensity:** A model to predict which donors and acceptors form hydrogen bonds in a crystal structure, based on the statistical analysis of hydrogen bonds in the Cambridge Structural Database (CSD).

- **Hydrogen-bond coordination:** A model indicating how many hydrogen-bond interactions a donor and acceptor can have.

- **Hydrogen-bond energies:** A term to calculate the interaction energies between a donor and acceptor.

- **Molecular electrostatic potentials:** The molecular electrostatic potential (MEP) at a given point $p(x,y,z)$ in the vicinity of a molecule is the force acting on a positive test charge (a proton) located at p through the electrical charge cloud generated through the molecules electrons and nuclei.

- **Synthon polymorphism:** Synthon polymorphism in a molecule occurs when the primary synthons in the forms are different.

- **Logistic regression model:** It is a statistical model that is usually taken to apply to a binary dependent variable. It is used to estimate the probability of a binary response based on one or more predictor (or independent) variables (features).

- **Full interaction maps:** The tool calculates regions around the molecule (maps) where chemical probe groups are likely to be found, based on pre-extracted Isostar interaction data from the CSD. The calculation procedure first identifies distinct functional groups in the molecules viewed, and then finds relevant interaction data in IsoStar. Next it pulls together the group-based interaction data and takes into account the environmental effects of combinative factors and steric exclusion to create a full 3D picture of molecular interaction preferences.

Acknowledgements

I would like to express my deepest gratitude to my advisor and mentor, Professor Christer Aakeröy for his tremendous support, patience and encouragement throughout my five years of Ph.D. at K-State. His continuous motivation, selfless attention helped me to become a better presenter, public speaker, writer, and overall a better scientist. I am so thankful to you for giving me an opportunity to work with you and making me a strong, and confident individual.

I would also like to thank my Ph.D. advisory committee; Professor Ryan Rafferty, Professor Tendai Gadzikwa, Professor Jennifer Anthony and Professor Vinod Kumarappan for their invaluable time and encouragement.

I would also like to extend special thank you to Professor Eric Maatta for serving on my committee for three years and providing invaluable support and encouragement.

I am very grateful to Dr. John Desper, Dr. Marijana Đaković and Dr. Abhijeet Sinha for solving all the crystal structures mentioned in this thesis and for always being very patient and helpful. A special thank you to Dr. Abhijeet Sinha for being such a great friend and always supportive and willing to help even when you were busy.

I also want to acknowledge Dr. Yasmin Patell (for being very supportive and caring friend), Dr. Leila Maurmann (for helping me with NMR), Mr. Ron Jackson (for fixing broken instruments and talks), Mr. Tobe Eggers (for fixing broke instruments), Mr. Michael Hinton (for teaching me lab skills), Mr. Jim Hodgson (for fixing broken glassware), Ms. Mary Dooley, Ms. Kimberly Ross and Ms. Lisa Percival for making my research life smooth at K-State.

A special thank you to my collaborators; Dr. Amy A. Sarjeant and Dr. Shyam Vyas, Dr. Jonas Nyman and Dr. Susan M. Reutzel-Edens. Thank you for your inputs, discussions and always willing to do skype meeting to help me advance in our crystallizability project.

Dedication

To my Grandfather, parents and brother!

For always loving and supporting me……..

Preface

Research carried out at Kansas State University for this dissertation led to the following publications in scientific journals:

1. Sandhu B., Sinha A., Desper J., Aakeroy C., "Modulating physical properties of solid forms of urea using co-crystallization technology" *Chem. Commun.***, 2018**, Advance Article.

2. Sandhu B., McLean A., Sinha A., Desper J., Sarjeant A., Vyas S., Reutzel-Edens S., Aakeroy C., "Evaluating competing intermolecular interactions through molecular electrostatic potentials and hydrogen-bond propensities" *Cryst. Growth Des.*, **2018**, *18* (1), pp 466–478.

Chapter 1 - Introduction

1.1 Crystal structure-property relationship

"Structure Determines Properties"[1] is a powerful concept in chemistry and plays an important role in all fields including environmental science, biology, biochemistry, polymer science, medicine, engineering, and nutrition.[2] Crystallization performed under different conditions can yield crystals of different sizes and morphologies, thus providing a way of modifying particles to desired specifications. The structure of a compound including form, size, shape, conformation, functional groups, intermolecular interactions have the potential to impact the purification, solubility, chemical and physical stability, melting point, particle size and mechanical properties of a molecule of interest (Figure 1.1). The desired properties in a compound can be achieved via two ways: covalent modification to design and synthesize molecule with desired properties (covalent synthesis) or supramolecular synthesis via non-covalent interactions.
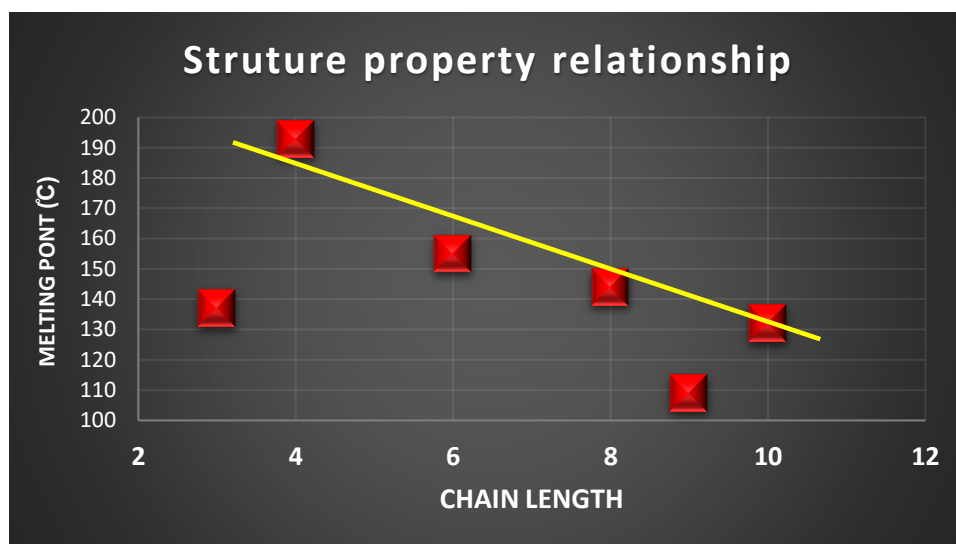
Figure 1.1 Correlation of the melting points with the dicarboxylic acid chain length[3]

1.2 Covalent chemistry vs supramolecular chemistry

Covalent chemistry involves breakage or formation of a covalent bond to perform multi-step reactions, and to isolate intermediates and products with desired yield. Organic synthesis can be defined in a simplistic way as – 'the conversion of A to B, using C as the reagent' and has been

studied over the years by organic chemists to better understand the structure, reactivity and reaction pathways (Scheme 1.1).[4]
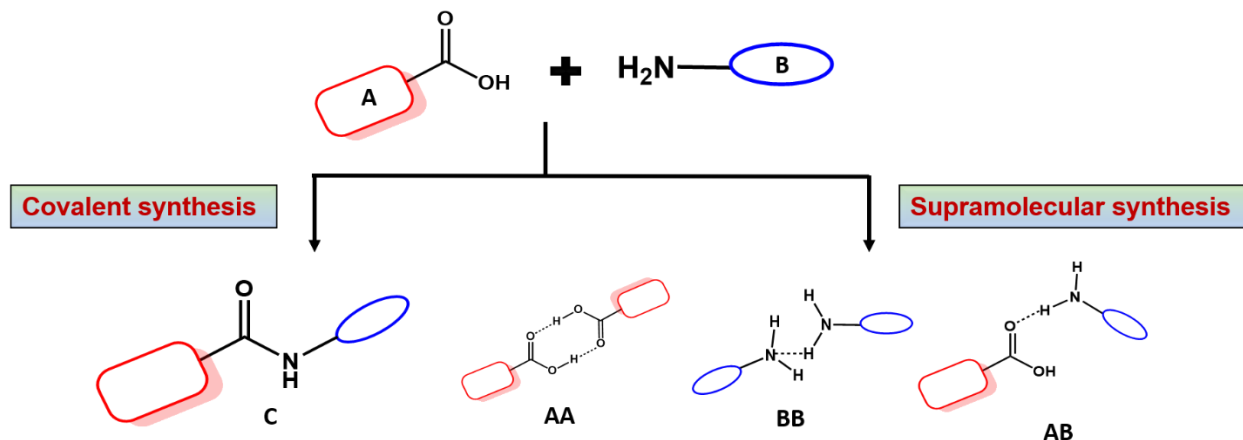
Scheme 1.1 Covalent vs supramolecular synthesis.

Supramolecular synthesis on other hand is based on non-covalent interactions and is largely unexplored, compared to organic covalent synthesis. Supramolecular chemistry is defined as "the chemistry of molecular assemblies and of the intermolecular bond," by Jean-Marie Lehn, who won the Nobel Prize in 1987 for his outstanding contributions to the field.[5] It can be achieved by the use of reversible intermolecular interactions such as hydrogen-bonds, halogen-bonds, vander waals and π-π interaction. The strength of each interaction is given in Table 1.1.

Table 1.1 Estimated strength of intermolecular interactions

Types of interaction	Strength (Kj/mol)
Covalent	100-4400
Hydrogen bond	10-70
Halogen bond	5-180
π- π	0-50
Vander Waals	<5

1.3 Supramolecular synthons

 The design of organic crystals relies on the concept of Supramolecular synthons,[6-7] a structural unit within a supermolecule which can be formed and/or assembled by known or conceivable

synthetic operations involving intermolecular interactions.[6] These synthons defines the structure of a molecule. However, robustness of these synthons is an issue because of weak, labile and deforming nature of non-covalent interactions. If a well-defined strategy is developed to generate robust synthons, then the protocol can be carried over from one molecule to another with generality. Supramolecular synthons are divided into two types; homomeric synthons (occurs within the individual molecules) are achieved by re-crystallization of individual molecules and heteromeric synthons (occurs between two different molecules) are achieved by co-crystallization of two or more molecules (Scheme 1.2).[8-9]
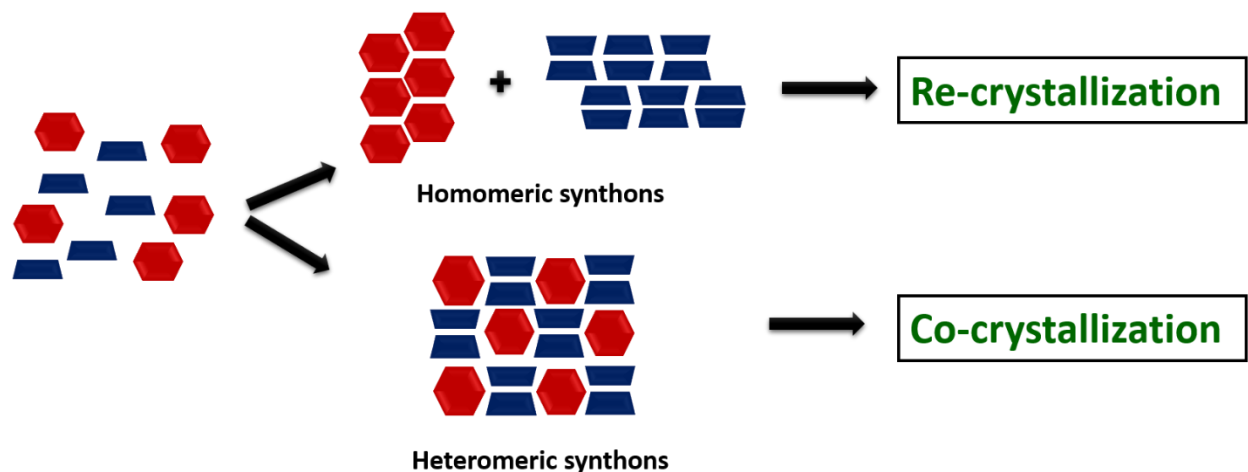
Scheme 1.2 Supramolecular synthons

1.4 Co-crystals

The definition of a co-crystal is still under debate. Thus, for the scope of this dissertation, a co-crystal will be defined as, solids that are crystalline single-phase materials composed of two or more different molecular compounds generally in a stoichiometric ratio.[10-11] A co-crystallization is a deliberate attempt at bringing together different molecular species within one periodic crystalline lattice without making or breaking covalent bonds via heteromeric interactions.[12] It is different from naturally occurring phenomenon called recrystallization in which homomeric interactions dominate in the crystal lattice (Scheme 1.2). The ability of a molecule to form a co-crystal depends on various factors, including the types of co-former, the stoichiometric ratio, the solvents, the temperature, the pressure, the crystallization technique, etc. Co-crystals have different crystal structures than the pure components, contain different intermolecular packing patterns, and as such they often exhibit widely different physical properties than the pure components.

3
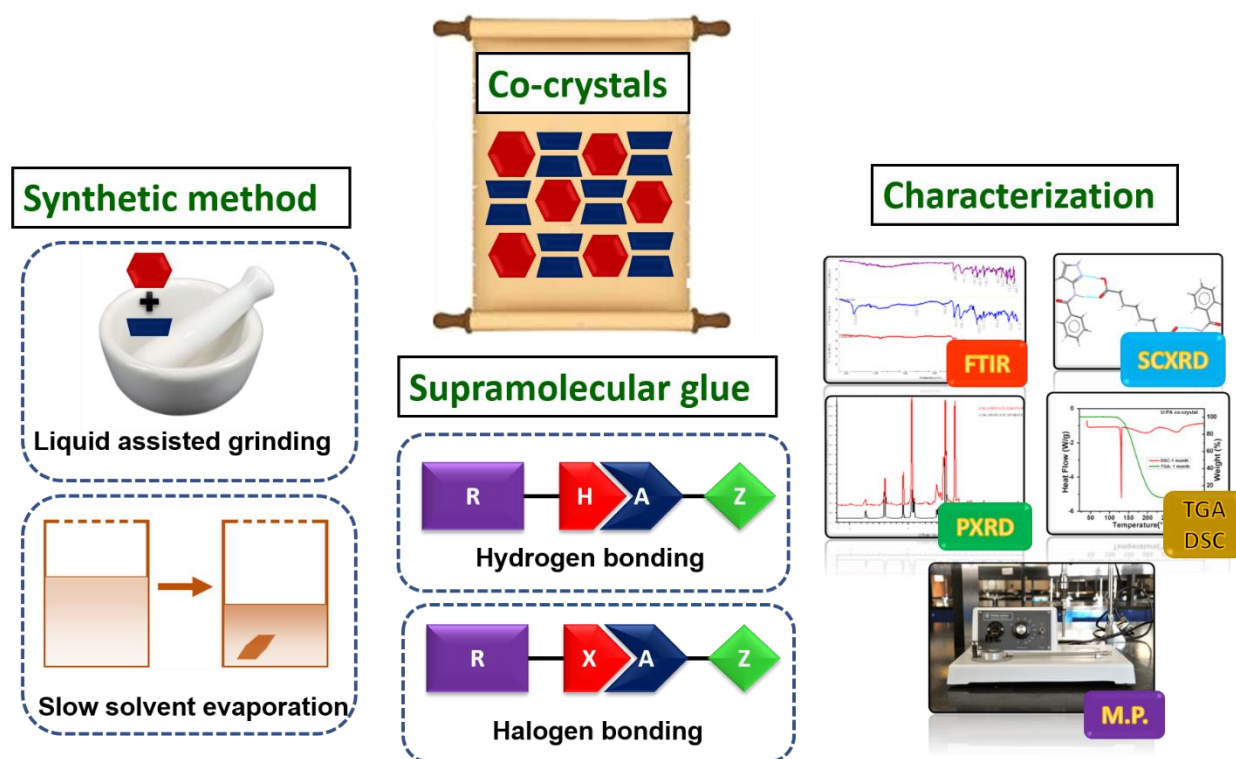
Figure 1.2 Summary of synthetic method, supramolecular interactions and characterization method used to study co-crystals in this thesis.

Co-crystal synthesis can be performed via several methods, including slow solvent evaporation, slurrying, liquid assisted grinding, melt (hot stage microscopy), solution crystallization and co-sublimation techniques.[13-16] Among these methods, liquid assisted grinding has been reported to be a cost-effective, green, and reliable method for discovery of new co-crystals as well as for preparation of existing co-crystals.[17] The components in a co-crystal exist in a definite stoichiometric ratio, and assemble via non-covalent interactions such as π-π stacking interactions,[18] dipole-dipole interactions,[19] halogen bonds,[20] and hydrogen bonds.[21] For this dissertation, only halogen and hydrogen bonds will be covered. The structure characterization of co-crystals can be done using infra-red spectroscopy, single crystal x-ray crystallography and powder x-ray diffraction and the physical properties using melting point apparatus, differential scanning calorimetry, thermogravimetric analysis (Figure 1.2).[22-24]

1.4.1 Hydrogen bonding

A hydrogen bond (HB) is defined as "an attractive interaction between a hydrogen atom from a molecule or a molecular fragment, R−H⋯A in which R is more electronegative than H, and an atom or a group of atoms in the same or different molecule, in which there is evidence of bond formation",[25] Figure 1.3.

Figure 1.3 A schematic of hydrogen-bonding

Hydrogen-bonds represent one of the most studied intermolecular interactions in supramolecular synthesis, due to the strength and directionality of such interactions, when compared to many other intermolecular interactions. They are the basis of molecular recognition and are responsible for the generation of families of molecular networks with the same molecular components (single component crystals and their polymorphs) or with different molecular components (multiple component crystals or co-crystals) in the crystalline state (Figure 1.4).[26]
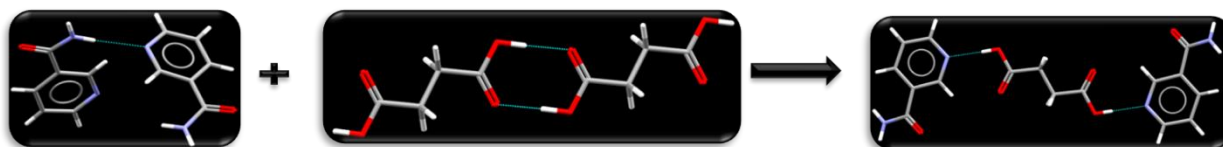
Figure 1.4 An example of hydrogen-bonding in co-crystals [27]

An important milestone to determine the supramolecular synthons is the hydrogen-bond rules was proposed by Margaret C. Etter in 1991.[28]

(1) All good acidic donors and acceptors in a molecule will be used in hydrogen bonding.

(2) Six-membered intramolecular hydrogen bonds are preferred over intermolecular hydrogen-bonds.

(3) The best acidic donors and acceptors remaining after intramolecular hydrogen-bond will form intermolecular hydrogen bonds to one another.

1.4.2 Halogen bonding

The definition of halogen bond (XB) states that, "A halogen bond R−X⋯A−Z occurs when there is evidence of a net attractive interaction between an electrophilic region on a halogen atom X belonging to a molecule or a molecular fragment R–X (where R can be another atom, including X, or a group of atoms) and a nucleophilic region of a molecule, or molecular fragment, A−Z",[29] Figure 1.5.

X= I, Br, Cl
A= N, O, S, Se, Cl, Br, I

Figure 1.5 A schematic of halogen-bonding

Halogen bonding is a non-covalent interaction, electrostatic and directional in nature. It is comparable in strength to hydrogen-bonds (~5-180 kJ/mol) and has gained increasing importance in crystal engineering because of its strength and effectiveness as a structure-directing force. The strength of the interaction increases in going from chlorine to bromine to iodine depending on the polarizability of the halogen atom. An interesting feature of the halogen-bond is that the strength of this interaction can be effectively increased by "activating" the halogen atom (also called positive "σ-hole" activation) with electron withdrawing substituents (Figure 1.6). Literature has shown that by introducing an *sp*-hybridized carbon next to the halogen atom[1] or by incorporating a fluorinated backbone[2] next to the halogen atom, the molecular electrostatic potential that determines the σ-hole is enhanced.[30-31]

σ hole
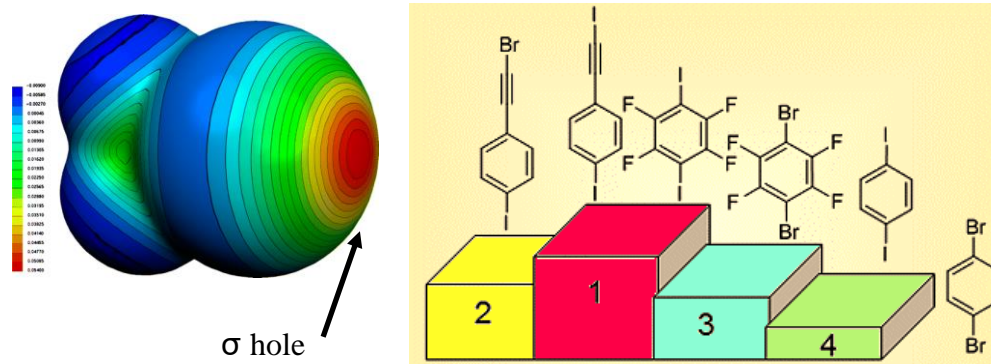
Figure 1.6 (a) Sigma hole on the halogen-bond donor (b) Activation of halogen-bond donors by fluorinating or introducing an sp-hybridized carbon.
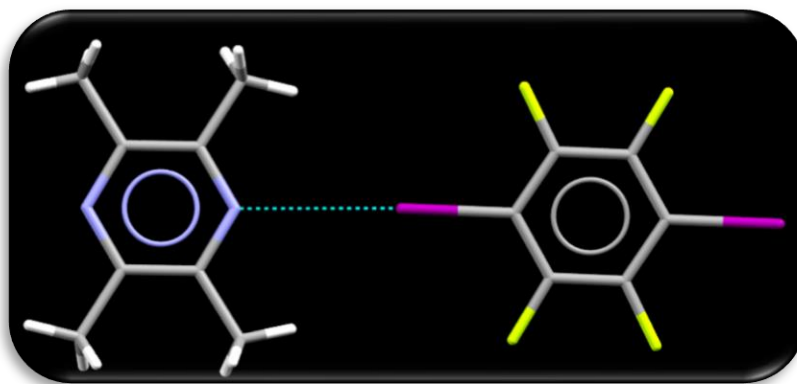
Figure 1.7 A halogen bond formed between 2,3,5,6-Tetramethylpyrazine and 1,4-bis(iodo)-2,3,5,6-tetrafluorobenzene.[32]

1.4.3 Applications of co-crystals

Co-crystals have been used to develop new anticancer,[33] antifungal,[34] antiviral,[35] anticonvulsant,[36] [37] and antidepressant[38] drugs with enhanced aqueous solubility. The same type of technology has also been explored for improving energetic materials,[39] enhancing agrochemicals and for stabilizing volatile chemicals.[40]

1.4.3.1 Pharmaceutical co-crystals

Co-crystals have been widely used in pharmaceutical industry to improve solubility, dissolution rates, stability and bioavailability. Several well-known commercial compounds have been targeted in cocrystal synthesis such as gabapentine,[36] aspirin, norfloxacin,[41] stanolone,[37] caffeine, saccharine, efavirenz,[42] and many others. An example of Figure 1.8 shows an increase in solubility of theophylline drug using careful choice of co-formers.[43]
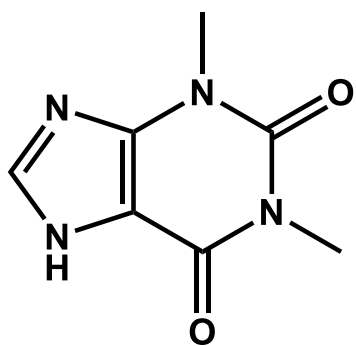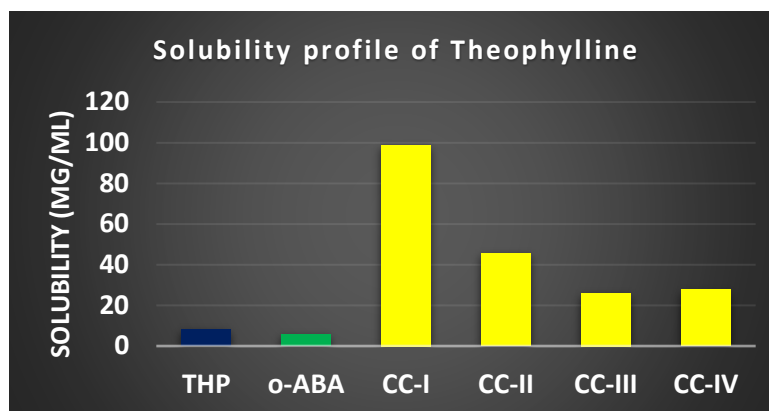
(a)   **(b)**

Figure 1.8 Solubility profile of co-crystals of theophylline[43]

1.4.3.2 Agrochemicals co-crystals

The term ─agrochemical refers to a collection of chemicals (pesticides) such as fungicides, herbicides and insecticides or acaricides Agrochemicals are used to protect crops from pests and thus help improve the quality and quantity of crop production. Two common agrochemicals; cyprodinil and terbuthylazine showed the potential to form co-crystals with a series of diacids. The co-crystals displayed melting points different from the parent active. In cyprodinil co-crystals, the solubility of the active could be altered using co-crystallization and a very good correlation between the co-crystal solubility and co-former solubility was reported (Figure 1.9).
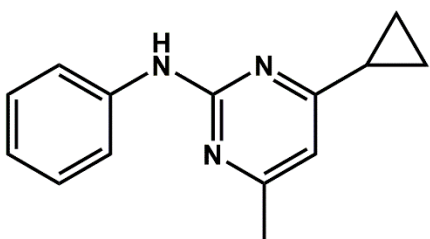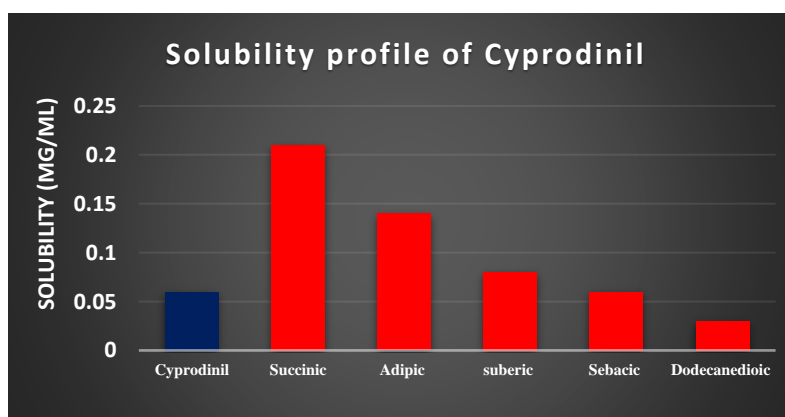
(a)   **(b)**

Figure 1.9 (a) chemical structure of cyprodinil and (b) solubility profile of co-crystals of cyprodinil[44]

1.4.3.3 Energetic co-crystals

An energetic material is a reactive substance that releases great amount of potential energy upon explosion, also accompanied by the production of light, heat, sound, and pressure.[45] Energetic materials include explosives, propellants and pyrotechnics. Co-crystallization is proving to be a powerful tool for creating less-sensitive explosives, as well as modifying and optimizing other properties of energetic materials.[39] While traditional strategies for energetic materials development have relied on the synthesis of novel energetic compounds and the optimization of their (polymorphic) solid forms, co-crystallization presents an elegant means to improve the performance of energetic materials without requiring new chemical synthesis.[46] Few examples of energetic co-crystals includes HMX, CL-20, DADP where properties such as sensitivity, density, detonation velocity, and oxygen balance are altered (Figure 1.10).
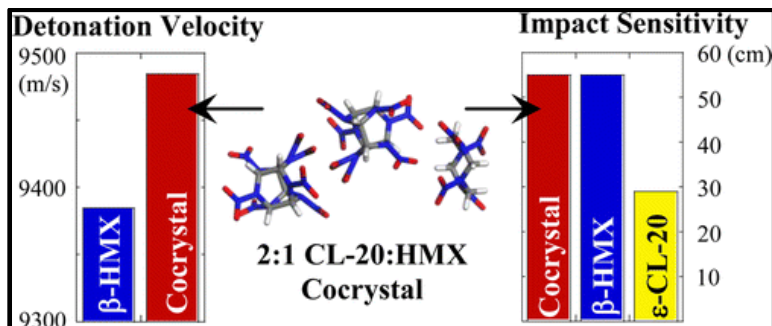
Figure 1.10 Detonation velocity and impact sensitivity of energetic-energetic co-crystal[47]

1.4.4 Co-crystal prediction

Co-crystal prediction has been reported to include the following steps: (1) determining whether a given set of two or more molecular components will undergo co-crystallization; (2) identifying the primary intermolecular interactions, or supramolecular synthons that will exist within a particular co-crystal structure. Both predictions are important form application stand point as co-crystals of APIs, agrochemicals, nutraceuticals and explosives represent a class of multi-component crystalline forms that are of interest for their advantageous physical properties and for intellectual property implications. The ability to predict a priori which compounds are likely to form co-crystals with a given API would provide a complementary tool to experimental screening and ability to predict the primary intermolecular interactions in the co-crystal will prevent synthon crossover and synthon polymorphism.

1.4.4.1 Current co-crystal screening prediction methods

Various approaches have been proposed in literature to elucidate the co-crystal screening outcome. Fabian employed molecular shape and polarity as benchmarks for predicting the formation of co-crystals.[48-49] Price and co-workers focused on lattice energy comparisons of co-crystals and pure components as a basis for predicting co-crystal formation,[50] and Velaga suggested that drugs and co-formers with similar Hanson solubility parameters are likely to form co-crystals with each other.[51] Jones and co-workers employed hydrogen-bond propensities as a way of quantifying homomeric and heteromeric interactions.[52-54] Hunter *et al.* and Aakeroy *et al.* used a combination of experimentally derived parameters and electrostatic potential surfaces to predict co-crystal screen and supramolecular synthons in the solid as well as solution phase.[55-59] Galek et al. demonstrated that hydrogen-bond propensity calculations,[60-61] which rely on a statistical analysis of the occurrence of hydrogen bonds in relevant structures present in the Cambridge Structural Database,[48, 62] could be used for predicting the outcome of co-crystallizations of the drug lamotrigine.[63-64] Validation studies performed so far indicate that a combination of methods is valuable when performing a knowledge based co-former screen.[65]

1.4.4.2 Current co-crystal synthon prediction methods

Co-crystal design is aided by the use of statistical tools such as knowledge-based hydrogen-bond propensity calculations[52-53] and computational methods such as molecular complementarity,[49] ΔpK_a calculations[66] and molecular electrostatic potential surfaces MEPS.[58-59] Despite the availability of such tools, and despite our evolving understanding of how molecules interact in the solid state, it still remains impossible to empirically and statistically predict the arrangement of individual or multiple molecules in a crystal lattice.[67-68] Likewise, it is extremely difficult to empirically predict the formation of supramolecular synthons in cocrystals composed of molecules containing a broad range of functional groups.[69]

In order to facilitate the design of cocrystals containing such complex entities, it will be necessary to considerably deepen our understanding of self-assembly processes in the solid state.[70] The prediction method knowledge will then be applied to large, flexible drug like molecules to solve an important issue in development: crystallizability.

1.5 Crystallization

Crystallization is the physical transformation (phase transition) of a liquid, solution, or gas to a crystal, a solid with an ordered internal arrangement of molecules, ions, or atoms.[71] It is different from amorphous materials, a solid that lacks the long-range order that is characteristic of a crystal (Figure 1.11).[72] The generation of the first crystalline seed from which a molecule can be crystallized is one of the most important, and sometimes most difficult, tasks in drug development. Crystallization is critical in that it provides a path forward for product isolation from the final step of the synthetic process, as well as sets the physical properties underpinning drug substance performance in the drug product. The crystallization process consists of two major events, nucleation and crystal growth which are driven by thermodynamic properties as well as kinetics.[73]

- Ordered structure
- Sharp melting point
- Definite heat of fusion
- Anisotropic
- Have covalent, ionic, metallic, non-covalent bonds

VS

Crystalline

Amorphous

- Disordered structure
- Do not have sharp melting point
- No definite heat of fusion
- Isotropic
- Have covalent bonded networks

- More stable
- Lower solubility
- Lower risk

- Less stable
- Prone to crystallization and degradation
- Higher solubility
- Higher risk

Figure 1.11 Properties of crystalline vs amorphous materials.[74]

Crystal nucleation is the step where the solute molecules or atoms dispersed in the solvent forms stable clusters on the microscopic scale under the current operating conditions.[75] These stable clusters then reaches a critical size to become a stable nuclei.[76] The critical size is dictated by many experimental and environmental factors (temperature, pressure, solvent, supersaturation, etc.). It is at the stage of nucleation that the atoms or molecules arrange in a defined and periodic manner that defines the crystal structure.[77-78]
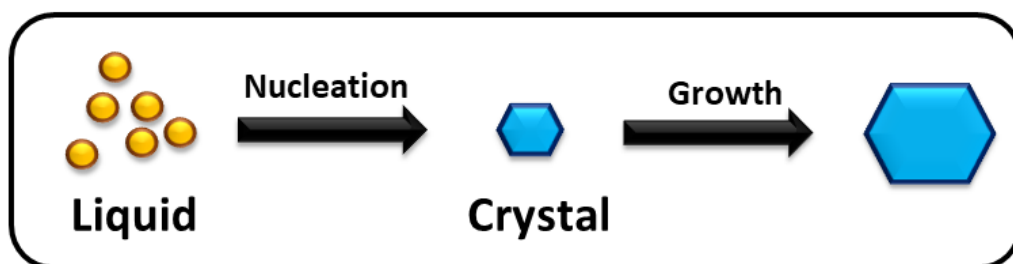
Figure 1.12 Schematics of crystal nucleation and crystal growth.[79]

The crystal growth is the subsequent size increase of the nuclei that has achieved the critical cluster size (Figure 1.12).[76] It is a dynamic process occurring in equilibrium where solute molecules or atoms precipitate out of solution, and dissolve back into solution. Supersaturation is one of the driving forces of crystallization, as the solubility of a species is an equilibrium process quantified by K_{sp}. Depending upon the conditions, either nucleation or growth may be predominant over the other, dictating crystal size.

1.5.1 Why is prediction of crystallizability important?

Predicting the crystallization propensity of drug-like molecules is one of the biggest challenges facing pharmaceutical scientists today. Specifically, understanding which molecules in a series of similar compounds will be the most difficult to crystallize would be extremely useful, as would any indication of the experimental conditions (such as temperature, solvent polarity, molar concentration, etc.) that might make crystallization take place more readily. There is also great interest in minimizing research and development costs, which are estimated to be about $1 billion per drug launch, taking about 6–10 years for drug development and only 10% of the compounds in development survive the efficacy and safety hurdles and become marketed drugs (Figure 1.13).[80] Therefore, predictions can save time, money and experimental efforts in both academic and industrial research as well as prevent any unexpected and unwanted discovery at the late stage of development.
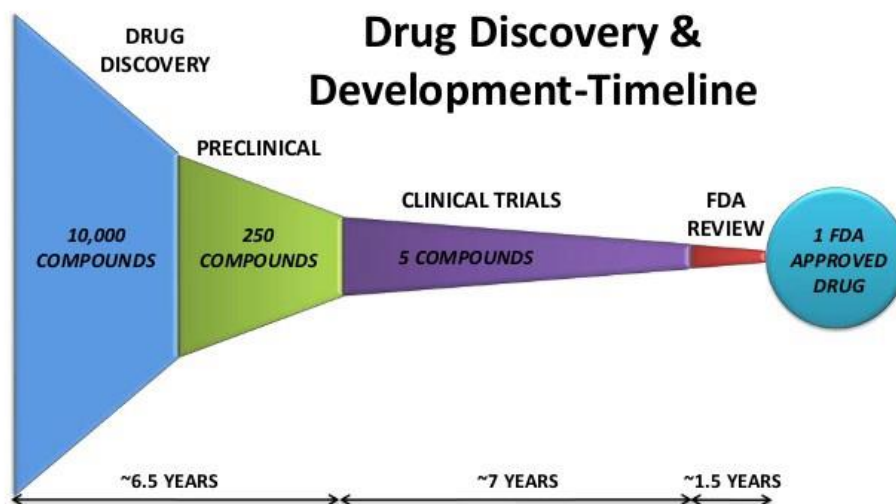
Figure 1.13 Timeline of drug development process from synthesis to market stage[81]

Different approaches have been taken in recent years to model and predict crystallization propensity. Machine learning approaches, statistical (random forest) modelling, DSC based heat/cool/heat methods have been employed to model and predict crystallizability in terms of giving single crystals suitable for X-ray diffraction.[82] These varied approaches may not give a consistent picture of crystallizability, as the difficulty in nucleation does not necessarily correlate to slow crystal growth rates or poor crystal quality.[83] More importantly, these approaches do not provide a path forward to overcome inherently poor crystallizability, i.e., crystallizing a molecule for the first time. Based on the information from various studies, it is concluded that combination of descriptors play an important role on the crystallization outcome. The following descriptors were considered important based on literature sources (Table 1.2).

Table 1.2 Important descriptors and literature sources

Important descriptors	References
Number of rotatable bond	Florescence *et. al*,[82] cooper *et. al* [84]
Molecular size	Cooper *et. al* [84]
Molecular weight	Taylor *et. al* [85]
Viscosity	Taylor *et. al* [85]
Thermodynamics parameters	Fischer *et. al* [86]
# of H-bond donors	Fischer *et. al* [86]
Carbon to heteroatom ratio	Fischer *et. al* [86]

1.6 Goals of the dissertation

Discovering the right solid form with optimal chemical/physical properties is of paramount importance in academic and industry research. This thesis will focus on the validating the prediction methods such as hydrogen-bond propensity (HBP), molecular electrostatic potentials (MEPs), hydrogen-bond energies (HBE), and hydrogen-bond coordination (HBC) for supramolecular synthon predictions in the target molecules and co-crystals. The second goal of the thesis will be to predict co-crystal screening of a library of small rigid molecules using combination of prediction methods. This knowledge will then be transferred to predict the crystallizability of large, flexible drug-like molecules. The basic concepts will then be used to alter bulk properties of functional materials in a desirable manner.

This dissertation will focus on following goals;

- Chapter 2 will focus on predicting the homomeric synthons in the twelve pyrazole molecules using four prediction methods: MEPs, HBE, HBP and HBC.

- Chapter 3 will focus on comparing the prediction (using HBE and HBP) vs experimental (using FTIR) co-crystal screening results of twelve pyrazole molecules with 20 carboxylic acids. The second part of chapter will focus on the predicting the supramolecular synthons in the co-crystals using MEPs, HBE, and HBP methods.

- Chapter 4 will focus on predicting the homomeric synthons in the twelve thiazole molecules using four prediction methods: MEPs, HBE, HBP and HBC.

- Chapter 5 will focus on comparing the prediction (using HBP) vs experimental (using FTIR) co-crystal screening results of twelve thiazole molecules with 20 carboxylic acids. The second part of chapter will focus on the predicting the supramolecular synthons in the co-crystals using MEPs, HBE, and HBP methods.[87]

- Chapter 6 will apply both hydrogen and halogen bonding as synthetic tools to design binary and ternary co-crystals.

- Chapter 7 will focus on developing a supramolecular hierarchy of activated amide containing halogen bond donors and co-crystallization to determine the strength of homomeric interactions in the target molecules.

- Chapter 8 will focus on using a newly developed prediction tool to predict the crystallization propensity of drug like molecules with diverse functionalities.

- Chapter 9 will address the major challenges with 19th century molecule urea and ways to improve the physicochemical properties such as solubility and stability using co-crystallization technology.[88]

1.7 References

1.   Seybold, P. G.; May, M.; Bagal, U. A., *Journal of Chemical Education* **1987**, *64* (7), 575.

2.   https://people.chem.umass.edu/mcdaniel/CHEM-267/Experiments/Recrystallization-Part-1/Structure%E2%80%93PropertyRelationships.pdf.

3.   Thalladi, V. R.; Nüsse, M.; Boese, R., *Journal of the American Chemical Society* **2000**, *122* (38), 9227-9236.

4.   L., M., *Recueil des Travaux Chimiques des Pays-Bas* **1996**, *115* (11-12), 552-552.

5.   Jean- Marie Lehn, N. L., December 8, **1987**.

6.   Thalladi, V. R.; Goud, B. S.; Hoy, V. J.; Allen, F. H.; Howard, J. A. K.; Desiraju, G. R., *Chemical Communications* **1996**, (3), 401-402.

7.   R., D. G., *Angewandte Chemie International Edition in English* **1995**, *34* (21), 2311-2327.

8.   Draguta, S.; Fonari, M. S.; Bejagam, S. N.; Storms, K.; Lindline, J.; Timofeeva, T. V., *Structural Chemistry* **2016**, *27* (4), 1303-1315.

9.   André, V.; Fernandes, A.; Santos, P. P.; Duarte, M. T., *Crystal Growth & Design* **2011**, *11* (6), 2325-2334.

10.   Dunitz, J. D., *CrystEngComm* **2003**, *5* (91), 506-506.

11.   Aakeröy, C. B.; Salmon, D. J., *CrystEngComm* **2005**, *7* (72), 439-448.

12.   Najar, A. A. A., Y. J. Indian I. Sci. **2014**, 94, 45−67.

13.   Friščić, T.; Jones, W., *Crystal Growth & Design* **2009**, *9* (3), 1621-1637.

14.   Trask, A. V.; van de Streek, J.; Motherwell, W. D. S.; Jones, W., *Crystal Growth & Design* **2005**, *5* (6), 2233-2241.

15.   Morissette, S. L.; Almarsson, Ö.; Peterson, M. L.; Remenar, J. F.; Read, M. J.; Lemmo, A. V.; Ellis, S.; Cima, M. J.; Gardner, C. R., *Advanced Drug Delivery Reviews* **2004**, *56* (3), 275-300.

16. Schultheiss, N.; Newman, A., *Crystal Growth & Design* **2009**, *9* (6), 2950-2967.

17. Weyna, D. R.; Shattock, T.; Vishweshwar, P.; Zaworotko, M. J., *Crystal Growth & Design* **2009**, *9* (2), 1106-1123.

18. Chang, Y.-C.; Chen, Y.-D.; Chen, C.-H.; Wen, Y.-S.; Lin, J. T.; Chen, H.-Y.; Kuo, M.-Y.; Chao, I., *The Journal of Organic Chemistry* **2008**, *73* (12), 4608-4614.

19. Madura, I. D.; Czerwińska, K.; Jakubczyk, M.; Pawełko, A.; Adamczyk-Woźniak, A.; Sporzyński, A., *Crystal Growth & Design* **2013**, *13* (12), 5344-5352.

20. Präsang, C.; Whitwood, A. C.; Bruce, D. W., *Crystal Growth & Design* **2009**, *9* (12), 5319-5326.

21. Buckingham, A. D. D. B., J.E.; McDowell, S.A.C. Chem. Phys. Lett. **2008**, 463, 1-10; Lehn, J.M. Science **2002**, 295, 2400-2403; Moulton, B.; Zaworotko, M.J. Chem. Rev. **2001**, 101, 1629-1658; Desiraju, G.R. Angew. Chem., Int. Ed. **1995**, 34, 2311-2327; Bosch, E. CrystEngComm **2007**, 9, 191-198; Aakeröy, C.B.; Desper, J.; Salmon, D.J.; Smith, M.M. Cryst. Growth Des. **2006**, 6, 1033-1042; Aakeröy, C.B.; Desper, J.; Urbina, J.F. CrystEngComm **2005**, 7, 193-201. .

22. Wenger, M.; Bernstein, J., *Crystal Growth & Design* **2008**, *8* (5), 1595-1598.

23. Basavoju, S.; Boström, D.; Velaga, S. P., *Pharmaceutical Research* **2008**, *25* (3), 530-541.

24. Lu, J.; Rohani, S., *Organic Process Research & Development* **2009**, *13* (6), 1269-1275.

25. Arunan, E. D., G.R.; Klein, R.A.; Sadlej, J.; Scheiner, S.; Alkorta, I.; Clary, D.C.; Crabtree, R.H.; Dannenberg, J.J.; Hobza, P.; Kjaergaard, H.G.; Legon, A.C.; Mennucci, B.; Nesbitt, D.J. Pure Appl. Chem. **2011**, 83, 1619-1636. 27.

26. Jayasankar, A.; Somwangthanaroj, A.; Shao, Z. J.; Rodríguez-Hornedo, N., *Pharmaceutical research* **2006**, *23* (10), 2381-2392.

27. Thompson, L. J.; Voguri, R. S.; Cowell, A.; Male, L.; Tremayne, M., *Acta Crystallographica Section C* **2010**, *66* (8), o421-o424.

28. Etter, M. C., *The Journal of Physical Chemistry* **1991**, *95* (12), 4601-4610.

29. Desiraju, G. R. H., S.; Kloo, L.; Legon, A.C.; Marquardt, R.; Metrangolo, P.; Politzer, P.A.; Resnati, G.; Rissanen, K. Pure Appl. Chem. **2013**, 85, 1711-1713.

30. Aakeröy, C. B.; Baldrighi, M.; Desper, J.; Metrangolo, P.; Resnati, G., *Chemistry – A European Journal* **2013**, *19* (48), 16240-16247.

31. Cavallo, G.; Metrangolo, P.; Milani, R.; Pilati, T.; Priimagi, A.; Resnati, G.; Terraneo, G., *Chemical Reviews* **2016**, *116* (4), 2478-2601.

32. Syssa-Magale, J.-L.; Boubekeur, K.; Palvadeau, P.; Meerschaut, A.; Schollhorn, B., *CrystEngComm* **2005**, *7* (50), 302-308.

33. Zhang, Y. N.; Yin, H.-M.; Zhang, Y.; Zhang, D.-J.; Su, X.; Kuang, H.-X., *Journal of Molecular Structure* **2017**, *1130*, 199-207.

34. Shevchenko, A.; Miroshnyk, I.; Pietilä, L.-O.; Haarala, J.; Salmia, J.; Sinervo, K.; Mirza, S.; van Veen, B.; Kolehmainen, E.; Nonappa; Yliruusi, J., *Crystal Growth & Design* **2013**, *13* (11), 4877-4884.

35. Jung, S.; Lee, J.; Kim, I. W., *Journal of Crystal Growth* **2013**, *373*, 59-63.

36. Reddy, L. S.; Bethune, S. J.; Kampf, J. W.; Rodríguez-Hornedo, N., *Crystal Growth & Design* **2009**, *9* (1), 378-385.

37. Takata, N.; Shiraki, K.; Takano, R.; Hayashi, Y.; Terada, K., *Crystal Growth & Design* **2008**, *8* (8), 3032-3037.

38. Sarkar, A.; Rohani, S., *Journal of Pharmaceutical and Biomedical Analysis* **2015**, *110*, 93-99.

39. Landenberger, K. B.; Bolton, O.; Matzger, A. J., *Journal of the American Chemical Society* **2015**, *137* (15), 5074-5079.

40. Aakeroy, C. B.; Wijethunga, T. K.; Benton, J.; Desper, J., *Chemical Communications* **2015**, *51* (12), 2425-2428.

41. Basavoju, S.; Boström, D.; Velaga, S. P., *Crystal Growth & Design* **2006**, *6* (12), 2699-2708.

42. Pawar Jaywant, N.; Amin Purnima, D., *Materials Today: Proceedings* **2016**, *3* (6), 1742-1751.

43. Saikia, B.; Bora, P.; Khatioda, R.; Sarma, B., *Crystal Growth & Design* **2015**, *15* (11), 5593-5603.

44. Neil, G. F., J.; Rebecca, B.; Aakeröy, C.B. US Pat. WO2011128618, **2011**.

45. J., M., *Water and Environment Journal* **1989**, *3* (5), 522-525.

46. Kent, R. V.; Wiscons, R. A.; Sharon, P.; Grinstein, D.; Frimer, A. A.; Matzger, A. J., *Crystal Growth & Design* **2018**, *18* (1), 219-224.

47. Bolton, O.; Simke, L. R.; Pagoria, P. F.; Matzger, A. J., *Crystal Growth & Design* **2012**, *12* (9), 4311-4314.

48. 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., *Journal of Applied Crystallography* **2008**, *41* (2), 466-470.

49.   Fábián, L., *Crystal Growth & Design* **2009**, *9* (3), 1436-1443.

50.   Issa, N.; Karamertzanis, P. G.; Welch, G. W. A.; Price, S. L., *Crystal Growth & Design* **2009**, *9* (1), 442-453.

51.   Mohammad, M. A.; Alhalaweh, A.; Velaga, S. P., *International Journal of Pharmaceutics* **2011**, *407* (1), 63-71.

52.   Delori, A.; Galek, P. T. A.; Pidcock, E.; Patni, M.; Jones, W., *CrystEngComm* **2013**, *15* (15), 2916-2928.

53.   Delori, A.; Galek, P. T. A.; Pidcock, E.; Jones, W., *Chemistry – A European Journal* **2012**, *18* (22), 6835-6846.

54.   Eddleston, M. D.; Arhangelskis, M.; Fábián, L.; Tizzard, G. J.; Coles, S. J.; Jones, W., *Crystal Growth & Design* **2016**, *16* (1), 51-58.

55.   Hunter, C. A., *Angewandte Chemie International Edition* **2004**, *43* (40), 5310-5324.

56.   McKenzie, J.; Feeder, N.; Hunter, C. A., *CrystEngComm* **2016**, *18* (3), 394-397.

57.   Musumeci, D.; Hunter, C. A.; Prohens, R.; Scuderi, S.; McCabe, J. F., *Chemical Science* **2011**, *2* (5), 883-890.

58.   B., A. C.; Kanishka, E.; Safiyyah, F.; Nathan, S.; John, D., *Chemistry – A European Journal* **2013**, *19* (44), 14998-15003.

59.   Aakeroy, C. B.; Wijethunga, T. K.; Desper, J., *New Journal of Chemistry* **2015**, *39* (2), 822-828.

60.   Galek, P. T. A.; Allen, F. H.; Fabian, L.; Feeder, N., *CrystEngComm* **2009**, *11* (12), 2634-2639.

61.   Chemburkar, S. R.; Bauer, J.; Deming, K.; Spiwek, H.; Patel, K.; Morris, J.; Henry, R.; Spanton, S.; Dziki, W.; Porter, W.; Quick, J.; Bauer, P.; Donaubauer, J.; Narayanan, B. A.; Soldani, M.; Riley, D.; McFarland, K., *Organic Process Research & Development* **2000**, *4* (5), 413-417.

62.   Allen, F., *Acta Crystallographica Section B* **2002**, *58* (3 Part 1), 380-388.

63.   Galek, P. T. A.; Chisholm, J. A.; Pidcock, E.; Wood, P. A., *Acta Crystallographica Section B* **2014**, *70* (1), 91-105.

64.   Galek, P. T. A.; Pidcock, E.; Wood, P. A.; Bruno, I. J.; Groom, C. R., *CrystEngComm* **2012**, *14* (7), 2391-2403.

65.   Wood, P. A.; Feeder, N.; Furlow, M.; Galek, P. T. A.; Groom, C. R.; Pidcock, E., *CrystEngComm* **2014**, *16* (26), 5839-5848.

66.   Cruz-Cabeza, A. J., *CrystEngComm* **2012**, *14* (20), 6362-6365.

67.   Day, G. M.; Motherwell, W. D. S., *Crystal Growth & Design* **2006**, *6* (9), 1985-1990.

68.   Bučar, D.-K.; Lancaster, R. W.; Bernstein, J., *Angewandte Chemie (International Ed. in English)* **2015**, *54* (24), 6972-6993.

69.   Bučar, D.-K.; Henry, R. F.; Lou, X.; Duerst, R. W.; Borchardt, T. B.; MacGillivray, L. R.; Zhang, G. G. Z., *Molecular Pharmaceutics* **2007**, *4* (3), 339-346.

70.   Corpinot, M. K.; Stratford, S. A.; Arhangelskis, M.; Anka-Lufford, J.; Halasz, I.; Judas, N.; Jones, W.; Bucar, D.-K., *CrystEngComm* **2016**, *18* (29), 5434-5439.

71.   Yu, L.; Reutzel-Edens, S. M., CRYSTALLIZATION | Basic Principles A2 - Caballero, Benjamin. In *Encyclopedia of Food Sciences and Nutrition (Second Edition)*, Academic Press: Oxford, **2003**; pp 1697-1702.

72.   Stachurski, Z. H., *Materials* **2011**, *4* (9), 1564.

73.   Erdemir, D.; Lee, A. Y.; Myerson, A. S., *Accounts of Chemical Research* **2009**, *42* (5), 621-629.

74.   https://chem.libretexts.org/Textbook_Maps/General_Chemistry_Textbook_Maps/Map%3A_Chemistry_(Averill_and_Eldredge)/12%3A_Solids/12.1%3A_Crystalline_and_Amorphous_Solids.

75.   Price, S.; Veesler, S.; Pan, H.; Lewtas, K.; Smets, M.; Rimez, B.; Myerson, A.; Hughes, C.; Hare, A.; Zhang, F.; Meekes, H.; Mazzotti, M.; Rosbottom, I.; Khamar, D.; van den Ende, J.; Fabian, L.; Black, S.; Taulelle, F.; Gich, M.; Vekilov, P.; Toroz, D.; Bertran, C. A.; Sefcik, J.; Schroeder, S.; Booth, S.; Rasmuson, A.; Breynaert, E.; Simone, E.; Hammond, R.; Sear, R.; de Yoreo, J.; Davey, R.; Anwar, J.; Ristic, R.; Camacho Corzo, D. M.; Roberts, K.; Harris, K.; Colfen, H.; Turner, T., *Faraday Discussions* **2015**, *179* (0), 155-197.

76.   Lee, B. H., *Fundamentals of Food Biotechnology*. Wiley: **2014**.

77.   Jones, A. G. C. P. S. O. B.-H. I.-.

78.   Mullin, J. W. C., 4th ed. Oxford: Butterworth-Heinemann. ISBN 978-0750648332.

79.   Sauter, A.; Roosen-Runge, F.; Zhang, F.; Lotze, G.; Feoktystov, A.; Jacobs, R. M. J.; Schreiber, F., *Faraday Discussions* **2015**, *179* (0), 41-58.

80.   Berger J, D. J. D., Johnson M.M., Karst K.R., Shear W.C., Am J Manag Care, **2017**, 22(16),S487–95.

81.   http://slideplayer.com/slide/10688957/.

82. Bhardwaj, R. M.; Johnston, A.; Johnston, B. F.; Florence, A. J., *CrystEngComm* **2015**, *17* (23), 4272-4275.

83. Hursthouse, M. B.; Huth, L. S.; Threlfall, T. L., *Organic Process Research & Development* **2009**, *13* (6), 1231-1240.

84. Wicker, J. G. P.; Cooper, R. I., *CrystEngComm* **2015**, *17* (9), 1927-1934.

85. Van Eerdenbrugh, B.; Baird, J. A.; Taylor, L. S., *Journal of Pharmaceutical Sciences 99* (9), 3826-3838.

86. Nurzyńska, K.; Booth, J.; Roberts, C. J.; McCabe, J.; Dryden, I.; Fischer, P. M., *Molecular Pharmaceutics* **2015**, *12* (9), 3389-3398.

87. Sandhu, B.; McLean, A.; Sinha, A. S.; Desper, J.; Sarjeant, A. A.; Vyas, S.; Reutzel-Edens, S. M.; Aakeröy, C. B., *Crystal Growth & Design* **2018**, *18* (1), 466-478.

88. Sandhu, B.; Sinha, A. S.; Desper, J.; Aakeroy, C. B., *Chemical Communications* **2018**.

Chapter 2 - Evaluating Homomeric Synthons in Pyrazole based molecules using Hydrogen-bond Energies and Propensities

2.1 Introduction

A common question in crystal engineering is, given a molecular structure, can we predict its crystal structure?[1] The core of a crystal structure is the supramolecular synthon[2-5] (introduced by G. R. Desiraju as a "*structural unit within supermolecules which can be formed and/or assembled by known or conceivable synthetic operations involving intermolecular interactions*"[2]) which encapsulates enough critical information so that it serves as a model for the entire crystal.[6],[7] So, the question is rather can we predict the supramolecular synthon given a molecular entity?

Synthon prediction is a difficult task especially in molecules with multiple binding sites which often leads to synthon polymorphism[8-10] or synthon crossover[11] (Scheme 2.1).[12] It is also difficult to predict synthons because of the complementary nature of certain functional groups and deviations in crystal structures from close packing. It also becomes more complicated in flexible molecules with large number of molecular conformations where both intra and intermolecular interactions are possible.
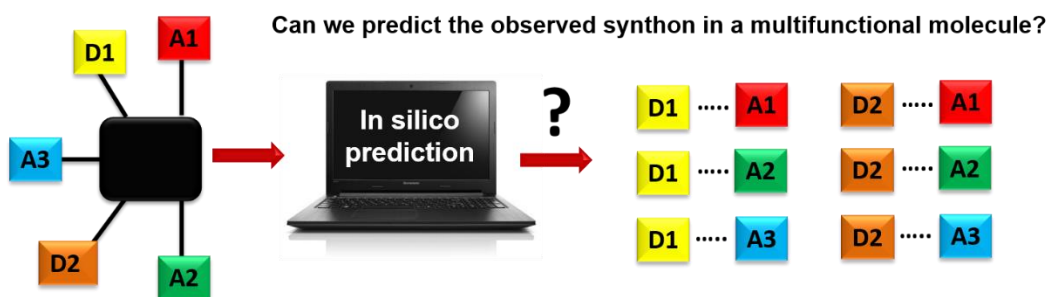
Scheme 2.1 Schematics of using an in-silico approach for predicting synthons in a multifunctional molecule.

The current methods for synthon predictions in organic molecules includes electrostatics, lattice energy predictions,[14-16] focusing only on the thermodynamic (enthalpic) factors.[17] For example, the preferred connectivity patterns in a molecule in the solid state can be determined using Margaret C. Etter rules which states that the best hydrogen bond acceptor binds to the best donor.[18] The ranking of best donor-acceptor can be determined using molecular electrostatic potential as

proposed by Aakeroy *et.al.*[19] Hunter *et al*. has converted potentials into hydrogen-bond energies to determine the synthon preference in a multicomponent systems.[20]

One way of accounting for kinetic factors is to analyze of a large amount of structural information encapsulated in the Cambridge Structural Database[21] (CSD), the world's repository of all published and patented small molecule crystal structures (900,000+). Of course, the crystal structure itself cannot provide information about kinetics of the seed formation and crystal growth, but it is, nevertheless, the final result of these processes. It is possible to suggest that if some structural array is abundant in a database, it may reflect thermodynamic stability of a given crystal packing array, as well as a kinetic preference of its formation. Therefore, this method used in conjunction with the electrostatics and energy may be useful in the predicting the outcome of supramolecular synthon in a complex multifunctional molecule.
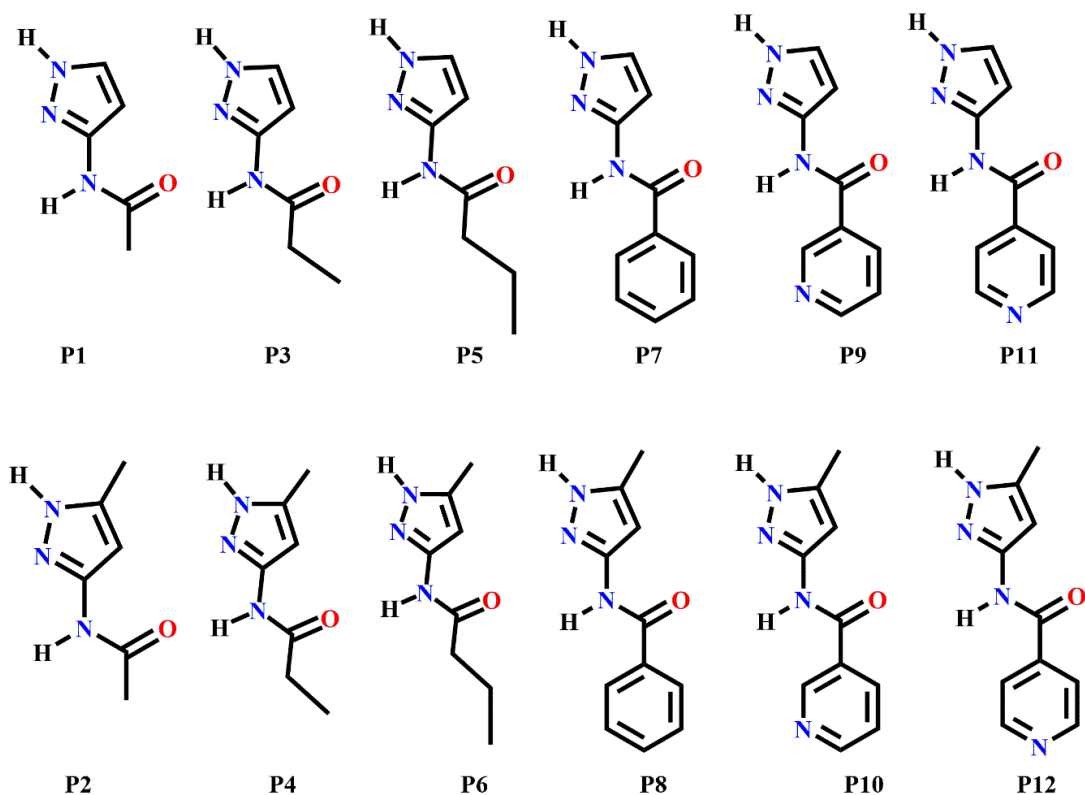
Scheme 2.2 Pyrazole molecules (**P1–P12**) employed in this study.

In order to explore the structural landscape of a single molecule capable of forming a few specific and competing interactions, a library of pyrazole containing molecules (**P1-P12**) were synthesized

and explained, Scheme 2.2. Pyrazoles are reported to possess a wide range of biological activities such as anti-microbial, anti-fungal, anti-tubercular, anti-inflammatory, anti-convulsant, anticancer, anti-viral, and so on.[22] The pyrazole-amide functionality is also present in some pharmaceutical related compounds such as Entrectinib, Graniseton, and Epirizol as well as antifungal compounds such as Furametpyr, Penthiopyrad and Tolfempyrad.[23] Due to the presence of multifunctional groups, these molecules are always at risk of synthon polymorphism. Therefore, knowledge gained from a successful use of tools such as molecular electrostatic potentials, hydrogen-bond propensity and coordination, could have significant practical applications.
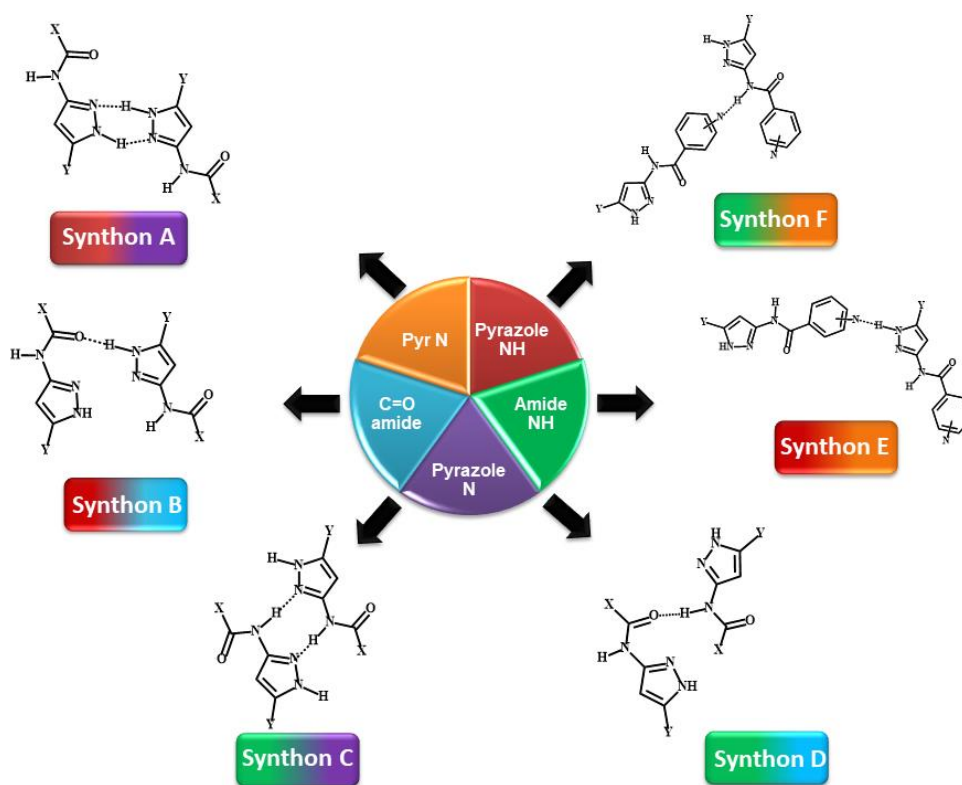
Scheme 2.3 Representations of six postulated synthons in a generic pyrazole-amide; (a) synthon **A**, N-H (pyrazole)…N(pyrazole); (b) synthon **B**, N-H (pyrazole)…C=O(amide); (c) synthon **C**, N-H (amide)…N(pyrazole); (d) synthon **D**, N-H (amide)…C=O(amide), synthon E, NH(pyrazole)…N(aromatic) and synthon F, NH(amide)…N(aromatic). (Y=H, CH$_3$; X= methyl, ethyl, or benzyl).

The target molecules, **P1-P12** can be divided into two groups: Group 1 (**P1-P8**) includes molecules with two donors (pyrazole NH and amide NH) and two acceptors (pyrazole N and carbonyl amide). In molecules **P1-P8**, four possible interactions based on two donors (pyrazole NH and amide NH) and two acceptors (pyrazole N and carbonyl C=O) can lead to four different synthons, scheme 3. The synthon **A** (NH (pyrazole)…N(pyrazole)) and synthon **C** (NH(amide)… N(pyrazole)) includes dimeric homo-synthon whereas synthon **B** (NH(pyrazole)…O=C(carbonyl)) and synthon **D** (NH(amide)…O=C(carbonyl)) include monomeric homo-synthon. It is likely that each molecule forms a combination of two synthons to satisfy all hydrogen-bond donors and acceptors and there are only two possibilities; Synthon **(A+D)** or synthon **(B+C))**. Group 2 (**P9-P12**) include molecules with two donors (pyrazole NH and amide NH) and three acceptor (pyrazole N, carbonyl C=O and pyridine N) groups. In additional to four synthons, synthon **E** (NH(pyrazole)…N(aromatic) and synthon **F** (NH(amide)…N(aromatic) are possible in group 2. Therefore, six different interactions are possible leading to six synthon possibilities; Synthon A-F, (Scheme 2.3).

The road map of this study is summarized in Scheme 3.4. We will use four different prediction methods; molecular electrostatic potentials (MEPs), hydrogen-bond energies (HBE), hydrogen-bond propensities (HBP) and hydrogen-bond coordination (HBC) tools to determine the most likely synthon in the target molecules **P1-P12** and compare our predictions with the experimental results. The overall goal is to develop robust and transferable protocols for identifying and predicting which hydrogen bonds are most likely to appear in homomeric molecular solids when there are numerous potential avenues for assembly, Scheme 2.1.
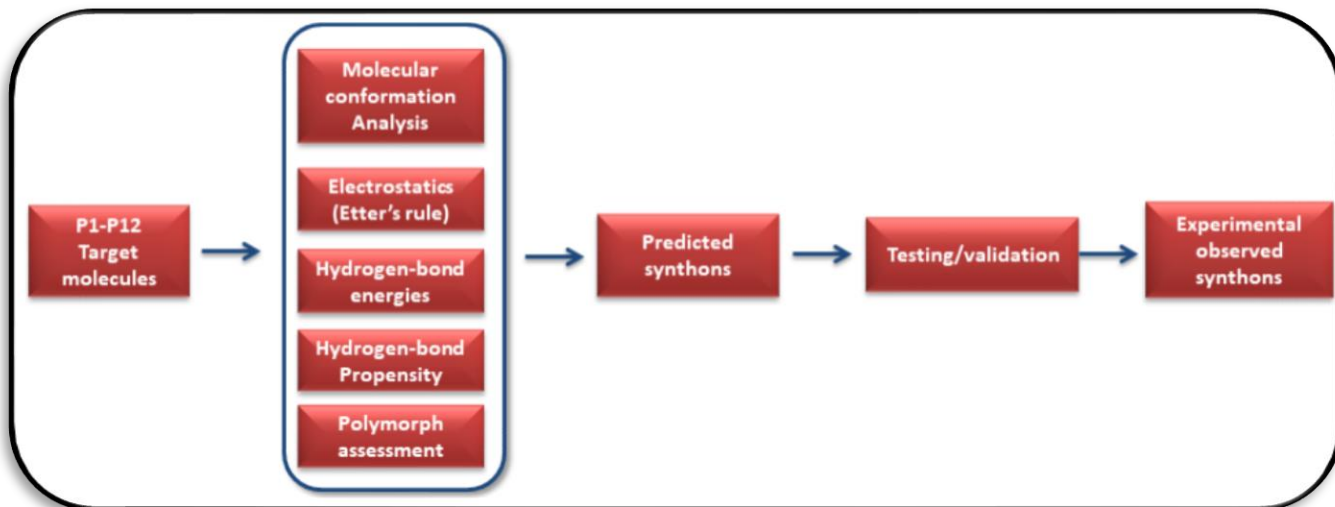
Scheme 2.4 The road map for synthon predictions

The study is undertaken to address the following questions,

1. Which method does a better job of predicting the synthon outcome in the target molecules; is it MEPs, HBE, HBP or HBC?
2. Is a combination of prediction methods preferred over individual methods?
3. Which synthon is the most optimal synthon in group **1** (**P1-P8**) and how does adding an acceptor group affect the choice of synthon in **P9-P12**.
4. Which molecules have the risk of forming synthon polymorphs and which method is most suitable for predicting polymorphism in this group of molecules?

2.2 Experimental

2.2.1 General

2-Amino-pyrazole, 2-amino-5-methyl-pyrazole, acetic anhydride, propionic anhydride and benzoyl chloride were purchased from Aldrich and utilized without further purification. Synthetic procedures and characterization of all molecules are provided in the Supporting Information (SI). Melting points were measured using Fisher-Johns melting point apparatus. ^{1}H NMR data were collected on a Varian Unity plus 400 MHz spectrophotometer in DMSO.

2.2.2 Synthesis

2.2.2.1 Synthesis of 3-acetamido-pyrazole, P1[24]
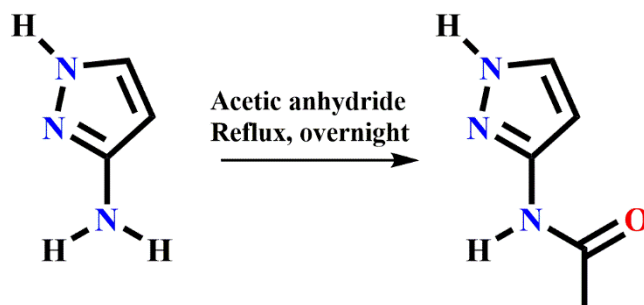
^{1}H-pyrazol-3-amine (0.486 g, 5.85 mmole) was dissolved in 50 mL of distilled water. NaHCO$_3$ (1.465 g, 17.4 mmole) was slowly added. Acetic anhydride (5ml) was then added dropwise and the resulting suspension was heated at reflux overnight. Then, the mixture was allowed to cool to room temperature and the solid obtained was filtered off and characterized as the title compound. After concentration of the filtrate, a second precipitate was obtained and also characterized as the title compound. 66% yield, m.p 218-220°C, ^{1}H NMR (400 MHz, DMSO-d$_6$) δ ppm: 12.23 (br s, 1H), 10.42 (br s, 1H), 7.54 (br s, 1H), 6.44 (br s, 1H), 1.97 (s, 3H).

2.2.2.2 Synthesis of 3-acetamido-5-methyl-pyrazole, P2
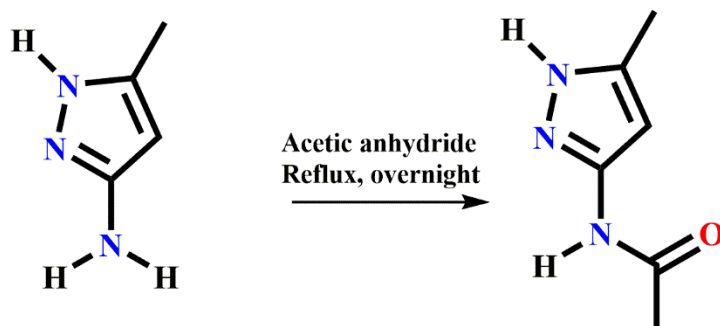
^{1}H-pyrazol-5-methyl-3-amine (2.25 g, 2.30mmole) was dissolved in 20 mL of distilled water. NaHCO$_3$ (5.8 g, 6.9 mmole) was slowly added. Acetic anhydride (8ml) was then added dropwise and the resulting suspension was heated at reflux overnight. Then, the mixture was allowed to cool down to room temperature and the solid obtained was filtered off and characterized as the title compound. 74 % yield, m.p. 215-218°C, ^{1}H NMR (300 MHz, DMSO-d$_6$) δ ppm: 11.89 (s, 1H), 10.16 (s, 1H), 6.22 (s, 1H), 2.16 (s, 3H), 1.94 (s, 3H).

2.2.2.3 Synthesis of 3-propamido-pyrazole, P3
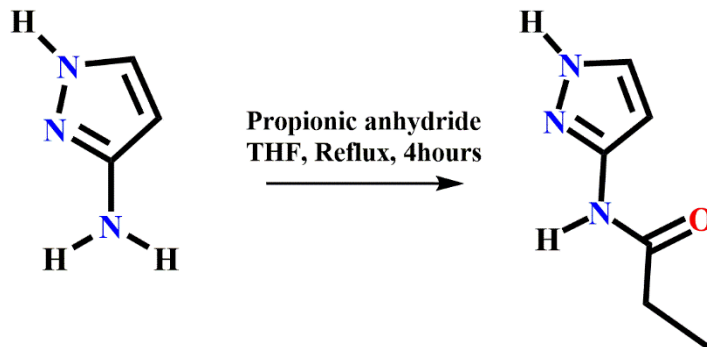
3-amino-[1]H-pyrazole (0.486 g, 5.85 mmole) was dissolved in 10ml tetrahydrofuran in a 50ml round bottomed flask. 1 to 1.3 equivalence of propionic anhydride was added to the mixture and the resulting mixture was refluxed at 60-65 °C for 4 hours, monitored with TLC and after completion the excess solvent was removed by rotatory evaporation. The product was recrystallized from methanol to obtain the white solid as the pure product. Yield: 85%; mp 189-190°C, ^{1}H NMR (δH; DMSO, 400MHz): 12.26 (s, 1H), 10.29 (d, 1H), 7.56 (d, 1H), 6.48 (d,1H), 2.26 (q, 2H), 1.04 (t, 3H).

2.2.2.4 Synthesis of 3-propamido-5-methyl-pyrazole, P4[24]
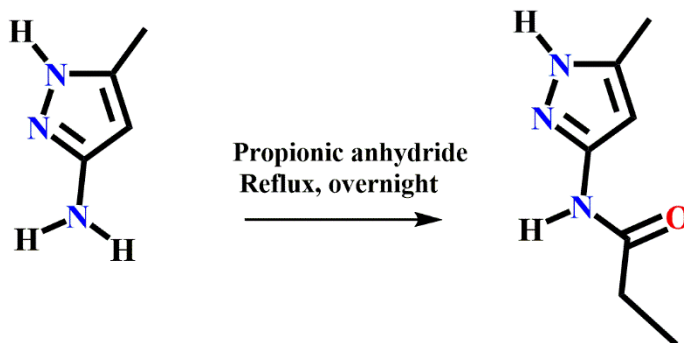
[1]H-pyrazol-5-methyl-3-amine (2.25 g, 2.30mmole) was dissolved in 20 mL of distilled water. NaHCO$_3$ (5.8 g, 6.9 mmol) was slowly added. Propionic anhydride (10ml) was then added dropwise and the resulting suspension was heated at reflux overnight. Then, the mixture was allowed to cool down to room temperature and the solid obtained was filtered off and characterized as the title compound. 74 % yield, m.p. 189-190°C, ^{1}H NMR (400 MHz, DMSO-d$_6$) δ ppm: 11.92 (s, 1H), 10.15 (s, 1H), 6.25 (s, 1H), 2.26 (s, 3H), 2.17 (q, 2H), 1.03 (s, 3H).

2.2.2.5 Synthesis of 3-butyramido pyrazole, P5
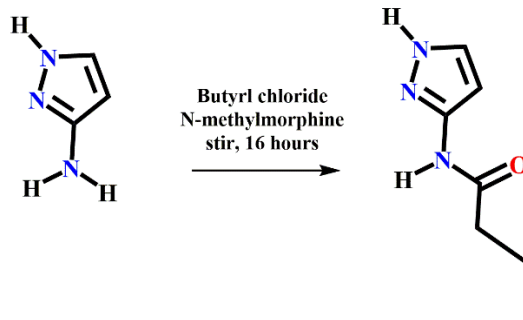
3-Aminopyrazole (0.499 g, 6.00 mmol) was dissolved in 30 mL of dichloromethane, 1.6 mL N-methylmorpholine (14.6 mmol), and 1.49 mL butyryl chloride (14.0 mmol). After stirring for 16 hours, the mixture was concentrated, and the residue was dissolved in 30 mL methanol. Then, 7 mL of 2.5M sodium hydroxide was added dropwise, followed by 10 mL of tetrahydrofuran. The mixture was allowed to sit for 15 minutes before being concentrated again and suspended in water. The tan precipitate was filtered, air dried, and finally recrystallized in methanol to yield white crystals. 66% yield, m.p. 130-141°C, ^{1}H NMR (400 MHz, DMSO-d$_6$) δ ppm: 12.24 (1H), 10.26 (1H), 7.55 (1H), 6.48 (1H), 2.24 (2H), 1.56 (2H), 0.88 (3H).

2.2.2.6 Synthesis of 3-butyramido 5-methyl pyrazole, P6
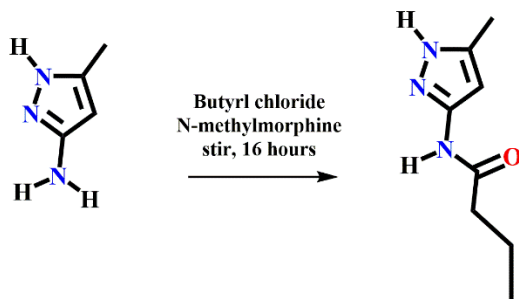
3-Amino-5-methylpyrazole (0.583 g, 6.00 mmol) was dissolved in 30 mL of dichloromethane, 1.6 mL N-methylmorpholine (14.6 mmol), and 1.49 mL butyryl chloride (14.0 mmol). After stirring for 16 hours, the mixture was concentrated, and the residue was dissolved in 30 mL methanol. Then, 7 mL of 2.5M sodium hydroxide was added dropwise, followed by 10 mL of tetrahydrofuran. The mixture was allowed to sit for 15 minutes before being concentrated again and suspended in water. The tan precipitate was filtered, air dried, and finally recrystallized in

methanol to yield white crystals. ~70% yield, m.p. 184-189°C, ^{1}H NMR (400 MHz, DMSO-d$_6$) δ ppm: 11.90 (1H), 10.12 (1H), 6.25 (1H), 2.22 (2H), 2.17 (3H), 1.56 (2H), 0.87 (3H).

2.2.2.7 Synthesis of 3-benzamido-pyrazole, P7[25]
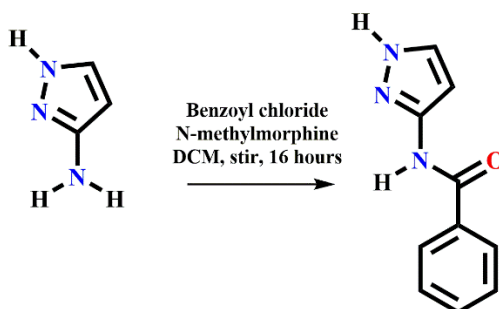
To a solution of 2.49 g of 3-amino-[1]H-pyrazole in 150ml of dichloromethane, 8ml of n-methylmorpholine and 8ml of benzoyl chloride were added at room temperature. After 16 hours of stirring, the mixture was concentrated, and the residue was dissolved in 150ml of methanol. An amount of 3.5 g of sodium hydroxide in 35ml of water was added dropwise and 100ml of THF was added to obtain homogenous solution. After 15 minutes under stirring, the solution was filtered to get rid of the solid and the solution was concentrated and the solid obtained was poured into water. The precipitate obtained was filtered and air-dried to obtain the pure product in 86% yield, m.p. 162-164°C, ^{1}H NMR (400 MHz, DMSO-d$_6$) δ ppm: 12.47 (s, 1H), 10.83 (s, 1H), 8.00 (s, 1H), 7.66 (s, 3H), 7.50 (q, 2H), 6.63(s, 1H).

2.2.2.8 Synthesis of 3-benzamido-5-methyl-pyrazole, P8[25]
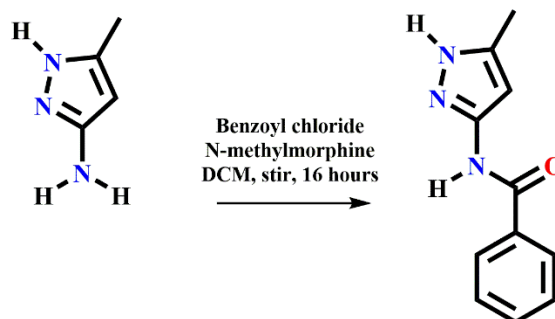
To a solution of 2.49 g of 3-amino-5methyl-pyrazole in 150ml of dichloromethane, 8ml of n-methylmorpholine and 8ml of benzoyl chloride were successfully added at room temperature. After 16 hours of stirring, the mixture was concentrated, and the residue was dissolved in 150ml

of methanol. An amount of 3.5 g of sodium hydroxide in 35ml of water was added dropwise and 100ml of THF was added to obtain homogenous solution. After 15 minutes under stirring, the solution was filtered to get rid of the solid and the solution was concentrated and the solid obtained was poured into the water. The precipitate obtained was filtered and air-dried to obtain the pure product in 89% yield, m.p. 216-217°C, ^{1}H NMR (400 MHz, DMSO-d$_6$) δ ppm: 12.12 (s, 1H), 10.67 (s, 1H), 7.98 (s, 1H), 7.53 (s, 3H), 7.48 (q, 2H), 6.40(s, 1H).

2.2.2.9 Synthesis of 3-pyridyl-pyrazole, P9
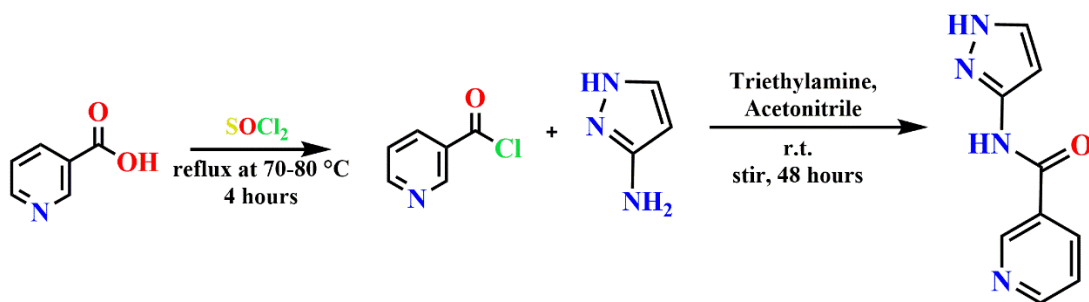
3-amino-1H-pyrazole (12mmol, 0.99g) was dissolved in 30 ml of acetonitrile. The resulting solution was treated with nicotinoyl chloride (12mmol, 1.70g) and triethylamine (2.0ml). The mixture was stirred at room temperature for 48 hours and monitored via TLC every 6 hours. Once, the reaction was completed, the organic phase was separated by filtration, washed with water and dried in air to get the white solid product. 84% yield, m.p. 226-227°C. ^{1}H NMR (400 MHz, DMSO-d$_6$) δ ppm: 12.52 (s, 1H), 11.10 (s, 1H), 9.13 (s, 1H), 8.73 (s, 3H), 8.32 (s, 1H), 7.69(s, 1H), 7.53 (s,1H) and 6.66 (s, 1H).

2.2.2.10 Synthesis of 3-pyridyl-5-methyl-pyrazole, P10
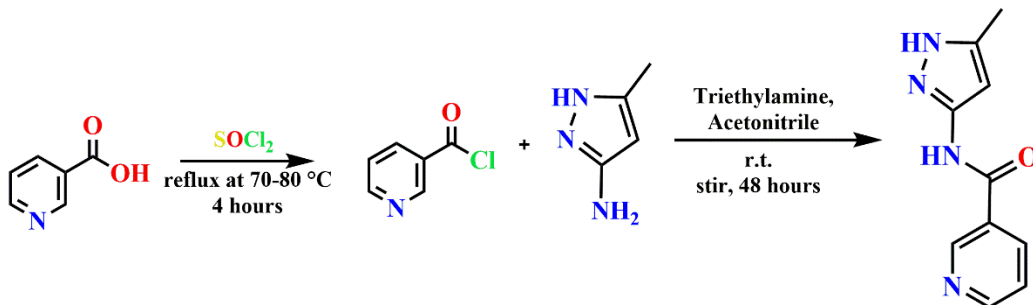
3-Amino-5-methyl-pyrazole (12mmol, 1.69g) was dissolved in 20 ml of acetonitrile. The resulting solution was treated with nicotinoyl chloride (12mmol, 1.70g) and triethylamine (1.6ml). The mixture was stirred at room temperature for 48 hours. The mixture was stirred at room temperature for 48 hours and monitored via TLC every 6 hours. Once, the reaction was completed, the organic phase was separated by filtration, washed with water and dried in air to get the yellow solid product in 85% yield, m.p. 205-207°C. ^{1}H NMR (400 MHz, DMSO-d$_6$) δ ppm: 12.18 (s, 1H), 10.96 (s, 1H), 9.11 (s, 1H), 8.72 (s, 1H), 8.30(s, 1H), 7.52(s, 1H), 6.41 (s, 1H) and 2.23 (s, 3H).

2.2.2.11 Synthesis of 4-pyridyl-pyrazole, P11
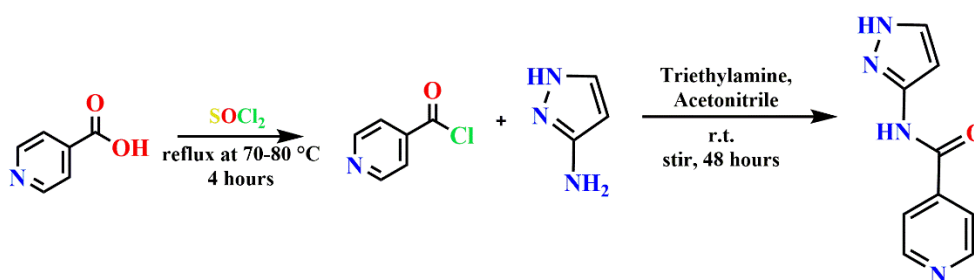
4-Amino-1H-pyrazole (12mmol, 0.99g) was dissolved in 20 ml of acetonitrile. The resulting solution was treated with nicotinoyl chloride (12mmol, 1.70g) and triethylamine (1.6ml). The mixture was stirred at room temperature for 48 hours. The mixture was stirred at room temperature for 48 hours and monitored via TLC every 6 hours. Once, the reaction was completed, the organic phase was separated by filtration, washed with water and dried in air to get the yellow solid product. 91% yield, m.p. 236-237°C. ^{1}H NMR (400 MHz, DMSO-d$_6$) δ ppm: 12.55 (s, 1H), 11.17 (s, 1H), 8.74 (s, 1H), 7.89 (s, 1H), 7.69 (s, 1H), 6.65(s, 1H).

2.2.2.12 Synthesis of 4-pyridyl-5-methyl-pyrazole, P12
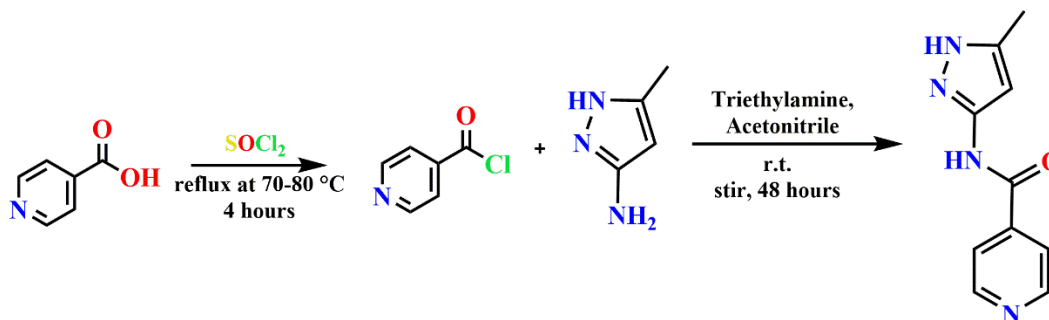
3-Amino-5-methyl-pyrazole (12mmol, 1.69g) was dissolved in 20 ml of acetonitrile. The resulting solution was treated with isonicotinoyl chloride (12mmol, 1.70g) and triethylamine (1.6ml). The mixture was stirred at room temperature for 48 hours. The mixture was stirred at room temperature for 48 hours and monitored via TLC every 6 hours. Once, the reaction was completed, the organic phase was separated by filtration, washed with water and dried in air to get the yellow solid product in 94% yield, m.p. 241-242°C. ^{1}H NMR (400 MHz, DMSO-d$_6$) δ ppm: 11.31 (s, 1H), 10.13(s, 1H), 9.10 (s, 1H), 8.72 (s, 1H), 8.30(s, 1H), 7.83(s, 1H), 6.96 (s, 1H), 5.51 (s, 1H) and 1.32 (s, 3H).

2.2.3 Molecular conformation analysis of P1-P12

(a)Amide functionality	(b) CSD structure analysis of amide conformation
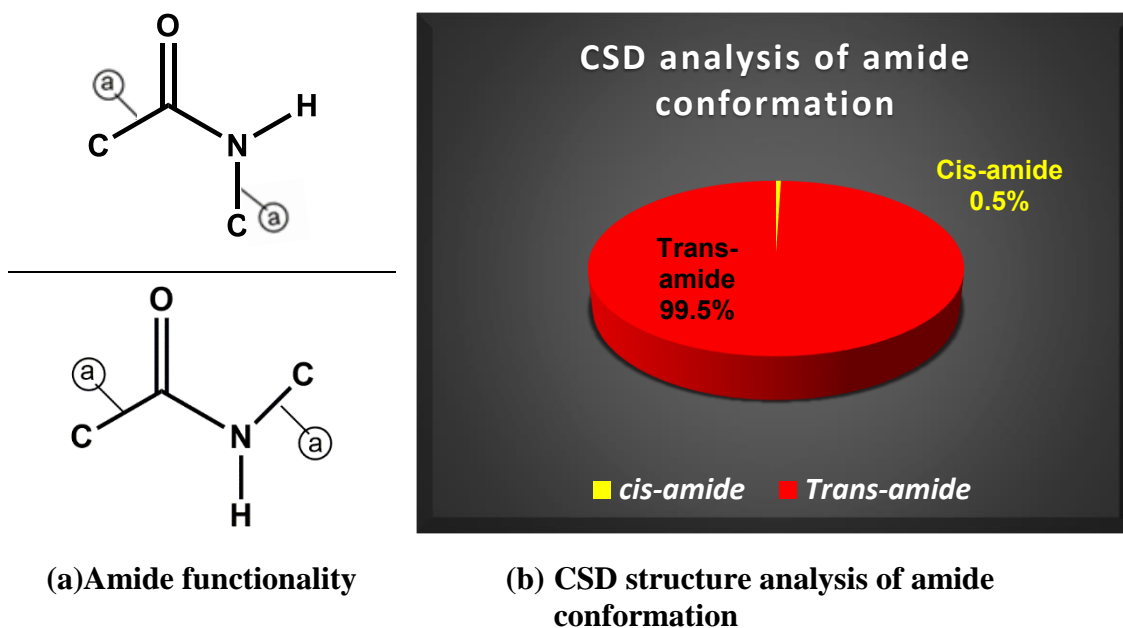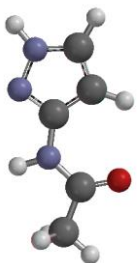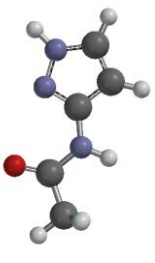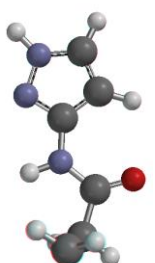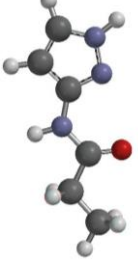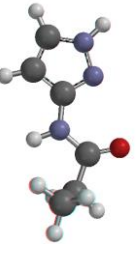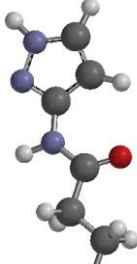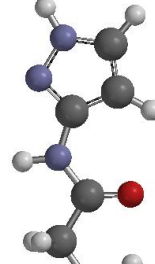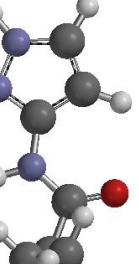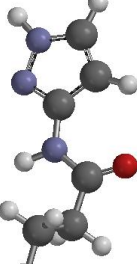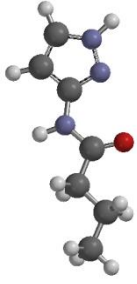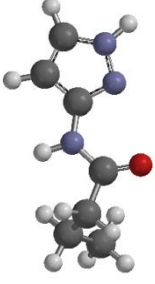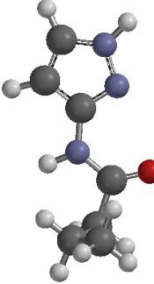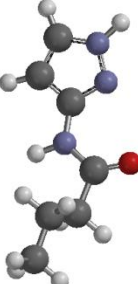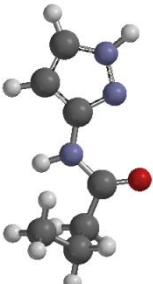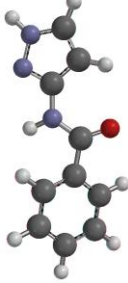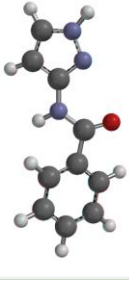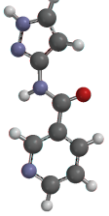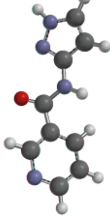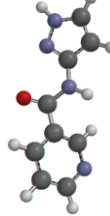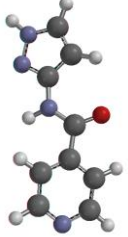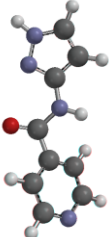
Figure 2.1 (a) *Cis* and *trans* amide functionality (both bonds are acyclic representing using symbol @) used to perform the torsion angle search. (b) Pie chart indicating number of structures with torsions for *cis* (yellow, ~32 structures, 0.5%) and *trans* (red, ~6303 Structures, 99.5%) conformations.

A CSD database search was performed to determine whether the molecules with amide functionality occur as *trans* conformation or *cis* conformation. A total of 6335 structures was obtained and approximately 30 (0.5%) of these structures have the *cis*-amide conformation and about 6300 structures (99%) have *trans*-amide conformation, Figure 2.1. Therefore, we determined the relative energy of all conformations with respect to most stable conformation with the restriction that all have a *trans*-configuration of the amide group (as suggested by the CSD

analysis). Geometry optimization was performed on each conformation of **P1-P12** and the most stable conformation is given a value of 0 kJ/mol; all other conformations energies are presented relative to this, Table 2.1.

Table 2.1 Energies of each trans amide conformation relative to most stable trans conformation is shown below in kJ/mol. The conformations with duplicate energies were ignored. Note: methyl-based target molecule conformations are not shown here.

	Conformation 1	Conformation 2	Conformation 3	Conformation 4	Conformation 5
P1					
ΔE (P1)	**0**	**+50**			
ΔE (P2)	**0**	**+49**			
P3					
ΔE (P1)	**0**	**0**	**+50**	**+50**	
ΔE (P2)	**0**	**+4**	**+50**	**+53**	
P5					
ΔE (P1)	**0**	**0**	**+2**	**+2**	**+3**
ΔE (P2)	**0**	**0**	**+2**	**+3**	**+5**

ΔE (P1)	+50	+51	+52	+53	+54
ΔE (P2)	+50	+50	+52	+53	+55
P7					
ΔE (P1)	0	+52			
ΔE (P2)	0	+51			
P9/P10					
ΔE (P1)	0	+4	+51	+55	
ΔE (P2)	0	+4	+51	+54	
P11					
ΔE (P1)	0	+52			
ΔE (P2)	0	+51			

2.2.4 Molecular electrostatic potential (MEPs) calculations

Molecular electrostatic potential surfaces (MEPS) of **P1**–**P12** were generated with DFT B3LYP level of theory using 6-311++G** basis set in vacuum. All calculations were carried out using Spartan'08 software. All molecules were geometry optimized with the maxima and minima in the electrostatic potential surface (0.002 e/au isosurface) determined using a positive point charge in the vacuum as a probe. The numbers indicate the interaction energy (kJ/mol) between the positive point probe and the surface of the molecule at that particular point. The Etter'r rule based on electrostatic potentials was used to determine the best donor-best acceptor interaction.

2.2.5 Hydrogen-bond energies (HBE) for synthon predictions

The synthon predictions for pure compounds **P1-P12** was made by calculating interaction energies to determine which of the three postulated synthons is most likely to appear in the crystal structures of the pure compounds. The hydrogen-bond parameters, α (hydrogen-bond donor) and β (hydrogen-bond acceptor) is determined using maxima and minima on the MEPS respectively (Equation 2.1 and 2.2), and the free energy of interaction is given by the product, $-\alpha\,\beta$.[20] The

$$\alpha = 0.0000162\, MEPmax^2 + 0.00962 MEP_{max} \quad \text{(Equation 2.1)}$$
$$\beta = 0.000146\, MEPmin^2 - 0.00930 MEP_{min} \quad \text{(Equation 2.2)}$$
$$E = -\sum_{ij} \alpha_i \beta_j \quad \text{(Equation 2.3)}$$
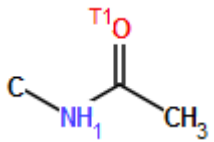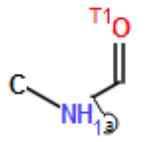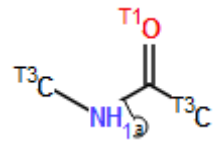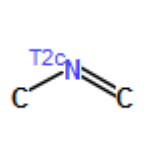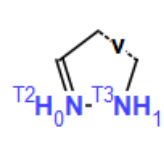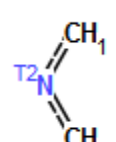
Table 2.2 Electrostatics (in kJ/mol) and α and β values for P1-P12 calculated using equations 2.1 and 2.2 for donor and acceptor group.

	NH amide (kJ/mol)	α (amide)	NH pyrazole	α (pyrazole)	cyclic N(kJ/mol)	β (N)	carbonyl O (kJ/mol)	β(O)	Pyridine N	B(N)
P1	210.00	2.73	255.00	3.51	-147.00	4.52	-198.00	7.57		
P2	205.00	2.65	244.00	3.31	-156.00	5.00	-202.00	7.84		
P3	206.00	2.67	252.00	3.45	-149.00	4.63	-195.00	7.37		
P4	200.00	2.57	242.00	3.28	-160.00	5.23	-199.00	7.63		
P5	207.00	2.69	253.00	3.47	-150.00	4.68	-196.00	7.43		
P6	201.00	2.59	242.00	3.28	-159.00	5.17	-199.00	7.63		
P7	196.00	2.51	255.00	3.51	-146.00	4.47	-191.00	7.10		
P8	189.00	2.40	244.00	3.31	-156.00	5.00	-195.00	7.37		
P9	209.00	2.72	250.00	3.42	-136.00	3.97	-179.00	6.34	-182.00	6.53
P10	209.00	2.72	250.00	3.42	-132.00	3.77	-167.00	5.62	-174.00	6.04
P11	190.00	2.41	240.00	3.24	-148.00	4.57	-173.00	5.98	-190.00	7.04
P12	204.00	2.64	244.00	3.31	-140.00	4.16	-171.00	5.86	-178.00	6.28

2.2.6 Hydrogen-bond propensities (HBP) for synthon predictions

To complement the electrostatic-based calculations, we used hydrogen-bond propensity calculations (CSD Version 5.38 and Mercury 3.9) to predict the synthons in the pure compounds **P1-P12**. Each compound was sketched and auto-edited, a careful selection of functional groups (Table 2.3) and training dataset (350-500) was made and the propensities were calculated with an ROC curve higher than 0.831 ("excellent discrimination"). The propensity was used to determine the most likely synthon in these molecules.

Table 2.3 Functional groups used to determine the hydrogen-bond propensities for the P1-P12 target molecules. The labels in the figures can be explained as follows: Tn = atom makes n bonds, c = atom is cyclic, ⓑ = bond is acyclic, and Hn = n bonded hydrogen atoms.

P1/P2	P3/P4/P5/P6	P7-P12	P1-P12	P1-P12	P9-P12

2.2.7 Hydrogen-bond coordination (HBC)

Because of the presence of multifunctional groups on each molecule, there is chance of synthon polymorphism and synthon crossover. Each molecule was run through a polymorph assessment analysis in the propensity tool and compared to experimentally obtained structure. In order to determine whether the target molecules **P1-P12** have the chance to form synthon polymorphs, propensity-coordination analysis was performed on each molecule and the most optimal hydrogen-bond motif was obtained.

2.2.8 Crystal structures

P1-P12 were kept in a vial for slow evaporation methanol solvent in order to obtain crystals suitable for single crystal X-ray diffraction. If suitable crystals were not obtained in methanol, then different solvents were tried to grow crystals. The molecules were categorized into crystallizable or non-crystallizable based on slow evaporation method. DSC heat/cool/heat method was also used

36

to determine the crystallizability class of molecules. X-ray experimental data and crystallographic data are given in the appendix D.

Table 2.4 Experimental details of crystals obtained in this study

	Molecules	Solvent used	Morphology	Melting point
Group 1	P1	Methanol	Block, colorless	218-220°C
	P2	Methanol	Plate, colorless	215-218°C
	P3	Methanol	Plates, colorless	189-190°C
	P4	Methanol	Block, colorless	189-190°C
	P5	Methanol	Crystals not solved	130-141°C
	P6	Methanol	Crystals not solved	184-189°C
	P7	Methanol	Block, colorless	162-163°C
	P8	Methanol	Prism, colorless	217-218°C
Group 2	P9	Methanol	N/A	227-228°C
	P10	Methanol	Block colorless	205-206°C
	T11	Methanol	Block colorless	236-237°C
	T12	Methanol	Crystals not solved	241-242°C

2.3 Results

2.3.1 Conformational analysis

Conformation analysis was performed using spartan 08' software to determine all the possible conformations of each molecule and the relative energy of each conformation was calculated respective to the most stable conformation. Molecule **P1**, **P2, P7, P8, P11 and P12** each has four possible conformations; **P3**, **P4**, **P9,** and **P10** each has eight possible conformations, and **P5** and **P6** has ten conformations each.

2.3.2 Molecular electrostatics potentials (MEPs)

MEPs values for each donor group (pyrazole NH and amide NH) and acceptor group (pyrazole N, C=O amide for **P1-P8** and additional acceptor as pyridine N for **P9-P12**) are presented in Figure 2.2 and 2.3. The ranking of the acceptor sites and a donor site is established using molecular

37

electrostatic potential surfaces. Higher negative potential (indicated by red clouds) on the acceptor suggests better hydrogen-bond accepting ability and higher positive potential (indicated by blue clouds) on the donor suggests better hydrogen-bond donor ability.
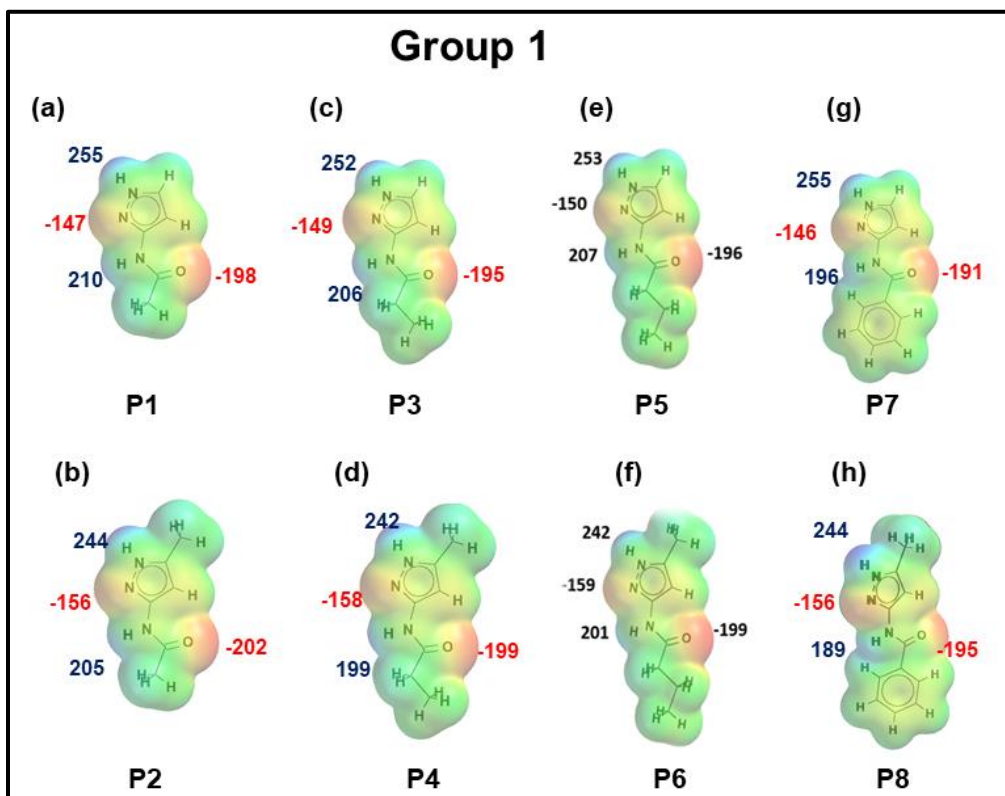
Figure 2.2 Electrostatic potentials values (in kJ/mol); a) P1 b) P2 c) P3 d) P4 e) P5 f) P6 g) P7 and h) P8.
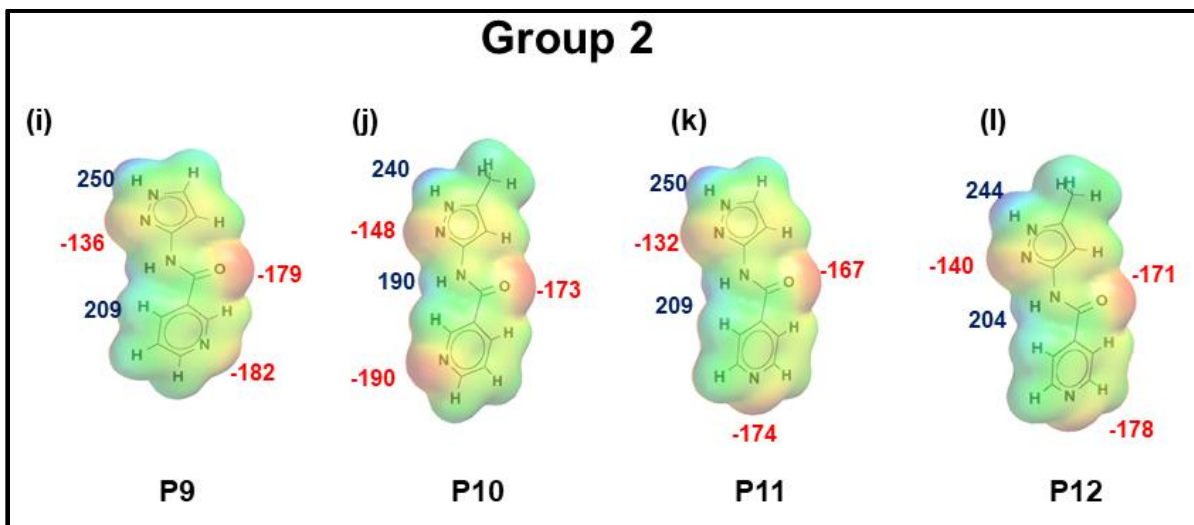
Figure 2.3 Electrostatic potentials values (in kJ/mol); i) P9 j) P10 k) P11 and l) P12.

2.3.3 Hydrogen-bond energies (HBE)

The hydrogen-bond energies calculated for each individual (Table 2.5) and combination synthon are summarized in Table 2.6.

Table 2.5 Hydrogen-bond energies (in kJ/mol) for each individual synthon for molecules P1-P12. Synthon A and C are dimeric synthons; therefore, energies are presented for pairs of molecules.

	Synthon A		Synthon B	Synthon C		Synthon D	Synthon E	Synthon F
	Monomeric	Dimeric	Monomeric	Monomeric	Dimeric	Monomeric		
P1	-15.86	-31.71	-26.53	-12.37	-24.73	-20.69		
P2	-16.57	-33.14	-25.95	-13.27	-26.55	-20.79		
P3	-15.98	-31.95	-25.43	-12.35	-24.70	-19.66		
P4	-17.12	-34.25	-25.01	-13.44	-26.88	-19.63		
P5	-16.24	-32.49	-25.79	-12.57	-25.14	-19.96		
P6	-16.94	-33.88	-25.01	-13.38	-26.76	-19.75		
P7	-15.67	-31.35	-24.91	-11.21	-22.42	-17.81		
P8	-16.57	-33.14	-24.39	-11.99	-23.99	-17.65		
P9	-13.55	-27.10	-21.68	-10.78	-21.56	-17.24	-22.31	-17.75
P10	-14.83	-29.66	-19.38	-11.04	-22.07	-14.42	-22.82	-16.98
P11	-12.89	-25.78	-19.22	-10.25	-20.50	-15.29	-20.64	-16.41
P12	-13.79	-27.58	-19.41	-10.98	-21.96	-15.45	-20.80	-16.56

Table 2.6 Hydrogen-bond energies (in kJ/mol) for each combination synthon for molecules **P1-P12**.

	Synthon (A+F)	Synthon (A+D)	Synthon (C+E)	Synthon C+B	Synthon (D+E)	Synthon (B+F)
	Dimeric	Dimeric	Dimeric	Dimeric		
P1	N/A	-52.40	N/A	-51.26	N/A	N/A
P2	N/A	-53.93	N/A	-52.50	N/A	N/A
P3	N/A	-51.61	N/A	-50.13	N/A	N/A
P4	N/A	-53.88	N/A	-51.89	N/A	N/A
P5	N/A	-52.44	N/A	-50.93	N/A	N/A
P6	N/A	-53.63	N/A	-51.77	N/A	N/A
P7	N/A	-49.16	N/A	-47.32	N/A	N/A
P8	N/A	-50.80	N/A	-48.38	N/A	N/A
AVG	N/A	-52.23	N/A	-50.52	N/A	N/A
S.Dev	N/A	1.67	N/A	1.81	N/A	N/A
P9	-44.85	-44.34	-43.87	-43.23	-39.55	-39.42
P10	-46.64	-44.08	-44.89	-41.45	-37.24	-36.36
P11	-42.19	-41.07	-41.14	-39.73	-35.93	-35.64
P12	-44.14	-43.03	-42.76	-41.36	-36.25	-35.97
AVG	-44.46	-43.13	-43.17	-41.44	-37.24	-36.85

S.Dev	1.84	1.49	1.61	1.43	1.64	1.74

2.3.4 Hydrogen-bond propensities (HBP)

The propensities calculations consider all possible interactions between two donors (pyrazole NH and amide NH) and two acceptors (pyrazole N and C=O) resulting in four propensity numbers for **P1-P8**. In molecules with an additional acceptor group **P9-P12,** six different combinations of propensities are calculated. The propensity of individual and combination synthon is presented in Table 2.6 and Table 2.7.

Table 2.7 Hydrogen-bond propensities for each individual synthon possible in molecules **P1-P12**.

	Synthon A	Synthon B	Synthon C	Synthon D	Synthon E	Synthon F
P1	0.65	0.69	0.49	0.54		
P2	0.63	0.76	0.39	0.55		
P3	0.68	0.75	0.48	0.57		
P4	0.69	0.73	0.51	0.56		
P5	0.61	0.69	0.45	0.54		
P6	0.62	0.73	0.44	0.56		
P7	0.57	0.51	0.38	0.32		
P8	0.51	0.49	0.34	0.32		
P9	0.48	0.48	0.27	0.27	0.69	0.47
P10	0.48	0.47	0.27	0.26	0.71	0.49
P11	0.45	0.48	0.25	0.27	0.69	0.47
P12	0.45	0.47	0.25	0.27	0.7	0.49

Table 2.8 Hydrogen-bond propensities for combination synthons possible in molecules **P1-P12**. Combination synthon propensities are calculated by multiplying the individual synthon propensities.

	Synthon (A+F)	Synthon (A+D)	Synthon (C+E)	Synthon C+B	Synthon (D+E)	Synthon (B+F)
P1	N/A	0.35	N/A	0.34	N/A	N/A
P2	N/A	0.35	N/A	0.30	N/A	N/A
P3	N/A	0.39	N/A	0.36	N/A	N/A
P4	N/A	0.39	N/A	0.37	N/A	N/A
P5	N/A	0.33	N/A	0.31	N/A	N/A
P6	N/A	0.35	N/A	0.32	N/A	N/A
P7	N/A	0.18	N/A	0.19	N/A	N/A
P8	N/A	0.16	N/A	0.17	N/A	N/A
P9	0.23	0.13	0.19	0.13	0.19	0.23
P10	0.24	0.12	0.19	0.13	0.18	0.23
P11	0.21	0.12	0.17	0.12	0.19	0.23
P12	0.22	0.12	0.18	0.12	0.19	0.23

2.3.5 Hydrogen-bond coordination (HBC)

The hydrogen-bond coordination tool was used to determine the coordination of each functional group which will guide us to determine which synthon is most likely to happen. The propensity-coordination chart and the corresponding coordination for each hypothetical motif is shown in Figure 2.4-2.7.

P1-P6
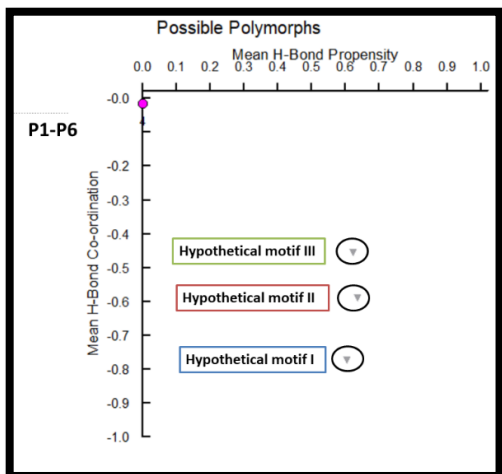
Motif #	Atom (D/A)	0	1	2	3
III	NH of pyrazole (d)	0.01	**0.61**	0.39	0.00
	NH of amide (d)	0.00	**0.96**	0.04	0.00
	N of pyrazole (a)	**0.12**	0.88	0.00	0.00
	O=C of amide (a)	0.02	0.74	**0.23**	0.01
I	NH of pyrazole (d)	0.01	**0.61**	0.39	0.00
	NH of amide (d)	0.00	**0.96**	0.04	0.00
	N of pyrazole (a)	0.12	**0.88**	0.00	0.00
	O=C of amide (a)	0.02	**0.74**	0.23	0.01
II	NH of pyrazole (d)	0.01	0.61	**0.39**	0.00
	NH of amide (d)	0.00	**0.96**	0.04	0.00
	N of pyrazole (a)	0.12	**0.88**	0.00	0.00
	O=C of amide (a)	0.02	0.74	**0.23**	0.01

(a) Propensity-coordination chart   **(b) Coordination table**

Figure 2.4 (a)Propensity-coordination chart of P1-P6 molecules and (b) coordination of each functional group in all predicted motifs.

P7-P8
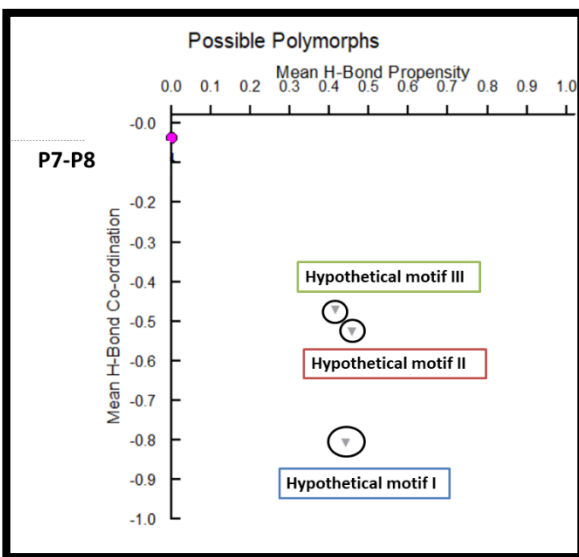
Motif #	Atom (D/A)	0	1	2	3
I	NH of pyrazole	0.01	**0.72**	0.27	0.00
	NH of amide	0.02	**0.95**	0.03	0.00
	N pyrazole	0.17	**0.83**	0.00	0.00
	O=C amide	0.04	**0.82**	0.14	0.00
III	NH of pyrazole	0.01	**0.72**	0.27	0.00
	NH of amide	0.02	**0.95**	0.03	0.00
	N pyrazole	**0.17**	0.83	0.00	0.00
	O=C amide	0.04	0.82	**0.14**	0.00
II	NH of pyrazole	0.01	0.72	**0.27**	0.00
	NH of amide	0.02	**0.95**	0.03	0.00
	N pyrazole	0.17	**0.83**	0.00	0.00
	O=C amide	0.04	0.82	**0.14**	0.00

Figure 2.5 Propensity-coordination chart of P7-P8 molecules and the coordination of each functional group in all predicted motifs.

P9-P10

Motif #	Atom (D/A)	0	1	2	3
I	NH of pyrazole (d)	0.00	0.62	0.38	0.00
	NH of amide (d)	0.02	0.95	0.03	0.00
	N of pyridine (a)	0.11	0.81	0.08	0.00
	N pyrazole (a)	0.40	0.60	0.00	0.00
	C=O amide (a)	0.17	0.80	0.03	0.00
III	NH of pyrazole (d)	0.00	0.62	0.38	0.00
	NH of amide (d)	0.02	0.95	0.03	0.00
	N of pyridine (a)	0.11	0.81	0.08	0.00
	N pyrazole (a)	0.40	0.60	0.00	0.00
	C=O amide (a)	0.17	0.80	0.03	0.00
II	NH of pyrazole (d)	0.00	0.62	0.38	0.00
	NH of amide (d)	0.02	0.95	0.03	0.00
	N of pyridine (a)	0.11	0.81	0.08	0.00
	N pyrazole (a)	0.40	0.60	0.00	0.00
	C=O amide (a)	0.17	0.80	0.03	0.00

(a) Propensity-coordination chart   **(b) Coordination table**

Figure 2.6 Propensity-coordination chart of P9-P10 molecules and the coordination of each functional group in all predicted motifs.
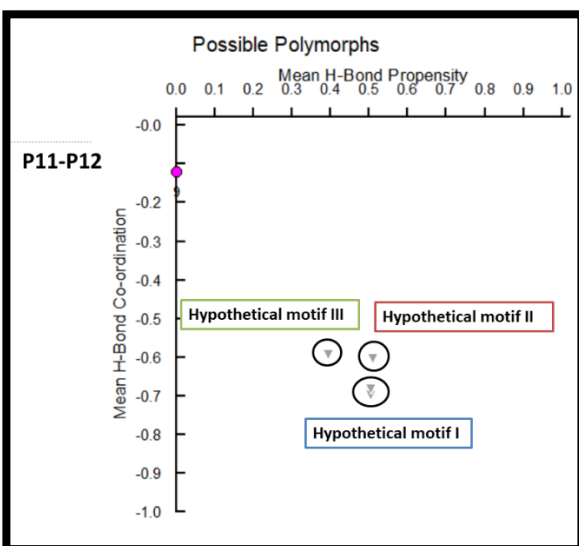
P11-P12

Motif #	Atom (D/A)	0	1	2	3
I	NH of pyrazole (d)	0.01	0.83	0.16	0.00
	NH of amide (d)	0.01	0.96	0.03	0.00
	N of pyridine (a)	0.11	0.82	0.07	0.00
	N of pyrazole (a)	0.35	0.65	0.00	0.00
	C=O of amide (a)	0.19	0.78	0.03	0.00
III	NH of pyrazole (d)	0.01	0.83	0.16	0.00
	NH of amide (d)	0.01	0.96	0.03	0.00
	N of pyridine (a)	0.11	0.82	0.07	0.00
	N of pyrazole (a)	0.35	0.65	0.00	0.00
	C=O of amide (a)	0.19	0.78	0.03	0.00
II	NH of pyrazole (d)	0.01	0.83	0.16	0.00
	NH of amide (d)	0.01	0.96	0.03	0.00
	N of pyridine (a)	0.11	0.82	0.07	0.00
	N of pyrazole (a)	0.35	0.65	0.00	0.00
	C=O of amide (a)	0.19	0.78	0.03	0.00

(a) Propensity-coordination chart   **(b) Coordination table**

Figure 2.7 Propensity-coordination chart of P11-P12 molecules and the coordination of each functional group in all predicted motifs.

2.3.6 Observed synthons in each crystal structure

We were able to obtain crystallographic data for nine out of twelve molecules, **P1-P12**. **P9** didn't grow crystals suitable for single crystal x-ray diffraction studies. Our structure determination of **P2** was consistent with the reported structure in the CSD(ARAGUV)[26] and the crystal structure of **P8** is also reported in the CSD (PESQEK)[27]. The experimentally observed synthon in each crystal structure is shown in Figure 2.8 and Figure 2.9.
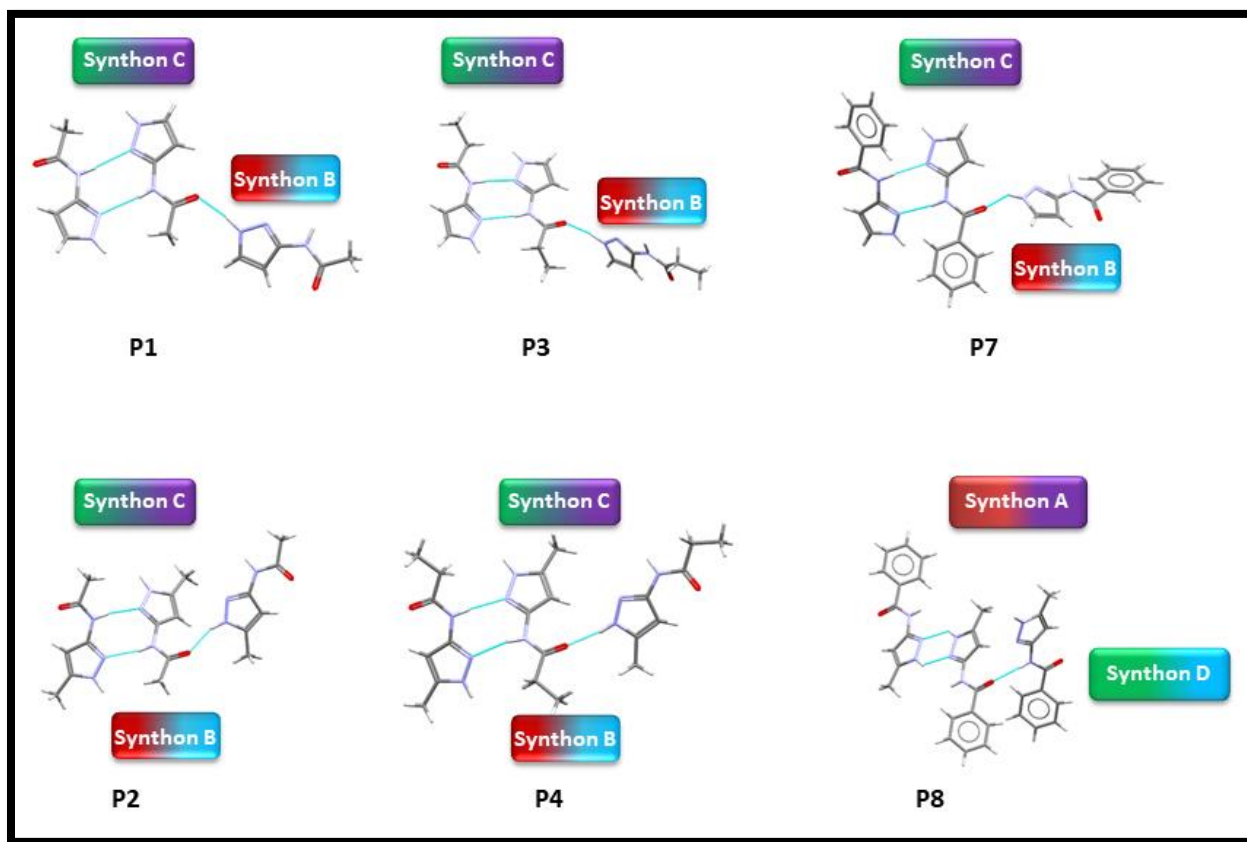
Figure 2.8 Synthon (B+C) observed in crystal structures of **P1**, **P2**, **P3**, **P4**, **P7** and **P8**.
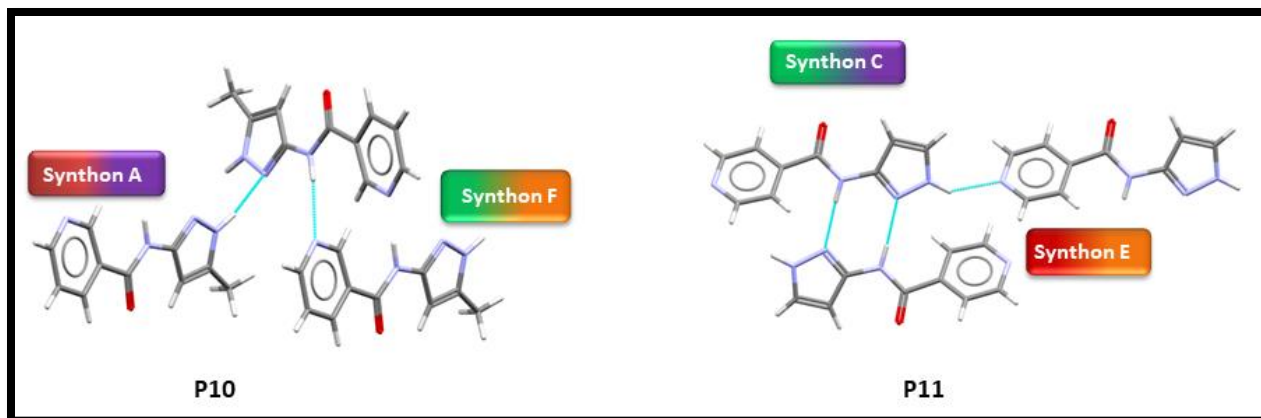
Figure 2.9 Synthon (A+D) observed in crystal structure of **P10** and **P11**.

2.4 Discussion

2.4.1 Molecular conformational analysis

Molecular conformational analysis shows that the most stable conformation in each molecule has the amide functionality *trans* to each other and cyclic N group *cis* to the amide NH (Table 2.1). Additionally, in molecule **P9** and **P10**, pyridine N is *trans* to the amide NH in the most stable conformation and the second conformation with pyridine N *cis* to amide NH is only ~4 kJ/mole in energy higher than the most stable conformation. Therefore, the main conformation in each molecule will have *trans* amide conformation as the most stable conformation. When compared to the experimental crystal structure, the *trans* amide conformation was indeed observed in all molecules. In the target molecule **P11**, second most stable conformation in which the pyridine N *cis* to the amide NH was observed, it is only ~4 kJ/mole higher than the most stable conformation where the pyridine N is *trans* to the amide NH. It is worth noting that the energy optimized conformations are not necessarily completely identical to those that may appear in the solid state, where a variety of close contacts and packing forces may distort some geometric parameters away from idealized gas phase values. However, these idealized conformations are likely to be most relevant in the solution phase at the point when target molecules begin to recognize and bind to each other during nucleation and growth.

2.4.2 What are the preferred synthons?

Four different methods (MEPs, HBE, HBP, and HBC) were used to predict synthons. For molecules **P1-P8**, Synthon (**B+C**) was predicted by all four methods. Also, synthon (**A+D**) was predicted by HBE and HBP method. For molecules **P9-P12**, Synthon (**D+E**) was predicted by MEPs and HBC. Synthon (**B+F**) was predicted by HBP and HBC. Synthon (**A+F**) was predicted by both HBE and HBP. Synthon (**C+E**) and synthon (**A+D**) were predicted by HBE only (Figure 2.10).
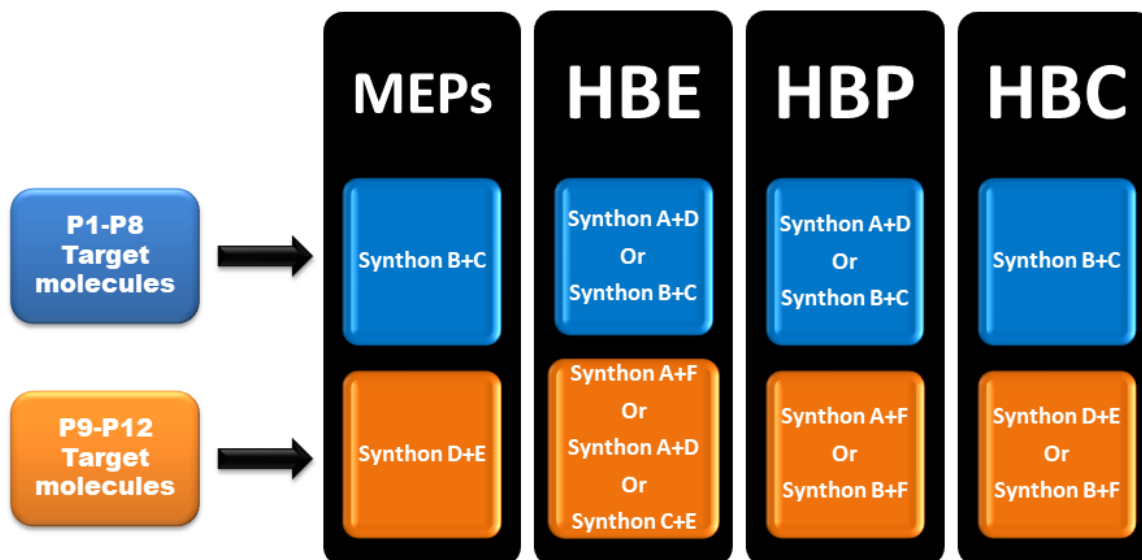
Figure 2.10 Summary of synthons predicted in the target molecules, **P1-P12**.

2.4.3 Molecular geometric complementarity

Our next goal was to rule out any synthons which are unfavorable based on geometric constraints. Based on our CSD search, we found that dimeric synthon **A** is less likely to happen due to geometric constraint of NH(pyrazole)…N(cyclic) hydrogen-bond angle as being non-linear. However, monomeric synthon **A** is still a possibility

2.4.4 Crystal structures

The summary of synthons observed in each crystal structure is presented in Figure 2.11. Synthon (**B+C**) was observed in five out six crystal structures obtained in group **1**. Synthon (**A+F**) was

observed in **P10** and synthon **(C+E)** was observed in **P11**. In total, eight crystal structures were obtained, and synthon **C** was the most common synthon (in six out of eight crystal structures) followed by synthon **B** (five out of eight crystal structures). Synthon **A** is observed in two crystal structures as a dimeric synthon in **P8** and monomeric synthon in **P10**.
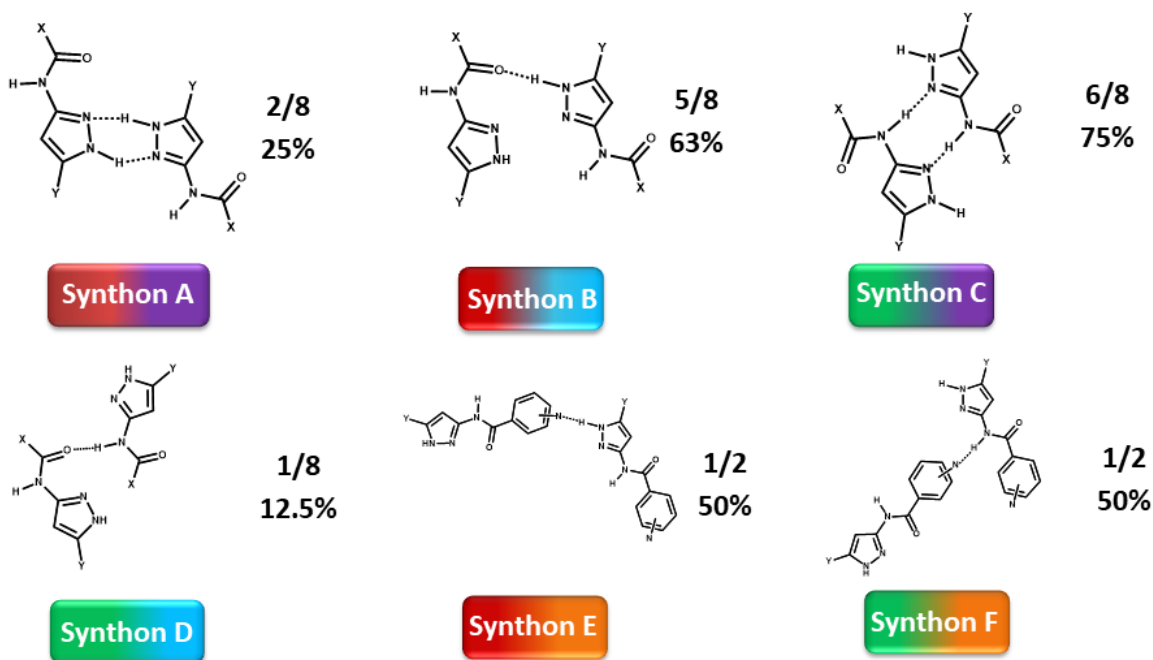
Figure 2.11 Summary of synthons observed in eight crystal structures obtained in this study.

2.4.5 Predicted vs experimental comparison

Group **1**: In the crystal structures of **P1-P4 and P7**, both monomeric synthon **B** (NH(pyrazole)…C=O(amide)) and dimeric synthon **C** (NH(amide)…N(pyrazole)) were observed thus following Etter's rule irrespective of substitution next to amide group or at the pyrazole ring. Therefore, in molecules **P1**-**P6**, five out of six times, synthon **(B+C)** is the dominant synthon as predicted by all three prediction methods. The best donor pyrazole NH binds to the best acceptor carbonyl oxygen (synthon **B**) and the second-best acceptor amide NH binds to the second-best acceptor pyrazole N (synthon **C**) as suggested by MEPs based on Etter's rule. Dimeric synthon **C** is favored over synthon A in these molecules because the hydrogen bond angle in the former is linear compared to latter due to geometric complementarity. The observed synthon in the crystal structure of **P1-P5** was predicted as the most preferred synthon by all four prediction methods; MEPs, HBE, HBP and HBC in the solid state informatic tools. The HBC results obtained by

importing the experimental crystal structure of **P1** shows that the observed experimental motif matches with the predicted hypothetical motif I in this molecule, Figure 2.12.
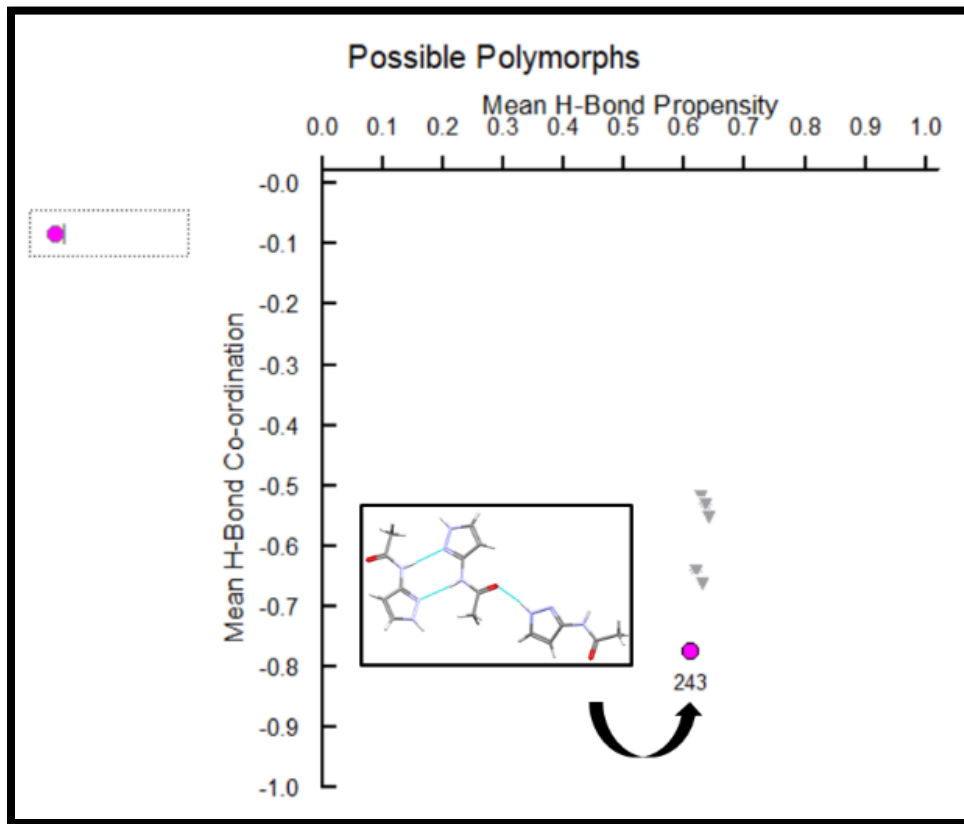
Figure 2.12 Polymorph assessment of **P1** molecule indicating the most optimal hypothetical structure matches with the experimental structure.

In **P8**, the pyrazole NH binds to cyclic N forming dimeric synthon **A** (NH(pyrazole)…N(pyrazole)) and the amide NH binds to carbonyl oxygen forming dimeric synthon **D** (NH(amide)…C=O(amide)). In one out of six crystal structures obtained, synthon **(A+D)** is the dominant synthon regardless of the geometric bias. It is not surprising to see such synthon as both hydrogen-bond energies and hydrogen-bond propensities predicted that either combination synthon (A+D) or (B+C) are possible.

Based on HBC tool, the observed synthon in **P8** doesn't match with the most stable hypothetical structure indicating there is a possibility of polymorphism in this molecule. The most stable structure involves synthon **(B+C)** where NH(amide)…N(pyrazole) and

NH(pyrazole)…C=O(amide) hydrogen-bond interactions are present and it was predicted by all three prediction methods, Figure 2.13.
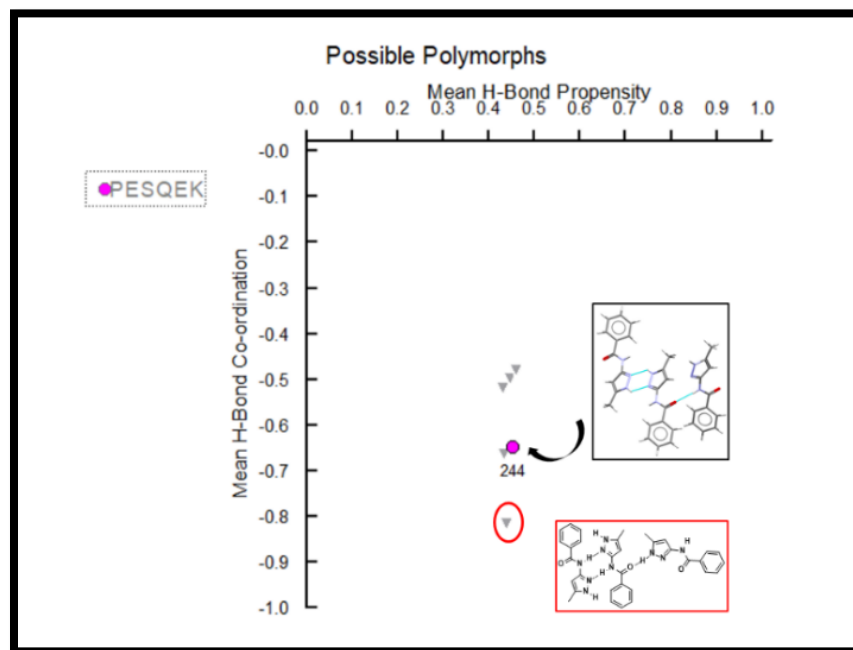
Figure 2.13 Polymorph assessment of P8 molecule indicating the second hypothetical structures matches with the experimental structure.

Group 2: When an extra acceptor group was added to the benzyl group, four new molecules were obtained; **P9-P12.** The crystal structures of **P10** and **P11** are reported in this study.

In the crystal structure of **P10**, the amide NH binds to pyridine N forming synthon **F** (NH(amide)…N(Pyridine) and pyrazole NH binds to cyclic N forming monomeric synthon **A** (NH(pyrazole)…N(pyrazole)). Therefore, synthon **(A+F)** is observed in **P10**. Synthon **(A+F)** was predicted by both HBE and HBP.

In the crystal structure of **P11**, the amide NH binds to cyclic N forming dimeric synthon **C** (NH(amide)…N(pyrazole) and pyrazole NH binds to pyridine N forming synthon **E** (NH(pyrazole)…N(pyridine), therefore synthon **(C+E)** was observed. Synthon **(C+E)** was predicted by hydrogen-bond energy but not by MEPs, HBP or HBC.

The HBC results of **P10** and **P11** shows that the experimentally observed synthon in these molecules matches with the hypothetical predicted motif II and there is chance of synthon polymorphism (Figure 2.14 and Figure 2.15). We haven't performed experimental polymorph screening on these molecules yet.
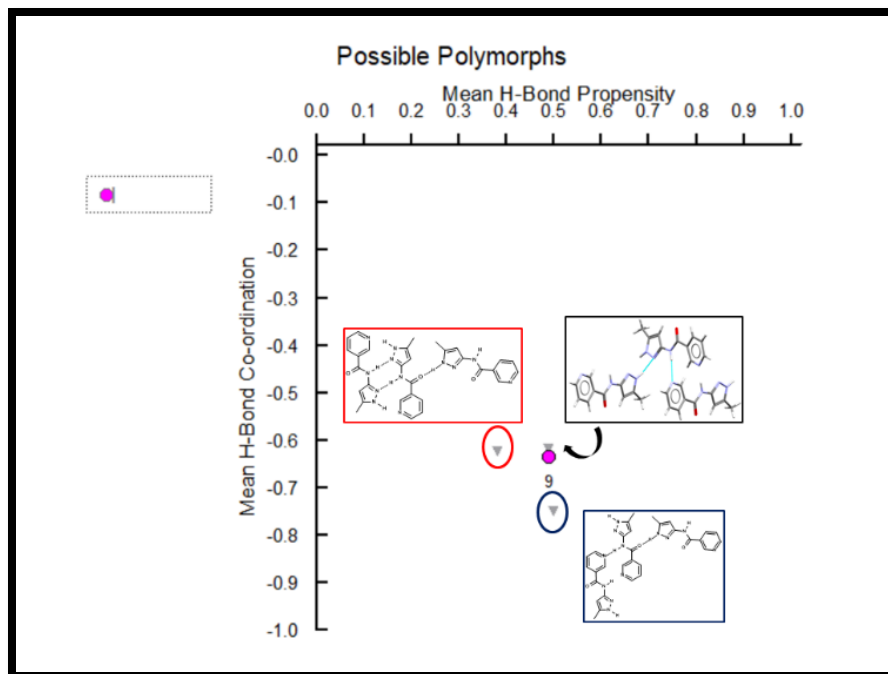
Figure 2.14 Polymorph assessment of P10 molecule indicating the experimental structure and more stable hypothetical structure.
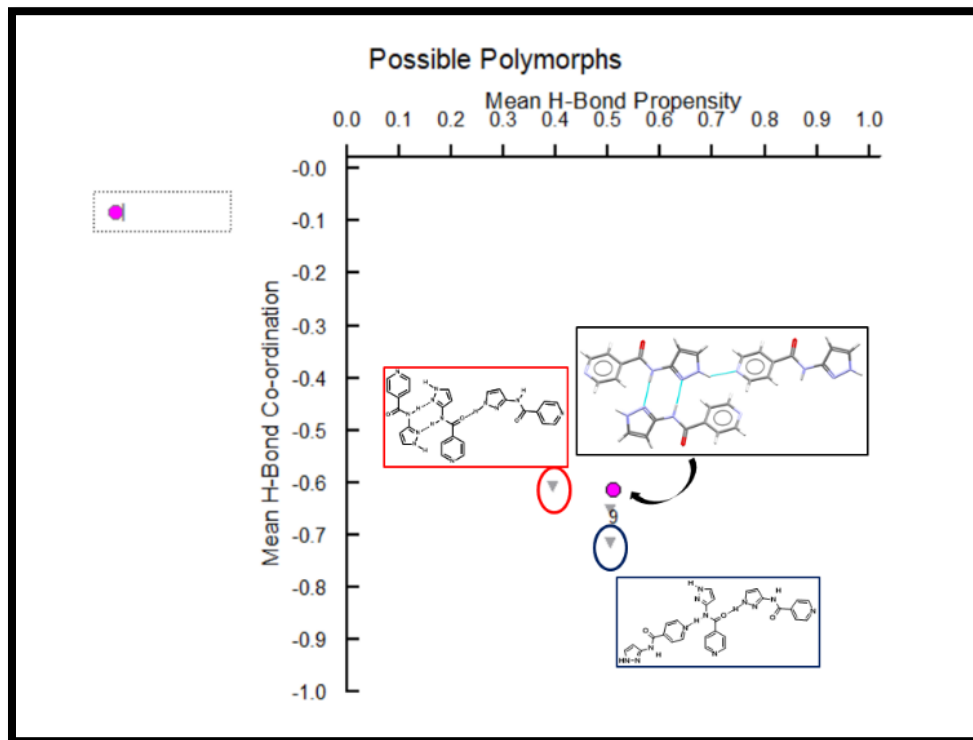
Figure 2.15 Polymorph assessment of **P11** molecule indicating the experimental structure and more stable hypothetical structure.

2.4.6 Validation studies

The comparison studies between the four predicted methods (MEPs, HBE, HBP and HBC) and experimentally observed results are mentioned here and tabulated in Table 2.9. In molecules **P1**-**P8** with two donors and two acceptor molecules, Synthon **(B+C)** as the most likely synthon was predicted by all four methods and in five out of six molecules (83%), synthon **(B+C)** was observed (4 out of 4 methods predicted correctly). In one molecule **P8**, synthon **(A+D)** was observed which was also predicted by HBE and HBP as the second possible synthon (2 out of 4 methods predicted correctly). As the competition was increased by increasing the number of acceptor groups to three in molecules **P9-P12** (two donors and three acceptors), synthon predictability became complex and it became difficult to predict the synthons because of the possibility of synthon crossover and synthon polymorphism. In molecule **P10**, synthon **(A+F)** was observed experimentally which was predicted correctly by HBE and HBP prediction methods (2 out of 4 methods predicted correctly). In molecule **P11**, synthon **(C+E)** was observed which was only predicted by HBE as the most stable synthon (1 out of 4 methods predicted correctly).

Table 2.9 Summary of experimental vs predicted synthons in target molecules **P1-P12**

	Experimental	MEPs	HBE	HBP	HBC
P1	Synthon **(B+C)**	Yes	Yes	Yes	Yes
P2	Synthon **(B+C)**	Yes	Yes	Yes	Yes
P3	Synthon **(B+C)**	Yes	Yes	Yes	Yes
P4	Synthon **(B+C)**	Yes	Yes	Yes	Yes
P7	Synthon **(B+C)**	Yes	Yes	Yes	Yes
P8	Synthon **(A+D)**	No	Yes	Yes	No
P9	Didn't crystallize	N/A			
P10	Synthon **(A+F)**	No	Yes	Yes	No
P11	Synthon **(C+E)**	No	Yes	No	No

2.5 Conclusions

The structural chemistry of twelve target molecules **P1**-**P12** has been analyzed using MEPS, HBE, HBP and HBC as a way of predicting which intermolecular interactions are likely to appear in the solid state and was compared to experimentally observed synthons.
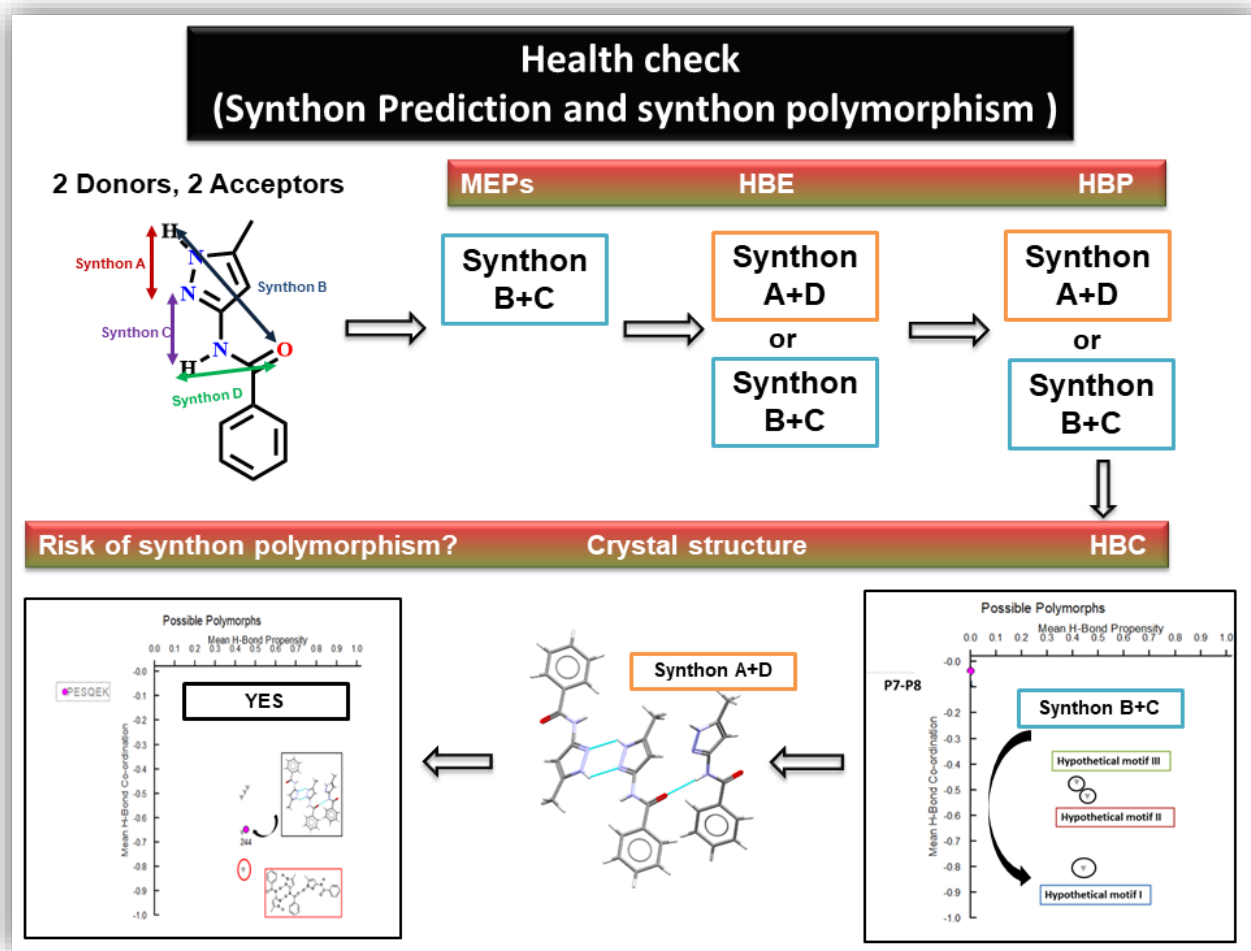
Figure 2.16 A representation of health check on a molecule P8 to determine the synthons and risk of synthon polymorphism.

1. In seven out of eight crystal structures obtained, a combination of HBE and HBP were the most effective prediction methods as they predicted the experimentally observed synthons correctly.

2. These results indicate that combination of methods is preferred over individual molecules.

3. In group 1(**P1-P8**), Synthon (**B+C**) was the most optimal synthon (observed in five out of six crystal structures) compared to synthon (**A+D**) thanks to the geometric complementarity, MEPs and supramolecular chelating effect.[28] However, when a strong acceptor group such as pyridine nitrogen was added as in **P9-P12**, synthon prediction became complex, so does the experimentally observed synthons. Synthon (**A+F**) and synthon (**D+E**) were the most optimal synthon in this category based on combination of atleast two prediction methods. In two crystal structures obtained, synthon (**A+F**) and synthon (**C+E**) were observed.

51

4. HBC was most effective in predicting the synthon polymorphs in this group of molecules. Three target molecules; **P8, P10** and **P11** have the risk of forming synthon polymorphs because the most optimal hypothetical synthon I is not observed in these molecules based on HBC tool.

Four prediction methods used in this study are valuable tool to determine which synthon is likely to form in the crystal structure of a molecule and if the molecule is at a risk of synthon polymorphism. Therefore, a simple health check as shown in Figure 2.16 on these molecules using structural informatics tools such as MEPs, HBE, HBP and HBC for mapping out the structural landscape of these types of molecules will have significant practical applications in various fields.

2.6 References

1.   Price, S. L., *Chemical Society Reviews* **2014,** *43* (7), 2098-2111.

2.   Desiraju, G. R., *Angewandte Chemie International Edition in English* **1995,** *34* (21), 2311-2327.

3.   Fedyanin, I. V.; Karnoukhova, V. A.; Lyssenko, K. A., *CrystEngComm* **2018,** *20* (5), 652-660.

4.   Thalladi, V. R.; Goud, B. S.; Hoy, V. J.; Allen, F. H.; Howard, J. A. K.; Desiraju, G. R., *Chemical Communications* **1996,** (3), 401-402.

5.   Nangia, A.; Desiraju, G. R., Supramolecular Synthons and Pattern Recognition. In *Design of Organic Solids*, Weber, E.; Aoyama, Y.; Caira, M. R.; Desiraju, G. R.; Glusker, J. P.; Hamilton, A. D.; Meléndez, R. E.; Nangia, A., Eds. Springer Berlin Heidelberg: Berlin, Heidelberg, 1998; pp 57-95.

6.   Dey, A.; Pati, N. N.; Desiraju, G. R., *CrystEngComm* **2006,** *8* (10), 751-755.

7.   Sarma, J. A. R. P.; Desiraju, G. R., *Crystal Growth & Design* **2002,** *2* (2), 93-100.

8.   Mukherjee, A.; Desiraju, G. R., *Chemical Communications* **2011,** *47* (14), 4090-4092.

9.   Sreekanth, B. R.; Vishweshwar, P.; Vyas, K., *Chemical Communications* **2007,** (23), 2375-2377.

10.   Pogoda, D.; Janczak, J.; Videnova-Adrabinska, V., *Acta Crystallographica Section B* **2016,** *72* (2), 263-273.

11.   Aakeröy, C. B.; Chopade, P. D.; Desper, J., *Crystal Growth & Design* **2011,** *11* (12), 5333-5336.

12.   Hofmann, D. W. M.; Kuleshova, L. N.; Antipin, M. Y., *Crystal Growth & Design* **2004,** *4* (6), 1395-1402.

13.   Braga, D.; Brammer, L.; Champness, N. R., *CrystEngComm* **2005,** *7* (1), 1-19.

14.   Singh, M. K., *Journal of Crystal Growth* **2014,** *396*, 14-23.

15.   Atahan-Evrenk, S.; Aspuru-Guzik, A., *Prediction and Calculation of Crystal Structures: Methods and Applications*. Springer International Publishing: 2014.

16.   Gharagheizi, F.; Sattari, M.; Tirandazi, B., *Industrial & Engineering Chemistry Research* **2011,** *50* (4), 2482-2486.

17.   Day, G. M., *Crystallography Reviews* **2011,** *17* (1), 3-52.

18.   Etter, M. C., *Accounts of Chemical Research* **1990,** *23* (4), 120-126.

19.   Aakeroy, C. B.; Wijethunga, T. K.; Desper, J., *New Journal of Chemistry* **2015,** *39* (2), 822-828.

20.   Musumeci, D.; Hunter, C. A.; Prohens, R.; Scuderi, S.; McCabe, J. F., *Chemical Science* **2011,** *2* (5), 883-890.

21.   Groom, C. R.; Allen, F. H., *Angewandte Chemie International Edition* **2014,** *53* (3), 662-671.

22.   Naim, M.; Alam, O.; Nawaz, F.; Alam, M.; Alam, P., *Journal of Pharmacy And Bioallied Sciences* **2016,** *8* (1), 2-17.

23.   Wu, J.; Wang, J.; Hu, D.; He, M.; Jin, L.; Song, B., *Chemistry Central Journal* **2012,** *6* (1), 51.

24.   Castro, P. J. F., M. R.; Fernandez, V. E. P.; Gonzalez, D. V. S.; and Mallo-Rubio, A.; Journal, 2012.

25.   Pevarello, P. B., M. G. ; Amici, R.; Orsini, P. ; Traquandi, G.; Corti, L.; Piutti, C.; Sansonna, P.; Villa, M. ; Pierce,B. S.; Pulici, M.; Giordano, P.; Martina, K.; Fritzen,E. L.; Nugent, R. A.; Casale, E.; Cameron, A.; Ciomei, M.; Roletto, F.; Isacchi, A.; Fogliatto, G.; Pesenti, E.; Pastori, W.; Marsiglio, A.; Leach, K. L.; Clare, P. M.; Fiorentini, F.; Varasi, M.; Vulpetti, A.; and Warpehoski,M. A. *Journal of Medicinal Chemistry*, **2004**, 47, 3367-3380.

26.   Yingshen Lu, H.-B. K., Inorg.Chim.Acta, **2004**, 357, 159.

27.   Daidone, G. M., B.; Raimondi, M.V.; Bombieri, G.; Marchini,N.; Artali, R. Heterocycles, **2005**, 65, 2753.

28.   D. Burrows, A.; Menzer, S.; Michael P. Mingos, D.; J. P. White, A.; J. Williams, D., *Journal of the Chemical Society, Dalton Transactions* **1997,** (22), 4237-4240.

Chapter 3 - Systematic Investigation of Knowledge based Prediction methods for Co-crystal Design in Pyrazoles

3.1 Introduction

The occurrence of polymorphism[1-5] in solid-state chemistry is one reason behind the need for robust guidelines and for versatile and dependable practical supramolecular synthesis. A knowledge of connectivity pattern between given functional groups is useful in order to design a desired solid-state material and various prediction methods have been employed to determine the composition, multiple forms, synthons and properties of a chemical compound.
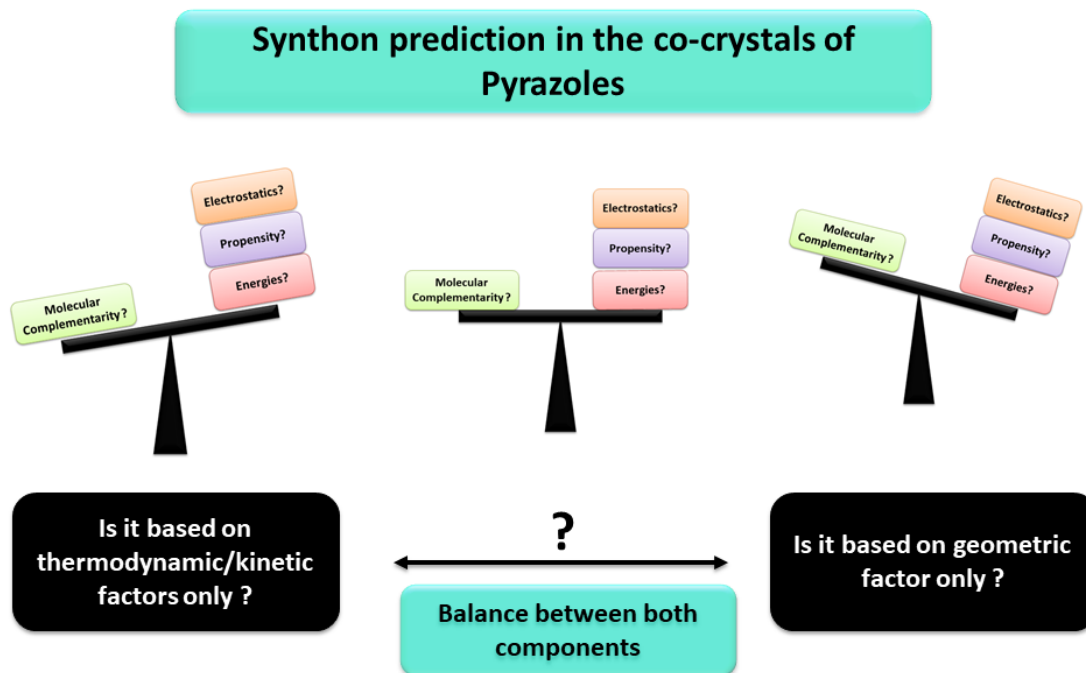
Scheme 3.1. Scheme showing different prediction methods to determine the supramolecular synthon in the co-crystals of pyrazole based molecules.

The preferred connectivity patterns in the solid state can be hypothesized using the Etter guidelines[6] which states that the best hydrogen bond acceptor binds to the best donor and the ranking of best donor-acceptor can be determined using molecular electrostatic potential as

proposed by Aakeroy *et.al.*[7] This connectivity pattern is also employed to predict the outcome of multicomponent systems such as co-crystals using interaction site pairing energies by Hunter *et.al.*[8-9] Recently, Price *et. al.* have used lattice energy,[10] Fabian *et.al.* have used molecular shape and polarity,[11] Galek [12]and jones *et. al.*[13] have used hydrogen-bond propensities as a kinetic tool to quantify and predict the connectivity patterns.

The geometric complementarity of the interacting species comes into play when the presence of multiple hydrogen bond functionalities in a molecule leads to multiple molecular recognition sites in the system.[14,15] The molecular geometry of the co-former, along with the precise orientation of binding sites in the co-crystallizing agent, control the resulting supramolecular synthon in reliable manner in a competitive environment.[16] Lehn et. al. has emphasized molecular recognition as a key component in the orientation/preorganization of binding sites to form synthons.[17] Hamilton et. al. studied the molecular recognition in the solid state for series of amino-pyridine with rigid spacers towards carboxylic acids.[18-19] They also emphasized that the directionality and orientation of two components involved in forming a co-crystal is a must to achieve the high yield supramolecular synthon.[20]

Despite the recent advancement in the field, it is still difficult to use the specific intermolecular interactions to construct architectures with desired dimensionality and high reproducibility due to challenges associated with one-pot supramolecular synthesis, and the presence of multifunctional groups.[21] These problems often lead to "synthon crossover"[18] and "synthon polymorphism" [2, 4, 7] in supramolecular synthesis due to reversible nature of weak hydrogen bonding.[22] However, a good synthetic strategy can be built to obtain high supramolecular yields with desired motifs. These motifs should be dominant within the crystal structure in such a way that different substitutions on the individual covalent building block will form similar lattice structures. This can be achieved by using intermolecular interactions that are *unlikely* to interfere with one another which will lead to strong but also discriminating interactions in the solid state.[23] Aakeröy *et al* have used both hydrogen and halogen bonding to avoid synthon crossover in 2-aminopyrazine molecules based on electrostatics and geometric complementarity. They have also addressed the appearance of synthon crossover in a study done to predict synthon preference in hydroxy-benzoic acids and list of hydrogen bond acceptors.[14]

Two different types of pyrazole based molecules were chosen for this study. Target molecules **P1-P6** has an alkyl next to the amide group, **P7-P8** has a phenyl group and **P9-P12** has a pyridyl nitrogen either at meta or para position (Scheme 3.1). Molecules **P2, P4**, **P6**, **P8, P10, and P12** have a methyl group on the pyrazole ring.
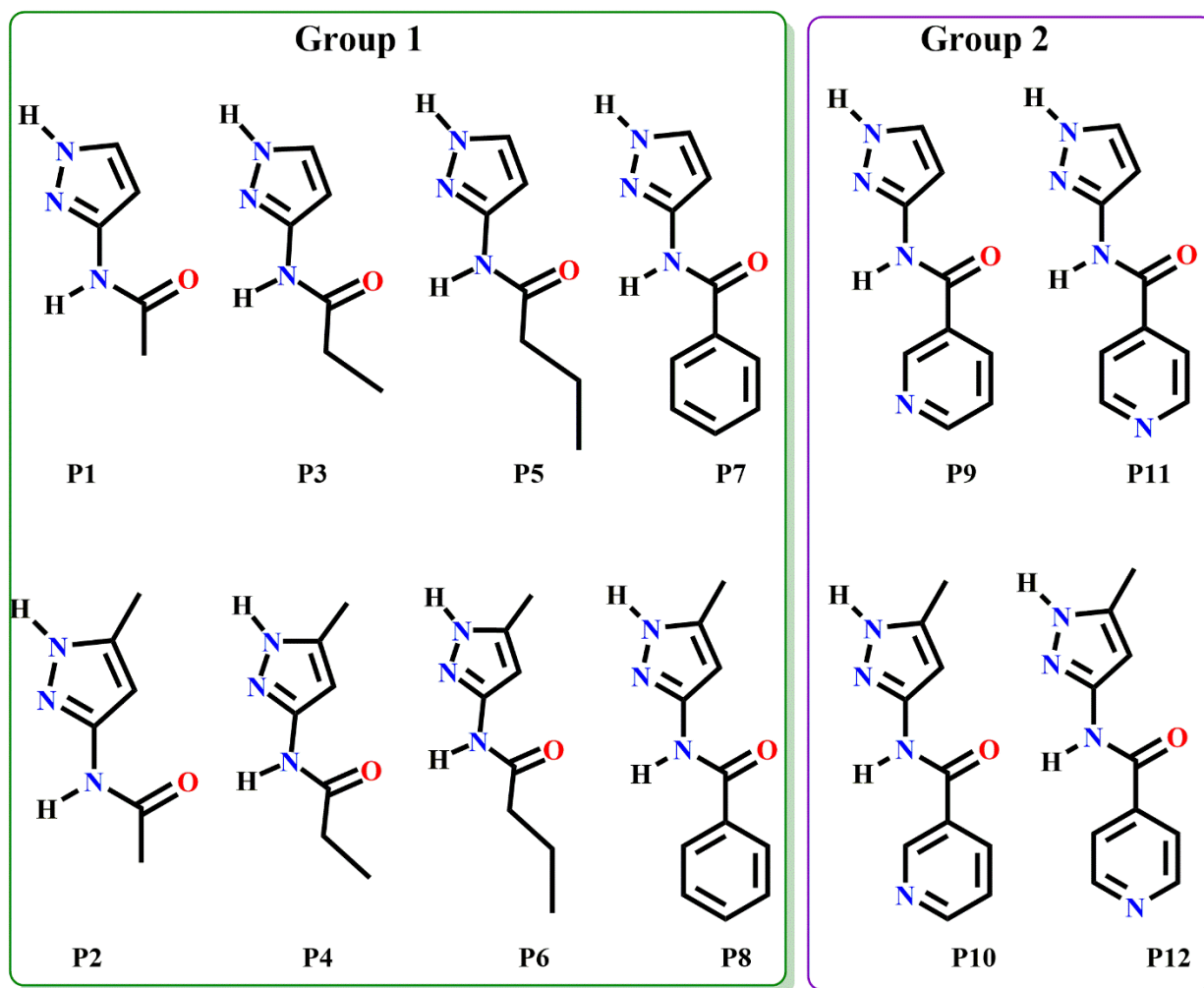
Scheme 3.2 Ten pyrazole based target molecules, **P1-P12** split into two groups; group-1: without pyridyl, and group-2: with pyridyl.

Twenty different carboxylic acids (10 diacids and 10 monoacids) were chosen as potential co-formers (Scheme 3.3).
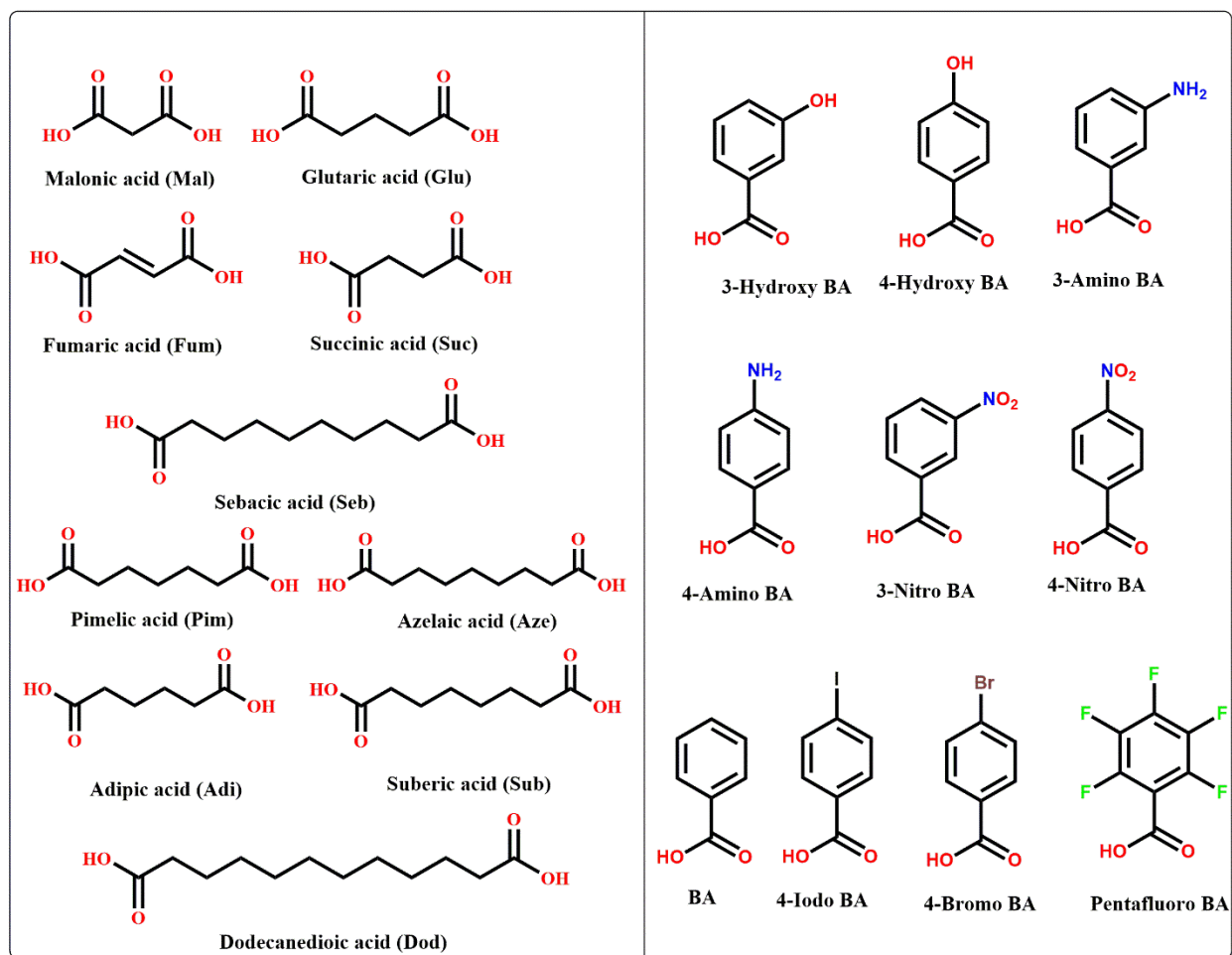
Scheme 3.3 Twenty carboxylic acid co-formers (10 aliphatics and 10 aromatics).

There are two donors (pyrazole NH and amide NH) and two acceptors (pyrazole N and carbonyl O) in **P1-P8**. There are three possibilities for the dicarboxylic acid to bind with **P1-P8**. The first possibility is the OH of acid binds to the pyrazole N via OH(acid)…N(pyrazole) and the carbonyl O=C of acid binds to pyrazole NH via NH(pyrazole)…O=C(acid), synthon **I**. The second possibility is that the OH of acid binds to the pyrazole N via OH(acid)…N(pyrazole) and the carbonyl O=C of acid binds to amide NH via NH(amide)…O=C(acid), synthon **II**. The third possibility is that the OH group of acid binds to O=C of target molecule via OH(acid)…O=C(amide) and the carbonyl group of acid binds to pyrazole NH of the target molecule via NH(pyrazole)…O=C(acid), synthon **III**. For **P9-P12,** an additional synthon is possible because of the presence of an additional acceptor group on each target molecule, the OH group of an acid can bind to pyridine N of target molecule via OH(acid)…N(pyridine), synthon

IV, therefore four different synthons are possible (synthon **I**, synthon **II**, synthon **III** and synthon **IV**) in **P9-P12** (Scheme 3.4). Additional synthons are possible with aromatic acids that contains additional donor group such as hydroxyBA and aminoBA, Scheme 3.7.
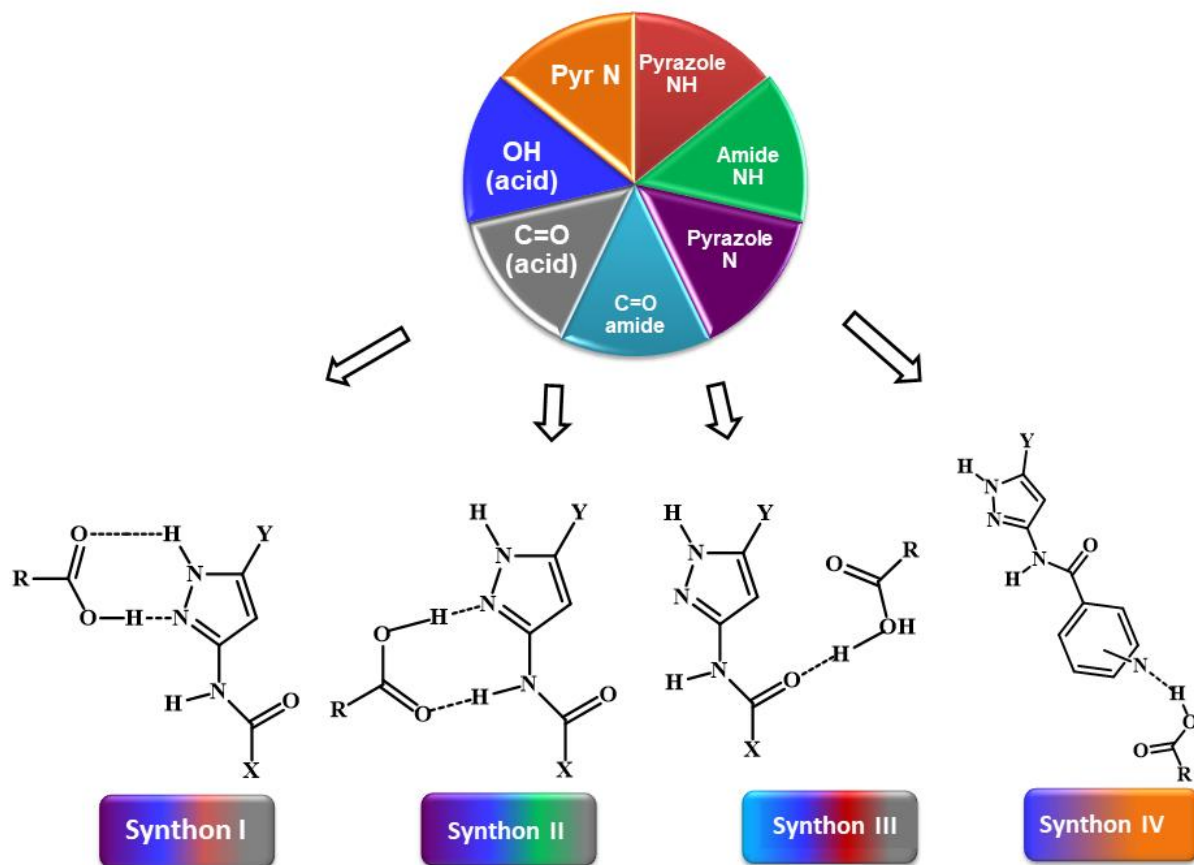
Scheme 3.4. Four postulated heteromeric hydrogen-bond based synthons between **P1-P12** and a carboxylic acid; (R$_1$=H, CH$_3$; R$_2$= methyl, ethyl, or phenyl).

In the work presented herein, first we want to determine if the chosen co-formers can break the homomeric interactions in **P1-P12** and form heteromeric interactions to form co-crystals. Also, what is the effect of changing alkyl side to phenyl ring and pyridyl ring on the resulting co-crystallization outcome? Can we predict the co-crystal outcome by already established prediction methods such as energies and propensities and is a combination of both methods useful in getting predicting the final outcome?
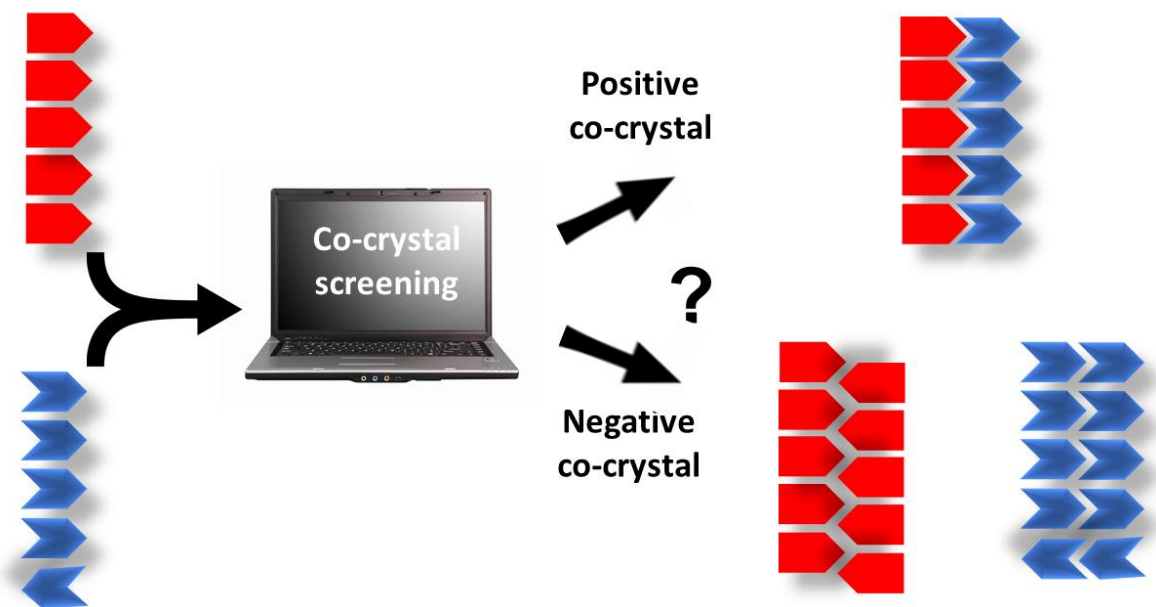
Scheme 3.5 Schematics of systematic investigation of knowledge based predictive methods for co-crystal screen.

Second, we want to determine if we can systematically predict which synthon is more likely to occur in each of these target molecules. We want to examine the competition of two binding pockets (pyrazole NH and amide NH binding pockets) on the pyrazole molecules, **P1-P8** to bind to a series of probe molecules containing carboxylic acid moieties and find out which of the three synthons; synthon **I** (using pyrazole binding pocket) or synthon **II** (using amide binding pocket) or synthon **III** (synthon crossover) is preferred in a competitive environment if a co-crystal is formed? Crystal structures containing two robust $R_2^2(8)$ based hydrogen-bond synthons; carboxylic acid---pyrazole (synthon **I**)[18,19,20] and carboxylic acid---amide (synthon **II**) [6, 4] have shown excellent structural control when they are present independently (Scheme 3.6). The presence of all three functional groups (carboxylic acid, pyrazole and amide) in the same system can lead to competition between carboxylic acid for either pyrazole or amide as well as synthon crossover in **P1-P8**. In fact, a comparatively less common carboxylic acid---pyrazole synthon-based crystal structures have been published even though the angle between N−H bond and the lone pair on the adjacent nitrogen atom is somewhat large to provide a perfect fit with a carboxylic acid.[18,19,20]

(a)	(b)
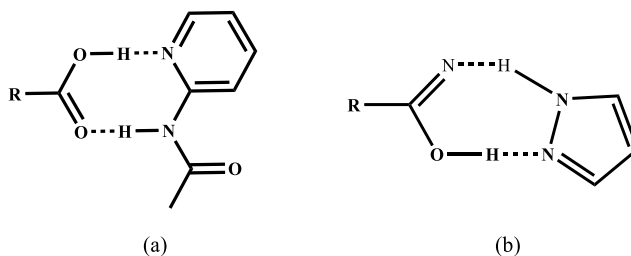
Scheme 3.6 Two robust hydrogen-bonded interactions; type I (carboxylic acid---amide) and type **II** (carboxylic acid---pyrazole) heterosynthons.

There are no systems where functional groups required to form both carboxylic acid---amide and carboxylic acid---pyrazole heterosynthons are combined on the same molecular backbone. This motivated us to do a competitive study by combining two robust supramolecular synthons in the presence of multiple functional groups on the same backbone, thereby eliminating important issues such as solubility and stability of the neutral components. We increased the competition by including a well-known pyridine acceptor in **P9-P12** to establish if the supramolecular chelate effect in hydrogen-bonding dominates the strong single point interaction of OH(acid)…N(pyridine). We also want to determine the role electrostatics, energies, propensities alone, geometric complementarity alone or combination of both on the outcome of co-crystal. The study is done to address following questions:

1. Can we break homomeric synthons observed in **P1-P12** and form heteromeric synthons with carboxylic acids?

2. Can we predict the co-crystal formation of **P1-P12** with 20 carboxylic acids?

3. Which synthon is most favored in heteromeric systems and which method is most reliable for predicting the correct synthon?

4. Is it electrostatics, propensity, energy, molecular complementarity or a combination of all that governs the synthon formation?
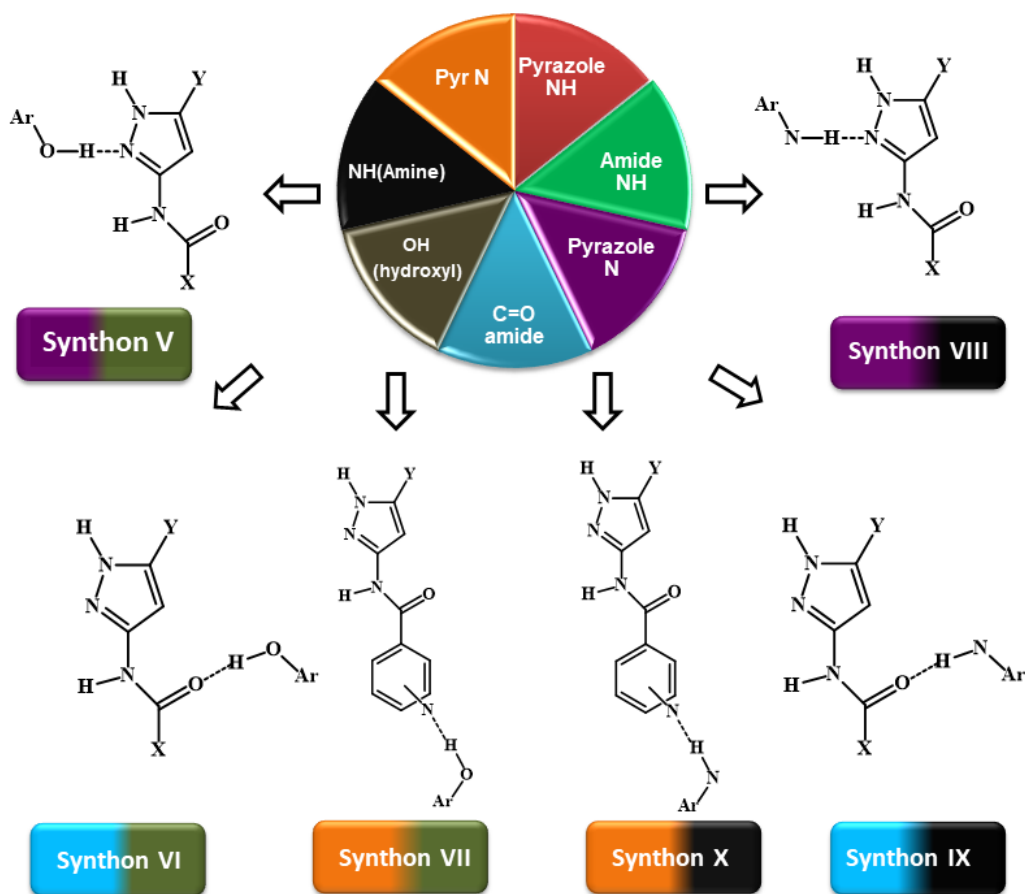
Scheme 3.7 Six postulated heteromeric hydrogen-bond based synthons between **P1-P12** and aromatic acids with an additional donor group (OH for hydroxyl and NH for amino substituents).

3.2 Experimental section

3.2.1 Materials

The synthesis and characterization of **P1-P12** is done in chapter 3. Carboxylic acids were purchased from commercial sources and used as received. Melting points were measured using a Fisher-Johns melting point apparatus. Solution ^{1}H NMR data were collected in DMSO-d$_6$ on a Varian Unity plus 400 MHz NMR spectrometer. IR spectra of co-crystal screening experiments were recorded with a Nicolet 380 FT-IR spectrometer using ATR (attenuated total reflection) technique and ZnSe as the crystal.

3.2.2 Molecular electrostatic potential (MEPs) calculations

See section 2.2.4 for molecular electrostatic potential surfaces (MEPS) of **P1**–**P12**. The electrostatics potentials of 20 carboxylic acids (See appendix E for details) are presented in this study using same methodology described in section 2.2.4.

3.2.3 Hydrogen-bond energies for predictions

See section 2.2.5 for the hydrogen-bond energy calculation methodology. The synthon predictions for heteromeric synthons of **P1-P12** was made by calculating interaction energies of the best donor of the acid with the best acceptor of the target and vice-versa.

The Multi-Component Energy (MCE) score is calculated by subtracting the best energy of homomeric interaction from the best energy of heteromeric interaction. The best hetero/homo interaction is determined based Etter's rule. Note that this calculation is different from Hunter's calculations of virtual co-crystal screening. We have only considered best hetero and homomeric interactions of conventional hydrogen bond donors and acceptors (equation 3.1) whereas in Hunter's calculations, all possible interactions in the homo and heteromeric systems are used for calculations.

$$MCE\ score = \binom{Energy\ of\ the\ best}{heteromeric\ interaction} - \binom{Energy\ of\ the\ best}{homomeric\ interaction} \qquad \text{(Equation 3.1)}$$
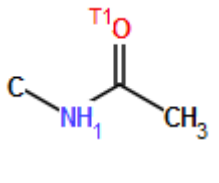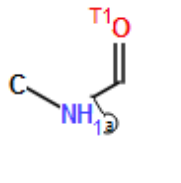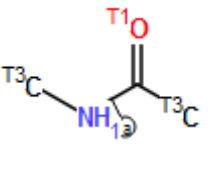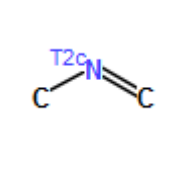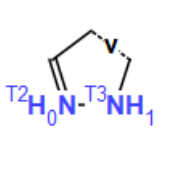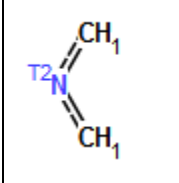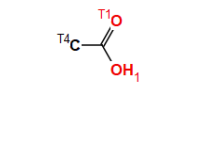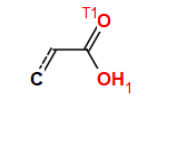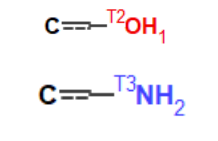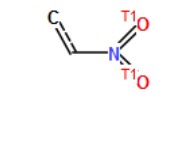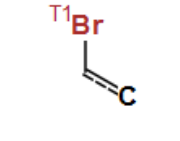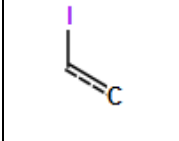
3.2.4 Hydrogen-bond propensities for synthon predictions

The hydrogen-bond propensities for **P1-P12** when screened against 20 different co-formers were calculated and the Multi-Component score (MC score) was obtained by subtracting the propensity of the highest probability homo-interaction (pure components) from the propensity of the highest probability hetero-interaction (co-crystal) obtained from same propensity chart, Equation 3.2. A positive MC score indicates that the hetero-interaction needed for co-crystal formation is favored, whereas a negative MC score indicates that homo-interactions are favored and therefore hydrogen-bond directed co-crystallization is unlikely to occur.

$$MCP\ score = \begin{pmatrix} Propensity\ of\ the \\ highest \\ probability \\ heteromeric\ interaction \end{pmatrix} - \begin{pmatrix} Propensity\ of\ the \\ highest \\ probability \\ homomeric\ interaction \end{pmatrix} \quad \text{(Equation 3.2)}$$

The functional groups of target molecules are listed in chapter 2 and the functional groups used for the acids are listed below in Table 3.1.

Table 3.1 Functional groups used to determine the hydrogen-bond propensities for the six target molecules. The labels in the figures can be explained as follows: T_n = atom makes n bonds, c = atom is cyclic, ⑥ = bond is acyclic, and H_n = n bonded hydrogen atoms.

P1/P2	**P3/P4/P5/P6**	**P7-P12**	**P1-P12**	**P1-P12**	**P9-P12**
Aliphatic	**Aromatic**	**HydroxyBA, AminoBA**	**3NitroBA, 4NitroBA**	**BromoBA**	**IodoBA**

3.2.5 Co-crystal screening and crystal growth

P1-P12 were put through a co-crystal screen using liquid-assisted grinding (LAG) in a drop of methanol against ten aliphatic dicarboxylic acids and ten aromatic acids (Scheme 3.3). In all 240 reactions the reactants were mixed in stoichiometric ratios and the solid resulting from each reaction was characterized using IR spectroscopy to determine if a co-crystal had formed. The IR analysis focused on the positions of readily identifiable vibrational modes (*cf.* O=C (amide), O=C (acid), *etc*) in the pyrazoles and acids, with the corresponding bands in the resulting solids. Shifts greater than seven wavenumbers were taken as an indication of co-crystal formation. In addition, the appearance of broad stretches around 2,300 cm^{-1} and 1,800 cm^{-1} as a result of intermolecular O-H…N hydrogen bonds, also support the formation of co-crystals since such interactions are not feasible in either of the pure compounds, see appendix C. Subsequently, the mixtures obtained from the grinding experiments were dissolved in a minimum amount of methanol or nitromethane

(1–2 ml) and kept in a vial for slow evaporation in order to obtain crystals suitable for single crystal X-ray diffraction. X-ray experimental data and crystallographic data are given in the appendix D.

Table 3.2 Experimental details of new co-crystal obtained in this study.

	Ligands/co-crystal	Ratio (A:D)	Melting point	Morphology
1	P1-PentafluoroBA	1:1	106-107°C	Colorless, blocks
2	P2-Fum	2:1	232-235 °C	Colorless blocks
3	P2-Suc	2:1	228-230 °C	Colorless plates
4	P2-Adi	2:1	164-165 °C	Colorless plates
5	P2-Pim	2:1	198-200 °C	Colorless prisms
6	P3-3nitroBA	1:1	156-159 °C	Colorless blocks
7	P3-4AminoBA	1:1	140-142 °C	Golden blocks
9	P4-Fum	1:1	195-200 °C	Colorless rhombohedra
10	P4-Adi	2:1	245-250 °C	Colorless blocks
11	P4-PentafluoroBA	1:1	104-106 °C	Colorless blocks
12	P7-Sub	2:1	120-122 °C	Colorless prisms
13	P8-Sub	2:1	134-136 °C	Colorless plates
14	P8-Aze	2:1	120-125 °C	Colorless needle

This chapter is divided into two sections:

1. Section 1 is focused on the prediction and experimental study of co-crystal screening of **P1-P12** with 20 co-formers.

2. Section 2 is focused on the prediction and experimental analysis of synthons in the co-crystals of **P1-P12** with 20 co-formers.

3.3 Results

3.3.1 Part 1: Co-crystal screening

The study is divided into two groups: group **1** contains molecules with 2 donors and 2 acceptors; **P1-P8** co-crystals (160 experiments). Group 2 contains molecules with 2 donors and 3 acceptors; **P9-P12** co-crystals (80 experiments).

3.3.1.1 Experimental co-crystal screening

The experimental screen produced co-crystals in 169 of the 240 experiments between **P1-P12** and aliphatic diacids and aromatic monoacids. The results of the IR grinding screen are shown in Table 3.3.

Table 3.3 Attempted co-crystallizations using LAG (methanol) of P1-P12 with aliphatic and aromatic acids.

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	S.Y.
Suc	red	green	green	green	red	red	green	green	green	green	green	green	75%
Adi	red	green	green	green	green	red	green	red	green	green	green	green	75%
Sub	red	green	green	green	green	green	green	green	green	green	green	green	92%
Seb	red	green	green	green	red	green	green	green	green	green	green	green	83%
Dod	red	green	green	red	red	red	green	green	green	green	green	green	67%
Fum	red	green	green	green	green	red	green	green	green	green	green	green	83%
Mal	red	green	green	green	green	green	green	green	green	green	green	green	92%
Glu	red	green	green	green	green	green	green	green	green	green	green	green	92%
Pim	red	green	green	green	green	red	green	red	green	green	green	green	75%
Aze	red	green	green	green	green	red	green	green	green	green	green	green	83%
	0/10	10/10	10/10	9/10	7/10	4/10	10/10	8/10	10/10	10/10	10/10	10/10	
3HydroxyBA	green	red	green	red	green	green	green	green	green	green	red	green	50%
4HydroxyBA	green	green	green	red	red	red	green	green	green	green	red	red	50%
3AminoBA	green	green	green	green	green	green	green	green	green	green	green	green	83%
4AminoBA	green	green	green	green	green	green	green	red	green	green	green	red	42%
3NitroBA	green	green	green	green	green	green	green	green	green	green	green	green	100%
4NitroBA	red	green	red	green	red	green	green	green	red	green	green	red	42%
BA	red	green	green	green	green	green	green	green	green	green	red	green	83%
4IodoBA	red	red	red	red	red	red	red	red	red	red	green	red	17%
4BromoBA	red	red	red	red	red	red	red	red	green	green	red	green	25%
PentaFBA	green	green	green	green	green	green	green	green	green	green	green	green	100%
	6/10	7/10	7/10	4/10	5/10	4/10	6/10	6/10	7/10	8/10	5/10	6/10	
	6/20	17/20	17/20	13/20	12/20	8/20	16/20	14/20	17/20	18/20	15/20	16/20	**169/240**
	30%	85%	85%	65%	60%	40%	80%	70%	85%	90%	75%	80%	**70%**

3.3.1.2 Energies as prediction method for co-crystal screening

The best hetero-meric interaction in **P1-P8** molecules is OH(acid)…C=O (amide) and in **P9-P12** is OH(acid)…. N(pyridine). The best homomeric interaction in **P1-P8** is NH(pyrazole)…C=O(amide) and in **P9-P12** is NH(pyrazole)…N(pyridine). The results of MCE are presented in Appendix G.

3.3.1.3 Propensities as prediction method for co-crystal screening

Hydrogen-bond propensity calculations were performed on **P1-P12** against 20 co-formers. The two highest propensity donor-acceptor interactions were taken into account for homomeric systems in the target molecules: NH(pyrazole)….C=O(amide) for group **P1-P12** and N-H(pyrazole)···N(pyridine) for group **P9-P12**. The two highest propensity donor-acceptor interactions were taken into account for hetero-meric systems: O-H(acid)···O=C(amide) for group **P1-P12** and OH(acid)…N(pyridine) for group **P9-P12**. An MCP score (as described in section 4.2.4) is used to predict whether a reaction (in this case, co-crystal formation) will or will not take place when two reactants are combined. A summary of the multi-component hydrogen-bond propensity (MCP) screening results for **P1-P12** with 20 co-formers is given in the Appendix G.

3.3.2 Part 2: Synthon Prediction study

3.3.2.1 Method 1: Electrostatics for synthon prediction in co-crystals

In order to determine which synthon is more likely to form in the co-crystals, prediction studies were performed using electrostatics best donor binding to best acceptor rule. The different synthons and the ranking of each donor and acceptor are shown in the Figure 3.1.

	Rank I' HO—R O Rank I'	Rank II' HO OH/NH₂ Rank I' O Rank I'
	Rank I, H, Y, N, Rank II, N, Synthon I, Synthon II, H—N, Rank II, O, Rank I, Synthon III, X, X= CH₃ , C₂H₅, C₆H₅, Y= H, CH₃	Rank I, H, Y, N, Rank III, N, Synthon I, Synthon II, H, N, Rank II, O, Synthon III, Rank II, N, Rank I, Synthon IV, Y= H, CH₃
Single point interaction	Synthon **III**	Synthon **IV**
Two-point interaction	Synthon **I**	Synthon **I**

Figure 3.1 Synthon prediction using electrostatics method for group 1 and group 2.

3.3.2.2 Method 2: Energies for synthon prediction in co-crystals

Hydrogen-bond energies of each synthon was calculated in each group and the average energy for each synthon is shown in Table 3.4-3.6.

Table 3.4 Hydrogen-bond energies for each synthon in each respective group (P1-P12 against 10 aliphatic acids and 6 aromatic acids).

P1-P12 against 20 acids					
		Synthon I	Synthon II	Synthon III	Synthon IV
GROUP 1	Aliphatic acids (10)	37.21±0.57	32.54±1.14	28.36±1.75	N/A
	Aromatic acids (6)	32.10±0.83	29.15±1.18	31.34±0.94	N/A
GROUP 2	Aliphatic acids (10)	34.32±0.99	29.88±0.54	22.64±0.56	26.90±1.74
	Aromatic acids (6)	29.02±1.10	26.23±0.91	24.96±1.23	29.26±1.89

Table 3.5 Hydrogen bond energies for synthons I, II, III and IV in each respective group (P1-P12 against aromatic acids with an additional donor group).

	Synthon I	Synthon II	Synthon III	Synthon IV
GROUP 1	29.68±0.60	26.52±1.00	26.30±0.79	N/A
GROUP 2	27.03±0.84	24.37±0.96	20.94±1.03	24.55±1.59

Table 3.6 Hydrogen bond energies for synthon V, VII, VIII and X in each respective group (P1-P12 against aromatic acids with an additional donor group).

		Synthon V/VIII	Synthon VI/IX	Synthon VII/X
GROUP 1	OH	20.64±1.25	31.75±0.96	N/A
	NH_2	13.35±0.81	20.53±0.62	
GROUP 2	OH	17.63±1.44	25.28±1.24	25.72±3.19
	NH_2	11.41±0.93	16.35±0.80	19.16±1.24

3.3.2.3 Method 3: Propensities for synthon prediction in co-crystals

Hydrogen-bond propensities were calculated for each group for different co-formers by looking at the heteromeric interactions in the propensity chart. The results of the propensity predictions are shown in the Table 3.7.

Table 3.7 Synthon prediction using propensity method for group 1 and group 2.

	Group 1(P1-P8)	Group 2 (P9-P12)
Aliphatic acid (10)	Synthon **I** (NH(pyrazole)….C=O(acid))	Synthon **IV**
Aromatic acids (w/o additional donor)	Synthon **I** (NH(pyrazole)….C=O(acid))	Synthon **IV**

HydroxyBA	Synthon **III**	Synthon **IV**
AminoBA	Synthon **IX**	Synthon **X**

3.3.2.4 Method 4: Molecular complementarity for synthon prediction

The molecular complementarity approach was used to determine which synthon is more likely to occur between synthon **I** and synthon **II** using CSD. The search was done by using fragments indicated in Figure 3.2 (a) and (c), and the results of each synthon are shown in Figure 3.2 (b) and (d).

(a) Synthon I   **(b) Synthon I results**

(c) Synthon II   **(d) Synthon II results**
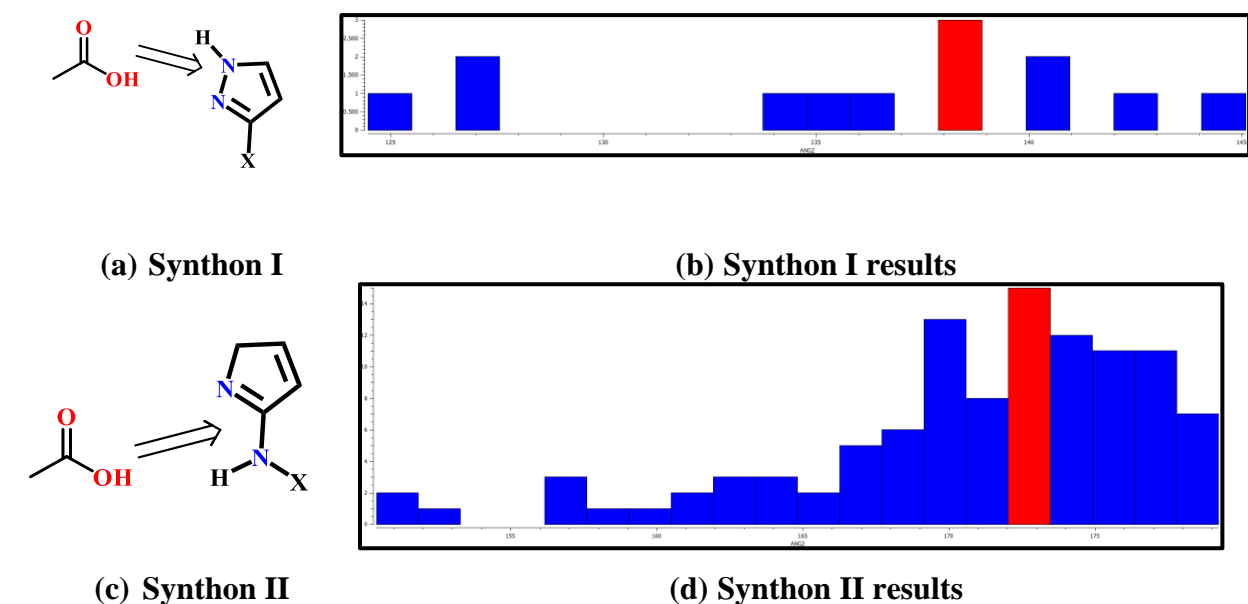
Figure 3.2 (a) Synthon I fragment used for search, (b) Bond angles distribution of NH(pyrazole)…O=C(acid) interaction in the CSD database, (c) Synthon II fragment used for search and (d)Bond angles distribution of NH(pyrazole)…O=C(acid) interaction in the CSD.

3.3.3 Experimentally observed synthons

The crystal structures are divided into three groups. Thirteen crystal structures (**P1-PentaFBA, P2-Fum, P2-Suc, P2-Adi, P2-Pim, P3-3nitroBA, P3-4aminoBA, P3-PentaFBA, P4-Fum, P4-Adi, P7-Sub, P8-Aze, and P8-Sub**) were obtained for group 1, and three crystal structures were obtained for group 2 (**P10-Fum, P10-Mal, P11-Aze and (d**) **P11-dod.**

(a) P2-Fum

Synthon II

(b) P2-Suc

Synthon II

(c) P2-Adi

Synthon II

(d) P2-Pim

Synthon II

(e) P4-Fum

Synthon II

(f) P7-Sub

Synthon II

(g) P8-Aze
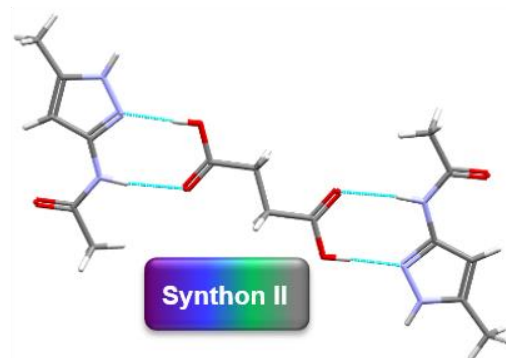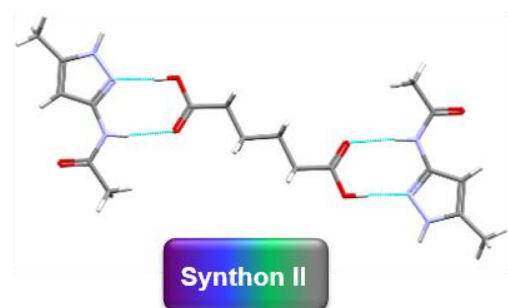
Figure 3.3 Supramolecular trimer formed via $R_2^2(8)$ heterosynthons in (a) P2-Fum, (b) P2-Suc, (c) P2-Adi, (d) P2-Pim, (e) P4-Fum, (f) P7-Sub and (g) P8-Aze leading in 2:1 stoichiometry

(a) P4-Adi
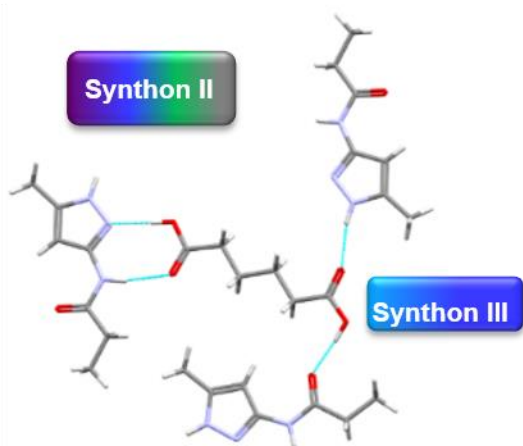
(b) P8-Sub
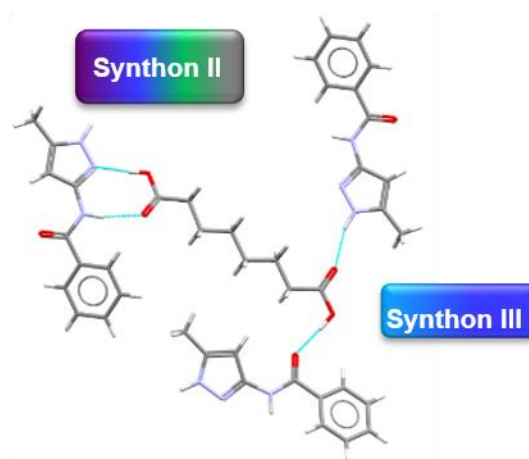
Figure 3.4 Synthon crossover in (a) P4-Adi and (b) P8-Sub leading to 1:1 stoichiometry.

(a) P1-PentaFBA

Synthon II

(b) P3-3nitroBA

Synthon II

(c) P3-4-aminoBA

Synthon II

(d) P3-PentaFBA

Synthon II
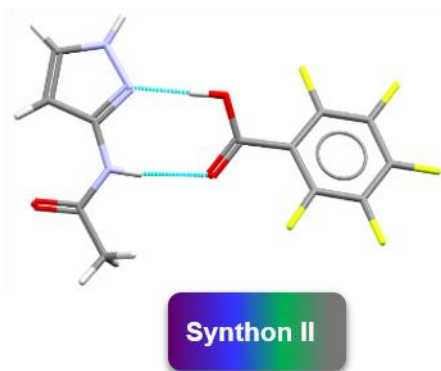
Figure 3.5 Supramolecular dimer in (a) P1-PentaFBA, (b) P3-3nitroBA, (c) P3-4-aminoBA and (d) P3-PentaFBA.

Four crystal structures were obtained in group 2 (**P10-Fum**, **P10-Mal**, **P11-Aze** and **P11-dod**) with di-aliphatic acids. No co-crystals were obtained with aromatic acids with and without additional donor group.
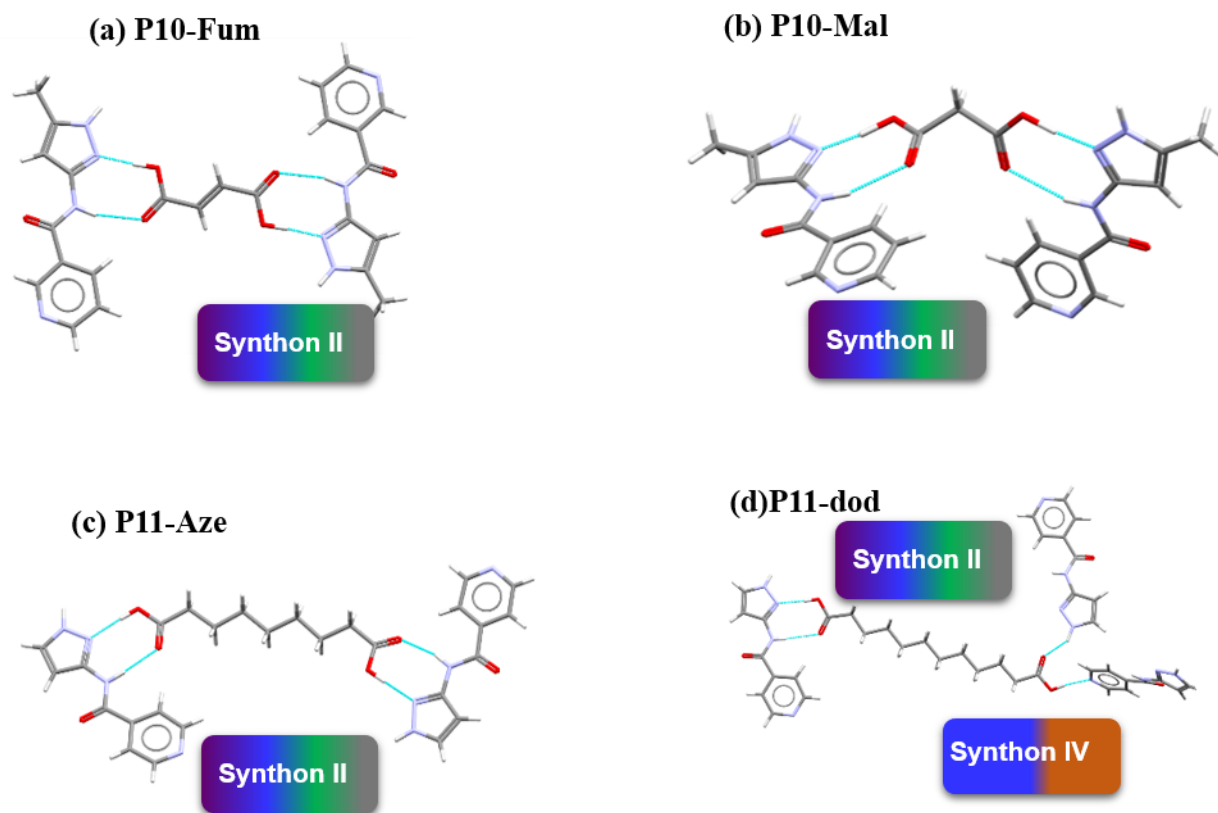
Figure 3.6 Supramolecular trimer in (a) P10-Fum, (b) P10-Mal and (c) P11-Aze in 2:1 stoichiometry. Synthon crossover in (d) P11-dod leading to 1:1 stoichiometry.

3.4 Discussion

3.4.1 Experimental co-crystal screening

3.4.1.1 Group 1(co-crystallization of P9-P12 with 20 carboxylic acids)

Target molecules **P1-P8** formed co-crystals in 58 out of 80 (73% supramolecular yield) experiments with aliphatic acids and in 45 out of 80 experiments (56% supramolecular yield) with aromatic acids. Overall, all target molecules except **P1**(with 30% yields) and **P6** (with 40% yields) has good success rate based on the grinding experiments with 20 co-formers (64% supramolecular yield). Adding a methyl group on **P1** to form **P6** increases the supramolecular yield from 30% to 85%, whereas adding a methyl group on **P5** to form **P6** decreases the supramolecular yield from 60% to 40%. Target molecules **P2**, **P3** and **P7** have the highest supramolecular yields in this group. The target molecules with an aliphatic side chain group (**P1-P6**) gave 73 out of 120 experimental

positive co-crystal outcomes indicating 61% success rate whereas the target molecules with an aromatic side group (**P7**-**P8**) gave 30 out of 40 positive co-crystal outcomes indicating 75% success rate. This indicates that aromatic group has higher probability of forming co-crystals compared to alkyl chains although the comparison might not be fair as the dataset of alkyl group molecules is larger compared to aromatic groups.

3.4.1.2 Group 2(co-crystallization of P9-P12 with 20 carboxylic acids)

Group **2**: **P9-P12** formed co-crystals in 40 out of 40 (100% supramolecular yield) experiments with aliphatic acids and in 26 out of 40 experiments (65% supramolecular yield) with aromatic acids. Overall, all target molecules, **P9-P12** have good success rate based on the grinding experiments with 20 co-formers (83% supramolecular yield), Figure 4.1. Adding a methyl group on **P9** and **P11** increases the supramolecular yield by 5% in **P10** and **P12**.

Overall, the experimental screen produced co-crystals in 169 of the 240 experiments (70% supramolecular yield) between **P1-P12** and 20 dicarboxylic acids, Table 4.2 (from IR). It is a rather high success rate considering the very limited scope of the experimental screen (only LAG from methanol was used to screen for co-crystals and only one solvent (methanol or nitromethane) was used to grow single crystals). With further experimentation, it is reasonable to assume that some pairs that did not succeed in co-crystallizing in this screen could eventually form. Still, it is necessary to employ consistent experimental conditions when trying to elucidate the success or suitability of a protocol for co-crystal prediction.
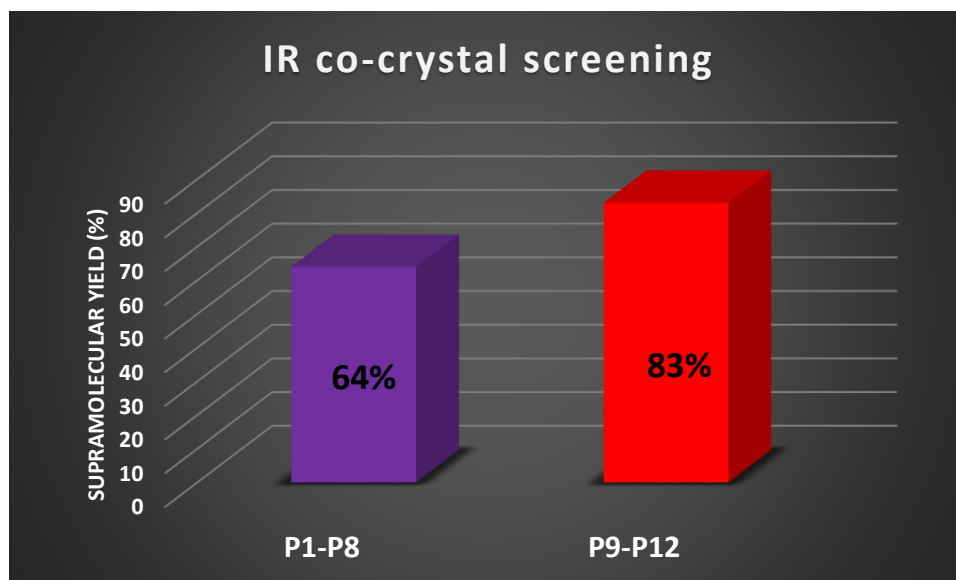
(a)

Figure 3.7 Graphs showing supramolecular yields with respect to each group.

3.4.2 Experimental co-crystal screening vs energy prediction

The multicomponent score for hydrogen-bond energies (MCE) was obtained by subtracting the best homo-meric interactions from the best hetero-meric interactions. The comparison between predicted and experimental co-crystal screen was made, Appendix G.

In **group 1**, 160 experiments were performed, and the energy prediction was compared with the experimental co-crystal screening. 92 out of 160 co-crystal screens were true positive, 51 were false positive, 11 were false negative and 6 were true negative. In total, there was an 61% agreement (98 out of 160 combinations were either true positive or true negative) between the predicted and experimental results for **P1-P8** with 20 acids, Figure 3.8a. A low success rate is because 51 out of 160 combinations were predicted as false positives, indicating that the energy method predicted them to be positive co-crystal but experimentally, they were designated as negative co-crystal outcome. Therefore, an energy method could not predict which combinations does not form co-crystals in group 1.

In **group 2**, 64 out of 80 experiments were true positive, 10 were false positive, 2 were false negatives and 4 were true negatives (Figure 3.8b). In total, there was an 85% agreement (68 out of

75

80 combinations were either true positive or true negative) between the predicted and experimental results for **P9-P12** with 20 acids. The energy method predicted the experimentally co-crystal outcome very well in group 2.

MCE(P1-P8)		Predicted outcome	
		Co-crystal	No co-crystal
Experiment	Co-crystal	92	11
	No co-crystal	51	6

98/160, 61%

(a) Group 1(**P1-P8**)

MCE(P9-P12)		Predicted outcome	
		Co-crystal	No co-crystal
Experiment	Co-crystal	64	2
	No co-crystal	10	4

68/80, 85%

(b) Group 2 (**P9-P12**)

Figure 3.8 Confusion matrices determined from multi-component energies, MCE results (cut off =-0.10) for P1-P12 molecules with co-formers, (a) Group 1(P1-P8) and (b) Group 2 (P9-P12).

3.4.3 Experimental co-crystal screening vs propensity prediction

The predictions made using the MCP model were compared to experimental co-crystal screening results, Appendix G. A confusion matrix was used to analyze the data as reported by Wood *et. al.*[24]

The confusion matrices determined from MCP results using cut off =0.00 and -0.10 for **P1-P12** are shown in Figure 3.9-3.10. In total, the number of false negatives reduced from 142 to 66 by changing the cut off from 0.0 to -0.10. The agreement between experimental and predicted also increased from 35% (89 out of 240 combinations) to 57% (136 out of 240 combinations).

MCE (P1-P8) Cut off = 0.0		Predicted outcome		MCE (P9-P12) Cut off = 0.0		Predicted outcome	
		Co-crystal	No co-crystal			Co-crystal	No co-crystal
Experiment	Co-crystal	24	79	Experiment	Co-crystal	3	63
	No co-crystal	10	47		No co-crystal	5	9
71/160, 44%				12/80, 15%			
(a) Group 1(**P1-P8**)				(b) Group 2 (**P9-P12**)			

Figure 3.9 Confusion matrices determined from multi-component propensity MCP results (cut off =0.0) for P1-P12 molecules with co-formers, (a) Group 1(P1-P8) and (b) Group 2 (P9-P12).

MCP(P1-P8) Cut off = -0.10		Predicted outcome		MCP(P9-P12) Cut off = -0.10		Predicted outcome	
		Co-crystal	No co-crystal			Co-crystal	No co-crystal
Experiment	Co-crystal	64	39	Experiment	Co-crystal	39	27
	No co-crystal	33	24		No co-crystal	5	9
88/160, 55%				48/80, 60%			
(a) Group 1(**P1-P8**)				(b) Group 2 (**P9-P12**)			

Figure 3.10 Confusion matrices determined from multi-component propensity MCP results (cut off = -0.10) for P1-P12 molecules with co-formers, (a) Group 1(P1-P8) and (b) Group 2 (P9-P12).

The prediction of co-crystal based on propensity in pyrazole-amide based molecules is difficult because of the presence of multiple donors and acceptors on the pyrazole. At-least two different interactions are possible in each co-crystal combination due to presence of two donor groups on the pyrazole molecules: both could be heteromeric interactions or one could be homomeric and other could be heteromeric interaction. The propensity MCP score is based on only best hetero and homo interaction observed. For examples, in group 1, The best homomeric interaction; NH(pyrazole)…C=O (amide) is a stronger interaction compared to the best heteromeric interaction; NH(pyrazole)…C=O(acid). Therefore, the outcome based on MCP is negative co-crystal. However, experimentally it is a positive co-crystal and the best homomeric interaction is observed in the co-crystal. This indicated that we must consider all the possible homomeric and heteromeric interactions in the molecule. Similar is true for group 2 in which the best homomeric interaction NH(pyrazole)…N(pyridine) or NH(pyrazole)….C=O(amide) is observed in the experimentally observed crystal structures.

3.4.4 Comparison between predicted and experimental co-crystal screening

Two prediction methods were used to analyze the co-crystal grinding experiment results in this study, Table 3.8. The data was divided in two groups based on the number of functional groups present on the target molecules.

Based on the MCE method, there was 61% agreement between prediction and experimental co-crystal outcome for group 1 and 85% for group 2. The method did well on group 2 and didn't do

that well on group 1. The possible reason for low success rate in group 1 could be due to the method not able to predict the experimentally observed negative co-crystal outcome. The best homomeric interaction is NH(pyr)...C=O(amide) and the best heteromeric interaction is OH(acid)…C=O(amide). The heteromeric interaction was stronger than homomeric interaction in most cases so gave positive MCE score. The method gave 51 false positives which means they were predicted as positive co-crystal outcome but experimentally formed negative co-crystal. Propensity method was used as an alternative to determine the MCP score which is determined by subtracting the best homomeric interaction from the best heteromeric interaction based on knowledge of Cambridge structural database. Based on MCP as 0.0 cut off, there was 44% and 15 % agreement between predicted and experimental for group **1** and group 2 respectively. This poor agreement is because the propensity predicted homo-meric NH(pyr)…C=O(amide) interaction in group **1** and NH(pyr)…N(pyr) in group 2 as the best interactions. Therefore, the MCP score was negative and 142 out of 240 co-crystals (59%) were predicted as false negative means they were predicted negative by propensity but they do form co-crystal experimentally. Since pyrazole target molecules has multiple functional groups including two donor groups, it is likely that more than one interaction happens in the co-crystal and the propensity calculation method need to be modified to include all possible interaction. Based on MCP cut off value of -0.10, the agreement was improved to 55% for group 1 and 60% for group 2.

Table 3.8 Summary of experimental co-crystal screening and energy and propensity prediction comparison for group 1 and group 2.

	Experimental co-crystal screening	**Energies (MCE)**	**Propensities (MCP= 0.0)**	**Propensities (MCP= -0.10)**
Group 1	64%	98/160, 61%	71/160, 44%	88/160, 55%
Group 2	83%	68/80, 85%	12/80, 15%	48/80, 60%

PART 2: SYNTHON PREDICION STUDY

Once, the co-crystal screen was done, synthon prediction studies were performed. There are two binding pockets in the pyrazole molecules, **P1-P8**. The acid functionality of carboxylic group (10 aliphatic acids and 10 aromatic acids w/o additional donor) can either bind to pyrazole binding pocket forming synthon **I**, or amide binding pocket forming synthon **II** or it can undergo synthon crossover leading to synthon **III**.
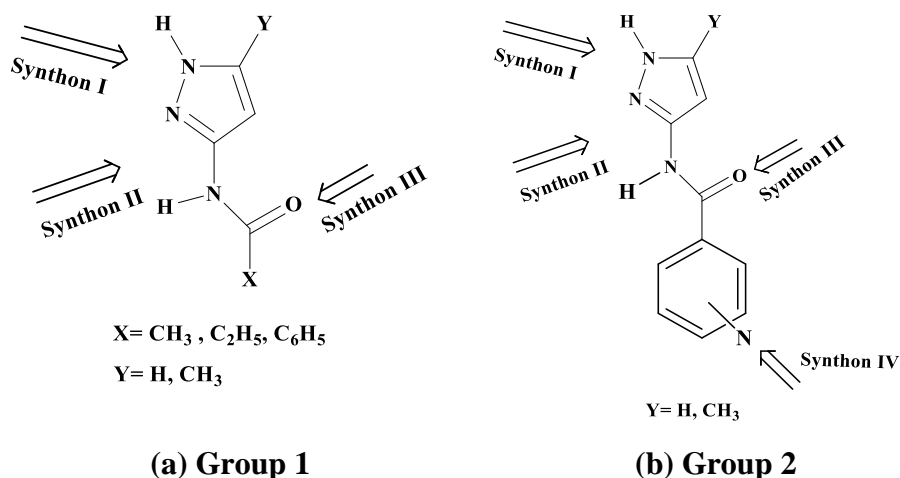
Scheme 3.8 Schematics showing different synthon possibilities.

In **P9-P12**, the additional pyridyl acceptor group can lead to synthon **IV**. Therefore, in **P1-P8**, synthon **I**, **II**, and **III** are possible and in **P9-P12**, synthon **I**, **II**, **III**, and **IV** are possible. The synthon complexity increases as an additional donor group is introduced on the aromatic acids (an additional synthon **V**, **VII** and V**III** for hydroxyBA and an additional synthon **VIII**, **IX** and **X** for aminoBA). To find out which of these synthons is more likely to happen in the crystal structure, predictions were made using electrostatics, energies and propensity method.

3.4.5 Synthon prediction in group 1

3.4.5.1 Method 1: Electrostatics

Synthon **I** is predicted to be the most optimal synthon as the NH(pyrazole) is the best donor and N (pyrazole) is the second best acceptor compared to synthon **II** in which NH(amide) and N(pyrazole) will form synthon in aliphatic acids and aromatic acids without additional donor. Synthon **VI** is predicted to be the most likely synthon in hydroxyBA and synthon **I** is the best predicted synthon in aminoBA, Figure 3.11.

3.4.5.2 Method 2: Energies

In group **1**, for aliphatic acids, synthon **I** was the most preferred synthon followed by synthon **II**. For aromatic acids without an additional donor group, **I** was the most preferred synthon followed

by synthon **III** and synthon **II**. In hydroxyBA, one-point synthon **VI** and two-point synthon **I** are most likely to happen. In aminoBA, synthon **I** is most likely to occur, Figure 3.11.
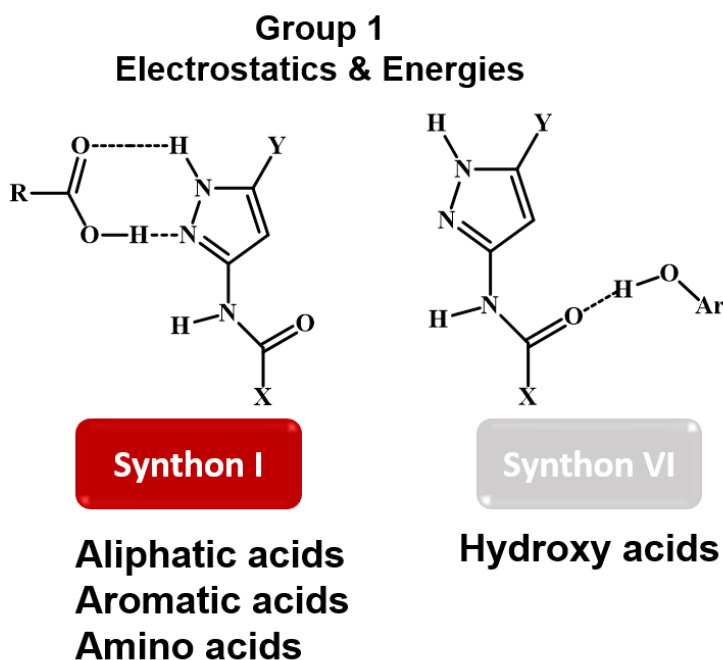
Figure 3.11 Synthon predicted in group 1 using electrostatics and energies.

3.4.5.3 Method 3: Propensities

The best heteromeric interaction was selected from the propensity chart and the synthon based on that interaction was chosen to be the most preferred synthon based on propensity. The preferred synthon in each category are shown in Figure 3.12.
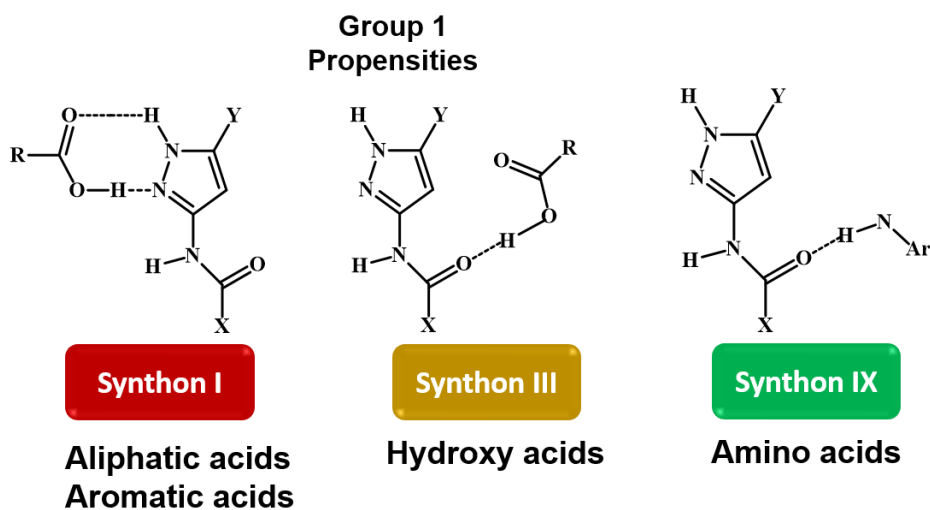
Figure 3.12 Synthon predicted in group 1 using propensities.

3.4.5.4 Molecular complementarity/geometrical constraints

The main difference between synthon **I** and synthon **II** is that in the former, pyrazole NH is used as the hydrogen-bond donor for ligand and in the latter, amide NH is used as the hydrogen-bond donor. The bond analysis of crystal structures in the CSD was done for NH(pyrazole)…O=C(acid) and NH(amide)...O=C(acid) interactions. Ten structures are reported in the CSD where pyrazole binding pocket binds to carboxylic acids via dimeric synthon **I**. 76 structures are reported in the CSD where amide pocket binds as dimer to the carboxylic acids visa synthon **II**. The bond angle analysis shows that most structures with NH(pyrazole)…O=C(acid) interaction has a bond angle of 138°and NH(amide)…. O=C(acid) interactions has a bond angle of 172°, Figure 3.13.
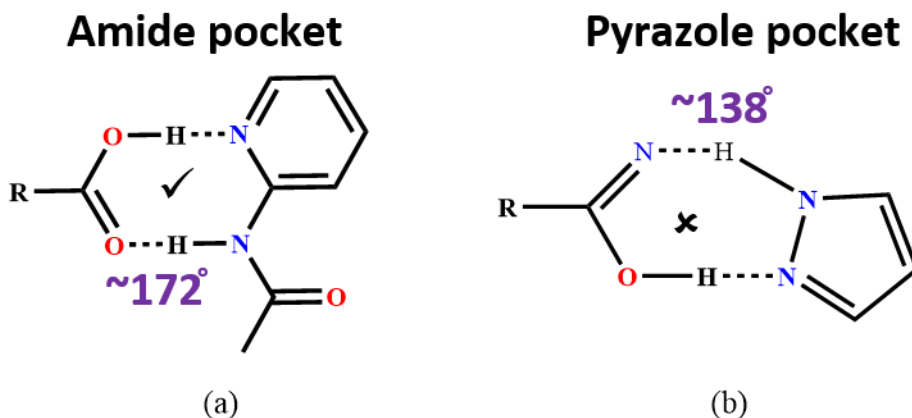
Figure 3.13 Differences in angles of amide and pyrazole binding pocket.

The second interaction OH(acid)…N(pyrazole) is common in both synthons and is almost linear. These results indicate that NH(amide)...O=C(acid) interaction is more linear compared to NH(pyrazole)…O=C(acid), therefore when a probe molecule has a choice between two binding pockets, it binds to the one which forms more linear hydrogen-bond interactions, i.e. the probe binds to pocket which is more geometrically favorable/compatible. In this study, amide binding pocket is indeed more favorable compared to pyrazole binding pocket.
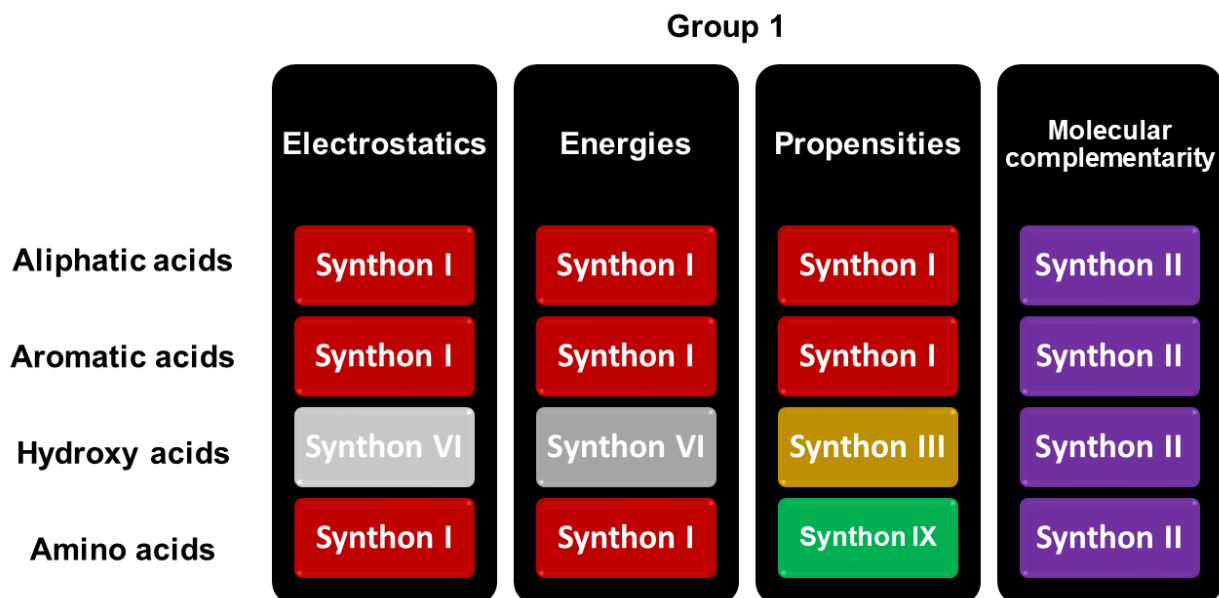
Figure 3.14 Summary of synthon predictions by different methods in group 1.

3.4.6 Synthon prediction in group 2

3.4.6.1 Method 1: Electrostatics

In group 2, synthon **I** is most likely to occur for two-point interaction and synthon **IV** based on one-point interaction. Note that the supramolecular chelate effect plays a role in the final determination of the synthon. For HydroxyBA, where additional single point interaction donor is present, synthon **VII** in group 2 are very likely to happen based on best donor-best acceptor Etter's rule. In AminoBA, synthon **IV** is predicted for group 2, Figure 3.15.

Group 2
Electrostatics
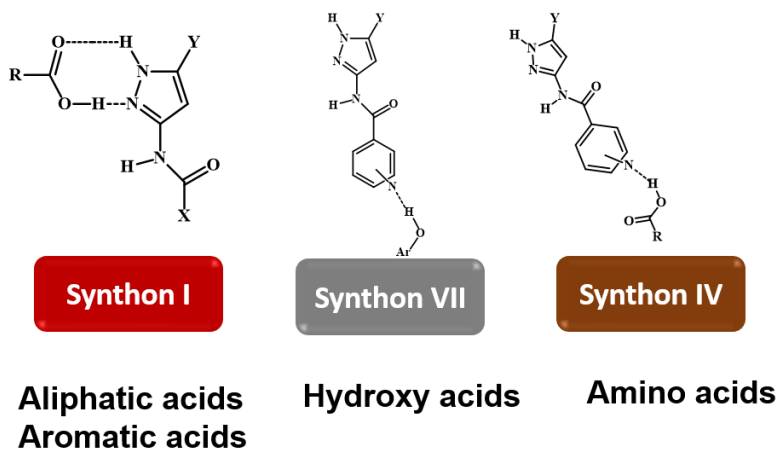
Figure 3.15 Synthon prediction in group 2 using electrostatics.

3.4.6.2 Method 2: Energies

In group 2, Synthon **I** was the most preferred synthon in aliphatic acids and synthon **III** was the most preferred synthon in aromatic acids (without additional donor group) followed by synthon **I** and synthon **II**. All synthons were close in energy; therefore, any synthon could be possible. In group 2 with hydroxyBA and aminoBA, there are seven synthon possibilities in each combination because of presence of an additional acceptor group on each target molecule. In hydroxyBA, synthon **II** (two-point interaction) and synthon **IV** (one-point interaction) are most likely to occur. In aminoBA, Synthon **II** is the preferred synthon, Figure 3.16a.
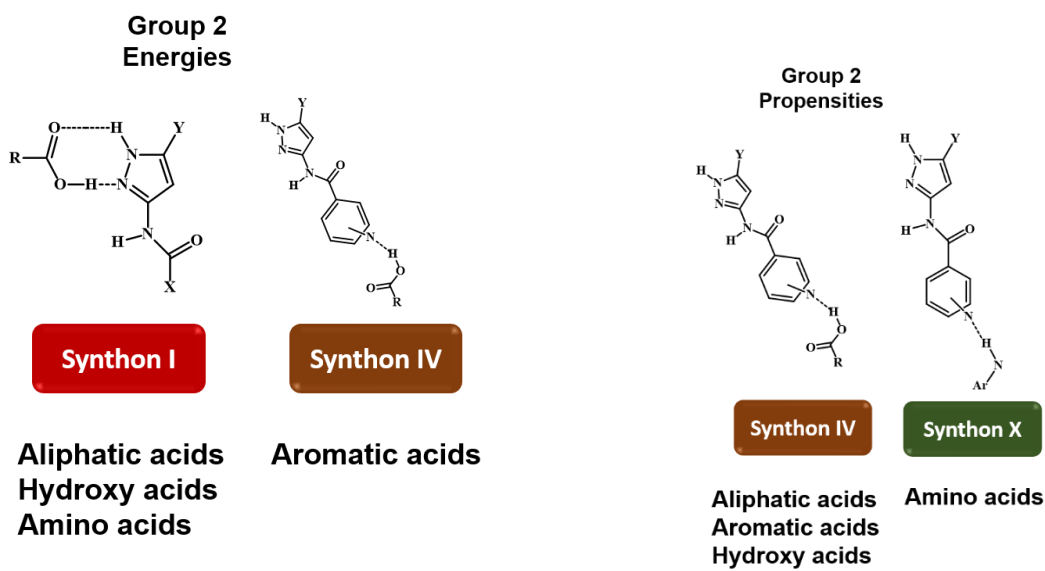
Figure 3.16 Synthon prediction in group 2 using (a) energies and (b) propensities.

3.4.6.3 Method 3: Propensities

Synthon **IV** is the most preferred synthon in aliphatic acids and aromatic acids without an additional donor group. In hydroxyBA, synthon **IV** is the most optimal synthon in group 2. In aminoBA, synthon **X** are the most preferred synthons in group 2, Figure 3.16b.

Figure 3.17 Summary of synthon predictions by different methods in group 2.

3.4.7 Synthon observed in the crystal structures

3.4.7.1 Crystal structures in group 1

In group **1**, nine crystal structures were obtained with di-aliphatic acids and four crystal structures were obtained with mono-aromatic acids. In seven out of nine cocrystals with di-aliphatic acids (78%), both ends of acids are involved in two complementary hydrogen bonds between carboxylic acid O-H(acid)…N(pyrazole) and N-H(amide)…O=C(acid) acid moieties gives rise to the robust $R_2^2(8)$ hydrogen-bonded synthon **II** in a 2:1 crystallizing agent: acid stoichiometry. These

heteromeric synthons are propagated into 2-dimensional sheets via hydrogen bonding between pyrazole N-H and carbonyl oxygen of the target molecule, forming homo-meric synthon.

There are two exceptions. The co-crystal of **P4**-Adi and **P8**-Sub forms dimeric synthon **II** on one side of acid (two-point interaction) but forms one point-interactions on the other end of acid; synthon **III** (OH(acid)…C=O(amide)) and half of synthon **I** (NH(pyrazole)….C=O(acid)) leading to synthon crossover. In aromatic acids without an additional donor group, three crystal structures were obtained and in all crystal structures, synthon **II** involving OH(acid)…N(pyrazole) and NH(amide)…O=C(acid) is the dominant interaction leading to 1:1 stoichiometry. In aromatic acids with an additional donor group, one crystal was obtained with **P3**-4minoBA and synthon **II** was observed as predicted by molecular complementarity approach. Therefore, in eleven out of thirteen crystal structures (85%) in group **1**, synthon **II** is preferred over synthon **I** or **III**. Synthon **II** was predicted based on molecular complementarity over synthon **I** even though the latter is predicted to more optimal synthon based on electrostatics, energies, and propensities in aliphatic diacids and aromatic acids with no additional donor, Figure 3.18.

Group 1

	Synthon II	Synthon III
Aliphatic acids	(7/9)	(2/9)
Aromatic acids	(3/3)	(0/3)
Hydroxy acids	N/A	N/A
Amino acids	(1/1)	(0/1)
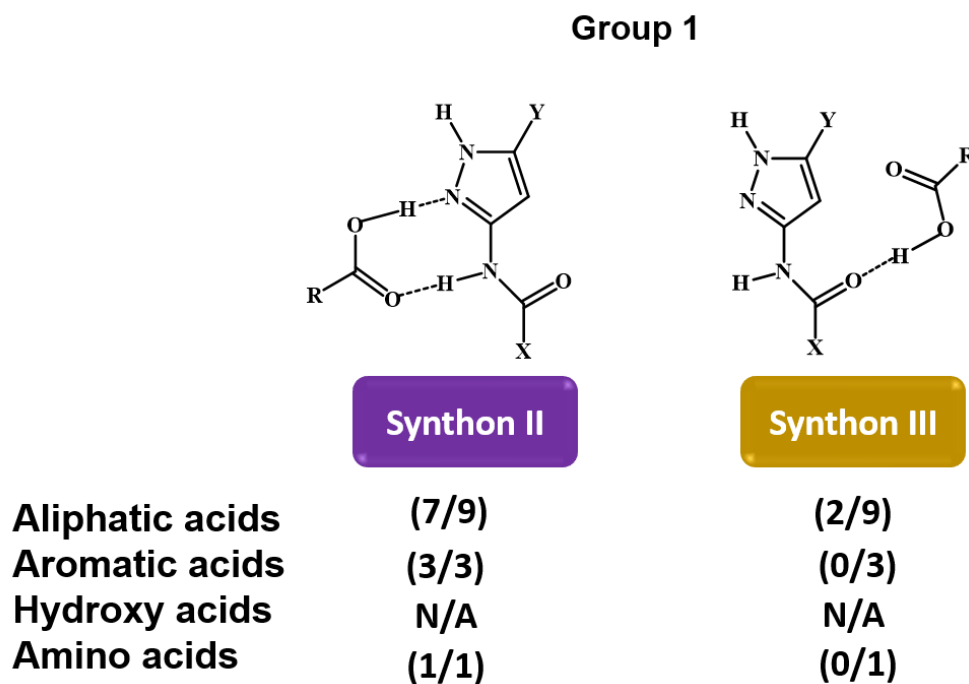
Figure 3.18 Summary of synthons observed in the crystal structures in group 1.

3.4.7.2 Crystal structures in group 2

Four crystal structures were obtained in group 2 and in three out of four co-crystals (**P10-Fum**, **P10-Mal**, and **P11-Aze**) with di-aliphatic acids, synthon **II** was observed. The crystal structure of **P11-Aze** forms robust synthon **II** on one side of suberic acid whereas the opposite end of acid participates in the synthon IV and half of synthon I leading to synthon crossover. No co-crystals were obtained with aromatic acids with and without additional donor group, Figure 3.19.

Group 2

	Synthon II	Synthon III
Aliphatic acids	(3/4)	(1/4)
Aromatic acids	N/A	N/A
Hydroxy acids	N/A	N/A
Amino acids	N/A	N/A
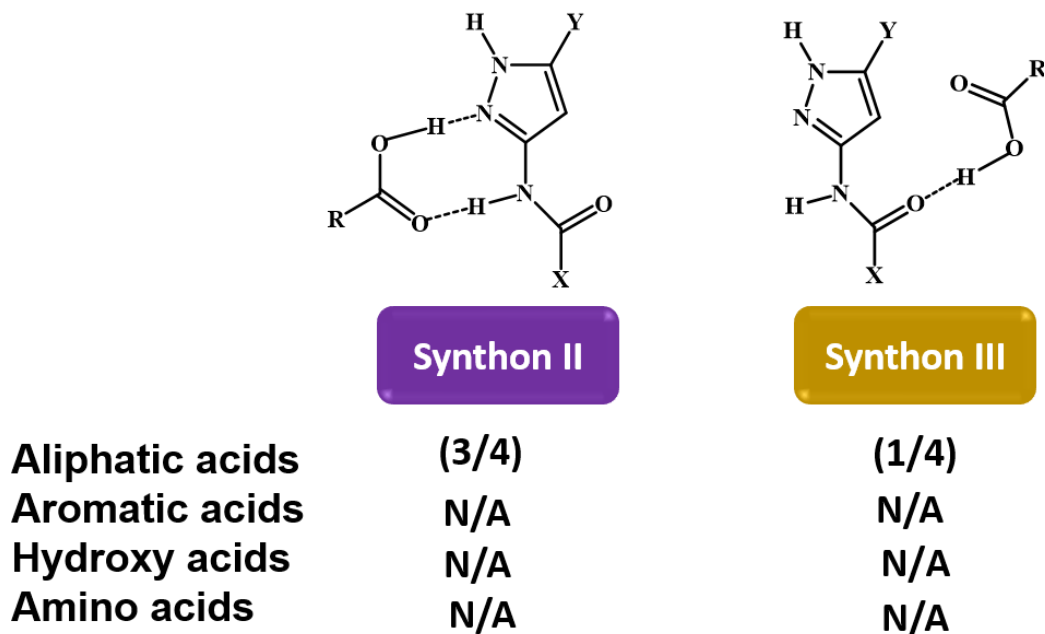
Figure 3.19 Summary of synthons observed in the crystal structures in group 2.

In total, we obtained 13 co-crystals with aliphatic acids and 4 co-crystals with aromatic acids. Therefore, 10 out of 13 (77%) outcomes showed dimeric synthon **II** and 2 out of 13 (15%) showed synthon **III**, 1 out of 13 (7%) showed synthon **IV** as the most preferred supramolecular outcome irrespective of different groups. The structure determinations reveal that synthon **II** is the most desired synthon and the primary intermolecular interactions, the O-H(acid)…N(pyrazole) and N-H(amide)…O=C(acid) hydrogen bonds, consistently remain intact. The supramolecular trimer was extended into 2-D sheets via one-point homomeric hydrogen bond between pyrazole N-H and carbonyl oxygen, N-H(pyrazole)…O=C(amide), resulting in homo-meric synthon. In each case, dimeric synthon **II** is preferred over synthon **I** in a competitive environment. Therefore, having methyl group on pyrazole backbone or introducing bulky substituents such as methyl, ethyl, phenyl

or pyridyl groups next to amide functionality didn't change the synthon preference in these co-crystals but there is chance of synthon crossover.

3.4.8 Comparison of experiment vs predictions

Comparing the predicted synthons with the observed synthons, it was found that the synthons observed in the crystal structures matches with the second possibility of predicted synthon (Synthon **II** or **III**) instead of the first possibility (synthon **I**) based on all three prediction methods (electrostatics, energies and propensities). Synthon **II** was also predicted based on molecular complementarity. Synthon **I** was the most preferred synthon based on prediction, but it wasn't observed in any of the crystal structures because of geometric constraints associated with this synthon. Based on CSD search, synthon **II** was more likely to occur compared to synthon **I** because in former, the hydrogen-bond angles are more linear compared to the former synthon. These results indicate that when a probe molecule has a choice between two binding pockets, it binds to the one which forms more linear hydrogen-bond interactions, i.e. the probe binds to pocket which is more geometrically favorable/compatible. Therefore, in this study even though synthon **I** is more energetically favored, synthon **II** is preferred because it is more geometrically compatible.

Table 3.9 Summary of predicted vs experimental synthons observed in each group

		Based on Electrostatics, energies, propensities		Molecular complementarity	
		1st possibility	2nd possibility	1st possibility	Experimentally Observed synthon
Group 1	**Aliphatic acids**	Synthon **I**	Synthon **II**	Synthon **II**	Synthon **II** (7/9) Synthon **III** (2/9)
	Aromatic acids	Synthon **I** or **III**			Synthon **II** (3/3)
	HydroxyBA	Synthon **VI**, **III**	Synthon **I**		N/A
	AminoBA	Synthon **I**, **III**, **IX**	Synthon **II**		Synthon **II** (1/1)
Group 2	**Aliphatic acids**	Synthon **I**, **IV**	Synthon **II**	Synthon **II**	Synthon **II** (3/4) Synthon **III** (1/4)
	Aromatic acids	Synthon **I**, **IV**			N/A
	HydroxyBA	Synthon **I**, **VII**, **IV**			N/A
	AminoBA	Synthon **IV**, **I**, **X**			N/A

3.5 Conclusions

We attempted to co-crystallize the twelve target molecules (**P1-P12**) with twenty different carboxylic acids to probe the strength and stability of the homomeric interactions.

1. The homomeric interactions were broken to form heteromeric synthons via experimental co-crystallization with 70% success rate.

2. Hydrogen bond energy and propensity were used as prediction methods to predict and compare the experimental co-crystal screening results comprising twelve target molecules and twenty carboxylic acids resulting in 240 data points. Hydrogen-bond energy resulted in 69% agreement and hydrogen bond propensity resulted in only 35% agreement.

3. Synthon prediction studies were done using electrostatics, energies, propensities and molecular complementarity. Synthon **I** was the preferred synthon over synthon **II** based on electrostatics, energies and propensities whereas Synthon **II** was preferred based on molecular complementarity.

4. Synthon **II** was observed experimentally in 14 out 17 crystal structures with 82% success rate. These results indicated the solid state supramolecular synthesis[25] depends not only upon the electrostatic or chemical environment around the binding site but also upon the size, shape, geometry and orientation of interacting molecular species to achieve robust and predictable supramolecular synthons.[26] Effective molecular recognition occurs when both factors, electrostatics and molecular geometry are compatible between the interacting partners.[27,28]

3.6 References

1. Babu, N. J.; Cherukuvada, S.; Thakuria, R.; Nangia, A., *Crystal Growth & Design* **2010,** *10* (4), 1979-1989.

2. Mukherjee, A.; Desiraju, G. R., *Chemical Communications* **2011,** *47* (14), 4090-4092.

3. Pogoda, D.; Janczak, J.; Videnova-Adrabinska, V., *Acta Crystallographica Section B* **2016,** *72* (2), 263-273.

4. Sarma, B.; Sanphui, P.; Nangia, A., *Crystal Growth & Design* **2010,** *10* (5), 2388-2399.

5. Tothadi, S., *CrystEngComm* **2014,** *16* (32), 7587-7597.

6. Etter, M. C., *Accounts of Chemical Research* **1990,** *23* (4), 120-126.

7. Aakeroy, C. B.; Wijethunga, T. K.; Desper, J., *New Journal of Chemistry* **2015,** *39* (2), 822-828.

8. Hunter, C. A., *Angewandte Chemie International Edition* **2004,** *43* (40), 5310-5324.

9.   Musumeci, D.; Hunter, C. A.; Prohens, R.; Scuderi, S.; McCabe, J. F., *Chemical Science* **2011,** *2* (5), 883-890.

10.   Issa, N.; Karamertzanis, P. G.; Welch, G. W. A.; Price, S. L., *Crystal Growth & Design* **2009,** *9* (1), 442-453.

11.   Fábián, L., *Crystal Growth & Design* **2009,** *9* (3), 1436-1443.

12.   Galek, P. T. A.; Allen, F. H.; Fabian, L.; Feeder, N., *CrystEngComm* **2009,** *11* (12), 2634-2639.

13.   Delori, A.; Galek, P. T. A.; Pidcock, E.; Jones, W., *Chemistry – A European Journal* **2012,** *18* (22), 6835-6846.

14.   Dubey, R.; Desiraju, G. R., *IUCrJ* **2015,** *2* (4), 402-408.

15.   Dubey, R.; Mir, N. A.; Desiraju, G. R., *IUCrJ* **2016,** *3* (2), 102-107.

16.   Aakeröy, C. B.; Panikkattu, S. V.; DeHaven, B.; Desper, J., *Crystal Growth & Design* **2012,** *12* (5), 2579-2587.

17.   Lehn, J.-M., *Angewandte Chemie International Edition in English* **1990,** *29* (11), 1304-1319.

18.   Garcia-Tellado, F.; Geib, S. J.; Goswami, S.; Hamilton, A. D., *Journal of the American Chemical Society* **1991,** *113* (24), 9265-9269.

19.   Fan, E.; Vicent, C.; Geib, S. J.; Hamilton, A. D., *Chemistry of Materials* **1994,** *6* (8), 1113-1117.

20.   Geib, S. J.; Hirst, S. C.; Vicent, C.; Hamilton, A. D., *Journal of the Chemical Society, Chemical Communications* **1991,** (18), 1283-1285.

21.   Aakeröy, C. B.; Beatty, A. M.; Helfrich, B. A., *Angewandte Chemie International Edition* **2001,** *40* (17), 3240-3242.

22.   Sanphui, P.; Bolla, G.; Das, U.; Mukherjee, A. K.; Nangia, A., *CrystEngComm* **2013,** *15* (1), 34-38.

23.   Tothadi, S.; Joseph, S.; Desiraju, G. R., *Crystal Growth & Design* **2013,** *13* (7), 3242-3254.

24.   Wood, P. A.; Feeder, N.; Furlow, M.; Galek, P. T. A.; Groom, C. R.; Pidcock, E., *CrystEngComm* **2014,** *16* (26), 5839-5848.

25.   Desiraju, G. R., *Current Opinion in Solid State and Materials Science* **1997,** *2* (4), 451-454.

26.   Braga, D., *Angewandte Chemie International Edition* **2012,** *51* (15), 3516-3516.

27.   Aakeröy, C. B.; Desper, J.; Leonard, B.; Urbina, J. F., *Crystal Growth & Design* **2005,** *5* (3), 865-873.

28.   Aakeroy, C. B.; Desper, J.; Urbina, J. F., *CrystEngComm* **2005,** *7* (31), 193-201.

Chapter 4 - Evaluating Homomeric Synthons in Thiazole based Molecules using Energies and Propensities

4.1 Introduction

Thiazole rings are present in many pharmaceutical compounds, and drugs such as FB,[24] AMG 517,[25] meloxicam,[26-27] and nitazoxanide,[28-29] contain amide-substituted thiazoles as part of their chemical make-up. This functionality has been successfully exploited to synthesize co-crystals of some of these active pharmaceutical ingredients (APIs). Knowing in advance which supramolecular pathways are accessible in molecular solids by themselves, and which of those possible options is the most likely route for a molecule, is thus clearly of importance for developing solid forms with optimum bulk properties.
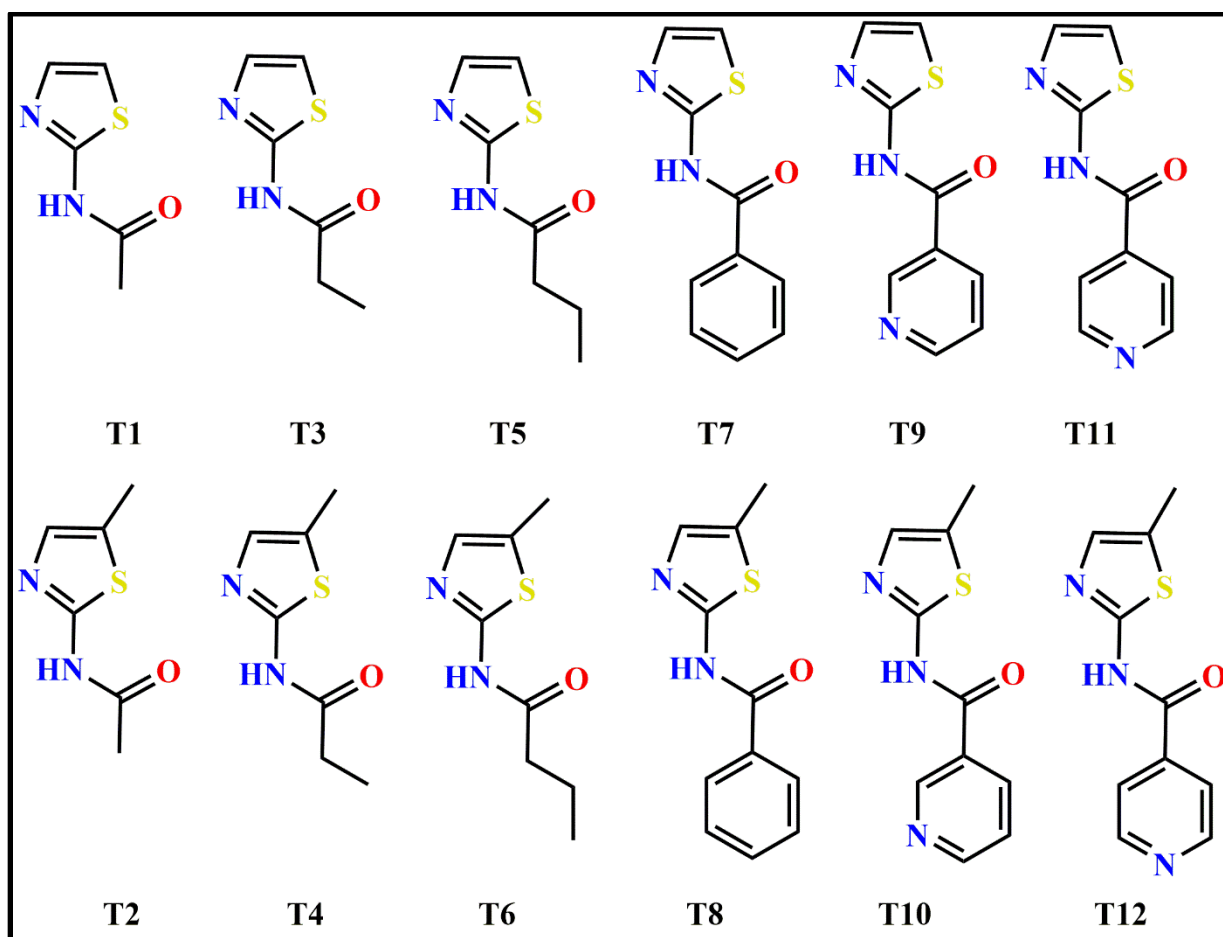
Scheme 4.1 Twelve thiazole amides used in this study.

(a) Synthon A   **(b) Synthon B**   **(c) Synthon C**   **(d) Synthon D**
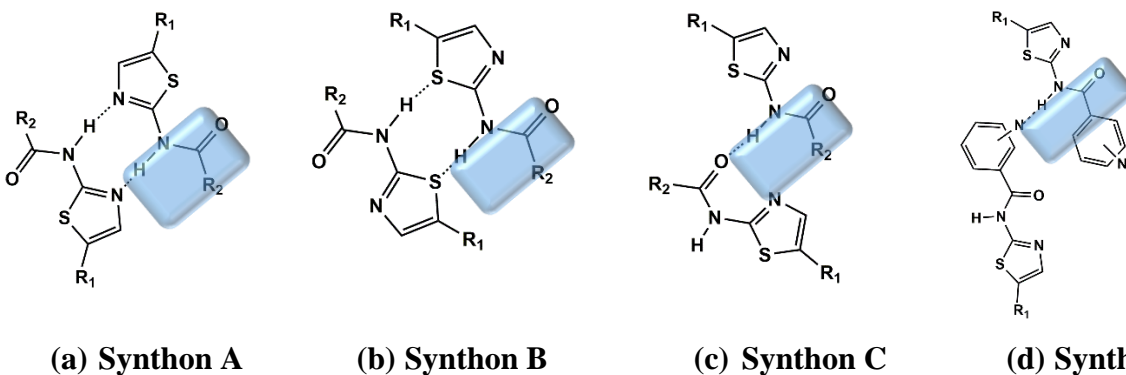
Scheme 4.2 Four postulated synthons in a generic thiazole-amide; (a) synthon **A**, N-H (amide)…(aromatic); (b) synthon **B**, N-H (amide)…S(aromatic); (c) synthon **C**, N-H (amide)…O=C(amide); and (d) synthon **D**, N-H (amide)…(pyridine N(R_1=H, CH$_3$; R_2= methyl, ethyl, or benzyl).

In this study, we want to establish what intermolecular preference, if any, the N-H group displays and whether we can rationalize this in a convincing manner. The target molecules **T1-T12** (scheme 4.2) are divided into two groups based on the number of functional groups present. **T1-T8** belongs to group1 and **T9-T12** belongs to group 2 molecules. The interaction between the hydrogen-bond donor and the three possible acceptors in group1 would lead to three different synthons (synthon **A**, synthon **B** and synthon **C**) and an additional synthon (synthon D) in group 2, Scheme 4.3, taking into account only the *trans* C=O-NH isomer and disregarding any catemeric versions of synthons **A** and **B**. We used structural informatics as a guideline to determine the most plausible conformation in these target molecules using data from the CSD. The search was restricted to structures without disorder, and with an R-factor below 5%. According to our CSD search (CSD database 5.38, Nov 2016 with updates from Feb and May 2017),[30-33] we found 6335 relevant structures containing the amide functionality and only about 0.5% of these displayed a *cis* CO-NH isomer. Therefore, we felt justified in focusing exclusively on the *trans* CO-NH geometry in these target molecules. In addition, structural data complemented by geometry optimizations (see Section 2.2), indicated that the amide is likely to be co-planar with the thiazole.
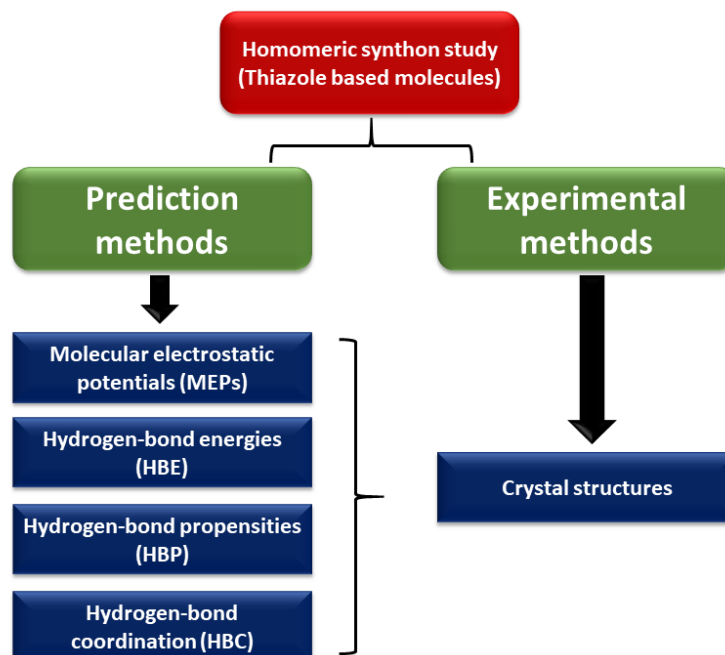
Scheme 4.3 An outline of prediction and experimental methods used in this study.

The study is done to answer the following questions,

1. Which method (MEPs, HBE, HBP or HBC) is successful in predicting the key hydrogen-bond interactions?

2. Is a combination of prediction methods preferred over individual methods?

3. Which synthon is the most optimal synthon in group **1** (**T1-T8**) and how does adding an acceptor group affect the choice of synthon in **T9-T12**.

4.2 Experimental

4.2.1 General

2-Amino-thiazole, 2-amino-5-methyl-thiazole, acetic anhydride, propionic anhydride, butyric anhydride, nicotinic acid, Isonicotinic acid and thionyl chloride and benzoyl chloride were purchased from commercial sources and utilized without further purification. Target molecules **T1-T6 and T9-T12** are synthesized following procedures below. Target molecules **T7-T8** are synthesized following literature methods.[34] Melting points were measured using a Fisher-Johns melting point apparatus. Solution ^{1}H NMR data were collected in DMSO-d_6 on a Varian Unity plus 400 MHz NMR spectrometer.

4.2.2 Synthesis

4.2.2.1 Synthesis of 2-acetamido-thiazole, T1
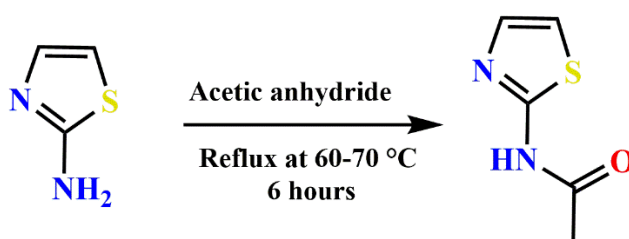
A solution of 2-amino-thiazole (2.46g, 24.6mmol) in 15ml of acetic anhydride was refluxed until acetylation was complete for approximately 6 hours. Thin layer chromatography (TLC) was performed to confirm the formation of only one product. Excess acetic anhydride and acetic acid were removed by rota-evaporation under vacuum. The product so obtained was washed with water and the solid precipitate was filtered. The solid was dissolved in a 1:1 ratio of methanol/diethyl-ether and the solution was allowed to stand at room temperature. Crystals of pure product were obtained in a week. Yellow solid; yield: 2.75 g (79%); mp 200-205 ℃ (lit. 200-202 ℃)[35] ^{1}H NMR (δH; DMSO, 400MHz): 11.48 (s, 1H), 7.85 (d, 1H), 7.53 (d, 1H), 2.69 (s, 3H).

4.2.2.2 Synthesis of 2-acetamido-5-methyl-thiazole, T2.
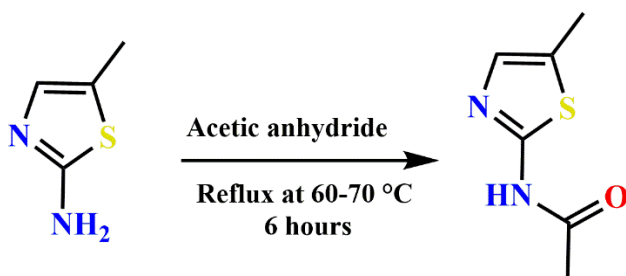
A solution of 2-amino-5-methyl-thiazole (2.6g, 17.4mmol) in 10ml of acetic anhydride was refluxed at 70 ℃ for approximately 6 hours until acetylation was complete. TLC was performed to confirm the formation of only one product. Excess acetic anhydride and acetic acid were removed by rota-evaporation under vacuum. The product so obtained was washed with water and the solid precipitate was filtered. The solid was dissolved in a 1:1 ratio of methanol/diethyl-ether and the

solution was allowed to stand at room temperature. Crystals of pure product were obtained in a week. White solid; yield: 3.55 g (98%); mp 219-225℃ (lit. 225℃)[i] H NMR (δH; DMSO, 400MHz): 11.32 (s, 1H), 7.49 (s, 1H), 2.80 (s, 3H), 2.65 (s, 3H).

4.2.2.3 Synthesis of 2-propamido-thiazole, T3.

Propionic anhydride

Reflux at 60-70 °C
2 hours
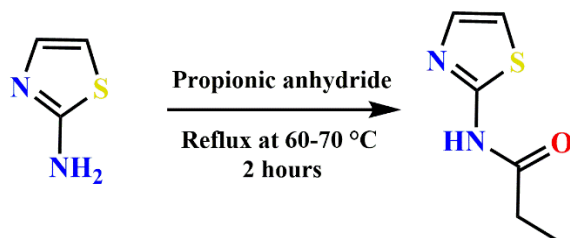
A solution of 2-amino-thiazole (2.46g, 24.6mmol) in 10ml of propionic anhydride was refluxed at 65℃ for 2 hours. TLC was performed to confirm the formation of only one product. Excess propionic anhydride and propionic acid were removed by rota-evaporation under vacuum. The brown product so obtained was washed with water and the solid precipitate was isolated by filtration. The solid was dissolved in methanol and the solution was allowed to stand at room temperature. Crystals of pure product were obtained in 5 days. Brown solid; yield: 3.00 g (78%); mp 158-162℃ (lit. 157-159℃)[ii]; ^{1}H NMR (δH; DMSO, 400MHz): 12.05 (s, 1H), 7.44 (d, 1H), 7.18(d, 1H), 2.42 (q, 2H), 1.08 (t, 3H).

4.2.2.4 Synthesis of 2-propamido-5-methyl-thiazole, T4.

Propionic anhydride

Reflux at 60-70 °C
2 hours
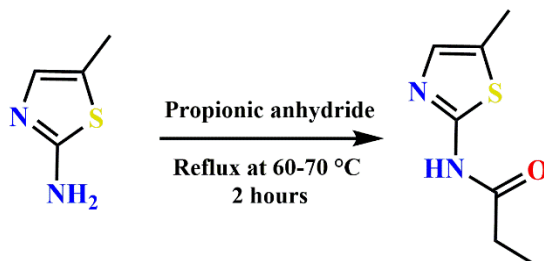
A solution of 2-amino-5-methyl-thiazole (2.6g, 17.4mmol) in 10ml of propionic anhydride was refluxed at 65℃ for 2.5 hours. TLC was performed to confirm the formation of only one product. Excess propionic anhydride and propionic acid were removed by rota-evaporation under vacuum. The yellow product so obtained was washed with water and the solid precipitate was isolated by

95

filtration. The solid was dissolved in methanol and the solution was allowed to stand at room temperature. Crystals of pure product were obtained in a week. Yield: 3.89 g (88%); mp 196-201℃; [^1]H NMR (δH; DMSO, 400MHz): 11.84 (s, 1H), 7.10 (s, 1H), 2.32 (q, 2H), 2.32 (s, 3H), 1.07 (t, 3H).

4.2.2.5 Synthesis of 3-butyramido thiazole, T5
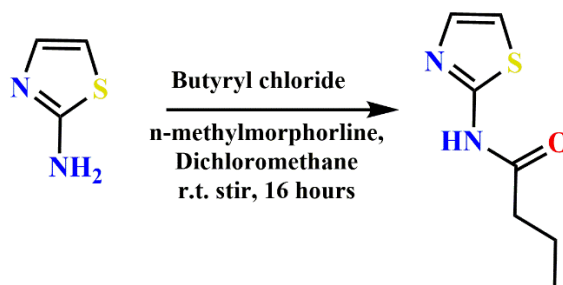
2-Aminothiazole (0.601 g, 6.00 mmol) was dissolved in 30 mL of dichloromethane, 1.6 mL N-methylmorpholine (14.6 mmol), and 1.49 mL butyryl chloride (14.0 mmol). After stirring for 16 hours, the mixture was concentrated and the residue was dissolved in 30 mL methanol. Then, 7 mL of 2.5M sodium hydroxide was added dropwise, followed by 10 mL of tetrahydrofuran. The mixture was allowed to sit for 15 minutes before being concentrated again and suspended in water. The tan precipitate was filtered, air dried, and finally recrystallized in methanol to yield dark, golden crystals. 61% yield, m.p. 143-146°C, [^1]H NMR (400 MHz, DMSO-d_6) δ ppm: 12.02 (1H), 7.42 (1H), 7.16 (1H), 2.38 (2H), 1.60 (2H), 0.88 (3H).

4.2.2.6 Synthesis of 3-butyramido 5-methyl thiazole, T6
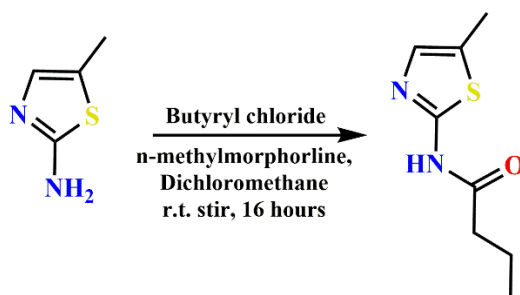
2-Amino-5-methylthiazole (0.685 g, 6.00 mmol) was dissolved in 30 mL of dichloromethane, 1.6 mL N-methylmorpholine (14.6 mmol), and 1.49 mL butyryl chloride (14.0 mmol). After stirring for 16 hours, the mixture was concentrated and the residue was dissolved in 30 mL methanol. Then, 7 mL of 2.5M sodium hydroxide was added dropwise, followed by 10 mL of tetrahydrofuran. The mixture was allowed to sit for 15 minutes before being concentrated again and suspended in water. The tan precipitate was filtered, air dried, and finally recrystallized in methanol to yield tan crystals. 45% yield, m.p. 126-135°C, [^1]H NMR (400 MHz, DMSO-d$_6$) δ ppm: 11.01 (1H), 6.28 (1H), 1.55 (2H), 1.51 (3H), 0.78 (2H), 0.07 (3H).

4.2.2.7 Synthesis of 2-benzamido-thiazole, T7
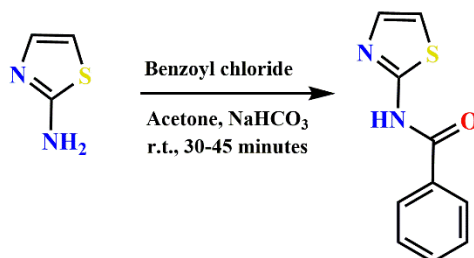
To a mixture of 2.5g of 2-amino-thiazole in 30ml acetone and 50ml brine solution (16.67g NaCl in 50ml water), 2.91ml of benzoyl chloride in 3ml acetone was added in increments over a period of 15 minutes with continuous stirring at room temperature. A total of 4.2g of sodium bicarbonate was dissolved in 50ml of water and the saturated solution was added to the mixture. The reaction mixture was further stirred for 15-30 minutes. The solid formed was separated by filtration and dried. It was dissolved in methanol and the solution was allowed to slowly evaporate at room temperature to grow single crystals. White solid; yield: 3.06 g (60%); mp 150-152℃ (lit. 152℃) [^1]H NMR (δH; DMSO, 400MHz): 12.65 (s, 1H), 8.10 (d, 2H), 7.63 (t, 1H), 7.54 (d, 2H), 7.52 (s, 1H), 7.29 (s, 1H).

4.2.2.8 Synthesis of 2-benzamido-5-methyl-thiazole, T8
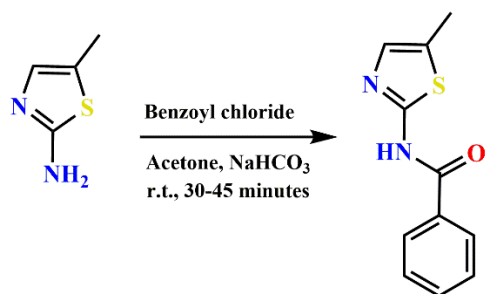
To a mixture of 1.42 g of 2-amino-5-methyl-thiazole in 15ml acetone and 25ml brine solution (8.34g NaCl in 50ml water), 1.45ml of benzoyl chloride in 1.5ml acetone was added in portions with continuous stirring during 15 minutes at room temperature. A total of 2.10 g of sodium bicarbonate was dissolved in 25ml of water and the saturated solution was added to the mixture. The reaction mixture was further stirred for 15-30 minutes. The solid formed was separated by filtration and dried. It was dissolved in methanol and the solution was allowed to slowly evaporate at room temperature to grow single crystals. White solid; yield: 2.01 g (73%); mp 160-162°C; ^{1}H NMR (δH; DMSO, 400MHz): 12.46 (s, 1H), 8.08 (d, 2H), 7.62 (t, 1H), 7.53 (d, 2H), 7.23 (s, 1H), 2.37 (S, 3H).

4.2.2.9 Synthesis of N-(thiazol-2-yl) nicotinamide, T9
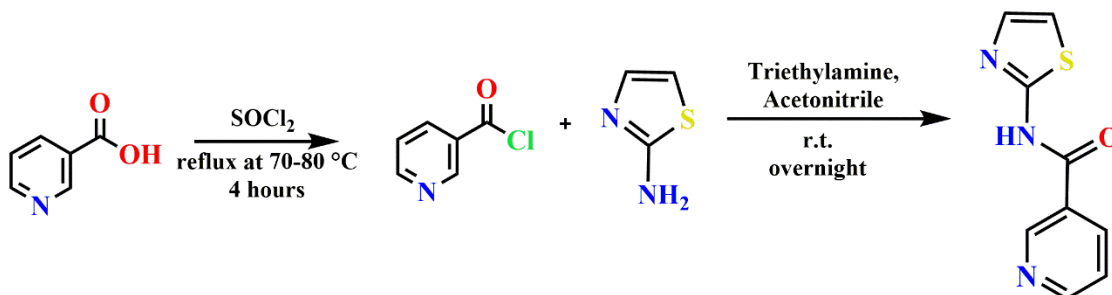
A solution of 3-pyridine-carboxylic acids (5 grams, 40.5 mmol) in thionyl chloride (50ml) was refluxed for 4 hours at 70-80 °C. Then, thionyl chloride was removed under vacuum to obtain white solid as nicotinoyl chloride (91.27%). 2-amino-thiazole (1.2g, 12 mmol) was added slowly into 50ml round bottom flask containing nicotinoyl chloride (1.70g, 12 mmole) and triethylamine (1.6 ml) in acetonitrile (20 ml). the mixture was stirred at room temperature overnight. The organic layer was washed with water and the solvent was removed by gravity filtration to yield a yellow powder as the crude product. The product was recrystallized from methanol. Yellow solid; yield

90%; mp 213-214℃, ^{1}H NMR (δH, 400 MHz, DMSO): 12.87 (br, 1H), 9.20 (s, 1H), 8.79 (s, 1H), 8.42 (s, 1H), 7.57 (s, 2H), 7.32 (s, 1H).

4.2.2.10 Synthesis of N-(5-methylthiazol-2-yl) nicotinamide, T10
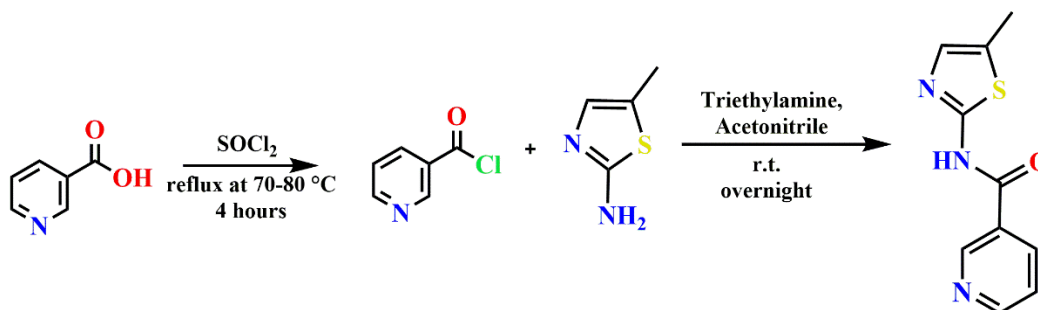
A solution of 3-pyridine-carboxylic acids (5 grams, 40.5 mmol) in thionyl chloride (50ml) was refluxed for 4 hours at 70-80 °C. Then, thionyl chloride was removed under vacuum to obtain white solid as nicotinoyl chloride (91.27%).2-amino-5methyl-thiazole (1.79g, 12 mmol) was added slowly into 50ml round bottom flask containing nicotinoyl chloride (1.70g, 12 mmole) and triethylamine (1.6 ml) in acetonitrile (20 ml). the mixture was stirred at room temperature overnight. The organic layer was washed with water and the solvent was removed by gravity filtration to yield a white powder as the crude product. white solid; yield 92%; mp 212-213℃, ^{1}H NMR (δH, 400 MHz, DMSO): 12.71 (br, 1H), 9.18 (s, 1H), 8.78 (s, 1H), 8.38 (s, 1H), 7.55 (s, 1H), 7.24 (s, 1H), 2.38 (s, 3H).

4.2.2.11 Synthesis of N-(thiazol-2-yl)isonicotinamide, T11
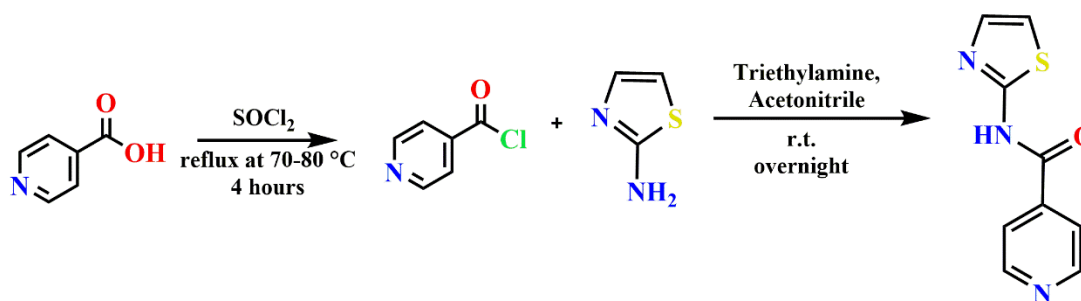
A solution of 4-pyridine-carboxylic acids (5 grams, 40.5 mmol) in thionyl chloride (50ml) was refluxed for 4 hours at 70-80 °C. Then, thionyl chloride was removed under vacuum to obtain

white solid as nicotinoyl chloride (93%). 2-amino-thiazole (1.2g, 12 mmol) was added slowly into 50ml round bottom flask containing isonicotinoyl chloride (1.70g, 12 mmole) and triethylamine (1.6 ml) in acetonitrile (20 ml). the mixture was stirred at room temperature overnight. The organic layer was washed with water and the solvent was removed by gravity filtration to yield a yellow powder as the crude product. The product was recrystallized from methanol. Yellow solid; yield 89%; mp 199-201 ℃, ^{1}H NMR (δH, 400 MHz, DMSO): 12.97 (br, 1H), 8.80 (s, 2H), 7.98 (s, 2H), 7.59 (s, 1H), 7.32 (s, 1H).

4.2.2.12 Synthesis of N-(5-methylthiazol-2-yl) isonicotinamide, T12
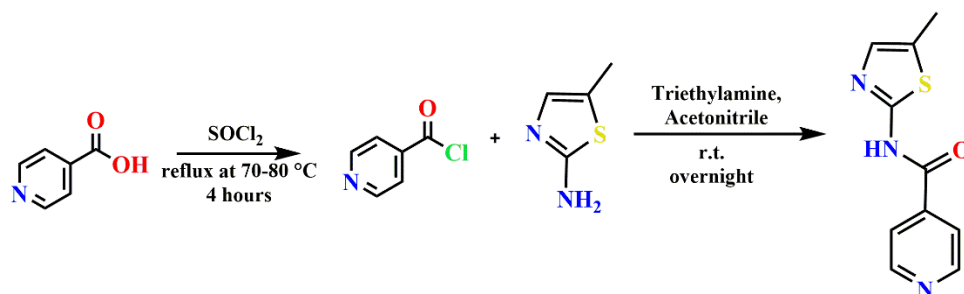
2-Amino-5methyl-thiazole (1.79g, 12 mmol) was added slowly into 50ml round bottom flask containing nicotinoyl chloride (1.70g, 12 mmole) and triethylamine (1.6 ml) in acetonitrile (20 ml). the mixture was stirred at room temperature overnight. The organic layer was washed with water and the solvent was removed by gravity filtration to yield a dark yellow powder as the crude product. Yellow solid; yield 85%; mp 213-214 ℃, ^{1}H NMR (δH, 400 MHz, DMSO): 12.78 (s, 1H), 8.79 (d, 2H), 7.97 (d, 2H), 7.26(d, 1H), 2.38 (s, 3H).

4.2.3 CSD analysis

(**a**)
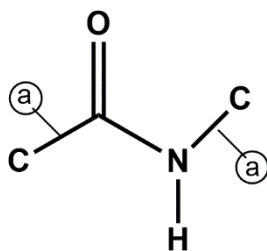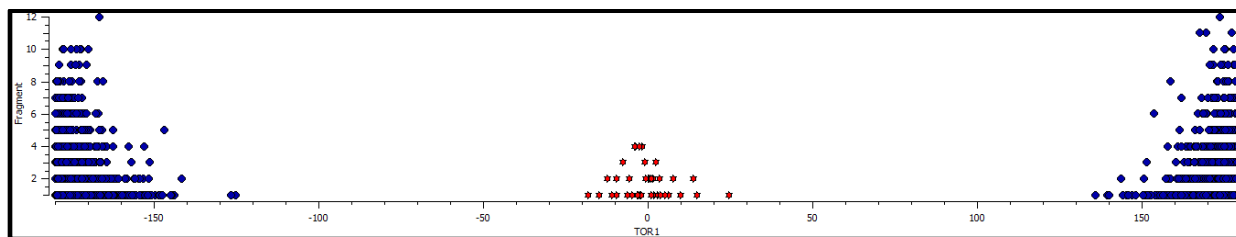
Figure 4.1 (a) Amide functionality (both bonds are acyclic) used to perform the torsion angle search. (b) Scatterplot indicating number of structures with torsions for *cis* (red, ~32 structures) and *trans* (blue, ~6303 Structures) conformations.
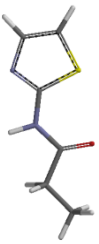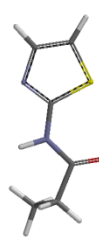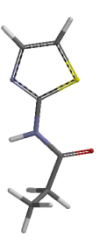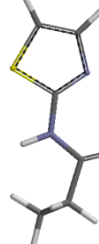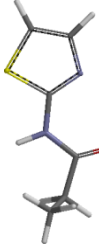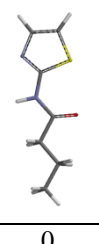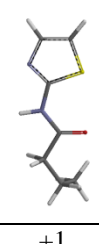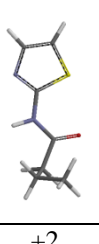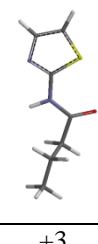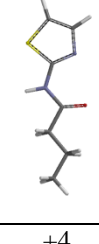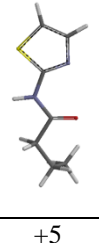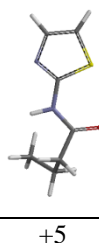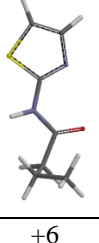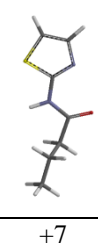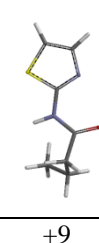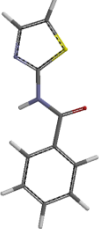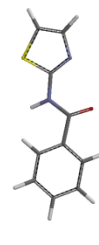
A search was done by drawing the amide group as shown in Figure 4.1 and CSD database 5.38 (Nov 2016) with updates from Feb and May 2017 was used to perform the torsion angle search on CSD database. The following filters were applied to the search (No disordered, No errors, Not polymeric, No ions, No powder patterns and only organics). The torsion angle on four atoms (OCNH) was defined from -180° to +180°. A total of 6335 structures was obtained and approximately 32 (0.5%) of these structures have the *cis*-amide conformation and about 6303 structures (99%) have *trans*-amide conformation.

4.2.4 Geometry optimizations for T1-T12

Geometry optimization was performed on each conformation of **T1-T12** and the most stable conformation is given a value of 0 kJ/mol; all other conformations energies are presented relative to this, Table 4.1.

Table 4.1 Energies of each *trans* conformation relative to most stable *trans* conformation is shown below in kJ/mol. Note: conformations of target molecules with methyl group on thiazole ring are not shown here.

	Conformation 1	Conformation 2	Conformation 3	Conformation 4	Conformation 5	Conformation 6
Group 1						
T1						

T1	**0**	+ 4				
T2	**0**	+4				
T3						
T3	0	+3	+3	+4	+7	+7
T4	0	+3	+3	+4	+7	+7
T5						
T5	0	+1	+2	+3	+4	
T6	0	+2	0	+3	+4	
T5	+5	+5	+6	+7	+9	
T6	+5	+5	+4	+7	+8	
T7						
T7	0	+7				
T8	0	+5				
Group 2						

T9					
T9	0	+4	+5	+8	
T10	+4	+8	0	+5	
T11					
T11	0	+5			
T12	0	+5			

4.2.5 Molecular electrostatic potential (MEPs) calculations

Molecular electrostatic potential surfaces (MEPS) of **T1**–**T12** were generated with density functional B3LYP level of theory using 6-311++G** basis set in vacuum, see section 2.2.4 for details.

4.2.6 Hydrogen-bond energies (HBE) for predicting structural outcomes

See section 2.2.5 for detailed description of hydrogen-bond energy calculations. The results are shown in Table 4.2.

Table 4.2 Electrostatics (in kJ/mol) and alpha and beta values for T1-T12 calculated using equations 1 and 2 for donor and acceptor group.
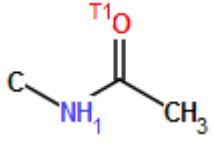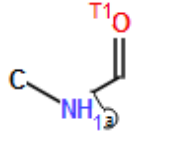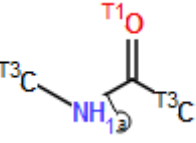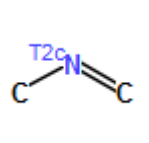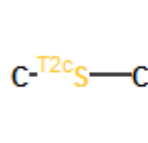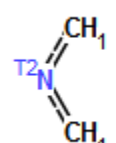
	NH amide (kJ/mol)	α (donor)	cyclic N (kJ/mol)	β(N acceptor)	cyclic S (kJ/mol)	β(S acceptor)	carbonyl O (kJ/mol)	β(O acceptor)	Pyridine N	β (pyrN)
T1	246.00	3.35	-146.00	4.47	-76.00	1.55	-178.00	6.28		
T2	237.00	3.19	-153.00	4.84	-80.00	1.68	-184.00	6.65		
T3	239.00	3.22	-149.00	4.63	-78.00	1.61	-172.00	5.92		

T4	231.00	3.09	-156.00	5.00	-88.00	1.95	-178.00	6.28		
T5	239.00	3.19	-150.00	4.84	-90.00	1.68	-173.00	6.65		
T6	232.00	3.22	-157.00	4.63	-70.00	1.61	-179.00	5.92		
T7	226.00	3.00	-149.00	4.63	-84.00	1.81	-172.00	5.92		
T8	221.00	2.92	-156.00	5.00	-82.00	1.74	-176.00	6.16		
T9	229.00	3.05	-143.00	4.32	-39.00	0.58	-152.00	4.79	-153.00	4.84
T10	225.00	2.98	-147.00	4.52	-60.00	1.08	-156.00	5.00	-158.00	5.11
T11	243.00	3.29	-134.00	3.87	-20.00	0.24	-149.00	4.63	-160.00	5.23
T12	241.00	3.26	-141.00	4.21	-28.00	0.37	-157.00	5.06	-169.00	5.74

4.2.7 Hydrogen-bond propensities (HBP) for predicting structural outcomes

The detailed description of HBP calculations is described in section 2.2.6. The different amide substituents in targets **T1/T2**, **T3/T4** and **T5/T6** were modeled separately as shown in Table 4.3.

Table 4.3 Functional groups used to determine the hydrogen-bond propensities for the twelve target molecules. The labels in the figures can be explained as follows: Tn = atom makes n bonds, c = atom is cyclic, $\odot$ = bond is acyclic, and Hn = n bonded hydrogen atoms.

T1/T2	T3/T4/T5/T6	T7-T12	T1-T12	T1-T12	T9-T12

4.2.8 Crystal growth

T1-T12 were kept in a vial for slow evaporation methanol solvent in order to obtain crystals suitable for single crystal X-ray diffraction. If suitable crystals were not obtained in methanol, then different solvents were tried to grow crystals. X-ray experimental data and crystallographic data are given in Table 4.4.

Table 4.4 Experimental details of crystals obtained

	Molecules	Solvent used	Morphology	Melting point
Group1	T1	Methanol	Block, colorless	200-205°C
	T2	Methanol	Plate, colorless	219-225°C
	T3	Methanol	Plate, colorless	158-162°C
	T4	Methanol	Block, colorless	196-201°C
	T5	Methanol	Blocks, colorless	143-146°C
	T6	Methanol	Blocks, colorless	126-135°C
	T7	Methanol	Plates, colorless	150-152°C
	T8	Methanol	Block, colorless	160-162°C
Group 2	T9	Methanol	Plate, yellow	213-214°C
	T10	Methanol	Plate, yellow	212-213°C
	T11	Methanol	Blocks, colorless	199-201°C
	T12	Methanol	Blocks, colorless	213-214°C

4.3 Results

4.3.1 Molecular electrostatic potentials (MEPs)

The twelve custom-designed thiazoles used in this study all have hydrogen-bond acceptors (amide carbonyl O, thiazole S, and thiazole N) in **T1-T8**, an additional hydrogen-bond acceptor (pyridine N) in **T9-T12** and one hydrogen-bond donor (an amide N-H). The ranking of the acceptor sites and a donor site is established using molecular electrostatic potential surfaces and the results from the density functional theory (DFT) calculations are shown in Figure 4.2 and 4.3. Higher negative potential on the acceptor suggests better hydrogen-bond accepting ability.
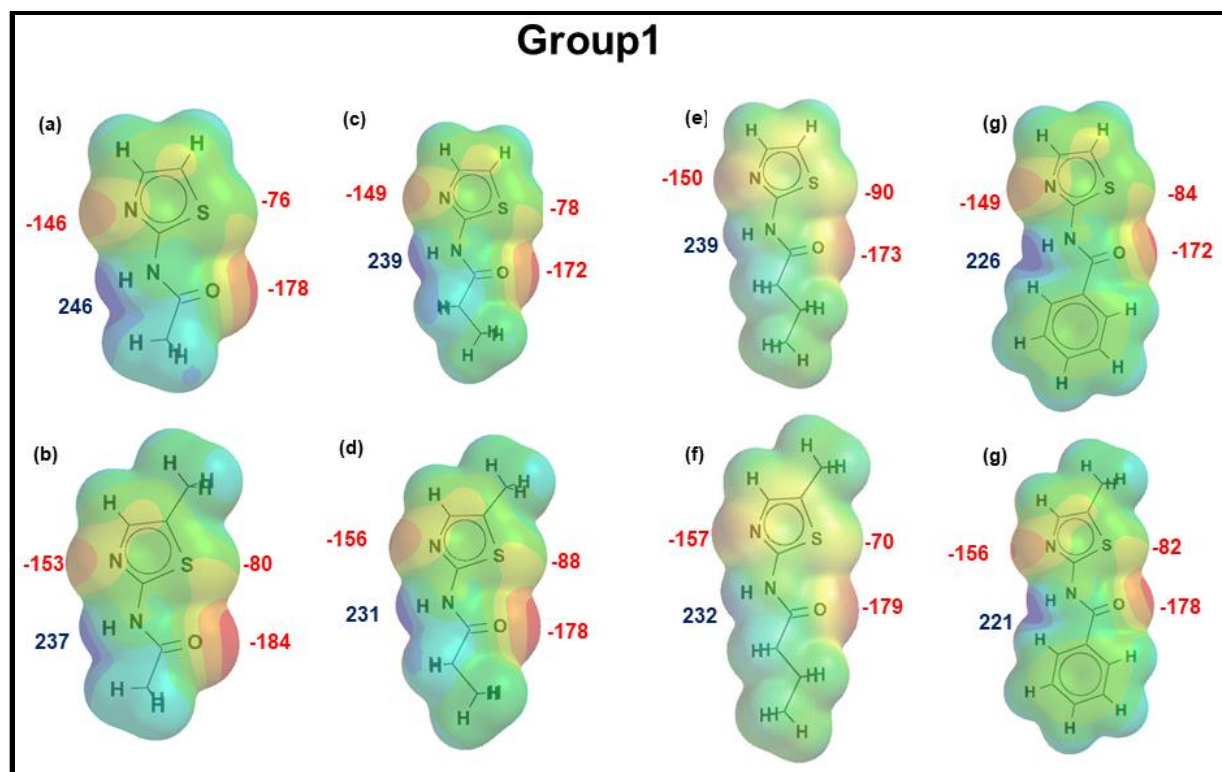
Figure 4.2 Electrostatic potentials values (in kJ/mol); a) **T1** b) **T2** c) **T3** d) **T4** e) **T5** f) **T6** g) **T7** and h) **T8**
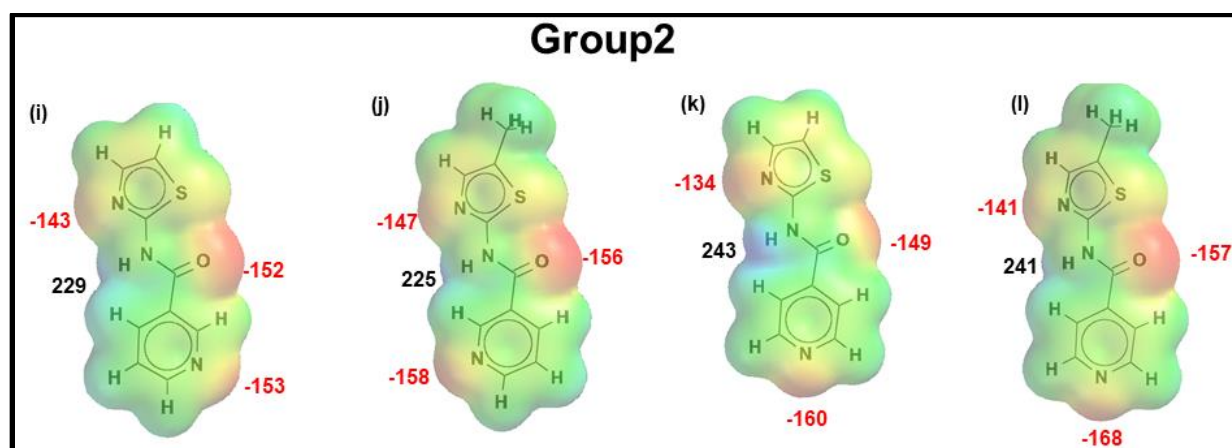
Figure 4.3 Electrostatic potentials values (in kJ/mol); i) **T9** j) **T10** k) **T11** and l) **T12**

4.3.2 Hydrogen-bond energies (HBE)

The calculated hydrogen-bond energy for individual synthons in **T1-T12** is tabulated in the table 4.5.

Table 4.5 Hydrogen-bond energies for monomeric (per hydrogen bond) and dimeric (per synthon) synthon A, synthon B, synthon C and Synthon D using equation 3.

Target molecules (T1-T12)		Synthon A		Synthon B		Synthon C	synthon D
		Monomeric	Dimeric	Monomeric	Dimeric	Monomeric	
		NH…N	NH…N	NH…S	NH…S	NH…C=O	NH…N
Group 1	T1	-14.96	-29.92	-5.19	-10.38	-21.02	N/A
	T2	-15.44	-30.88	-5.35	-10.71	-21.23	N/A
	T3	-14.92	-29.84	-5.20	-10.41	-19.09	N/A
	T4	-15.45	-30.89	-6.02	-12.03	-19.39	N/A
	T5	-15.44	-30.88	-5.35	-10.71	-21.23	N/A
	T6	-14.92	-29.84	-5.20	-10.41	-19.09	N/A
	T7	-13.89	-27.78	-5.44	-10.87	-17.77	N/A
	T8	-14.60	-29.19	-5.09	-10.18	-17.97	N/A
	AVG.	**-14.95**	**-29.90**	**-5.36**	**-10.71**	**-19.60**	**N/A**
	Std.dev	**0.53**	**1.06**	**0.29**	**0.58**	**1.41**	**N/A**
Group 2	T9	-13.17	-26.35	-1.79	-3.57	-14.61	-14.78
	T10	-13.50	-26.99	-3.23	-6.47	-14.93	-15.26
	T11	-12.74	-25.48	-0.81	-1.61	-15.24	-17.21
	T12	-13.73	-27.47	-1.22	-2.44	-16.49	-18.71
	AVG.	**-13.29**	**-26.57**	**-1.76**	**-3.52**	**-15.32**	**-16.49**
	Std.dev	**0.43**	**0.86**	**1.06**	**2.12**	**0.82**	**1.82**

4.3.3 Hydrogen-bond propensities (HBP)

The propensity calculations for the twelve target molecules consider all possible pairwise interactions that may take place between one of the three acceptors in group1 (four acceptors in group 2) and the sole hydrogen-bond donor, N-H (amide) are shown in Table 4.6.

Table 4.6 Hydrogen-bond propensities (larger value indicates increased likelihood of formation) for each postulated synthon in T1-T12. All propensities are given for individual hydrogen bonds whereas hydrogen-bond energies are given per synthon.

Hydrogen-bond propensities (Lower bound, upper bound)					
		Synthon A	Synthon B	Synthon C	Synthon D
Group1	T1	0.66 (0.54, 0.76)	0.02 (0.01, 0.05)	0.50 (0.39, 0.61)	N/A
	T2	0.60 (0.47, 0.71)	0.02 (0.01, 0.04)	0.47 (0.36, 0.59)	N/A
	T3	0.65 (0.56, 0.74)	0.03 (0.01, 0.06)	0.42 (0.30, 0.55)	N/A
	T4	0.61 (0.50, 0.71)	0.02 (0.01, 0.05)	0.43 (0.31, 0.56)	N/A
	T5	0.66 (0.56, 0.74)	0.03 (0.01, 0.06)	0.37 (0.26, 0.51)	N/A
	T6	0.60 (0.49, 0.70)	0.02 (0.01, 0.05)	0.37 (0.26, 0.51)	N/A
	T7	0.46 (0.29, 0.64)	0.01 (0.00, 0.03)	0.12 (0.05, 0.24)	N/A
	T8	0.42 (0.24, 0.61)	0.01 (0.00, 0.02)	0.12 (0.05, 0.24)	N/A
Group 2	T9	0.58 (0.43, 0.72)	0.02 (0.01, 0.06)	0.17(0.09, 0.32)	0.64 (0.45, 0.80)
	T10	0.50 (0.35, 0.65)	0.02 (0.01, 0.05)	0.17(0.08, 0.30)	0.63 (0.44, 0.78)
	T11	0.57 (0.42, 0.71)	0.02 (0.01, 0.06)	0.16 (0.08, 0.30)	0.67 (0.49, 0.81)
	T12	0.51 (0.36, 0.65)	0.02 (0.01, 0.05)	0.16 (0.08, 0.29)	0.66 (0.49, 0.80)

4.3.4 Hydrogen-bond coordination (HBC)

Hydrogen-bond coordination tool was used to determine the coordination of each functional group for the most optimal predicted synthon, Table 4.7. In group1, the most optimal motif is predicted to have synthon **A** and in group2, both synthon **A** and synthon **D** are possible, Figure 4.4.

Table 4.7 Hydrogen-bond coordination for each functional group T1-T12.

		Hydrogen-bond coordination					
		NH (donor)	N (thiazole)	S (thiazole)	C=O	N (pyridine)	which synthon is preferred?
Group1	T1-T8	1	1	0	0	N/A	Synthon A
Group 2	T9-T12	1	1	0	0	0	Synthon A
		1	0	0	0	1	Synthon D

(a)   **(b)**
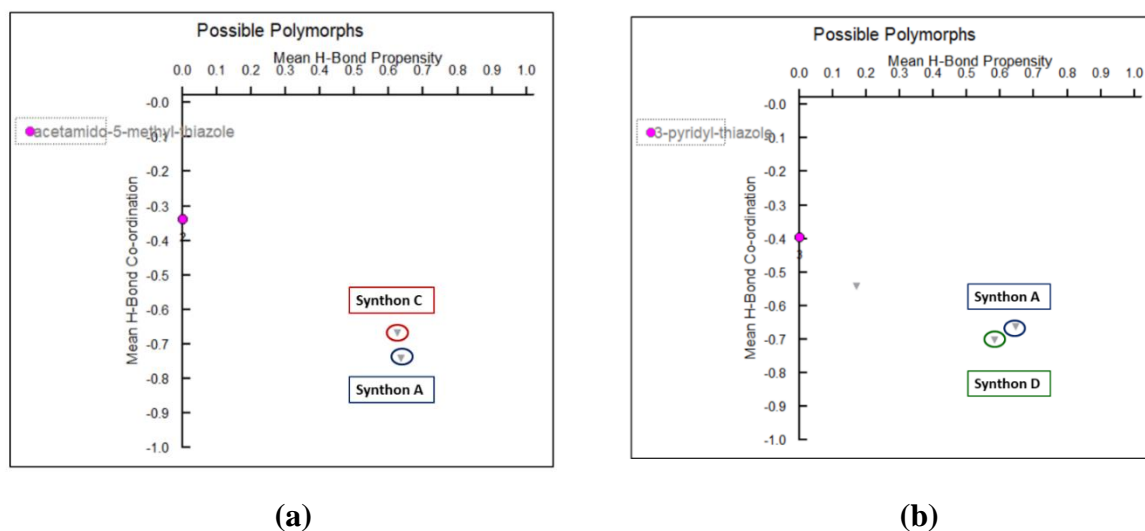
Figure 4.4 Synthon predictions using hydrogen-bond coordination for (a) T2 (group 1) and (b) T9 (group 2).

4.3.5 Experimentally observed crystal structures

We obtained crystal structures for **T2, T4** and **T6** (crystal structures for **T1** (YODJAD)[39] and **T5** (NORLAI)[40] have previously been reported) in group 1 and in all five cases, synthon **A** (with a graph-set notation of $R_2^2(8)$) was observed, Figure 4.5. The only conventional hydrogen-bond donor in this group of molecules, the N-H amide moiety, selects the N(thiazole) acceptor site as

its preferred partner. In group 2, two types of synthons are observed, synthon **D** is observed for meta substituted pyridine (**T9, T10**) and synthon **A** is observed for para substituted pyridine (**T11**), Figure 4.6.

Group 1

| T1 | T2 | T4 | T5 | T6 |
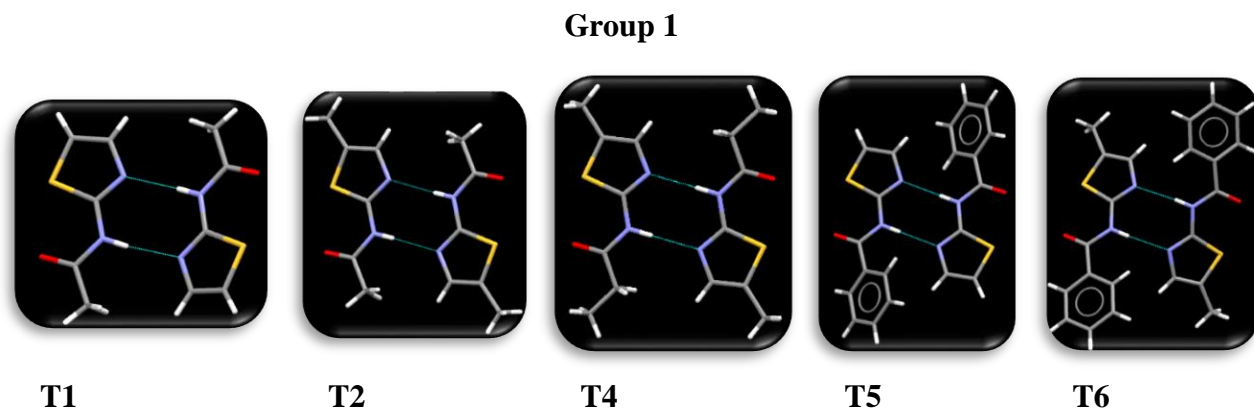
Figure 4.5 Synthon A appears in the crystal structures of T1-T2, T4, and T7-T8.

Group 2

| T9 | T10 | T11 |
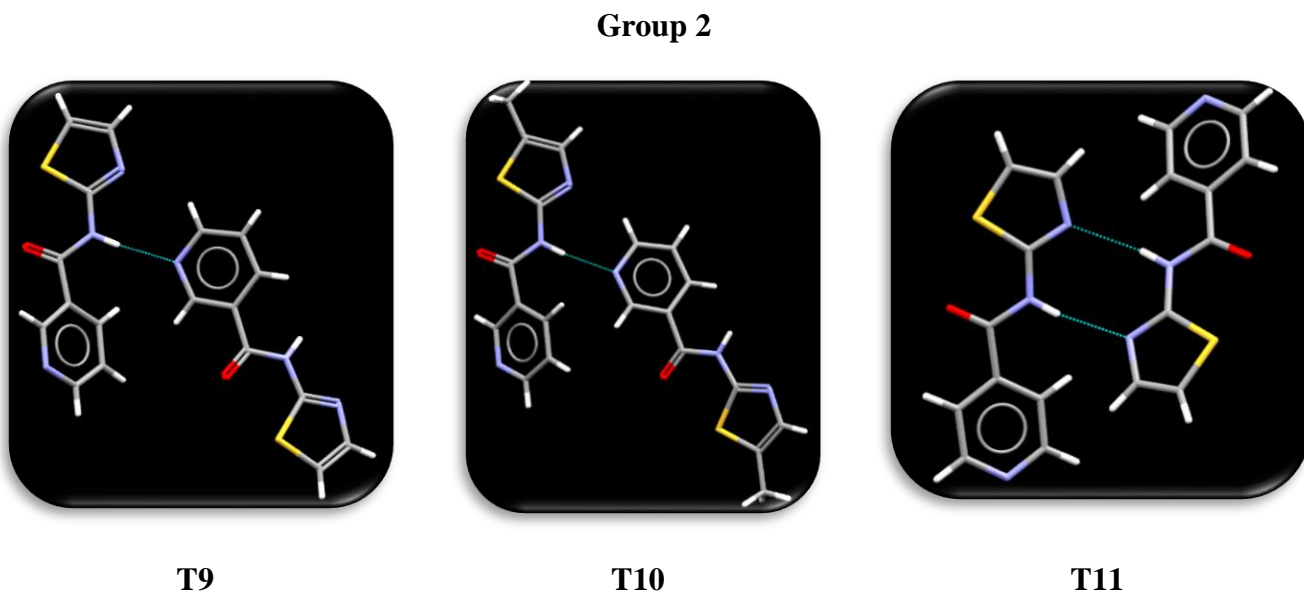
Figure 4.6 Synthon D appears in the crystal structures of T9, T10 and synthon A appears in the crystal structure of T11.

4.4 Discussion

4.4.1 What are the predicted synthons?

Based on four methods (MEPS, HBE, HBP, and HBC) used to predict synthons, three (HBE, HBP and HBC) out of four methods predicted synthon **A** in group1 (Figure 4.7a). In group 2, two out

of four methods (HBE and HBC) predicted synthon **A** and three out of four methods (MEPS, HBP and HBC) predicted synthon **D** as the most optimal synthons (Figure 4.7b).

Group1

MEPs Synthon C	**HBE** Synthon A
HBP Synthon A	**HBC** Synthon A

Group 2

MEPs Synthon D	**HBE** Synthon A
HBP Synthon D	**HBC** Synthon A or Synthon D
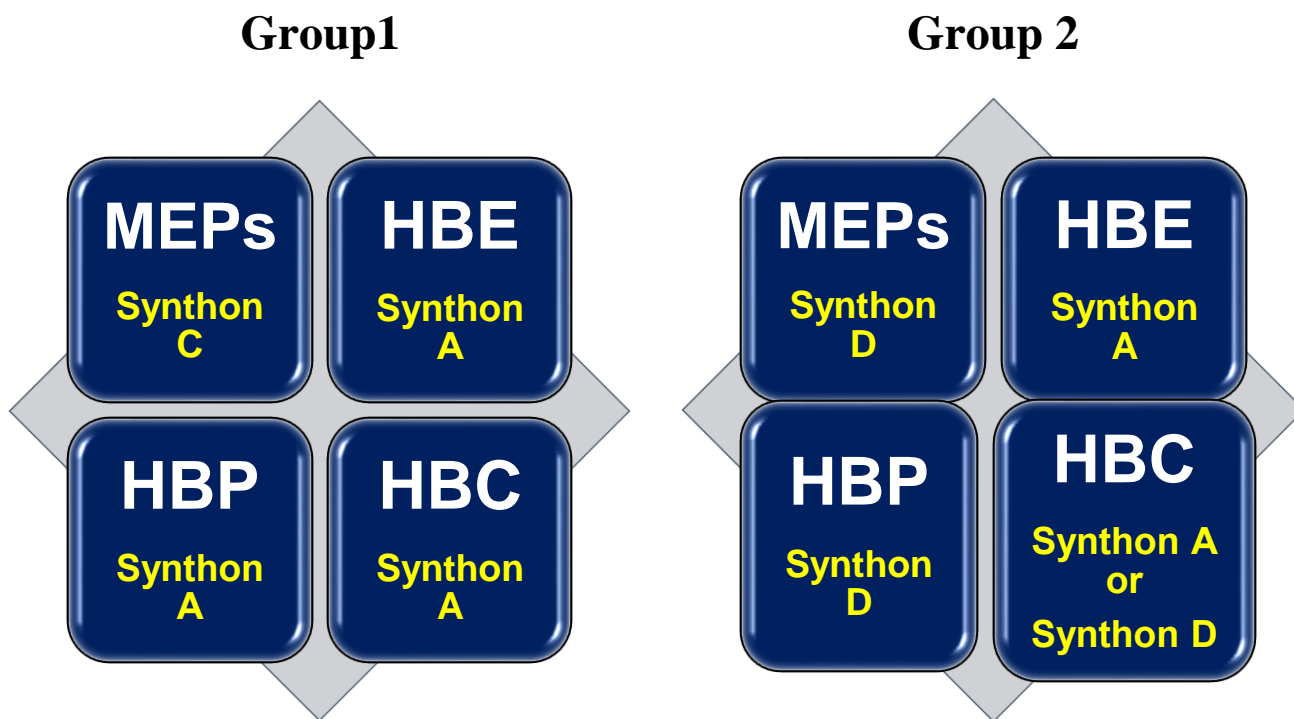
Figure 4.7 Synthon predictions by each method in (a) Group 1 and (b) Group 2.

4.4.2 Predicted vs experimental data

Group 1: In total, five out of eight crystal structures (**T1**, **T2**, **T4**, **T7**, and **T8**) of target molecules were obtained in group 1 and each crystal structure showed synthon **A** as the most preferred binding site. The preference for synthon **A** (NH(amide)**….**N(aromatic))over synthon **B** (NH(amide)…S(aromatic) in group 1 is readily explained on the basis of electrostatics. As there is no geometric bias for either motif, the outcome appears to be determined by the fact that the sulfur acceptor atom with its significantly lower (less negative) electrostatic potential value, is not competitive relative to the thiazole ring nitrogen atom. This is also born out in the propensity analysis which shows very a low probability for the occurrence of the hydrogen bonds in synthon **B**, implying that is not as frequently observed in crystal structures. Even though individual N-H**…**N interactions (in synthon **A**) are weaker than individual N-H**…**O=C hydrogen bonds (in synthon **C**) based on hydrogen-bond energies, synthon **A** gains an advantage as a result of the

geometry and relative position of the donors/acceptors on these molecules which facilitates the appearance of two hydrogen bonds through a supramolecular 'chelate' effect. HBE, HBP and HBC predicted the observed synthon correctly (Figure 4.8).

Group 1 (T1-T8)

MEPs	HBE	HBP	HBC	Experimental
Synthon C 0/5	Synthon A 5/5	Synthon A 5/5	Synthon A 5/5	Synthon A
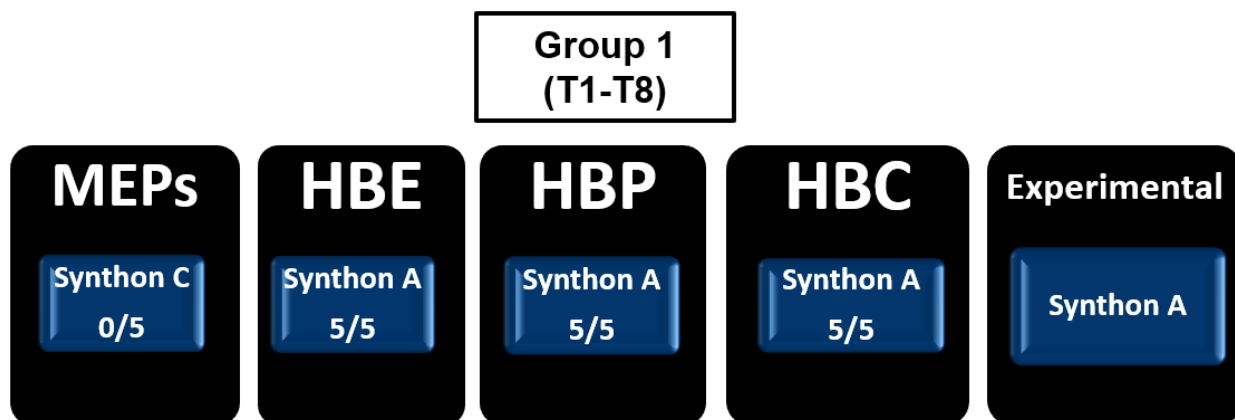
Figure 4.8 Predicted vs experimental comparison of homomeric synthons in group 1 molecules using different prediction methods.

Group 2: Three out of four crystal structures (**T9, T10,** and **T11**) were obtained in this group. The meta substituted **T9-T10** forms monomeric synthon **D** (predicted based on MEPS, HBP and HBC) and para substituted **T10** molecule forms dimeric synthon **A** (predicted based on HBE and HBC). Both synthon **A** (preferred due to supramolecular chelating effect) and synthon **D** (preferred due to strong pyridine site) are preferred over synthon **B** or **C** because of above reasons.

Both synthons are predicted by four methods therefore, its wasn't surprising to see these synthons (Figure 4.9).

Group 2 (T9-T12)

MEPs	HBE	HBP	HBC	Experimental
Synthon D 2/3	Synthon A 1/3	Synthon D 2/3	Synthon A Synthon D 3/3	Synthon A and Synthon D

111
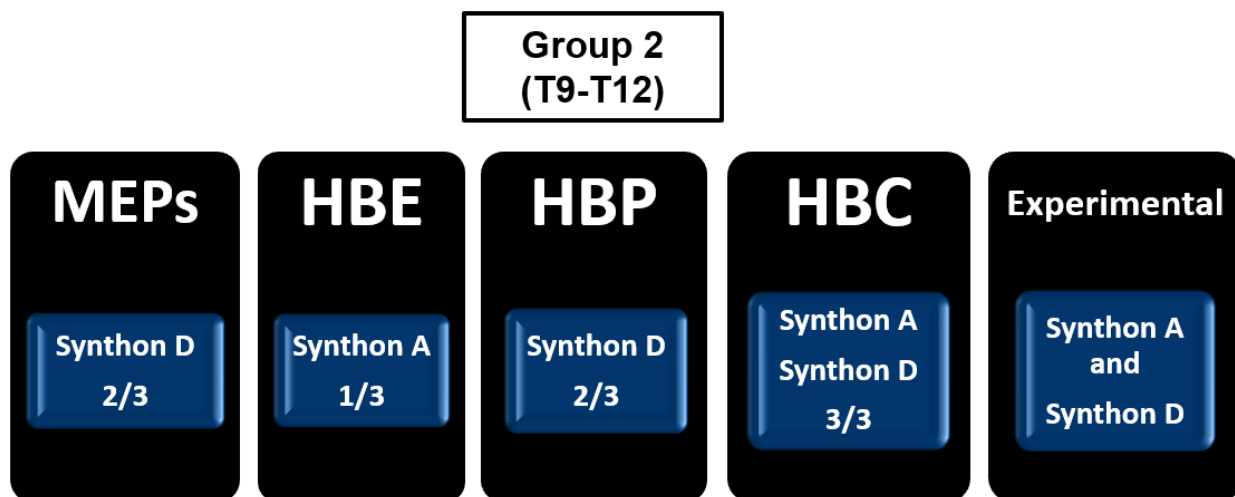
Figure 4.9 Predicted vs experimental comparison of homomeric synthons in group 2 molecules using different prediction methods.

We hypothesize that such dimeric assemblies can most likely compensate for individually weaker interactions in a substantial number of cases but only if the electrostatic differences between the possible acceptors is not too large. It is also possible that this energetic bias is most important during crystal growth, where a dimeric synthon **A** growth unit would be preferred over synthon **C**. It is of course not too surprising that a pair of weak interactions can overcome a single interaction that is thermodynamically more favorable. However, it is a real challenge to predict exactly where the energetic 'tipping point' may reside and this is why a combined MEPS/HBP approach is potentially very powerful. If/when both methods agree (*i.e.* they rank the order of interactions in the same manner), this likely means that combinations of interactions have limited ability to overcome the stabilization of isolated stronger interactions. If the HBP shows that weaker interactions are represented in the statistics, it would indicate that crystal packing forces can tilt the balance in favor of the weaker interactions.

4.5 Conclusions

The structural chemistry of twelve target molecules has been analyzed using MEPS, HBE, HBP and HBC as a way of predicting which intermolecular interactions are likely to appear in the solid state. The molecules were divided into two groups based on hydrogen-bond functionalities.

1. In group 1, based on HBE, HBP and HBC, synthon **A** was predicted to be favored over synthons **B** and **C**. Based on group 2, synthon **A** and synthon **D** were predicted by combination of four methods.
2. The combination of prediction methods is successful over individual methods.
3. In group 1, Synthon **A** was observed experimentally in all five crystal structures. In group 2, both predicted synthons **A** and **D** were observed experimentally. The four methods accurately predicted the experimentally observed intermolecular hydrogen-bond interactions in three crystal structures; Synthon **A** was observed in **T11** and synthon **D** was observed in **T9-T10**.
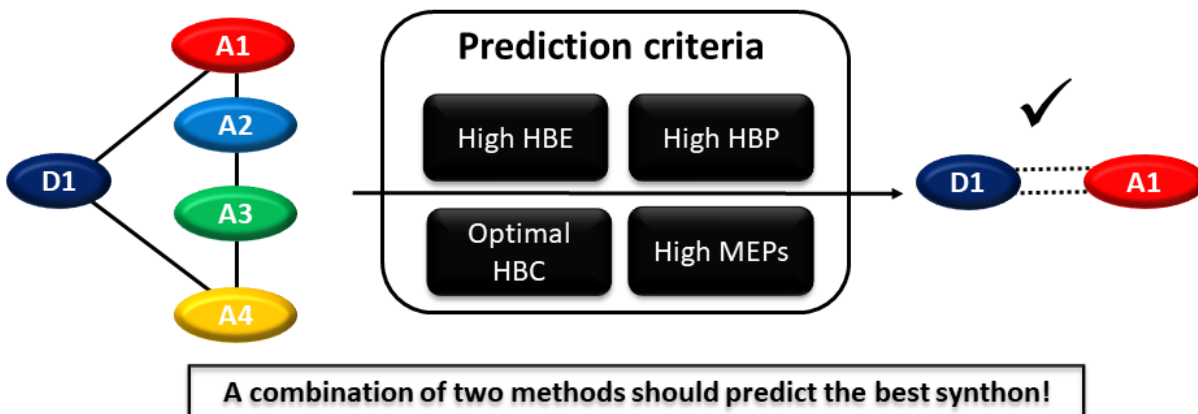
Figure 4.10 Results of four prediction methods suggesting combination of atleast two methods is valuable in making predictions.

The ability to make informed predictions about how different molecules recognize and bind to each other, and how they subsequently assemble into solid-state architectures is of critical importance in areas such as drug design and formulation, and the recently published guidance by the FDA on the regulatory classification of pharmaceutical co-crystals[41] will continue to keep this fundamental science in sharp focus. Therefore, the successful use of structural informatics tools such as hydrogen-bond propensity and calculated molecular electrostatic potential surfaces for mapping out the structural landscape of these types of molecules will have significant practical applications.

4.6 References

1.   Desiraju, G. R., *Angewandte Chemie International Edition in English* **1995,** *34* (21), 2311-2327.

2.   Shattock, T. R.; Arora, K. K.; Vishweshwar, P.; Zaworotko, M. J., *Crystal Growth & Design* **2008,** *8* (12), 4533-4545.

3.   Dey, A.; Pati, N. N.; Desiraju, G. R., *CrystEngComm* **2006,** *8* (10), 751-755.

4.   Aakeröy, C. B.; Epa, K., In *Electronic Effects in Organic Chemistry*, Kirchner, B., Ed. Springer Berlin Heidelberg: Berlin, Heidelberg, 2014; pp 125-147.

5.   Nangia, A.; Desiraju, G. R., Supramolecular Synthons and Pattern Recognition. In *Design of Organic Solids*, Weber, E.; Aoyama, Y.; Caira, M. R.; Desiraju, G. R.; Glusker, J. P.;

Hamilton, A. D.; Meléndez, R. E.; Nangia, A., Eds. Springer Berlin Heidelberg: Berlin, Heidelberg, 1998; pp 57-95.

6. Thalladi, V. R.; Goud, B. S.; Hoy, V. J.; Allen, F. H.; Howard, J. A. K.; Desiraju, G. R., *Chemical Communications* **1996,** (3), 401-402.

7. Reddy, D. S.; Craig, D. C.; Desiraju, G. R., *Journal of the American Chemical Society* **1996,** *118* (17), 4090-4093.

8. Lehn, J.-M., *Science* **2002,** *295*.

9. Graham Smith, K. E. B., Karl A. Byriel, Colin H.L. Kennard, *Aust. J. Chem.* **1997,** *50*, 727-736.

10. Bosch, E., *CrystEngComm* **2007,** *9* (3), 191-198.

11. A. D. Buckingham, J. E. D. B. a. S. A. C. M., *Chem. Phys. Lett.* **2008,** *463*, 1–10.

12. Merz, K.; Vasylyeva, V., *CrystEngComm* **2010,** *12* (12), 3989-4002.

13. Bedeković, N.; Stilinović, V.; Piteša, T., *Crystal Growth & Design* **2017,** *17* (11), 5732-5743.

14. Lehn, J.-M., *Rep. Prog. Phys.* **2004** (67), 249.

15. Moulton, B.; Zaworotko, M. J., *Chemical Reviews* **2001,** *101* (6), 1629-1658.

16. Aitipamula, S.; Nangia, A., Polymorphism: Fundamentals and Applications. In *Supramolecular Chemistry*, John Wiley & Sons, Ltd: 2012.

17. Nauha, E.; Bernstein, J., *Crystal Growth & Design* **2014,** *14* (9), 4364-4370.

18. Etter, M. C., *Accounts of Chemical Research* **1990,** *23* (4), 120-126.

19. Desiraju, G. R., *Current Opinion in Solid State and Materials Science* **1997,** *2* (4), 451-454.

20. Braga, D., *Angewandte Chemie International Edition* **2012,** *51* (15), 3516-3516.

21. Aakeröy, C. B.; Desper, J.; Leonard, B.; Urbina, J. F., *Crystal Growth & Design* **2005,** *5* (3), 865-873.

22. Aakeroy, C. B.; Desper, J.; Urbina, J. F., *CrystEngComm* **2005,** *7* (31), 193-201.

23. Wood, P. A.; Feeder, N.; Furlow, M.; Galek, P. T. A.; Groom, C. R.; Pidcock, E., *CrystEngComm* **2014,** *16* (26), 5839-5848.

24. Peterson, M. L.; Stanton, M. K.; Kelly, R. C.; Staples, R.; Cheng, A., *CrystEngComm* **2011,** *13* (4), 1170-1180.

25. Stanton, M. K.; Bak, A., *Crystal Growth & Design* **2008,** *8* (10), 3856-3862.

26. Cheney, M. L.; Weyna, D. R.; Shan, N.; Hanna, M.; Wojtas, L.; Zaworotko, M. J., *Crystal Growth & Design* **2010,** *10* (10), 4401-4413.

27. Tumanov, N. A.; Myz, S. A.; Shakhtshneider, T. P.; Boldyreva, E. V., *CrystEngComm* **2012,** *14* (1), 305-313.

28. Suresh, K.; Mannava, M. K. C.; Nangia, A., *Chemical Communications* **2016,** *52* (22), 4223-4226.

29. Félix-Sonda, B. C.; Rivera-Islas, J.; Herrera-Ruiz, D.; Morales-Rojas, H.; Höpfl, H., *Crystal Growth & Design* **2014,** *14* (3), 1086-1102.

30. Taylor, R.; Macrae, C. F., *Acta Crystallographica Section B* **2001,** *57* (6), 815-827.

31. Bruno, I. J.; Cole, J. C.; Edgington, P. R.; Kessler, M.; Macrae, C. F.; McCabe, P.; Pearson, J.; Taylor, R., *Acta Crystallographica Section B* **2002,** *58* (3 Part 1), 389-397.

32. Macrae, C. F.; Edgington, P. R.; McCabe, P.; Pidcock, E.; Shields, G. P.; Taylor, R.; Towler, M.; van de Streek, J., *Journal of Applied Crystallography* **2006,** *39* (3), 453-457.

33. 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., *Journal of Applied Crystallography* **2008,** *41* (2), 466-470.

34. Chattopadhyay, G. C., S.; Saha, C. Synth. Commun., 2008, 38:23, 4068-4075.

35. Wang, X. J. Y., Q.; Liu, F.; You, Q.D.Synth. Commun., 2008, 38:7, 1028-1035.

36. Musumeci, D.; Hunter, C. A.; Prohens, R.; Scuderi, S.; McCabe, J. F., *Chemical Science* **2011,** *2* (5), 883-890.

37. Galek, P. T. A.; Chisholm, J. A.; Pidcock, E.; Wood, P. A., *Acta Crystallographica Section B* **2014,** *70* (1), 91-105.

38. Galek, P. T. A.; Allen, F. H.; Fabian, L.; Feeder, N., *CrystEngComm* **2009,** *11* (12), 2634-2639.

39. Yunus, U.; Tahir, M. K.; Bhatti, M. H.; Wong, W.-Y., *Acta Crystallographica Section E* **2008,** *64* (8), o1516.

40. Zonouzi, A.; Mirzazadeh, R.; Rahmani, H.; Ng, S. W., *Acta Crystallographica Section E* **2009,** *65* (4), o817.

41. Gadade, D. D.; Pekamwar, S. S., *Advanced Pharmaceutical Bulletin* **2016,** *6* (4), 479-494.

Chapter 5 - Evaluating Competing Heteromeric Interactions through Electrostatics, Energies and Propensities in Thiazole based molecules

5.1 Introduction

In Chapter 4, we analyzed the possible homomeric interactions in the thiazole based molecules using four different methods to predict and compare the homomeric synthons in these molecules with the experimental crystal structures. In this chapter, we want to try to break the homomeric interactions (synthons **A, D, E**) by introducing a suitable carboxylic acid that may favor heteromeric interactions that can drive co-crystal formation (synthons **F-I**, Scheme 5.1).
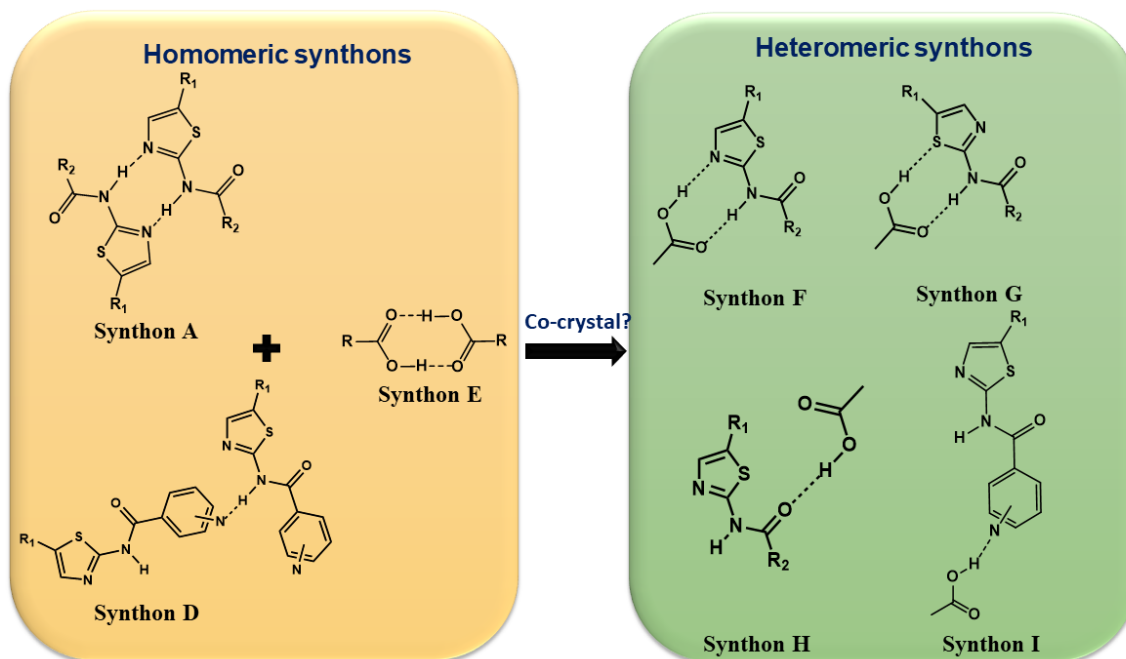
Scheme 5.1. Three postulated heteromeric hydrogen-bond based synthons between a thiazole-amide and a carboxylic acid; (R$_1$=H, CH$_3$; R$_2$= methyl, ethyl, or benzyl).

Our working strategy for the synthesis of co-crystals is based on optimizing the affinity between different molecules bearing complementary hydrogen-bonding moieties, thereby controlling the balance between homomeric (re-crystallization) and heteromeric (co-crystallization) outcomes. The complete sets of targets (thiazole amides) and probes (carboxylic acids) are listed in Scheme 5.2 and 5.3. Molecules **T1**-**T12** have systematically increasing molecular weight and they also display specific and controllable changes to the charge density on the synthon. By exploring the structural landscape of this family of closely related molecules, we may be able to highlight the relative influence of small structural differences. Likewise, for the co-formers, the aliphatic carboxylic acids show increasing chain length and molecular flexibility. Similarly, the aromatic acids explore the addition of competing hydrogen-bonding groups and the influence of substituent groups on the aromatic ring.
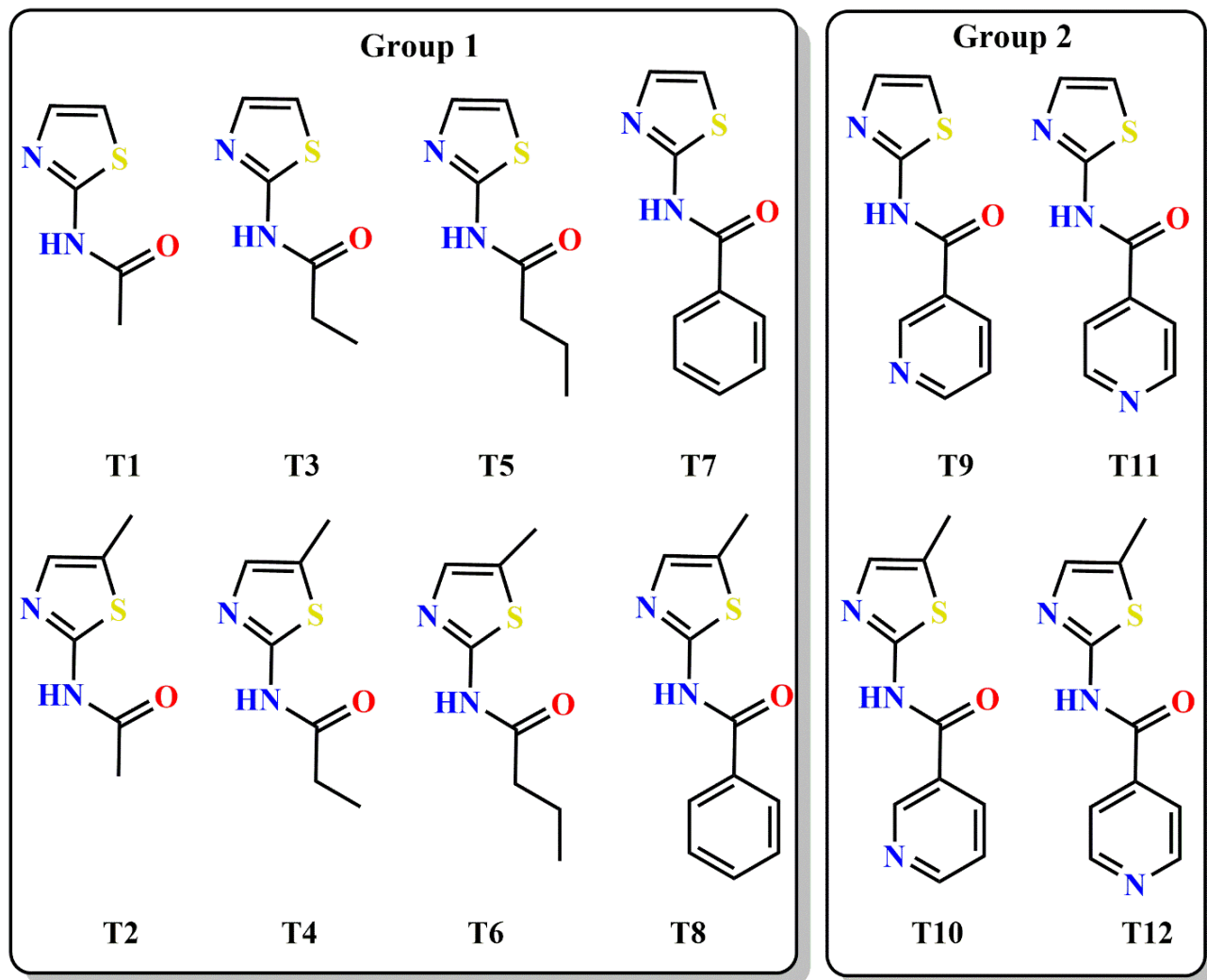
Scheme 5.2 Twelve thiazole amides used in this study.

Malonic acid (Mal)   Glutaric acid (Glu)   Pimelic acid (Pim)   Azelaic acid (Aze)

Fumaric acid (Fum)   Succinic acid (Suc)   Adipic acid (Adi)   Suberic acid (Sub)

Sebacic acid (Seb)   Dodecanedioic acid (Dod)

3Hydroxy BA   4Hydroxy BA   3Amino BA   4Amino BA   3Nitro BA

4Nitro BA   BA   4Iodo BA   4Bromo BA   PentaFBA
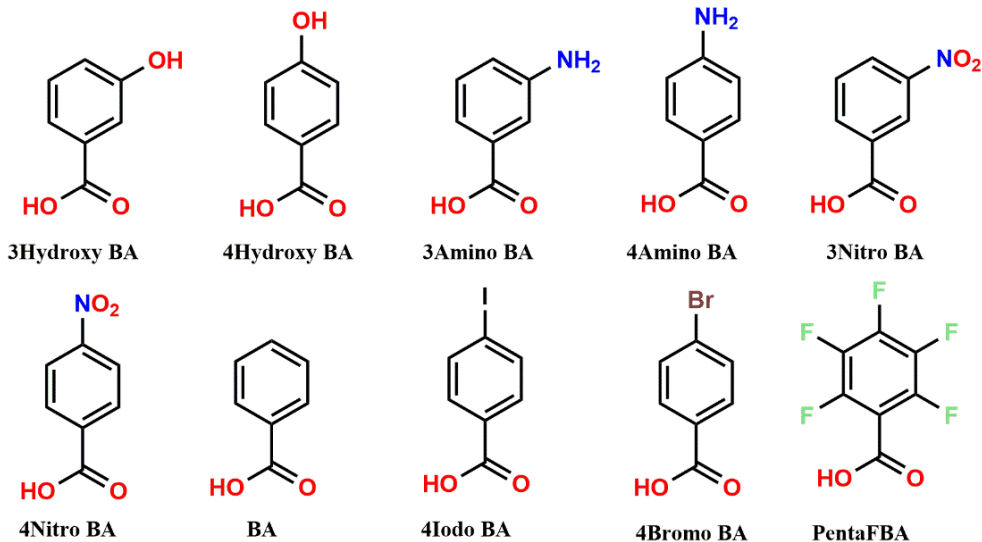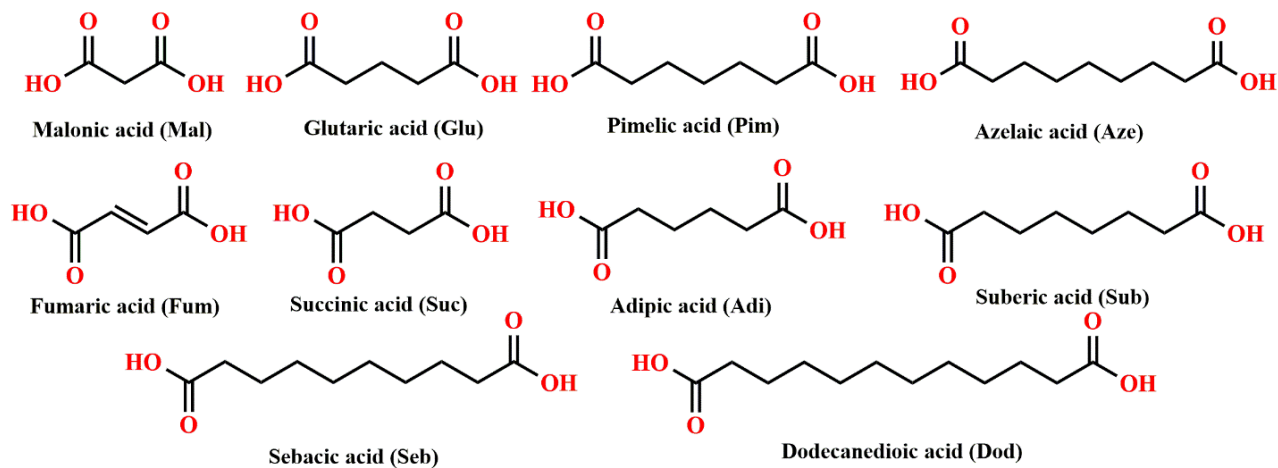
Scheme 5.3 Twenty carboxylic acid co-formers used in this study.

The study is undertaken to answer three important questions.

1. Can we find a reliable method for predicting co-crystallization in a set of target molecules **T1-T12** with 20 carboxylic acids?
2. Can we predict the correct heteromeric synthons in the co-crystals?
3. Which method is best for predicting heteromeric-synthons?

5.2 Experimental

5.2.1 Materials

The synthesis and characterization of target molecules (**T1-T12**) were described chapter 5. Carboxylic acids were purchased from commercial sources and used as received. Melting points were measured using a Fisher-Johns melting point apparatus. Solution ^{1}H NMR data were collected in DMSO-d$_6$ on a Varian Unity plus 400 MHz NMR spectrometer. IR spectra of co-crystal screening experiments were recorded with a Nicolet 380 FT-IR spectrometer using ATR (attenuated total reflection) technique and ZnSe as the crystal.

5.2.2 IR Co-crystal screening

T1-T12 were put through a co-crystal screen as explained in section 3.2.5. The detailed IR analysis is provided in Appendix C.
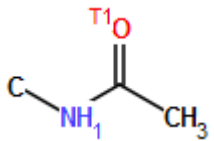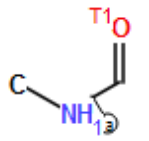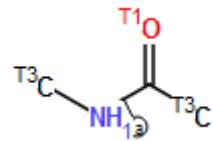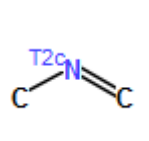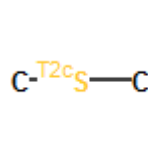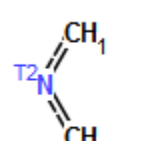
5.2.3 Molecular electrostatic potential (MEPs) calculations
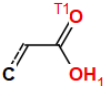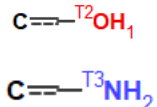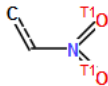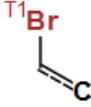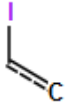
See section 2.2. for details of molecular electrostatic potential surfaces (MEPs) calculations. The results of MEPs of **T1**–**T12** are shown in section 4.3.1. The electrostatics potentials of 20 carboxylic acids are shown in appendix E.

5.2.4 Hydrogen-bond propensities (HBP) for predicting co-crystal outcome

The detailed description of HBP calculations of co-crystal screening is described in section 3.2.4. The description of functional groups used are shown in Table 5.1.

Table 5.1 Functional groups used to determine the hydrogen-bond propensities for the six target molecules. The labels in the figures can be explained as follows: Tn = atom makes n bonds, c = atom is cyclic, ⑥ = bond is acyclic, and Hn = n bonded hydrogen atoms.

T1/T2	T3/T4/T5/T6	T7-T12	T1-T12	T1-T12	T9-T12

Aliphatic	Aromatic	HydroxyBA, AminoBA	3NitroBA, 4NitroBA	BromoBA	IodoBA

5.2.5 Hydrogen-bond energies (HBE) for predicting structural outcomes

See section 2.2.5 for methodology for the hydrogen-bond energy calculation for synthon prediction and section 3.2.3 for co-crystal screen.

5.2.6 Single crystal growth and X-ray crystallography

All 240 combinations from the grinding experiments were dissolved in a minimum amount of methanol or nitromethane (1–2 ml, both reactants were readily soluble in these solvents) and kept in a vial for slow evaporation in order to obtain crystals suitable for single crystal X-ray diffraction. Of 240 LAG reactions, 151 produced co-crystals according to IR data and fourteen of those yielded crystals suitable for single-crystal X-ray diffraction (Table 5.2). X-ray experimental data and crystallographic data are given in the Appendix.

Table 5.2 Experimental details for crystal structures obtained.

	Stoichiometric ratio	Solvent	Morphology/color	Melting point
T1-3HydroxyBA	1:1	Methanol	Rod, bronze	135-138 C
T2-4HydroxyBA	1:1	Methanol/ Nitromethane	Plate, colorless	176-178° C
T4-3NitroBA	1:1	Methanol	Plate, colorless	126-128° C
T8-Suc	2:1	Methanol	Plate, colorless	178-181° C
T8-Sub	2:1	Methanol	Plate, colorless	121-123° C
T8-Seb	2:1	Methanol	Plate, colorless	135-137° C
T8-Aze	2:1	Methanol	Prism, colorless	99-102° C
T8-Dod	2:1	Methanol	Prism, colorless	106-108° C
T9-3HydroxyBA	2:1	Methanol	Plates, yellow	185-195° C
T11-Sub	2:1	Methanol	Colorless, rectangular	135-137° C

5.3 Results

5.3.1 IR co-crystal screening experiments

The experimental screen produced co-crystals in 151 of the 240 experiments between **T1-T12** and aliphatic diacids and aromatic monoacids. The results of the IR grinding screen are tabulated in Table 5.3. An example of how the IR was analyzed by comparing the grinded mixture to the individual pure components is shown in Figure 5.1.

Table 5.3 Attempted co-crystallizations using LAG (methanol) of T1-T12 with aliphatic and aromatic acids (red indicates negative co-crystal outcome and green indicates positive co-crystal outcome).

ACID	Group 1								Group 2				Success rate
	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T12	T12	
Suc	🟩	🟥	🟥	🟥	🟥	🟥	🟩	🟩	🟩	🟩	🟩	🟩	58%
Adi	🟥	🟥	🟥	🟥	🟥	🟩	🟩	🟩	🟩	🟩	🟩	🟩	58%
Sub	🟥	🟥	🟥	🟥	🟥	🟩	🟩	🟩	🟩	🟩	🟩	🟩	58%
Seb	🟥	🟥	🟥	🟥	🟥	🟥	🟩	🟩	🟩	🟩	🟩	🟩	50%
Dod	🟥	🟥	🟥	🟥	🟥	🟥	🟩	🟩	🟩	🟩	🟩	🟩	50%
Fum	🟥	🟩	🟥	🟥	🟥	🟩	🟩	🟩	🟩	🟩	🟩	🟩	67%
Mal	🟥	🟩	🟥	🟩	🟩	🟥	🟩	🟩	🟩	🟩	🟩	🟩	75%
Glu	🟥	🟥	🟥	🟥	🟥	🟩	🟩	🟩	🟩	🟩	🟩	🟩	58%
Pim	🟥	🟥	🟥	🟥	🟥	🟥	🟩	🟩	🟩	🟩	🟩	🟩	50%
Aze	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟩	🟩	🟩	🟩	🟩	42%
3-HydroxyBA	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	100%
4-HydroxyBA	🟩	🟩	🟥	🟩	🟥	🟩	🟩	🟩	🟩	🟩	🟩	🟩	83%
3-AminoBA	🟩	🟩	🟩	🟥	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	92%
4-AminoBA	🟩	🟩	🟥	🟩	🟥	🟥	🟥	🟩	🟩	🟩	🟩	🟩	67%
3-NitroBA	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	100%
4-NitroBA	🟥	🟥	🟥	🟥	🟥	🟥	🟩	🟩	🟩	🟩	🟩	🟩	50%
BA	🟥	🟥	🟥	🟩	🟥	🟩	🟩	🟩	🟩	🟩	🟩	🟩	67%
4-IodoBA	🟥	🟥	🟥	🟥	🟥	🟥	🟩	🟥	🟥	🟥	🟥	🟥	8%
4-BromoBA	🟥	🟥	🟥	🟥	🟥	🟥	🟩	🟥	🟩	🟥	🟥	🟥	16%
PentaFBA	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	100%
Positive outcomes	7	8	4	5	6	11	19	18	19	18	18	18	**151/240**
Success rate	35%	40%	20%	25%	30%	55%	95%	90%	95%	90%	90%	90%	**63%**
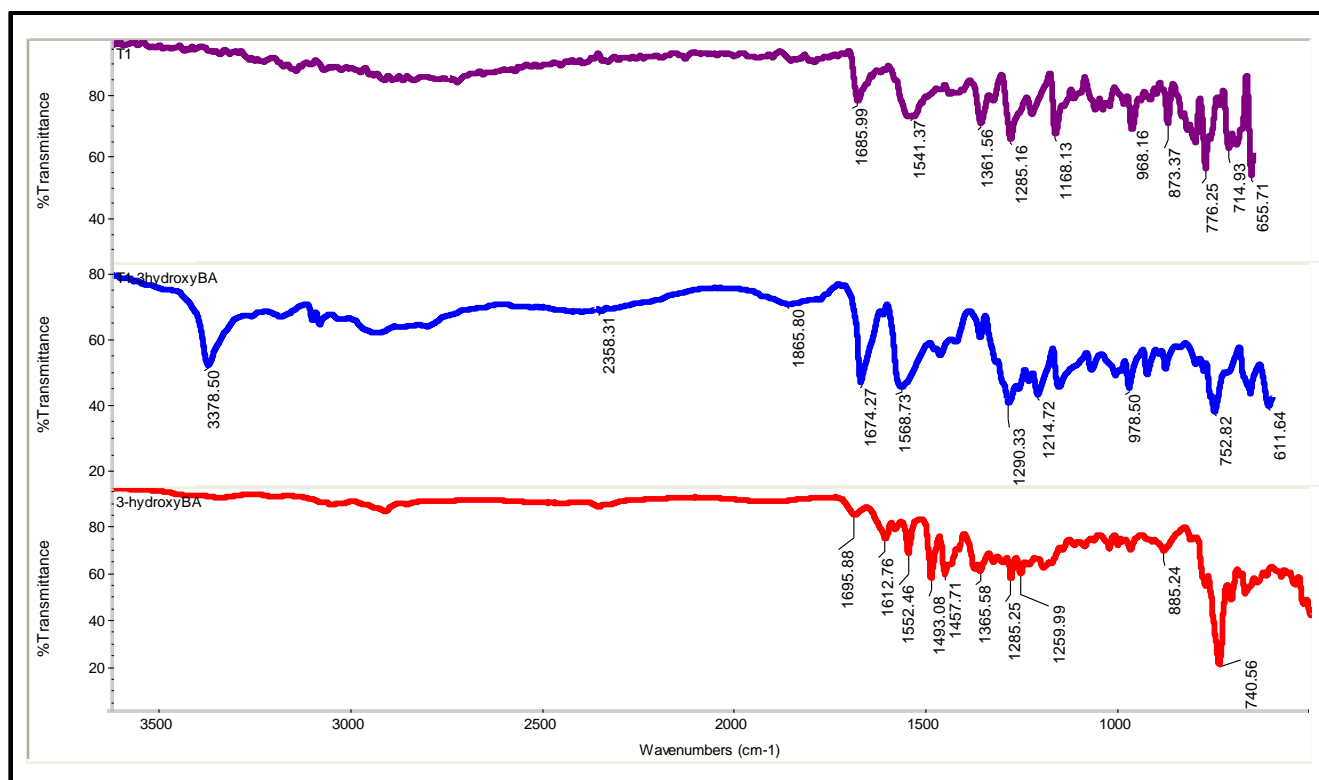
Figure 5.1 IR spectra of T1-3HydroxyBA **and the component starting materials** (T1 – **top;** T1-3HydroxyBA **co-crystal – middle;** 3HydroxyBA – **bottom).**
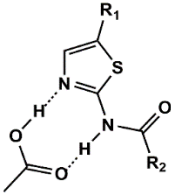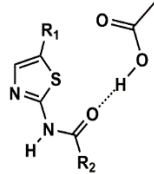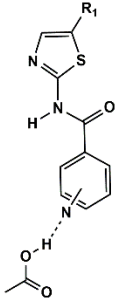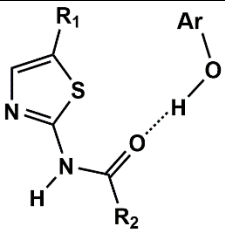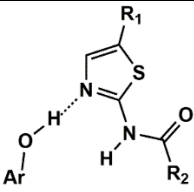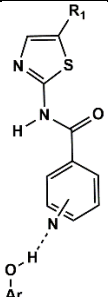
5.3.2 Hydrogen-bond propensities for co-crystal screening
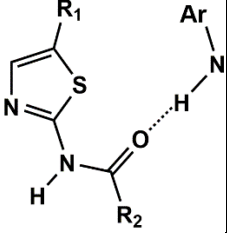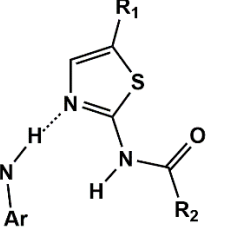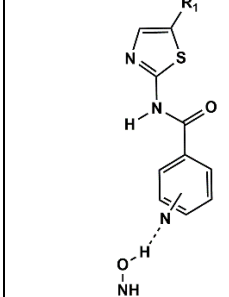
A hydrogen-bond propensity model was used to predict the outcome of co-crystal. The two highest propensity donor-acceptor interactions were taken into account for homomeric systems: N-H⋯N for the target molecule designated as T:T (target donor : target acceptor) and O-H⋯O=C or N-H⋯O=C (in the case of aminoBA) for the acid co-former designated as C:C (co-former donor : co-former acceptor). The interaction with the maximum propensity for heteromeric systems was designated as T:C (target donor : co-former acceptor) or C:T (co-former donor : target acceptor). An MC score (as described in section 2.4) is used to predict whether a reaction (in this case, co-crystal formation) will or will not take place when two reactants are combined. A summary of the multi-component hydrogen-bond propensity (HBP) screening results for **T1-T12** with 20 co-formers is given in the Table 5.4.

5.3.3 Using hydrogen-bond energies to predict heteromeric synthons

The hydrogen-bond energy for synthon **F, H** and synthon **I** in each co-crystal combination (240 combinations) was calculated and is tabulated in the appendix (only heteromeric interactions are considered) and the average energies are provided in Table 5.4. The larger (more negative) the hydrogen-bond energy, the stronger the interaction between the two different components favoring co-crystallization.

Table 5.4 Hydrogen -bond energies of each hetero-synthon in T1-T12 combined with aliphatic and aromatic acids.

All energies are displayed in kJ/mol			
	(a) Synthon F	**(b) Synthon H**	**(c) Synthon I**
Group 1 (w/o additional h-bond donor)	-35.54 ±0.63	-24.52±1.21	n/a
Group 2 (w/o additional h-bond donor)	-33.79±0.93	-19.94±1.26	-26.87±1.18
Group 1 (with additional h-bond donor)	-34.52±0.62	-21.53±1.06	n/a
Group 2 (with additional h-bond donor)	-33.19±0.81	-17.52±1.10	-19.37 ±0.85
	(d) Synthon J	**(e) Synthon K**	**(f) Synthon L**
Group 1 (with additional h-bond donor)	25.42 ± 1.72	19.43± 1.31	n/a
Group 2 (with additional h-bond donor)	20.68±1.40	16.72±1.13	22.87±1.54

		(g) Synthon M	(i) Synthon N	(j) Synthon O
Group 1 (with additional h-bond donor)		17.27±1.36	13.20±1.04	n/a
Group 2 (with additional h-bond donor)		14.05±1.11	11.36±0.90	15.53±1.23

5.3.4 Using electrostatics to predict heteromeric synthons

Based on Etter's rule guidelines, the best donor of target molecule binds to best acceptor of acid and the best donor of acid binds to best acceptor of target molecule and, therefore following synthons are preferred in each group. The electrostatic potentials of target molecules are provided in Chapter 4 and acids are provided in the appendix. The details of synthon preferred in each group are listed in Table 5.5.

Table 5.5 Synthon preferred in each group based on electrostatics.

		Best acceptor of target	Best donor of acid	Preferred synthon
Group 1	(w/o h-bond donor)	C=O	OH(acid)	Synthon H
	(hydroxyBA)	C=O	OH (Hydroxylic)	Synthon J
	(AminoBA)	C=O	OH(acid)	Synthon H
Group 2	(w/o h-bond donor)	N pyridine	OH(acid)	Synthon I
	(hydroxyBA)	N pyridine	OH (Hydroxylic)	Synthon L
	(AminoBA)	N pyridine	OH(acid)	Synthon H

5.3.5 Experimentally observed crystal structures

5.3.5.1 Group 1: T1-T8 against 20 carboxylic acids

In group 1, we obtained single-crystal XRD data for five co-crystals with aliphatic acids, and in all five crystal structures the carboxylic acid O-H····N thiazole and amide N-H····O=C acid moieties

give rise to the robust $R_2^2(8)$ hydrogen-bonded synthon **F** in a 2:1 target:acid stoichiometry, Figure 5.2d-h. Three crystal structures were obtained with aromatic acids with varying stoichiometry (1:2 ratio for **T1**-**3HydroxyBA**, 2:1 for **T2**-**4HydroxyBA** and 1:1 for **T4**-**3NitroBA**). As with the aliphatic acid co-crystals, a robust $R_2^2(8)$ hydrogen-bonded synthon **F** is formed using O-H····N and N-H····O hydrogen bonds in each case (Figure 5.2a-c). In **T1-3HydroxyBA**, the additional hydrogen-bond donor on 3hydroxyBA binds to the carbonyl oxygen of the acid via homomeric interaction, OH(hydroxyl) ····O=C(acid), Figure 5.2a. In **T2-4HydroxyBA**, the extra hydrogen-bond donor interacts with a carbonyl oxygen atom of **T2** to form synthon **J**, Figure 5.2b.

Group 1

(a)

(b)

(c)

(d)

(e)

(f)

(g)

(h)
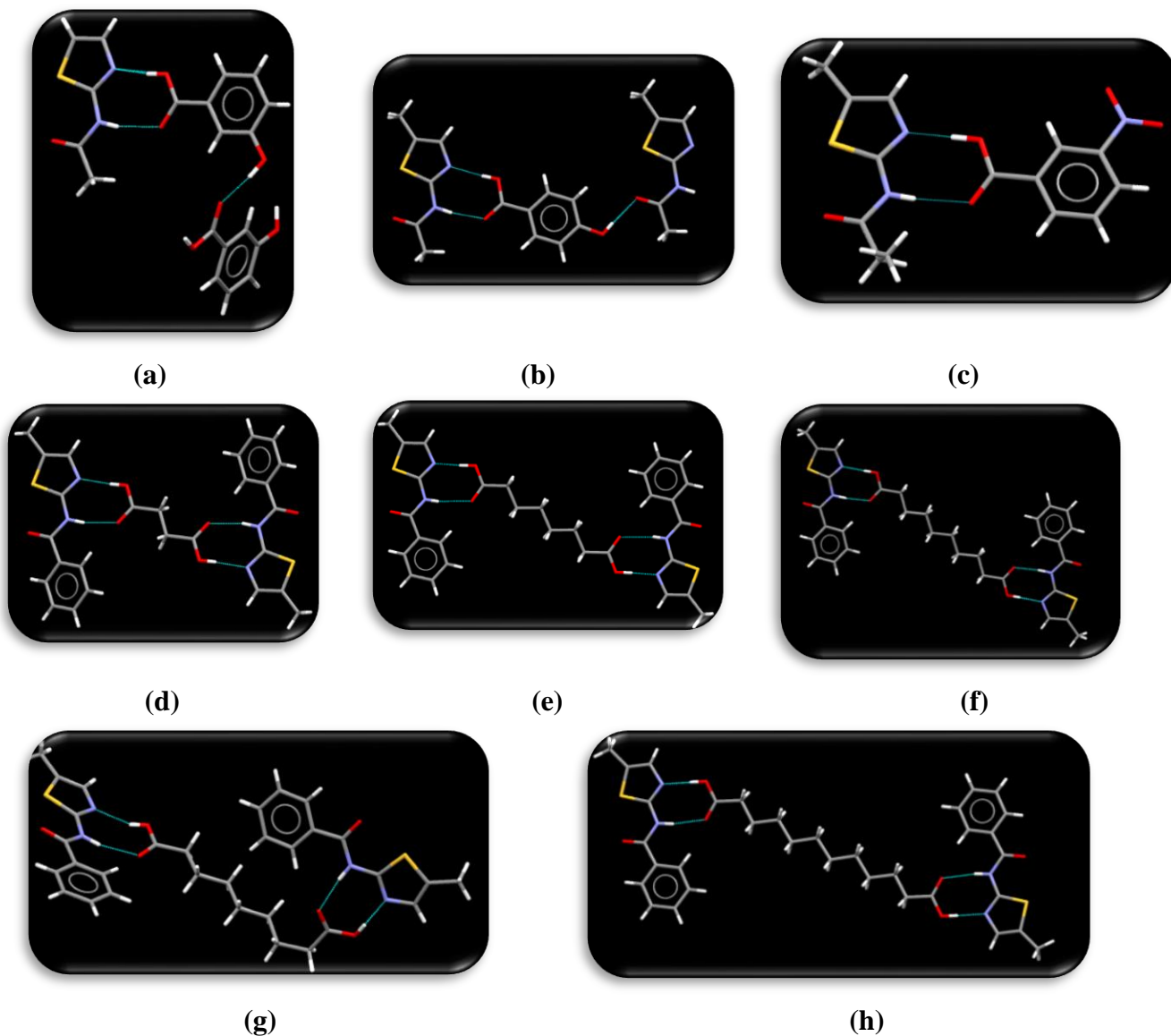
Figure 5.2 Main hydrogen bonds in the crystal structures of (a) **T1-3HydroxyBA**, (b) **T2-4HydroxyBA**, (c) **T4-3NitroBA**, (d) **T8-Suc**, (e) **T8-Sub**, (f) **T8-Seb**, (g) **T8-Aze** and (h) **T8-Dod**.

5.3.5.2 Group 2: T9-T12 against 20 carboxylic acids

In group 2, we obtained single-crystal XRD data for four co-crystals with aliphatic acids, and in all four crystal structures, either one end or both ends of the carboxylic acid is involved in the O-H····N thiazole and amide N-H····O=C acid moieties giving rise to the robust $R_2^2(8)$ hydrogen-bonded synthon **F** and in three out of four crystal structures, one end of acid is also involved in OH---N pyridine hydrogen bond synthon **I** (Figure 6.3b, d, e, and f). Two crystal structures were obtained with aromatic acids (2:1 ratio for **T9-3HydroxyBA**, 2:1 for **T12-4HydroxyBA).** As with the aliphatic acid co-crystals, a robust $R_2^2(8)$ hydrogen-bonded synthon **F** is formed using O-H····N and N-H····O hydrogen bonds in each case (Figure 5.3 a, c). In both **T9-3HydroxyBA** and **T12-4HydroxyBA**, the extra hydrogen-bond donor interacts with a pyridine N of each target molecule to form synthon **L**, Figure 5.3a, c.

Group 2
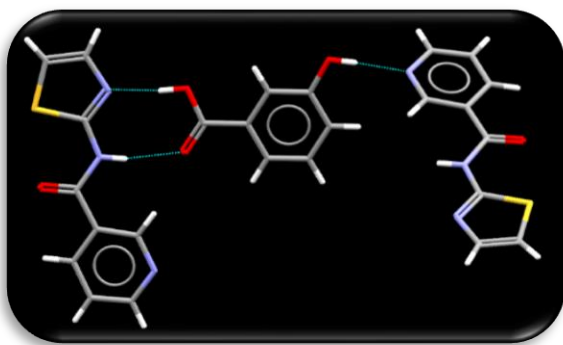
(a)
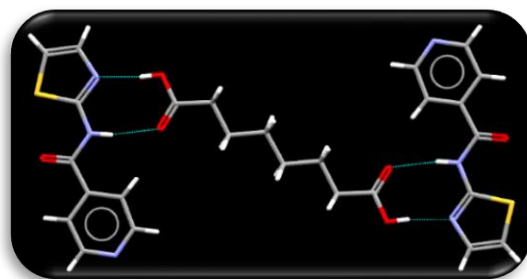
(b)
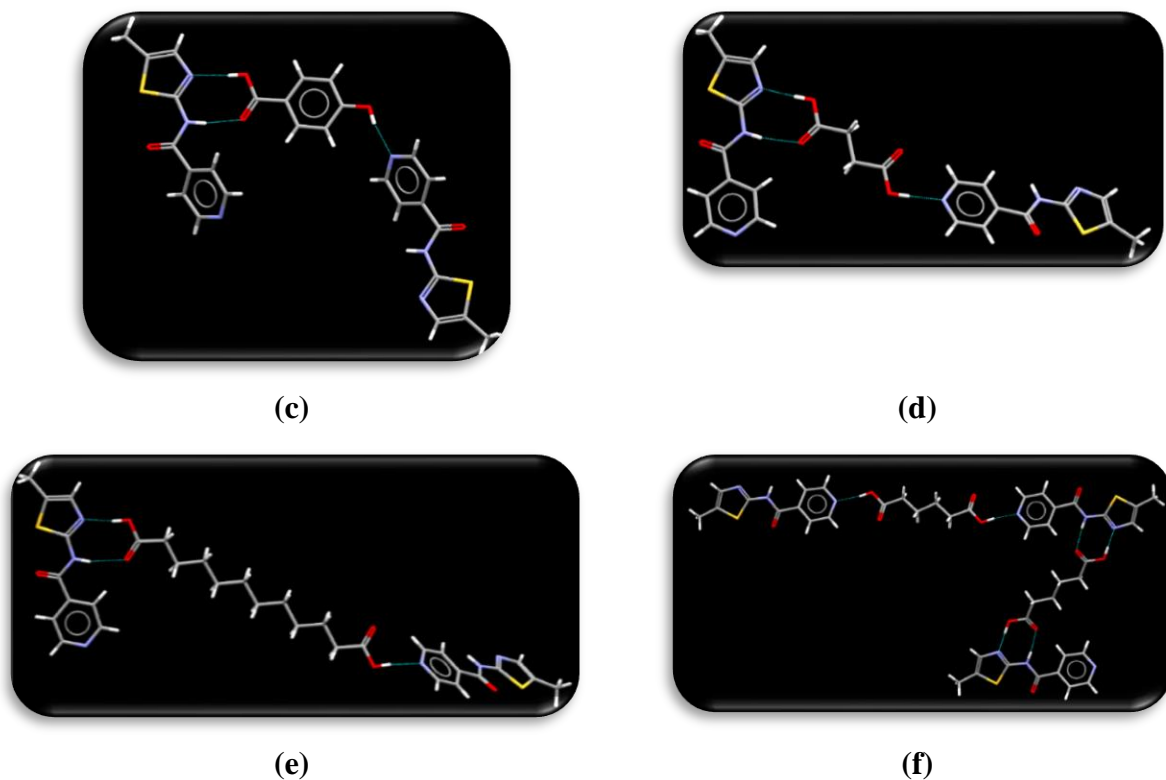
Figure 5.3 Main hydrogen bonds in the crystal structures of **(a) T9-3HydroxyBA, (b)T11-Sub, (c) T12-4HydroxyBA, (d) T12-Suc, (e) T12-dod,** and **(f) T12-Adi.**

5.4 Discussion

5.4.1 Experimental vs predicted co-crystal screening

The predictions made using the HBP model were compared to experimental co-crystal screening results. A confusion matrix was used to analyze the data as reported by Wood *et. al.*[17] The entries in the matrix are labelled true positive (TP; positive for co-crystal via both prediction and experimental), false positive (FP; positive for co-crystal via prediction but negative via experimental), false negative (FN; negative for co-crystal via prediction but positive via experimental), and true negative (TN; negative for co-crystal via both prediction and experimental), Figure 5.4. Any MC value > 0.00 was considered as a prediction of co-crystal formation and an MC value below or equal to 0.00 was considered as a prediction against co-crystal formation.

5.4.1.1 Group 1: T1-T8 against 20 carboxylic acids

Hydrogen-bond propensity calculations were performed on **T1-T8** against 20 co-formers giving a total of 160 prediction results. When compared to experimental screening results, 112 out of the 160 combinations (70%), the HBP model accurately predicted the outcome of the reaction; 49 out of 112 combinations were true positive and 63 of the 112 combinations were true negative, respectively (Table 5.7, Figure 5.4).

Experimentally, **T1-T6** displayed a 15% and **T7-T8** a 95% success rate of co-crystal formation with aliphatic acids. Such a discrepancy between molecules that carry the same principle functionality is difficult to predict or foresee using chemical intuition, but we were delighted that the HBP model delivered predictions that reflected these highly unexpected experimental results very well. Using an $MC_{cutoff} > 0.00$, only twelve out of sixty reactions (20%) between an aliphatic diacid and **T1-T6** was predicted to yield a co-crystal, whereas 18 of 20 combinations (90%) between a diacid and **T7-T8** were predicted to produce a co-crystal. In **T1-T6** with aliphatic acids, homomeric interactions; N-H(amide)····N(thiazole) or O-H(acid) ····O=C(acid) are more probable than the heteromeric interaction, O-H(acid)····N(thiazole), whereas in **T7**-**T8**, the heteromeric interaction, O-H(acid)····N(thiazole), dominates over homomeric interactions which suggests that descriptors such as "steric density" and "aromaticity" play a significant role in the co-crystallization outcome. There was an 80% agreement (64 out of 80 combinations were either true positive or true negative) between the predicted and experimental results for aliphatic acids with **T1-T8**. Among the sixteen unsuccessful predictions, nine were false positives (red box) and seven were false negatives (yellow box).

The experimental co-crystal screening of ten aromatic acids against **T1-T8** produced 49 positive results out of eighty attempts. **T5-T6** showed a higher success rate than **T1-T6**, but almost all target molecules were able to form co-crystals with a success rate greater than 50% (the one exception, **T3**, had a 40% success rate). The HBP models predicted mostly negative outcomes for **T1-T6**, whereas reactions of **T5-T6** were predicted as positive for co-crystal formation (3-aminoBA and 4-aminoBA co-formers were always predicted to produce co-crystals, regardless of target). Overall, the outcome of 47 of the 80 reactions (58%) was accurately predicted by the HBP model (either true positives or true negatives) using an MC > 0.00 cutoff. Additionally, 24 of the 80-aromatic acid/**T1-T8** combinations were false negatives (yellow box) and nine were false positives (red box).

128

Overall, in group 1, in 112 out of 240 (70% success rate) combinations, there was agreement between experimental and predicted results, Table 5.8.

5.4.1.2 Group 2: T9-T12 against 20 carboxylic acids

Hydrogen-bond propensity calculations were performed on **T9-T12** against 20 co-formers giving a total of 80 prediction results. When compared to experimental screening results, 55 out of the 80 combinations (68%), the HBP model accurately predicted the outcome of the reaction; 48 out of 55 combinations were true positive and 7 of the 55 combinations were true negative, respectively, Table 5.8, Figure 5.4.

The experimental co-crystal screening of **T9-T12** against 10 aliphatic acids produced 40 positive results out of 40 attempts (100%). The HBP model predicted 36 of 40 co-crystal screens (90%, either true positive or true negative) correctly using an MC>0.00 cutoff. There were 0 false positives and four were false negatives.

The experimental co-crystal screening of **T9-T12** against 10 aromatic acids produced 33 positive results out of 40 co-crystal attempts (83%). The HBP model predicted only 19 of them correctly (47%, either true positive or true negative). There were 0 false positives and 21 false negatives.

Overall, in group 2, 55 out of 80 (69% success rate) combinations, there was agreement between experimental and predicted results, Table 5.6.

Table 5.6 True positive, false positive, false negative and true negative propensity comparison results for each target molecule (T1-T12) against 20 carboxylic acids and the % yields

	Group 1						
	True positive	False positive	False negative	True negative	TP/TN	FP/FN	%
T1	2	0	6	12	14	6	
T2	2	0	6	12	14	6	
T3	1	1	3	15	16	4	
T4	1	2	4	13	14	6	
T5	3	5	4	8	11	9	112/ 160
T6	5	6	6	3	8	12	70%
T7	19	1	0	0	19	1	
T8	16	2	2	0	16	4	
Total	49	17	31	63	112	48	
	Group 2						

	True positive	False positive	False negative	True negative	TP/TN	FP/FN	
T9	12	0	7	1	13	7	55/ 80 68%
T10	12	0	6	2	14	6	
T12	12	0	6	2	14	6	
T12	12	0	6	2	14	6	
Total	48	0	25	7	55	25	

As indicated by Wood *et al*[17] false negatives are a concern with prediction tools when it comes to creating a list of potential co-formers for a systematic search for new solid forms, as these co-formers will not be included in the experimental study for being seen as unlikely to co-crystallize. On the other hand, a false positive can also be detrimental as extensive experimental resources and time may then be wasted on a co-former that is expected to produce a co-crystal but does not. Although the predictions suggest that the pairs are favored to form co-crystals, they do not indicate how much effort is required to find them.

		Predicted outcome	
		Co-crystal	No co-crystal
Experiment	Co-crystal	True positive	False Negative
	No co-crystal	False Positive	True Negative

(a)

Group1 (T1-T8)

MC_{cutoff}: >0.00		Predicted outcome	
		Co-crystal	No co-crystal
Experiment	Co-crystal	49	31
	No co-crystal	17	63

(b)

Group 2 (T9-T12)

MC_{cutoff}:>0.00		Predicted outcome	
		Co-crystal	No co-crystal
Experiment	Co-crystal	48	25
	No co-crystal	0	7

(c)

Figure 5.4 Confusion matrices determined from multi-component HBP results for T1-T12 molecules with co-formers, (a) explanation of matrix (b) MC_{cutoff}: > 0.00.

5.4.2 Comparison of predicted vs experimental synthons in group 1 and group 2

5.4.2.1 Group 1 (without an additional hydrogen-bond donor)

Two synthons (synthon **F** and synthon **H**) are possible in aliphatic acids and aromatic acids without additional hydrogen-bond donor. The dimeric synthon **F** is favored over synthon **H** by 11 kJ/mol.
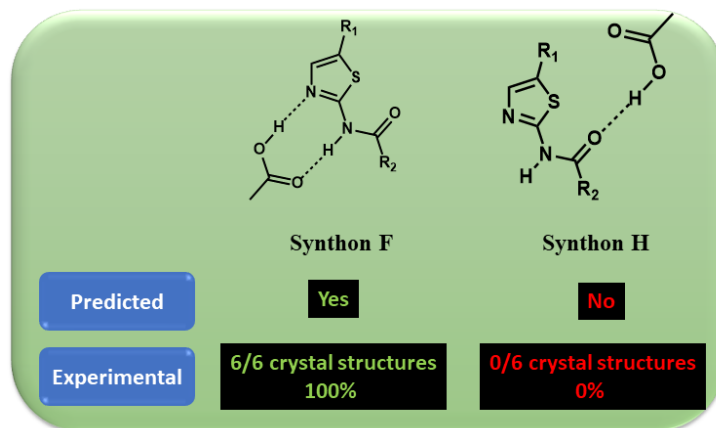
Figure 5.5 Predicted vs experimental results of group 1 (without an additional hydrogen-bond donor).

Six crystal structures (5 with aliphatic acids and 1 with aromatic acids) were obtained in this group, synthon **F** was observed experimentally irrespective of the aliphatic or aromatic acids. The energy results were also supported by propensity results, Table 5.7.

5.4.2.2 Group 1(with an additional hydrogen-bond donor)

Based on hydrogen-bond energies, synthon **F** is favored for both hydroxybenzoic acid co-crystals (over synthons **H, J** and **K**) and aminobenzoic acid co-crystals (over synthons **H, M**, and **N**). Two crystal structures of hydroxybenzoic acid co-crystals were obtained; in **T1**-3HydroxyBA, synthon **F** was observed as predicted. The additional hydrogen-bond donor on 3hydroxyBA binds to the carbonyl oxygen of the acid, OH(hydroxyl)····O=C(acid), Figure 5.2a. In the crystal structure of **T2-4HydroxyBA**, both synthons **F** and synthon **J** were observed. In **T2-4HydroxyBA**, the extra hydrogen-bond donor interacts with a carbonyl oxygen atom of **T2** to form 1-D chains. All conventional hydrogen-bond donors and acceptors are involved in heteromeric interactions (thus following Etter's rule). Without geometric constraints, the best donor of the acid (hydroxylic O-H) binds to the best acceptor of **T2** (C=O amide), the second-best donor of the acid (carboxylic O-

H) binds to the second-best acceptor of **T2** (aromatic N) and the best donor of **T2** (N-H amide) binds to the best acceptor of the acid (O=C). Overall, the predictions made using electrostatics agree very well with the experimentally observed crystal structures, Table 5.7.
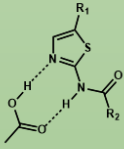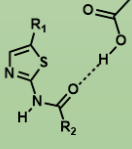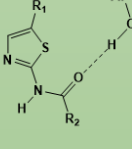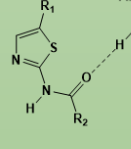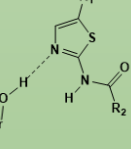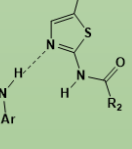
Figure 5.6 Predicted vs experimental results of group 1 (with an additional hydrogen-bond donor).

Moreover, all synthons observed in these crystal structures were either predicted by electrostatics or energy or propensity. Unfortunately, we have not been able to grow crystals suitable for single-crystal X-ray diffraction for aminobenzoic acids, so we do not yet have the data to experimentally determine the synthons observed in this category.

5.4.2.3 Group 2 (without an additional hydrogen-bond donor)

Three synthons (synthon **F**, **H** and **I**) are possible in aliphatic and aromatic acids without an additional hydrogen-bond donor. The dimeric synthon **F** is favored over monomeric synthon **I** and synthon **H** by 8 kJ/mol and 14 kJ/mol respectively. Four crystal structures were obtained in this group. In the crystal structure of **T11**-Sub, synthon **F** was observed on the both side of acid. In the crystal structure of **T12**-suc and **T12**-Dod, one end of the acid forms a robust $R_2^2(8)$ hydrogen-bonded synthon **F** using O-H(acid)····N(thiazole) and N-H(amide)····O=C(acid) hydrogen bonds in each case and the other end forms monomeric synthon **I** using O-H(acid) ····N(aromatic) hydrogen bonds.
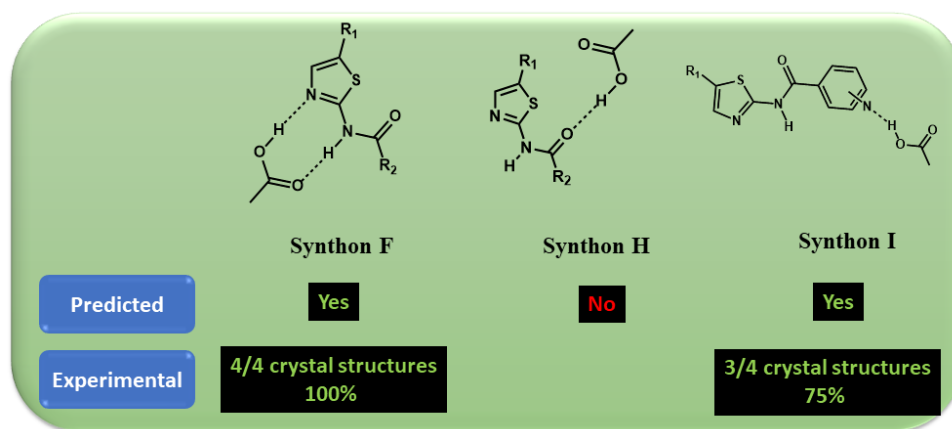
Figure 5.7 Predicted vs experimental results of group 2 (without an additional hydrogen-bond donor).

In the crystal structure of **T12**-Adi, both synthon **F** and **I** are observed in the crystal structure, but both ends of same acid forms same synthon (either **F** or **I**). This type of interaction is different to what observed in the **T12**-Suc and **T12**-dod. Both synthon **F** and synthon **I** only differs by 8 kJ/mol in energy therefore both type of synthons are highly likely to happen as the pyridine N is a stronger acceptor compared to thiazole N based on electrostatics charges and therefore, taking into balance between electrostatics and energies, both synthon **F** and **I** are competitive and has equal chance to form the synthon. Synthon **I** was predicted to be the most optimal synthon according to hydrogen-bond propensity, Table 5.7.

5.4.2.4 Group 2 (with an additional hydrogen-bond donor)

Six different synthons (synthon **F, H, I, J, K** and **L** for hydroxybenzoic acids) and (synthon **F, H, I, M, N** and **O** for aminobenzoic acids) are possible for aromatic acids with additional hydrogen-bond donors and in each category, synthon **F** is the most preferred synthon. Two crystal structures are observed with hydroxybenzoic acids; **T9-3HydroxyBA** and **T12-4HydroxyBA.** Two synthons (Synthon **F** and synthon **L**) were observed in each crystal structure. The best donor of the acid (hydroxylic O-H) binds to the best acceptor of **T9 and T12**(Pyridine N) following Etter'r rule. The second-best donor of the acid (carboxylic O-H) binds to the thiazole N and the C=O (best acceptor of acid) binds to the amide NH forming synthon **F** which is dominant due to its supramolecular chelating effect. Both synthon **F** and **L** are ranked 1 and 2 based on energies in this group out of

133

six different synthon possibilities and differs by about 10 kJ/mo based on hydrogen-bond energies.
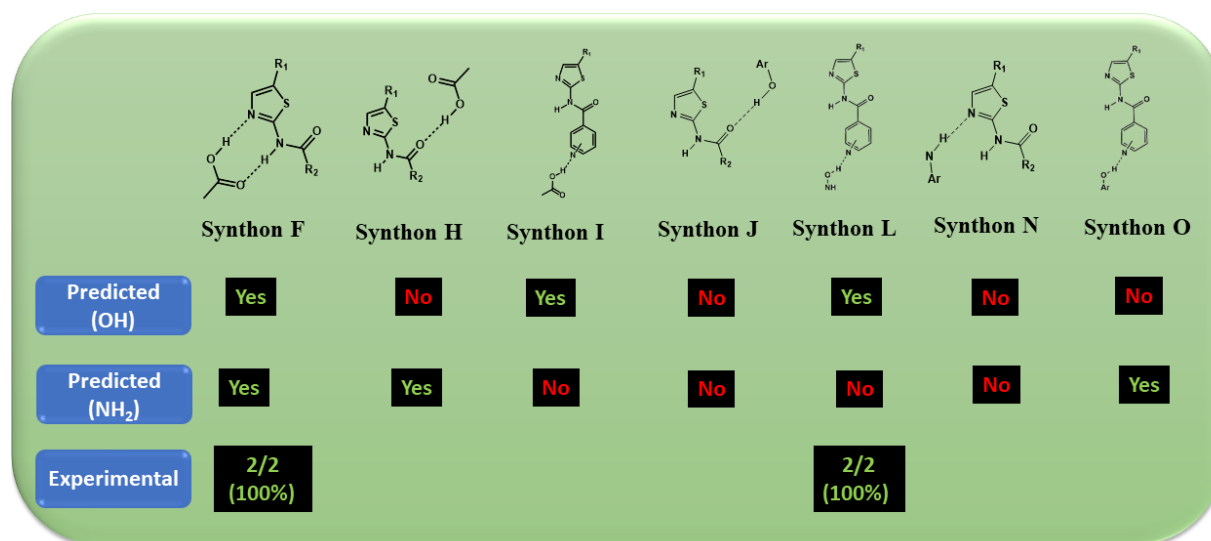
Figure 5.8 Predicted vs experimental results of group 2 (with an additional hydrogen-bond donor).

The presence for additional donor group on the acid binds to best acceptor on the target molecule as expected based on electrostatics, Table 5.8. Unfortunately, we have not been able to grow crystals suitable for single-crystal X-ray diffraction for aminobenzoic acids, so we do not yet have the data to experimentally determine the synthons observed in this category.

Table 5.7 Summary of predicted vs experimental synthons in co-crystals of T1-T12 with 20 dicarboxylic acids.

	Electrostatics	Energies	Propensities	Experimental synthon	# data points
Group 1					
(w/o h-bond donor)	Synthon H	Synthon F	Synthon F	Synthon F	6
(HydroxyBA)	Synthon J	Synthon F	Synthon F	Synthon F and J	2
(AminoBA)	Synthon H	Synthon F	Synthon N	N/A	0
Group 2					
(w/o h-bond donor)	Synthon I	Synthon F	Synthon I	Synthon F and I	4
(HydroxyBA)	Synthon L	Synthon F	Synthon I	Synthon F and L	2
(AminoBA)	Synthon H	Synthon F	Synthon O	N/A	0

5.5 Conclusions

- The HBP model was used to predict and compare the experimental co-crystal screening results comprising twelve target molecules and twenty carboxylic acids resulting in 240 data points. An MC score produced from the HBP calculations using a cut-off value of > 0.0 resulted in a 69% agreement between prediction and experiment (70% success for group 1 and 68% for group 2). The HBP model made 80% and 90% predictions correctly for group 1 and group 2 respectively with aliphatics acids. The HBP model made 58% and 47% predictions correctly for group 1 and group 2 respectively with aromatic acids.

- The electrostatics and propensity along with hydrogen-bond energies were used to predict synthons in positive co-crystal outcomes. The HBE predicted synthon **F** to be more favorable in both groups. The HBP calculations also predicted synthon **F** to be more favorable in group 1 but in the aminoBA co-crystals. Synthon **N** is favored over synthons **F** in aminoBA co-crystals. In group2, HBP predicted synthon **I** to be more favorable and synthon O is favored in aminoBA co-crystals. Based on electrostatics, synthon **H** was favored in group 1 but synthon **J** in hydroxyBA. In group 2, different synthons were predicted in each category.

- Hydrogen-bond energy was the best method to predict heteromeric synthons.
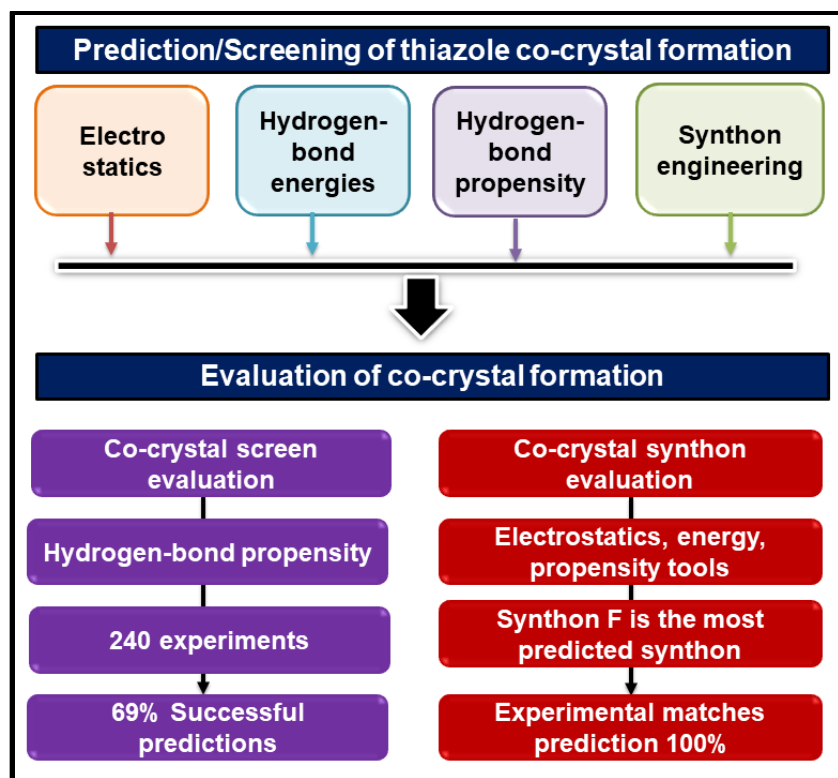
Figure 5.9 Summary of results obtained in this study.

The ability to make informed predictions about how different molecules recognize and bind to each other, and how they subsequently assemble into solid-state architectures is of critical importance in areas such as drug design and formulation, and the recently published guidance by the FDA on the regulatory classification of pharmaceutical co-crystals[34] will continue to keep this fundamental science in sharp focus. Therefore, the successful use of structural informatics tools such as hydrogen-bond propensity and calculated molecular electrostatic potential surfaces for mapping out the structural landscape of these types of molecules will have significant practical applications.

5.6 References

1.   Sun, C. C., *Expert Opinion on Drug Delivery* **2013,** *10* (2), 201-213.

2.   Fala, L., *American Health & Drug Benefits* **2015,** *8* (6), 330-334.

3.   Thipparaboina, R.; Kumar, D.; Chavan, R. B.; Shastri, N. R., *Drug Discovery Today* **2016,** *21* (3), 481-490.

4.   Chavan, R. B.; Thipparaboina, R.; Yadav, B.; Shastri, N. R., *Drug Delivery and Translational Research* **2018**.

5.   Hunter, C. A., *Angewandte Chemie International Edition* **2004,** *43* (40), 5310-5324.

6.   McKenzie, J.; Feeder, N.; Hunter, C. A., *CrystEngComm* **2016,** *18* (3), 394-397.

7.   Musumeci, D.; Hunter, C. A.; Prohens, R.; Scuderi, S.; McCabe, J. F., *Chemical Science* **2011,** *2* (5), 883-890.

8.   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., *Journal of Applied Crystallography* **2008,** *41* (2), 466-470.

9.   Fábián, L., *Crystal Growth & Design* **2009,** *9* (3), 1436-1443.

10.   Issa, N.; Karamertzanis, P. G.; Welch, G. W. A.; Price, S. L., *Crystal Growth & Design* **2009,** *9* (1), 442-453.

11.   Mohammad, M. A.; Alhalaweh, A.; Velaga, S. P., *International Journal of Pharmaceutics* **2011,** *407* (1), 63-71.

12.   Galek, P. T. A.; Allen, F. H.; Fabian, L.; Feeder, N., *CrystEngComm* **2009,** *11* (12), 2634-2639.

13.   Chemburkar, S. R.; Bauer, J.; Deming, K.; Spiwek, H.; Patel, K.; Morris, J.; Henry, R.; Spanton, S.; Dziki, W.; Porter, W.; Quick, J.; Bauer, P.; Donaubauer, J.; Narayanan, B. A.; Soldani, M.; Riley, D.; McFarland, K., *Organic Process Research & Development* **2000,** *4* (5), 413-417.

14.   Allen, F., *Acta Crystallographica Section B* **2002,** *58* (3 Part 1), 380-388.

15.   Mapp, L. K.; Coles, S. J.; Aitipamula, S., *Crystal Growth & Design* **2017,** *17* (1), 163-174.

16.   Vologzhanina, A. V.; Sokolov, A. V.; Purygin, P. P.; Zolotarev, P. N.; Blatov, V. A., *Crystal Growth & Design* **2016,** *16* (11), 6354-6362.

17.   Wood, P. A.; Feeder, N.; Furlow, M.; Galek, P. T. A.; Groom, C. R.; Pidcock, E., *CrystEngComm* **2014,** *16* (26), 5839-5848.

18.   Delori, A.; Galek, P. T. A.; Pidcock, E.; Patni, M.; Jones, W., *CrystEngComm* **2013,** *15* (15), 2916-2928.

19.   Delori, A.; Galek, P. T. A.; Pidcock, E.; Jones, W., *Chemistry – A European Journal* **2012,** *18* (22), 6835-6846.

20.   Eddleston, M. D.; Arhangelskis, M.; Fábián, L.; Tizzard, G. J.; Coles, S. J.; Jones, W., *Crystal Growth & Design* **2016,** *16* (1), 51-58.

21.   Tilbury, C. J.; Chen, J.; Mattei, A.; Chen, S.; Sheikh, A. Y., *Crystal Growth & Design* **2017**.

22.   Bosch, E. C., 9, 191-198.

23.   Desiraju, G. R. A. C., Int. Ed. 199, 534, 2311-2327

24.   C.B.; Desper, J. U., J.F. CrystEngComm 2005, 7, 193-201.

25.   Lehn, J. M. S., 295, 2400-2403.

26.   Buckingham, A. D. D. B., J.E.; McDowell, S.A.C. Chem. Phys. Lett. 2008, 463, 1-10.

27.   Etter, M. C., *Accounts of Chemical Research* **1990,** *23* (4), 120-126.

28.   Peterson, M. L.; Stanton, M. K.; Kelly, R. C.; Staples, R.; Cheng, A., *CrystEngComm* **2011,** *13* (4), 1170-1180.

29.   Stanton, M. K.; Bak, A., *Crystal Growth & Design* **2008,** *8* (10), 3856-3862.

30.   Cheney, M. L.; Weyna, D. R.; Shan, N.; Hanna, M.; Wojtas, L.; Zaworotko, M. J., *Crystal Growth & Design* **2010,** *10* (10), 4401-4413.

31.   Tumanov, N. A.; Myz, S. A.; Shakhtshneider, T. P.; Boldyreva, E. V., *CrystEngComm* **2012,** *14* (1), 305-313.

32.   Suresh, K.; Mannava, M. K. C.; Nangia, A., *Chemical Communications* **2016,** *52* (22), 4223-4226.

33.   Félix-Sonda, B. C.; Rivera-Islas, J.; Herrera-Ruiz, D.; Morales-Rojas, H.; Höpfl, H., *Crystal Growth & Design* **2014,** *14* (3), 1086-1102.

34.   Gadade, D. D.; Pekamwar, S. S., *Advanced Pharmaceutical Bulletin* **2016,** *6* (4), 479-494.

Chapter 6 - Constructing Binary and Ternary Co-crystals Using Hydrogen and Halogen Bonds as Synthetic Vectors

6.1 Introduction

Intermolecular interactions such as hydrogen bonding,[1-5] halogen bonding,[6-9] or $\pi \cdots \pi$ interactions represent the primary tools in supramolecular chemistry and crystal engineering.[2] It is important to have a better understanding of the fundamental nature of these interactions in order to successfully design complex supramolecular systems in a predetermined and effective manner. The hydrogen bond is a key synthetic tool in assembly of molecules into well-defined architectures.[12] Hydrogen bonded binary co-crystals have been constructed through a variety of heterosynthons such as carboxylic acid$\cdots$pyridine,[10-12] oxime$\cdots$N-heterocycles,[13] and carboxylic acid$\cdots$amide.[14-15] Halogen bonding is a relatively recent addition to the tool box of supramolecular chemistry.[4] According to IUPAC, "*A halogen bond R−X⋯Y−Z occurs when there is evidence of a net attractive interaction between an electrophilic region on a halogen atom X belonging to a molecule or a molecular fragment R–X (where R can be another atom, including X, or a group of atoms) and a nucleophilic region of a molecule, or molecular fragment, Y−Z*" (Scheme 6.1).[5] The halogen bond, due to its relative strength and highly directional nature, has been employed in a similar manner to hydrogen bonding for the assembly and synthesis of binary co-crystals.[16] The most common halogen bonded synthons includes I…N type interactions.

Scheme 6.1 Schematic diagram showing the formation of a halogen bond (X is a halogen atom and Y is a nucleophilic atom)

The successful synthesis of multi-component supramolecular systems[17] is challenging because of the relatively weak and reversible nature of non-covalent forces. If the synthetic target contains more than two different molecules, such as a ternary or quaternary co-crystal, the number of

binding sites also increases which expands the possibilities of unwanted synthon crossover.[18] [14] As a result, the exact chemical composition of the product, as well as the relative orientation of the participating molecular constituents, becomes much less predictable. Against this background, it is not surprising that it is still very challenging to produce ternary systems since the co-crystallization process can lead to a very diverse set of outcomes; simple recrystallizations of all starting materials, binary co-crystals, solvates/hydrates of any of those combinations, or the targeted ternary co-crystal (Scheme 6.2).[13] The role of the solvent can obviously play a key role in the outcome which means that both thermodynamic and kinetic factors need to be balanced,[19] in order to successfully synthesize ternary co-crystals.
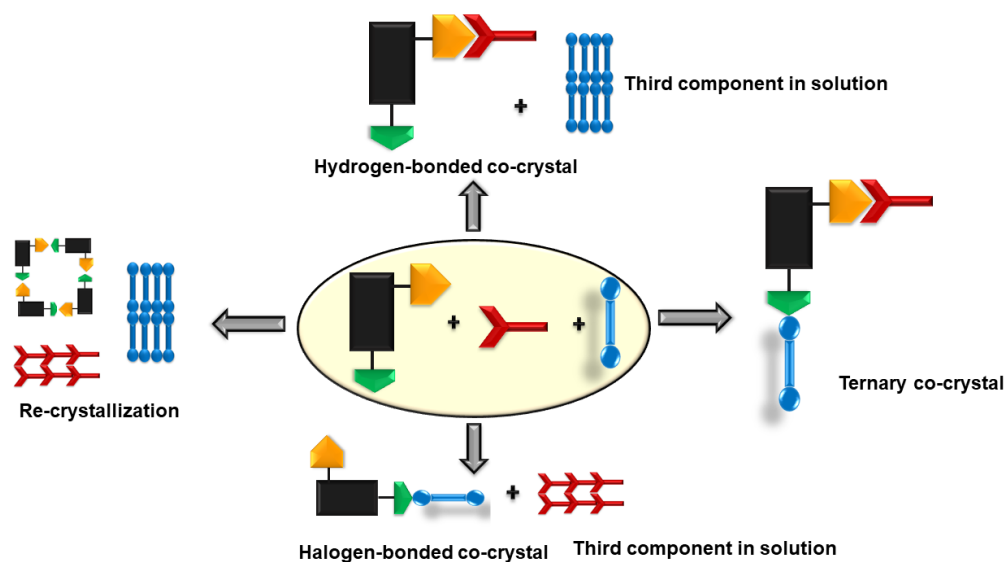
Scheme 6.2 Different possibilities for an attempt to make ternary co-crystals.

Clearly both hydrogen and halogen bonding offer effective synthetic avenues for binary systems[20-21] but only a relative limited number of examples of the assembly of ternary co-crystals have been reported.[22-23] A variety of strategies have been pursued for ternary co-crystal synthesis many of which have been inspired by Etter's observation that "the best hydrogen-bond donor and the best hydrogen-bond acceptor will preferentially form hydrogen bonds to one another,"[24] and Aakeröy and co-workers have employed molecular electrostatic potential calculations as a way of ranking acceptors/donors. The combination of these ideas have been found to be of considerable practical use for determining the connectivity patterns in molecules with multiple binding pockets,[24] as well

as for synthesizing ternary co-crystals with desired composition and intermolecular connectivity.[3, 19, 22] Desiraju and co-workers have presented size, shape and template based approaches to the synthesis of co-crystals, [25] and these protocols have also been modified and extended to the assembly of very rare examples of quaternary, quintenary co-crystals[26-27] and six-component molecular solids.[28] Recently, Rissanen and co-workers used the orthogonality of two supramolecular interaction modes: hydrogen bonding between crown ethers and thioureas, and halogen bonding between thioureas and perfluorohalocarbons to construct ternary co-crystals.[29]

The combination of hydrogen- and halogen bonding in the same supramolecular synthetic scheme may offer more efficient and reliable avenues for the preparation of complex multi-component assemblies as these two interactions, although similar in some ways, can be tailored independently in order to minimize synthon crossover.[23, 30-32]

We have previously explored binding preferences of thiazole-based ligands with two competing binding pockets and shown that the N-H/N site is preferred over the N-H⋯S site when the ligand is presented with a carboxylic acid, Scheme 6.3. This selectivity is readily explained on the basis of a simple electrostatic argument which favors the nitrogen atom over the sulfur atom as a hydrogen-bond acceptor due to the higher negative electrostatic potential of the former
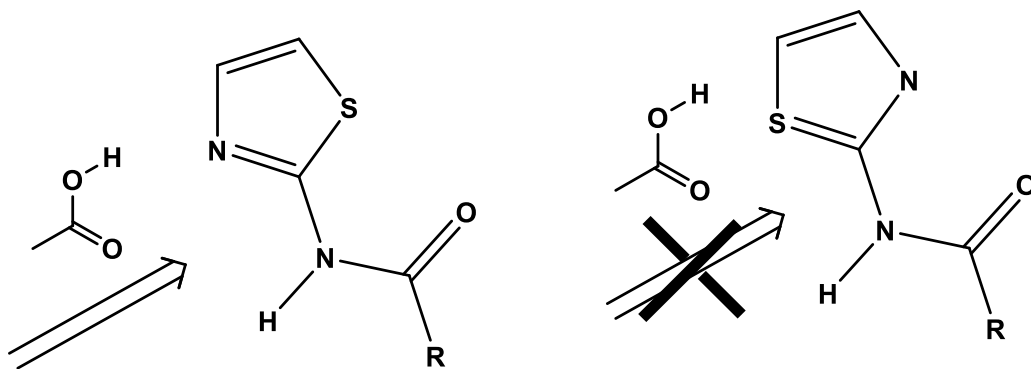
Scheme 6.3 Hydrogen-bond preferences of a carboxylic acid vis-à-vis a thiazole-based ligand[33]

In order to now build a new molecule capable of acting as an assembly point for ternary co-crystals, we decided to add a suitable acceptor capable of reliably forming halogen bonds with a suitable donor without interfering with the thiazole N-H⋯N binding site.
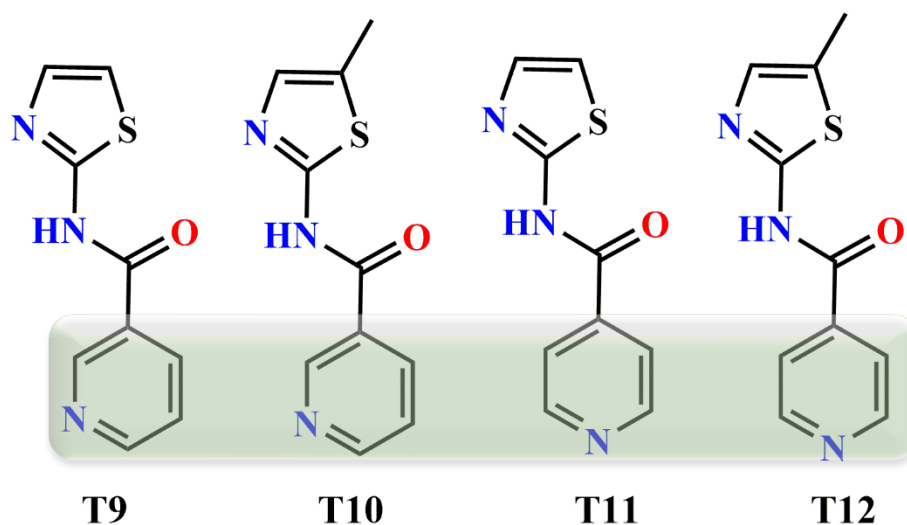
Scheme 6.4 Four target molecules for ternary co-crystals.

6.1.1 Proposed hypothesis

Our choice of acceptors, **T9-T12** was guided by the fact that there is considerable structural data in the CSD to suggest that activated halogen atoms are prone to forming C-X···N halogen bonds with pyridyl moieties. Recognition point 1 is a two-point interaction designed for a suitable hydrogen-bond donor (Scheme 6.5a) and recognition point 2 is a one-point interaction designed for a suitable halogen-bond donor (Scheme 6.5b).
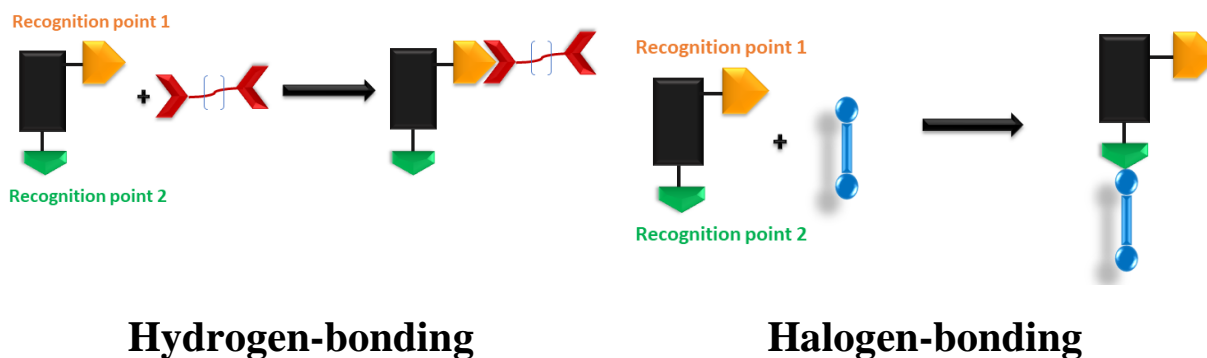
Scheme 6.5 Schematics of hypothesized (a) hydrogen-bonding and (b) halogen-bonding interactions in **T9-T12**.

Our strategy is that when binary hydrogen-bond co-crystallization is attempted, recognition point 1 will be occupied by the donor due to supramolecular chelating effect leaving the recognition

point 2 vacant. One the other hand, when binary halogen-bond co-crystallization is attempted, recognition point 2 will be occupied by the suitable halogen-bon donor leaving the recognition point 1 vacant or form homomeric interactions. We also hypothesized that as the halogen-bond only involves a single-point interaction, whereas a carboxylic acid generates a heterosynthon with two hydrogen bonds with thiazole pocket, the halogen-bond donor would not be competitive for the thiazole nitrogen atom. Similarly, we would not expect the self-complementary thiazole homosynthon to be disrupted in a binary halogen-bonded co-crystal. These ideas form the basis for a synthetic strategy of ternary co-crystals driven by structurally independent hydrogen- and halogen bonds, Scheme 6.6.
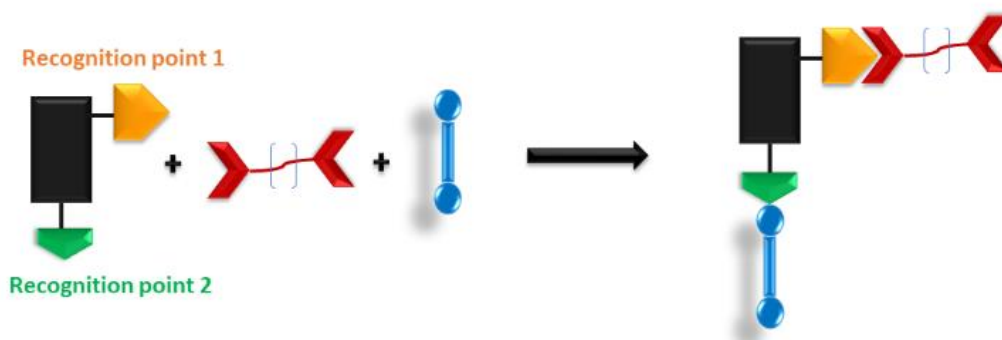
Scheme 6.6 Proposed supramolecular design strategy for the assembly of ternary co-crystals

In this study we want to address following questions;

1. Can we make hydrogen-bond binary co-crystals at recognition point 1 without using recognition point 2?
2. Can we form halogen-bond binary co-crystals at recognition point 2 without using recognition point 1?
3. Can we use both recognition points to form ternary co-crystals by using recognition point 1 for hydrogen-bonding and recognition point 2 for halogen bonding?

Fifteen different hydrogen-bond donors were chosen for this study (10 aliphatic acids and 5 aromatic acids), Scheme 6.7.
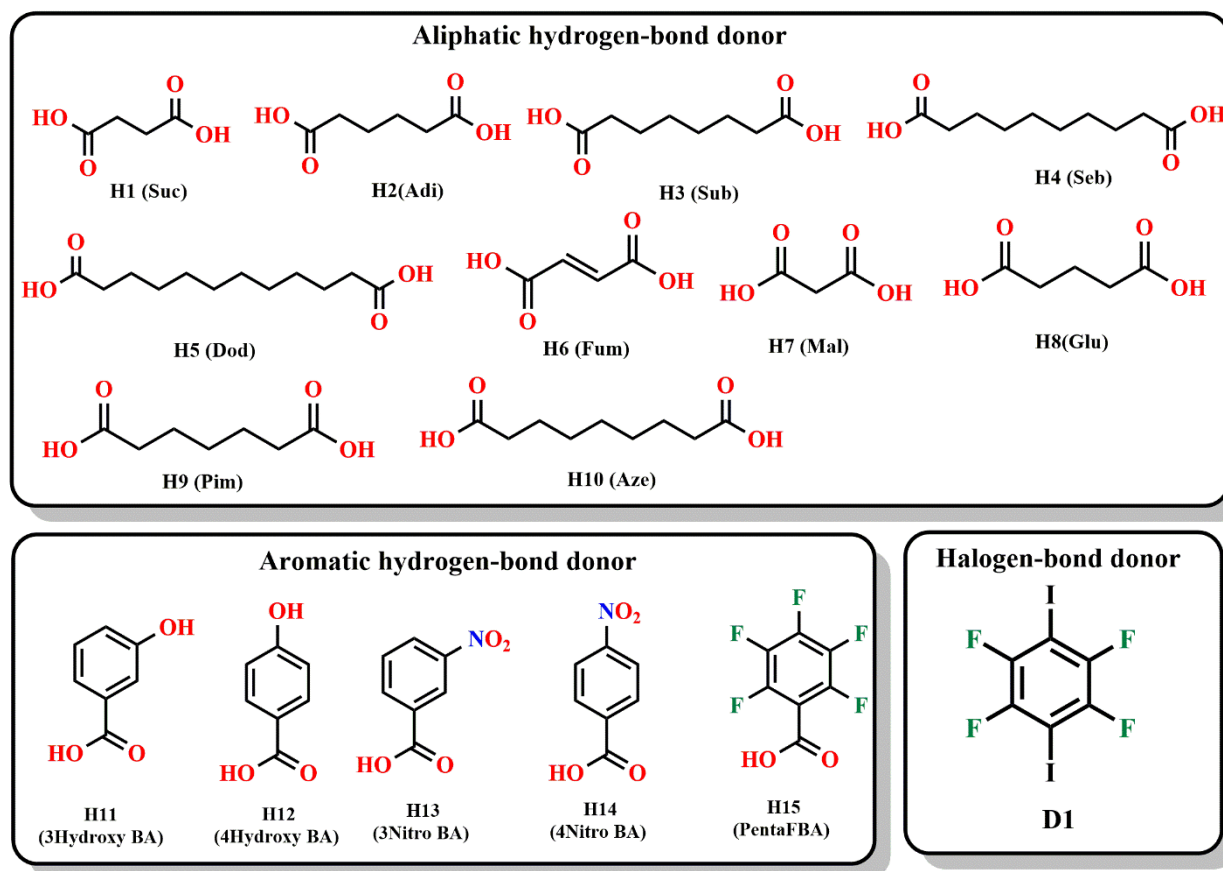
Scheme 6.7 Fifteen hydrogen-bond donors and one halogen-bond donor chosen for ternary co-crystallization.

To get a better understanding of how we may control not just synthon reproducibility but also the three-dimensional consequences thereof, we have also explored how the same type of molecular recognition events may lead to supramolecular architectures of different dimensionality and connectivity, Scheme 6.8 with hydrogen-bond donors. First talking about aliphatic diacids, three different synthons (synthon I, synthon II and synthon III) are possible, Scheme 6.8. Synthon I represent both ends of same acid binding to recognition point 1; N-H(amide)…O=C(acid) and O-H(acid)…N(thiazole) hydrogen bonds. Synthon II represent both ends of same acids binds to recognition point 2; O-H(acid)…N(pyridine) hydrogen bonds. Synthon III represents one end of acids binds to recognition point 1 (N-H(amide)…O=C(acid), O-H(acid)…N(thiazole)) and other end binds to recognition point 2 (O-H(acid)…N(pyridine)). A possible complication (synthon II and synthon III), would result if the carboxylic diacid were to bind to the thiazole pocket at one

end and to the pyridyl moiety (O-H…N(py) at the other end which could prevent the halogen-bond donor from being included in the product. The most desired synthon with aliphatic diacids is synthon I which will leave the pyridyl moiety vacant for halogen bonding.

Synthon I	**Synthon II**
Synthon III	**Synthon IV**
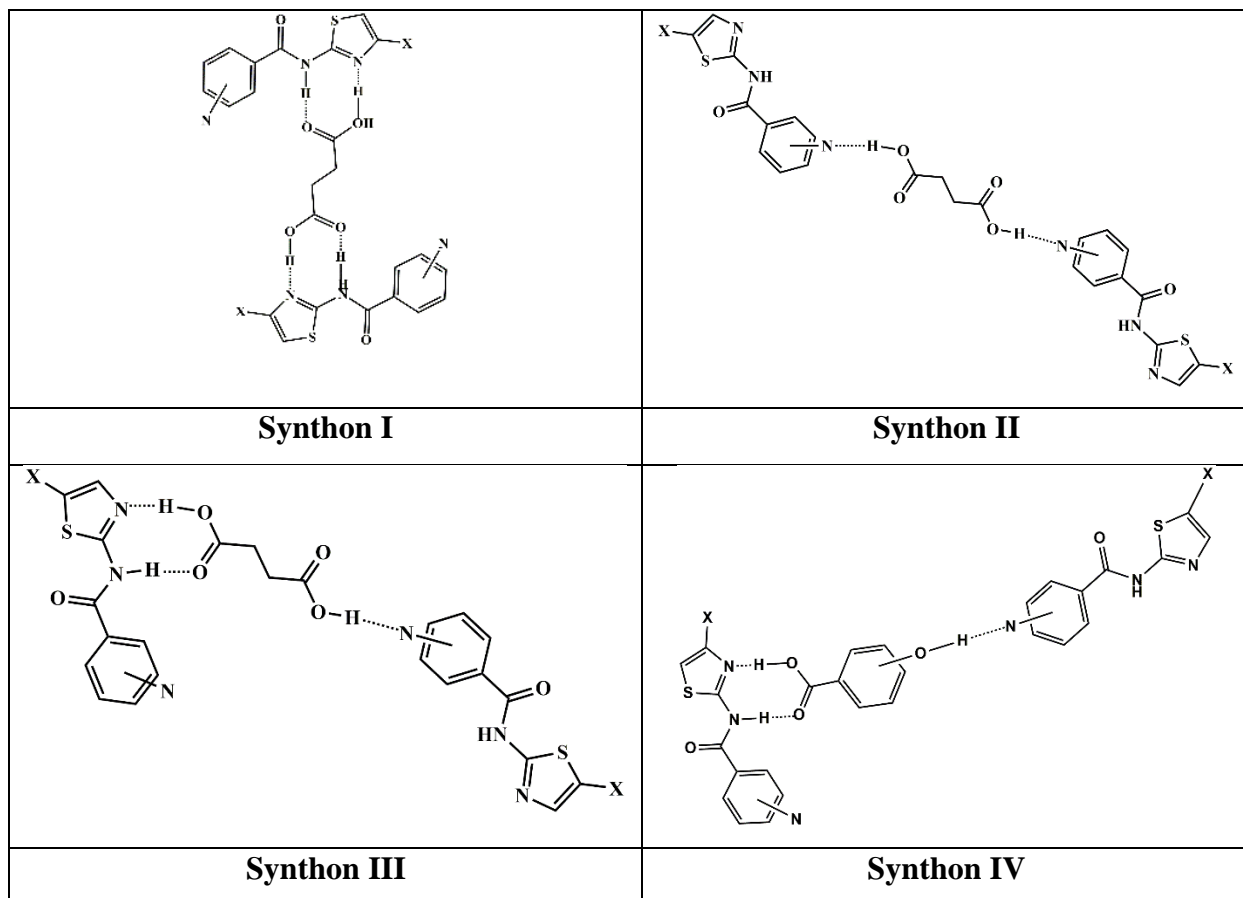
Scheme 6.8 Possible synthons with ditopic acceptors leading to either 0-D discrete or 1-D infinite chains in binary hydrogen bonded cocrystals with aliphatic and aromatic acids.

In order to understand how having additional functional group on the co-former affects the resulting synthons, we included five aromatic acids with functional groups such as hydroxy, and nitro groups available. We postulate that in the case of H11 and H12, the carboxylic acid group would bind to recognition point 1 via a two-point interaction and the hydroxy group would bind to recognition point 2 via one-point interaction, synthon **IV**. If synthon **IV** is to form in the hydrogen-bonding, then the formation of halogen bonding or ternary co-crystals would not be possible because the pyridyl nitrogen would be occupied. Therefore, co-formers H13-H15 would be ideal candidate for ternary co-crystallization.

6.2 Experimental

6.2.1 Materials

All precursors, solvents and aliphatic dicarboxylic acids, and halogen-bond donors were purchased from commercial sources and used without further purification. **T9-T12** synthetic procedures are provided in chapter 4. ^{1}H NMR spectra were recorded on a Varian Unity plus 400 MHz spectrometer. Infrared spectra were recorded with a Nicolet 380 FT-IR. Melting points were determined using Fischer-Johns Mel-Temp melting point apparatus and are uncorrected.

6.2.2 Hydrogen-bonded co-crystals

The binary hydrogen-bond synthesis of **T9-T12** with 10 aliphatic acids and 5 aromatic acids is described in detail in chapter 5. In all 40 combinations with aliphatics and 20 combinations with aromatics, the ditopic acceptors and donors were mixed in 2:1 and 1:1 stoichiometric ratio respectively and the solid resulting from each reaction was characterized using IR spectroscopy to determine if a co-crystal had formed.

6.2.3 Halogen-bonded co-crystals

Synthesis of halogen bonded co-crystals was performed by mixing each acceptor with the halogen-bond donor in a 2:1 stoichiometric ratio in a drop of methanol solvent. The mixture was ground, and the mixture was analyzed using infra-red spectroscopy to determine if a co-crystal has formed. Once the co-crystal was confirmed, the grinded mixture was dissolved in the minimum amount of solvent to grow crystals using slow evaporation method.

6.2.3.1 T9-D1 (2:1)

N-(thiazol-2-yl) nicotinamide-1,4-diiodotetrafluorobenzene (2:1): **T9** (41mg, 0.10 mmol) and 1,4-diiodotetrafluorobenzene, **D1** (40mg, 0.05mmol) were dissolved in minimum amount of methanol and the solution was allowed to evaporate.

6.2.3.2 T10-D1 (2:1)

N-(5-methylthiazol-2-yl) nicotinamide-1,4-diiodotetrafluorobenzene (2:1): **T10** (43mg, 0.10 mmol) and 1,4-diiodotetrafluorobenzene, **D1** (40mg, 0.05mmol) were dissolved in minimum amount of methanol and the solution was allowed to evaporate.

6.2.3.3 T11-D1 (2:1)

N-(thiazol-2-yl) isonicotinamide-1,4-diiodotetrafluorobenzene (2:1): **T11** (41mg, 0.10 mmol) and 1,4-diiodotetrafluorobenzene, **D1** (40mg, 0.05mmol) were dissolved in minimum amount of methanol and the solution was allowed to evaporate.

6.2.3.4 T12-D1 (2:1)

N-(5-methylthiazol-2-yl) isonicotinamide-1,4-diiodotetrafluorobenzene (2:1): **T12** (43mg, 0.10 mmol) and 1,4-diiodotetrafluorobenzene, **D1** (40mg, 0.05mmol) were dissolved in minimum amount of methanol and the solution was allowed to evaporate.

6.2.4 Ternary co-crystals

A total of 60 different combination experiments were performed using **T9-T12**, 15 hydrogen-bond donors (10 aliphatic diacids and 5 aromatic monoacids) and one halogen bond donor. A 2:1:1 stoichiometric ratio of acceptor: hydrogen bond donor: halogen-bond donor was ground together in a mortar and pestle in a drop of methanol. Once confirmed positive via infra-red spectroscopy, the mixture was dissolved in the minimum amount of solvent to grow crystals using slow evaporation.

Table 6.1 Stoichiometric ratios of each donor and acceptor and solvent used for ternary co-crystallization attempts.

Ternary co-crystals	Acceptor	Hydrogen-bond donor	Halogen-bond donor	Solvent used
T9-H1-D1	41mg, 0.10 mmol	6 mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T9-H2-D1	41mg, 0.10 mmol	7.3 mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T9-H3-D1	41mg, 0.10 mmol	8.7 mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T9-H4-D1	41mg, 0.10 mmol	10 mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T9-H5-D1	41mg, 0.10 mmol	12mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T9-H6-D1	41mg, 0.10 mmol	8 mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T9-H7-D1	41mg, 0.10 mmol	5.8 mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T9-H8-D1	41mg, 0.10 mmol	6.6mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol

T9-H9-D1	41mg, 0.10 mmol	8 mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T9-H10-D1	41mg, 0.10 mmol	9.4mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T9-H11-D1	41mg, 0.10 mmol	13mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T9-H12-D1	41mg, 0.10 mmol	13mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T9-H13-D1	41mg, 0.10 mmol	16mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T9-H14-D1	41mg, 0.10 mmol	16mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T9-H15-D1	41mg, 0.10 mmol	21mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T10-H1-D1	43mg, 0.10 mmol	6 mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T10-H2-D1	43mg, 0.10 mmol	7.3 mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T10-H3-D1	43mg, 0.10 mmol	8.7 mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T10-H4-D1	43mg, 0.10 mmol	10 mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T10-H5-D1	43mg, 0.10 mmol	12mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T10-H6-D1	43mg, 0.10 mmol	8 mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T10-H7-D1	43mg, 0.10 mmol	5.8 mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T10-H8-D1	43mg, 0.10 mmol	6.6mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T10-H9-D1	43mg, 0.10 mmol	8 mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T10-H10-D1	43mg, 0.10 mmol	9.4mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T10-H11-D1	43mg, 0.10 mmol	13mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T10-H12-D1	43mg, 0.10 mmol	13mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T10-H13-D1	43mg, 0.10 mmol	16mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T10-H14-D1	43mg, 0.10 mmol	16mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T10-H15-D1	43mg, 0.10 mmol	21mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T11-H1-D1	41mg, 0.10 mmol	6 mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T11-H2-D1	41mg, 0.10 mmol	7.3 mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T11-H3-D1	41mg, 0.10 mmol	8.7 mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T11-H4-D1	41mg, 0.10 mmol	10 mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T11-H5-D1	41mg, 0.10 mmol	12mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T11-H6-D1	41mg, 0.10 mmol	8 mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T11-H7-D1	41mg, 0.10 mmol	5.8 mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T11-H8-D1	41mg, 0.10 mmol	6.6mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T11-H9-D1	41mg, 0.10 mmol	8 mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T11-H10-D1	41mg, 0.10 mmol	9.4mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T11-H11-D1	41mg, 0.10 mmol	13mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T11-H12-D1	41mg, 0.10 mmol	13mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T11-H13-D1	41mg, 0.10 mmol	16mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T11-H14-D1	41mg, 0.10 mmol	16mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T11-H15-D1	41mg, 0.10 mmol	21mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T12-H1-D1	43mg, 0.10 mmol	6 mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T12-H2-D1	43mg, 0.10 mmol	7.3 mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T12-H3-D1	43mg, 0.10 mmol	8.7 mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T12-H4-D1	43mg, 0.10 mmol	10 mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T12-H5-D1	43mg, 0.10 mmol	12mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T12-H6-D1	43mg, 0.10 mmol	8 mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T12-H7-D1	43mg, 0.10 mmol	5.8 mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T12-H8-D1	43mg, 0.10 mmol	6.6mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T12-H9-D1	43mg, 0.10 mmol	8 mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T12-H10-D1	43mg, 0.10 mmol	9.4mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol

T12-H11-D1	43mg, 0.10 mmol	13mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T12-H12-D1	43mg, 0.10 mmol	13mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T12-H13-D1	43mg, 0.10 mmol	16mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T12-H14-D1	43mg, 0.10 mmol	16mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol
T12-H15-D1	43mg, 0.10 mmol	21mg, 0.05mmol	40mg, 0.05mmol	Methanol, Ethanol

6.3 Results

6.3.1 Hydrogen-bonding: Grinding

The solvent-assisted grinding experiments of hydrogen-bonding were analyzed through IR spectroscopy and 60 of the 60 experiments resulted in a co-crystal formation, (the details of grinding experiment is provided in chapter 6 and appendix C). All co-crystals displayed broad bands in the 1,850 and 2,500 cm^{-1} region (as a result of O–H…N hydrogen bonds) and significant changes in the C=O stretch of the carboxylic acid, see chapter 6 for detailed results.

Table 6.2 Grinding experiment results with aliphatic and aromatic hydrogen-bond donors.

	Aliphatic diacids	**Aromatic monoacids**	**Halogen bonding**
T9	10/10	5/5	1/1
T10	10/10	5/5	1/1
T11	10/10	5/5	1/1
T12	10/10	5/5	1/1

The energies of each possible synthon in the hydrogen-bonded binary systems is listed in the table 6.3.

Table 6.3 Hydrogen-bond energies for synthon I-IV in a binary system

	Synthon I	**Synthon II**	**Synthon III**	**Synthon IV**
Energies (kJ/mol)	-70.76±1.85	-67.81±2.21	-64.87±2.86	-55.28±2.93

6.3.2 Halogen-bonding: Grinding and TGA analysis

The solvent-assisted grinding experiments of halogen-bonding were analyzed through IR spectroscopy and 4 of the 4 experiments resulted in a co-crystal. The IR analysis for identifying co-crystal formation focused on the C−F and C-I stretch of the halogen-bond donor (Figure 6.1, Table 6.4). A shift of three wave numbers or more was considered to be significant and indicative

of a positive result, i.e., co-crystal formation (the subsequent single-crystal X-ray diffraction analyses confirmed that the assignment based on IR data was correct).
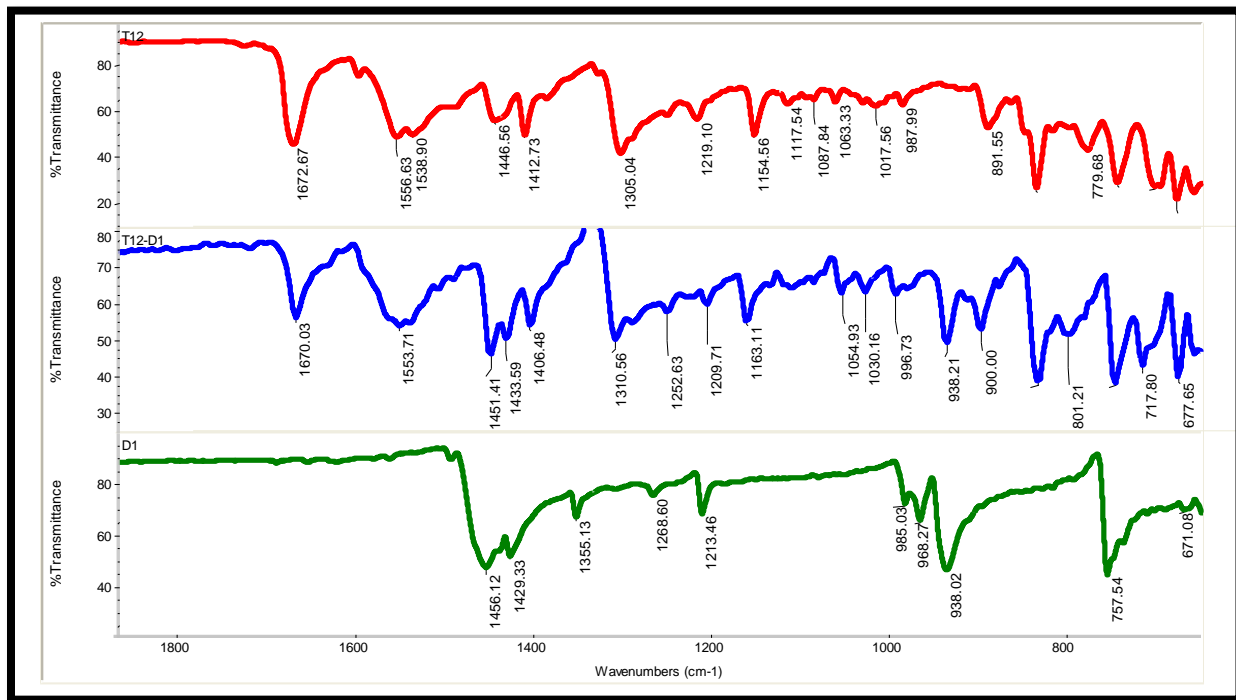
Figure 6.1 IR spectra of T12:D1 and the respective starting materials (**T12** – top; **T12:D1** co-crystal – middle; **D1** – bottom).

A TGA analysis was performed to confirm the formation of halogen-bond co-crystals (Figure 6.2).
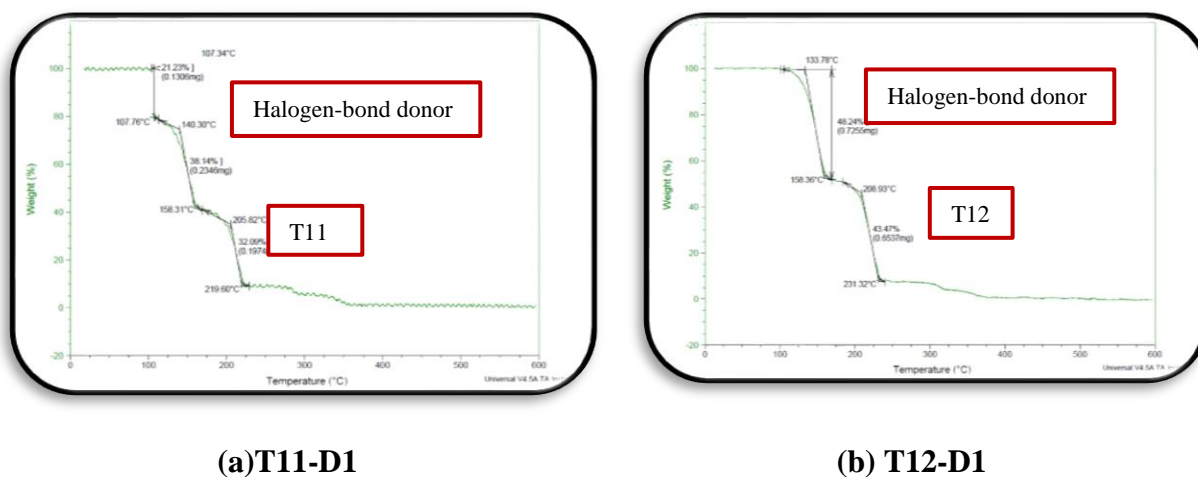
(a)T11-D1   **(b) T12-D1**

Figure 6.2 TGA on crystals of halogen-bonded co-crystals (a) **T11-D1** and (b) **T12-D1**.

Table 6.4 Grinding experiment details and co-crystal assessment of halogen-bonded co-crystals.

	Acceptor	Halogen bond donor	Grinded mixture	Co-crystal?
T9-D1	1668	1456	1648	Yes
	1598	1429	1548	
	1549	938	1457	
	1483	757	1295	
	1427		913	
	1298		745	
T10-D1	1670	1456	1681	Yes
	1541	1429	1548	
	1292	938	1453	
	1056	757	1283	
			889, 797	
T11-D1	1668	1456	1666	Yes
	1558	1429	1557	
	1534	938	1452	
	1311	757	1415	
	1195		933, 750	
T12-D1	1672	1456	1670	Yes
	1556	1429	1553	
	1412	938	1451	
	1305	757	1310	
	837		938	
			748	

6.3.3 Ternary co-crystal: Grinding and DSC analysis

A total of 60 ternary co-crystallizations were performed and analyzed using infra-red spectroscopy and 44 out of 60 (73% success rate) showed positive outcome based on solid-state grinding analysis. All positive co-crystals displayed broad bands in the 1850 and 2500 cm^{-1} region (as a result of O–H…N hydrogen bonds) as well as significant shifts in the C-F and C-I stretch of the halogen-bond donor indicting both hydrogen and halogen bonding is observed (Figure 6.3, Appendix C). Also, The IR of ternary co-crystals was compared to binary hydrogen and halogen bonded IRs and pure components as well. The DSC analysis was also done on the crystals obtained from ternary attempts to confirm single melting exothermic peak.
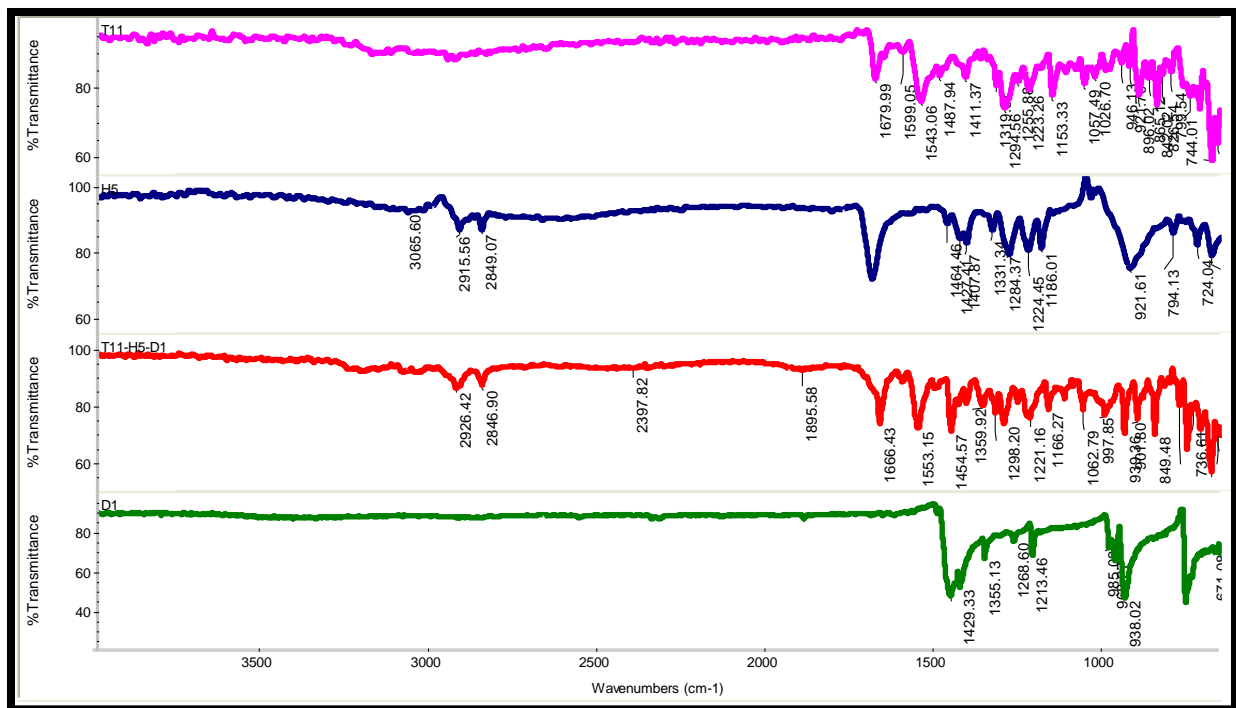
Figure 6.3 IR spectra of T11-H5-D1 co-crystal and the respective starting materials (**T11** – pink, H5-blue, **T11-H5-D1** co-crystal – red; and **D1** – green).

The melting point was obtained for ternary co-crystals and compared with the melting point of individual and binary components, Table 6.5.

Table 6.5 Melting point comparisons of ternary systems and corresponding individual and binary components.

	Target molecule	Hydrogen-bond donor	Halogen bond donor	Binary XB	Ternary system
T11-H5-D1	199-201℃	127-129℃	107-109℃	208-210℃	155-157℃
T11-H13-D1	199-201℃	139-141℃	107-109℃	208-210℃	140-141℃
T11-H15-D1	199-201℃	100-102℃	107-109℃	208-210℃	150-152℃
T12-H14-D1	213-214℃	237-238℃	107-109℃	205-206℃	185-186℃
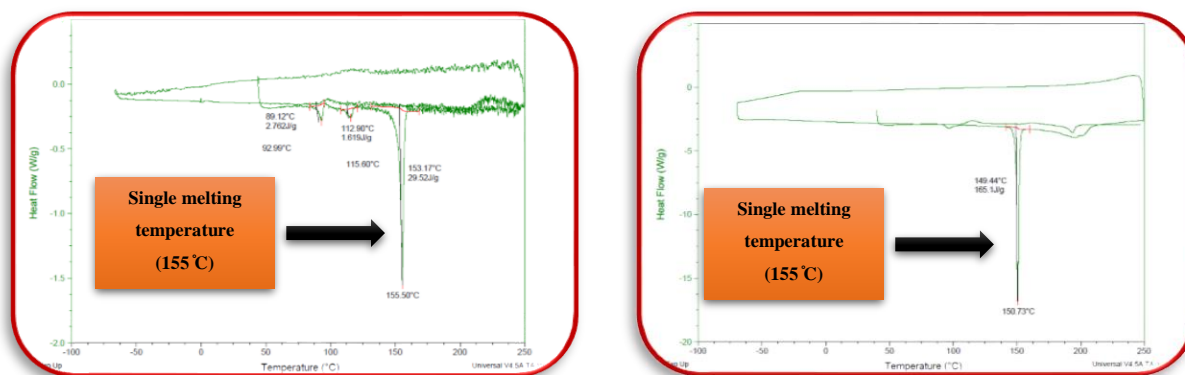
(a) T11-H5-D1

(b) T11-H15-D1
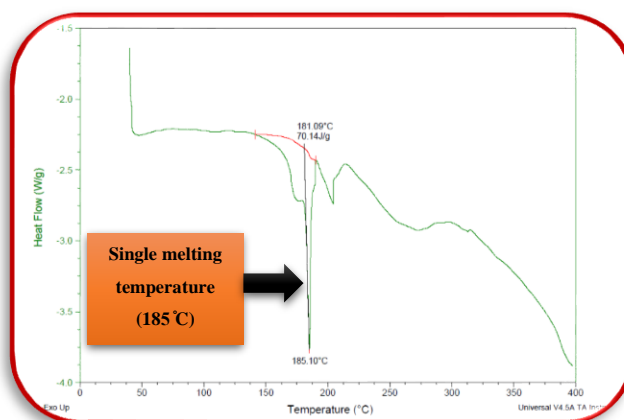
(c) T12-H14-D1

Figure 6.4 DSC analysis of ternary systems representing single melting temperatures different from individual or binary systems.

6.3.4 Crystal structures

6.3.4.1 Hydrogen bonded co-crystals

Although vibrational spectroscopy provides unambiguous information about whether a co-crystal has formed or not, it does not reveal which acceptor site(s) is/are involved and thus in order to examine any hydrogen-bond preferences, single crystal data is required. Four crystal structures were obtained from combinations of **T9-T12** with ten aliphatic acids, one for **T10**-based co-crystals, three for **T12**-based co-crystals (few highlights of crystal structures listed in Figure 6.5). Despite repeated efforts suitable crystals were not available for any co-crystal of **T9** and **T11**, even though IR spectroscopy indicates that both target molecules forms co-crystals. Two crystal

structures were obtained from combinations of **T9-T12** with five aromatic acids, one for **T9** and one for **T12**, see chapter 5 for detailed results.

(a)	(b)	(c)
Pyridyl site is vacant	**Both sites are occupied**	**Both sites are occupied**
Optimal for ternary systems	**Less optimal for ternary systems**	**Less optimal for ternary systems**
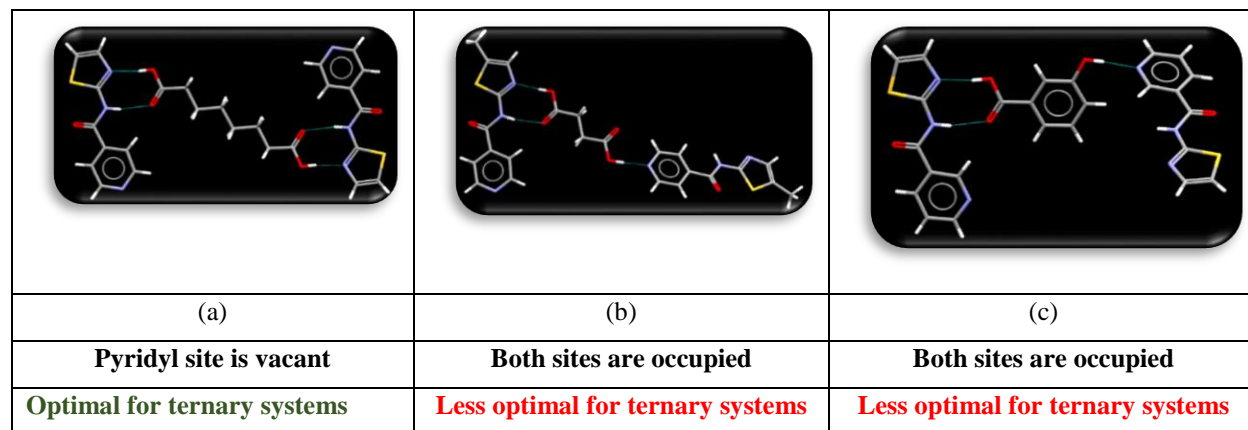
Figure 6.5 Hydrogen-bonding in the crystal structures of **(a) T11-Sub, (b) T12-Suc** and **(c) T9-3HydroxyBA.**

6.3.4.2 Halogen bonded co-crystals

Based on the values of MEPS for the acceptor atoms on **T9-T12**, the pyridine nitrogen atoms represent the best acceptors, while the thiazole nitrogen is the second-best acceptor. Three crystal structures were obtained with **D1**. One crystal structure was obtained with **T10-D1**, where the pyridine nitrogen atom (best acceptor site) of **T10** forms a halogen bond with **D1** (Figure 6.6). The halogen bond formation takes place at both ends of **D1**, resulting in a supramolecular ribbon. The second-best acceptor site on T10 forms homomeric interactions with adjacent molecule via NH…N hydrogen bonding. Same interactions were observed with **T11-D1** and **T12-D1** leading to 1D supramolecular ribbons (Figure 6.7 and 6.8). The ribbons were extended into 2D architecture via CH----F interactions (Figure 6.6b).
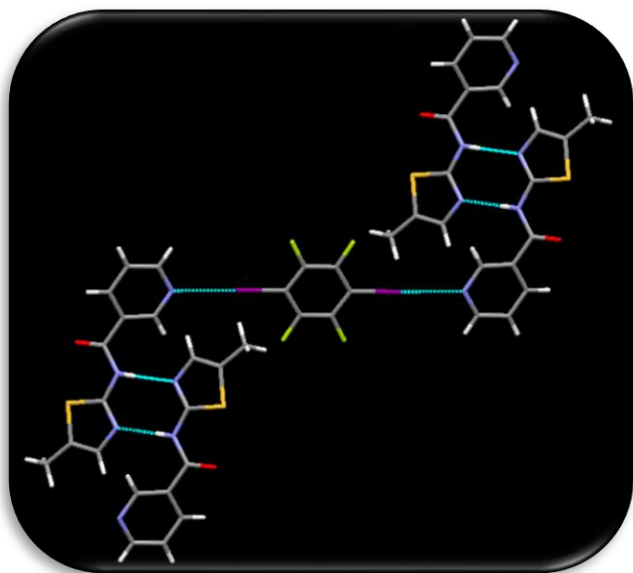
(a)
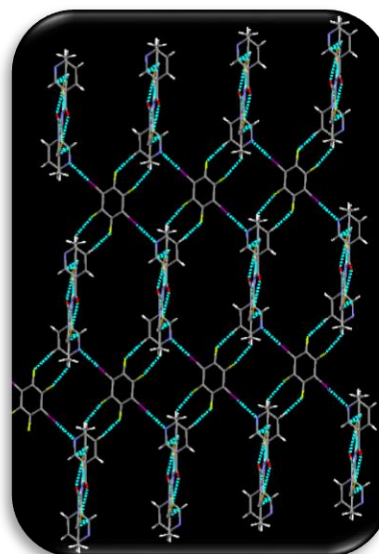
(b)

Figure 6.6 (a) One-dimensional halogen-bonded (heteromeric) and hydrogen-bonded (homomeric) chain in the crystal structure of **T10-D1**, (b) crystal packing in **T10-D1**.
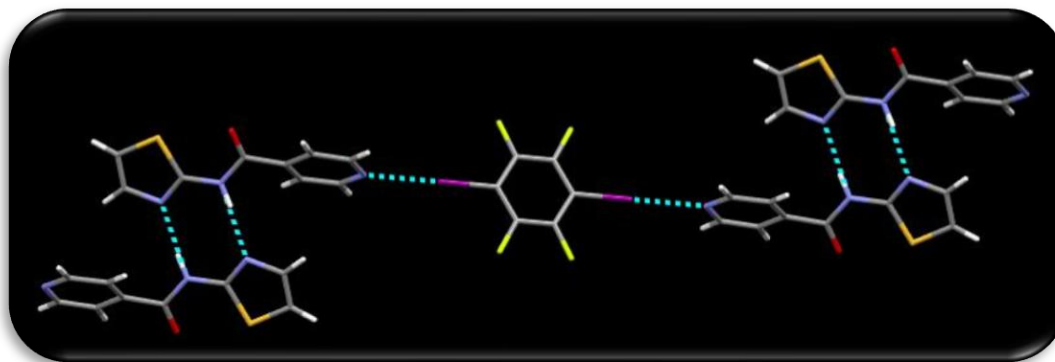
Figure 6.7 One-dimensional halogen-bonded (heteromeric) and hydrogen-bonded (homomeric) chain in the crystal structure of **T11-D1**.
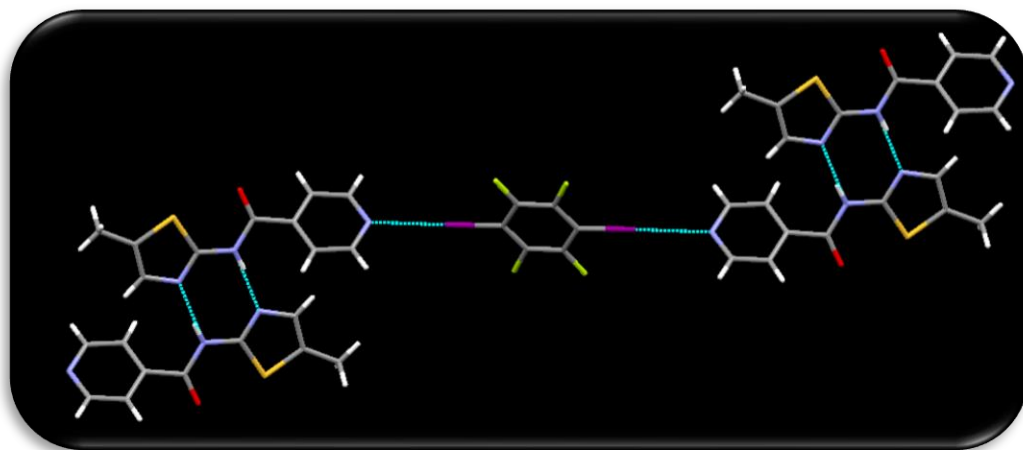
Figure 6.8 One-dimensional halogen-bonded (heteromeric) and hydrogen-bonded (homomeric) chain in the crystal structure of **T12-D1.**

6.3.4.3 Ternary co-crystals

We were able to obtain good quality single crystals for two ternary co-crystals, **T11-H5-D1** and **T11-H13-D1**(Figure 6.9 and 6.10).

(a)

(b)
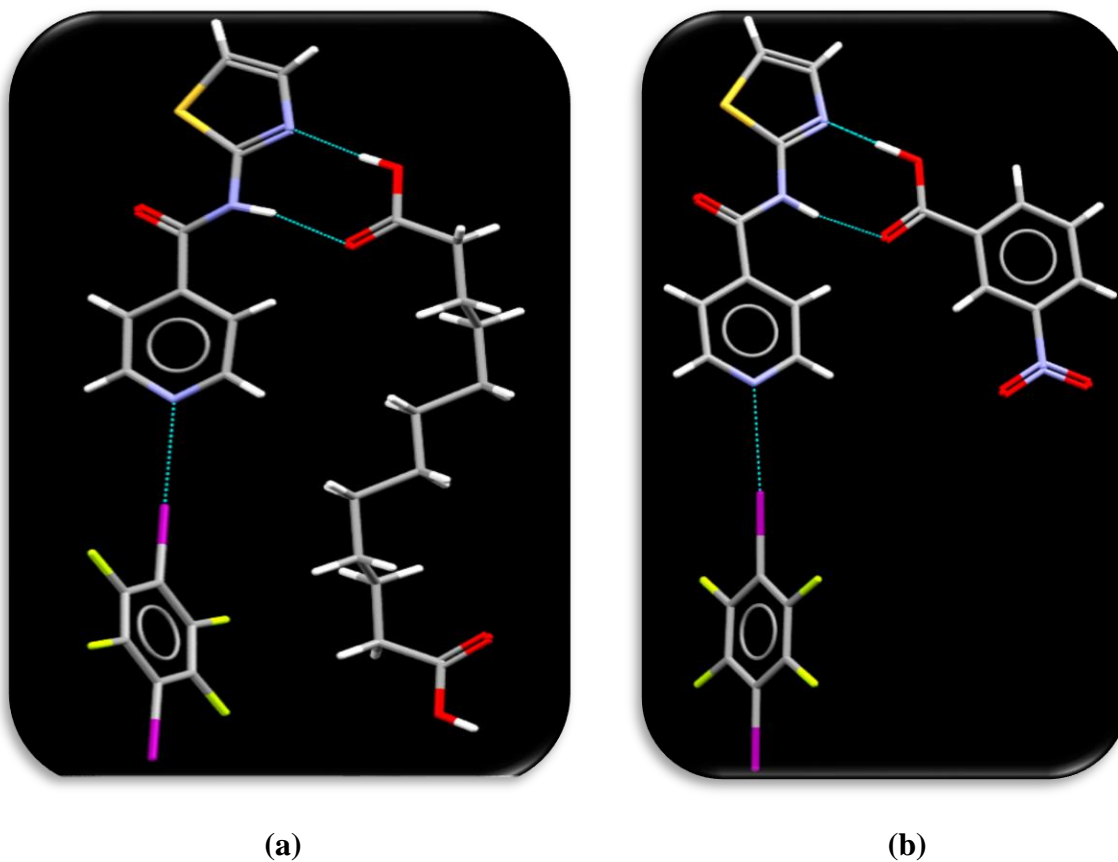
Figure 6.9 Hydrogen-bonded and halogen-bonded based ternary co-crystals (a) **T11-H5-D1** and (b) **T11-H13-D1.**
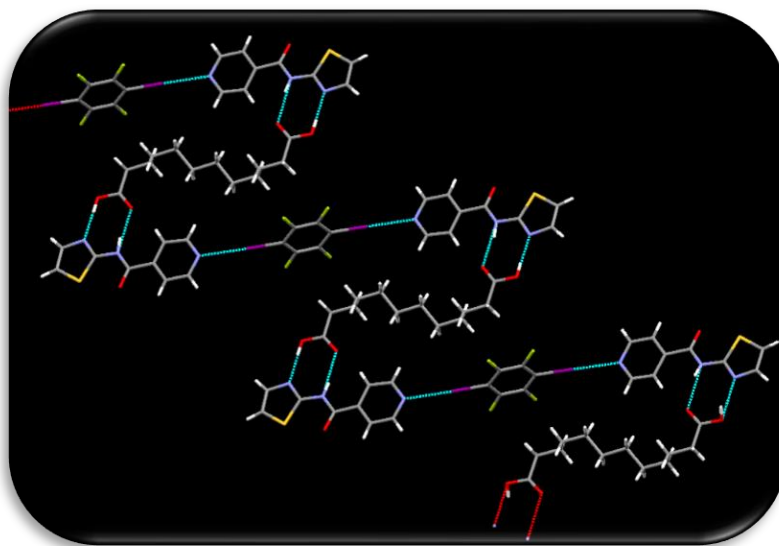
(a)
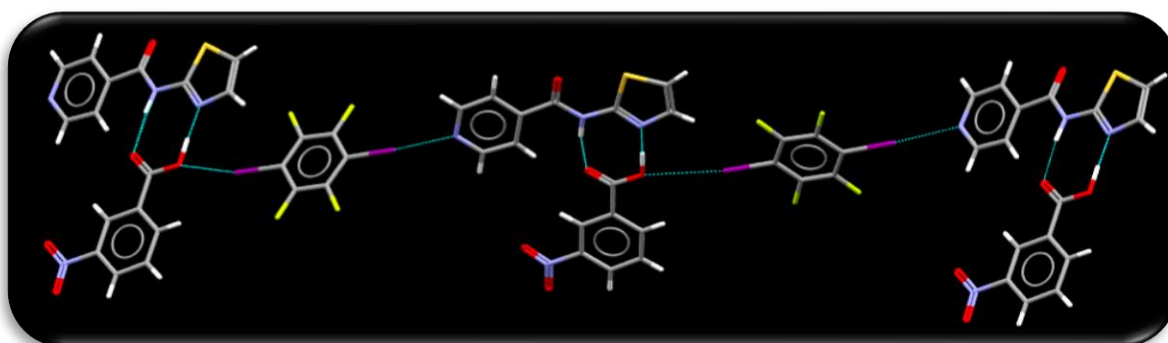
(b)

Figure 6.10 Crystal packing in the ternary co-crystals **(a) T11-H5-D1** and **(b) T11-H13-D1.**

6.4 Discussion

The starting point of our design of a ternary cocrystal was to first obtain binary co-crystals of **T9-T12** with hydrogen-bond donors and halogen bond donors. If the hydrogen-bond donor binds to only recognition point 1 (synthon I), and the halogen bond donor can bind to recognition point 2, ternary co-crystals are possible. However, the following three possibilities could cause a possible complication towards the success of ternary systems:

1. If the hydrogen-bond donor binds to both recognition 1 and recognition 2 (synthon **III**), then the halogen bonding in a ternary system is unlikely.
2. If the hydrogen-bond donor binds to only recognition point 2 on both sides of acid, then halogen bonding in a ternary system is unlikely.
3. If one binary system is stronger than other or have lower solubility in solution than other, it will co-crystallize first, then only binary co-crystals are possible (hydrogen or halogen bonding) and ternary would be less likely.

Aromatic acids with an additional donor group such as HydroxyBA were chosen as reference co-formers which allowed us to include a monotopic donor, hydroxyl OH on the acid. Having a monotopic donor on the acid would mean it can bind to recognition point 2 while blocking the halogen bond donor from binding and not resulting in ternary co-crystals.

6.4.1 Hydrogen-bonded binary co-crystals

The binary co-crystals with hydrogen-bond donors were successfully obtained. There was 100% success rate with the chosen hydrogen-bond donors (60 grinding and solution experiments), and we were able to obtain crystal structures of six co-crystals (four with aliphatic acids and 2 with aromatic acids), Figure 6.11. All co-crystals are described in detail in chapter 5.
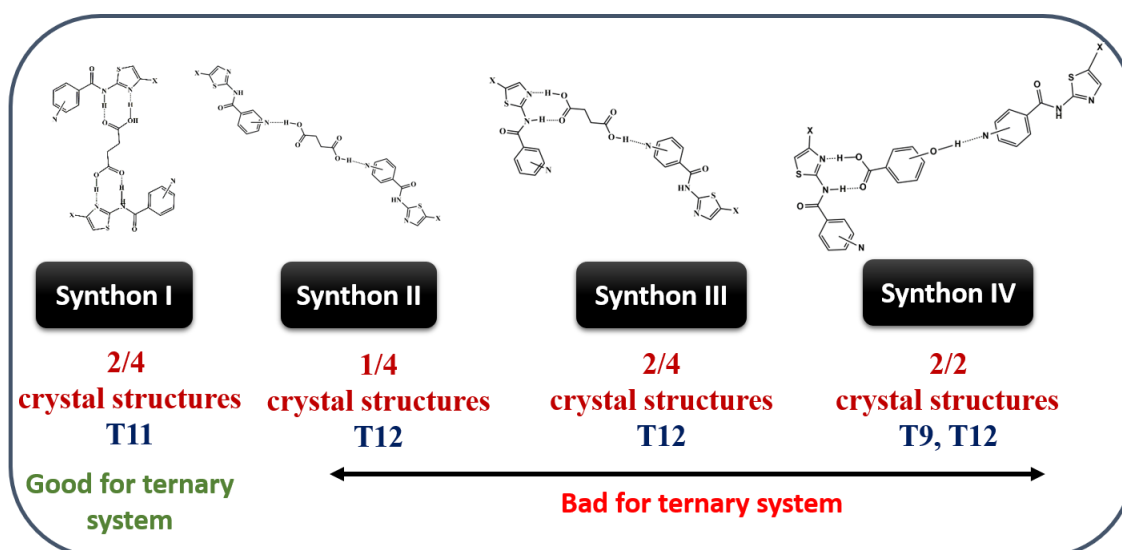
Figure 6.11 Summary of hydrogen-bonded binary co-crystals.

The results indicated that **T11** might be better acceptor compared to **T9, T10** and **T12** to make ternary co-crystals because in hydrogen-bonding, the recognition point 2 (pyridine N) was vacant.

6.4.2 Halogen-bonded binary co-crystals

Three 2:1 stoichiometric binary halogen bonded co-crystals were obtained; **T10-D1**, **T11-D1** and **T12-D1**. **T9** also formed halogen-bond co-crystal with 1,4-diiodotetrafluorobenzene but the crystal quality wasn't good for X-ray diffraction studies. The iodine of halogen bond donor, 1,4-diiodotetrafluorobenzene binds to pyridine N via I…N single point interaction. The recognition point 1 on each acceptor (**T10-T12**) forms homomeric $R_2^2(8)$ NH(amide)…N(cyclic) hydrogen-bond interaction and the weak CH…F interactions expands the structure into 2-D architecture, Figure 7.12.
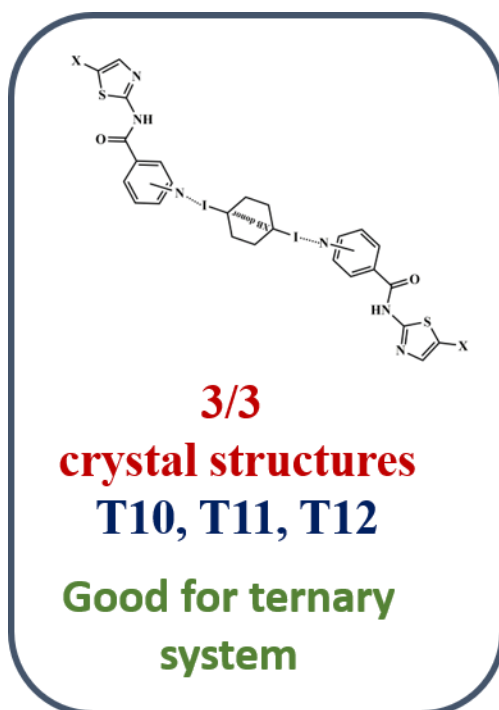
Figure 6.12 Summary of halogen-bonded binary co-crystals.

6.4.3 Ternary co-crystals

Once the binary co-crystals were obtained, the next step was to use the combination of hydrogen and halogen bonding to form ternary co-crystals. Hydrogen bonding in acceptor **T11** was most optimal for ternary system because the pyridine nitrogen site was vacant which means the halogen bond donor can bind to the pyridine nitrogen. Once the binary cocrystals were obtained with

hydrogen and halogen bond donors, we attempted ternary co-crystallization with all four target molecules, 15 hydrogen-bond donors and one halogen-bond donor giving 60 ternary grinding experiments. Based on grinding experiment results, 44 out of 60 co-crystallization attempts (74% success rate) formed ternary co-crystals as confirmed by IR spectroscopy by carefully looking at peak shifts of both donor groups. All 60 co-crystallizations were put through solution crystallization via slow solvent evaporation method in methanol and ethanol solvent. We were able to successfully obtain four ternary complexes crystals as confirmed by DSC. The melting point of ternary co-crystals was compared with the melting temperatures of individual and binary components (Figure 6.13) and it was found to be different from halogen-bonded co-crystals confirming IR results that it is indeed a three-component system.

In ternary co-crystal **T11-H5-D1,** an aliphatic hydrogen-bond donor, dodecanedioic acid binds to recognition point 1 and the 1,4-diiodotetrafluorobenzene binds to the recognition point 2 on both ends of halogen bond donor. In ternary co-crystal **T11-H13-D1**, an aromatic hydrogen-bond donor, 3-nitrobenzoic acid binds to recognition point 1 whereas the halogen-bond donor 1,4-diiodotetrafluorobenzene binds to the recognition point 2 on one end and oxygen of OH of aromatic acid on the other end.
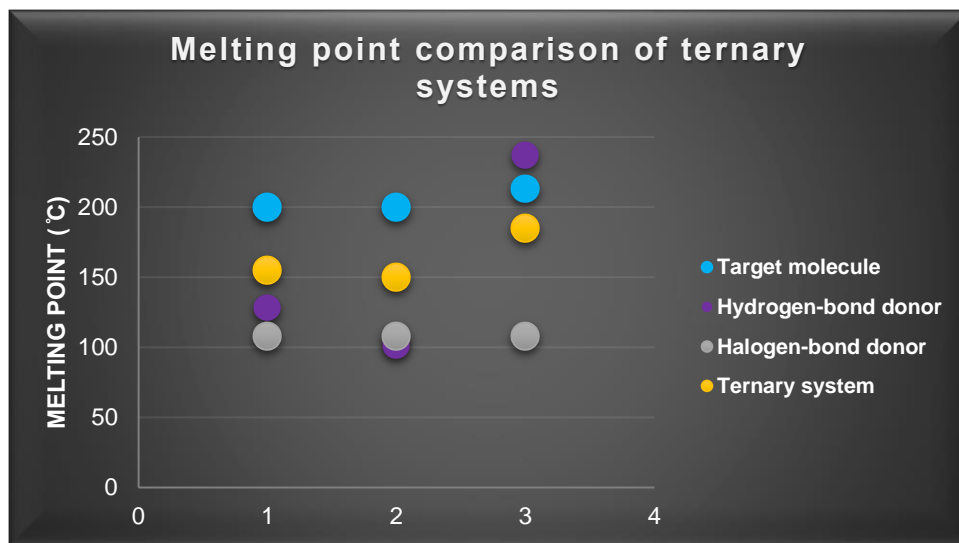
Figure 6.13 Melting point analysis of ternary co-crystals and individual components.

160

6.4.4 Ternary crystal growth issues

The main criteria for co-crystal formation include comparable solubility, stable synthons, and shape/size mimicry among the co-formers[29, 31, 34-36]. Most solution crystal growth attempts in this study resulted in binary halogen-bonded co-crystals with few hydrogen-bond binary systems indicating that there were some issues with the formation of crystals of ternary complexes from solution.

1. Relative solubility: The path toward crystal growth depends on the difference in solubility of the co-formers in a certain solvent. It is believed that the lower the solubility difference between the co-formers, the greater the probability of crystal formation with desired components[34]. However, if there is solubility differences between the individual components, or formation of binary systems, then it is possible that low solubility component will crystallize out first. In this study, halogen-bond binary co-crystal crystallizes out first indicating it was the less soluble co-crystal.
2. Presence of multiple functional groups: The presence of multiple functional groups in the solution can cause competition between donor and acceptors groups to choose the right binding site and delay the nucleation and crystal growth phenomenon and could result in crystallization of either individual or binary components.
3. Stoichiometric ratios: The different stoichiometric ratios of target molecules, hydrogen-bond donor and halogen-bond acceptor could play an important role in the outcome of binary or ternary systems. In this study, we have analyzed 2:1:1 (target: aliphatic hydrogen-bond donor: halogen-bond donor) and 1:1:0.5 (target: aromatic hydrogen-bond donor: halogen-bond donor) stoichiometries only. It is possible that varying stoichiometric ratios could give us more co-crystals.

6.5 Conclusions

This study shows that ternary systems are easy to form with a well-defined design strategy of target molecules in solid state. We were successfully able to form binary hydrogen-bond co-crystals with target molecules (**T9-T12**) forming all four synthons in the crystal structures. Also, these target molecules were good candidates for halogen-bonding particularly with 1,4-diiodotetrafluorobenzene (100% supramolecular yield). Finally, 60 out of 60 ternary co-

crystallization attempts resulted in the crystal structure of two ternary co-crystal from solution state.

1. We successfully made hydrogen-bond binary co-crystals and obtained six new co-crystals but only **T11** target molecule has the optimal binding pattern at recognition point 1 without using the recognition point 2.

2. We successfully formed crystal structure of three halogen-bond binary co-crystals at recognition point 2 without using the recognition point 1.

3. We were able to make 60 ternary co-crystals and crystals of two of these ternary systems were obtained.
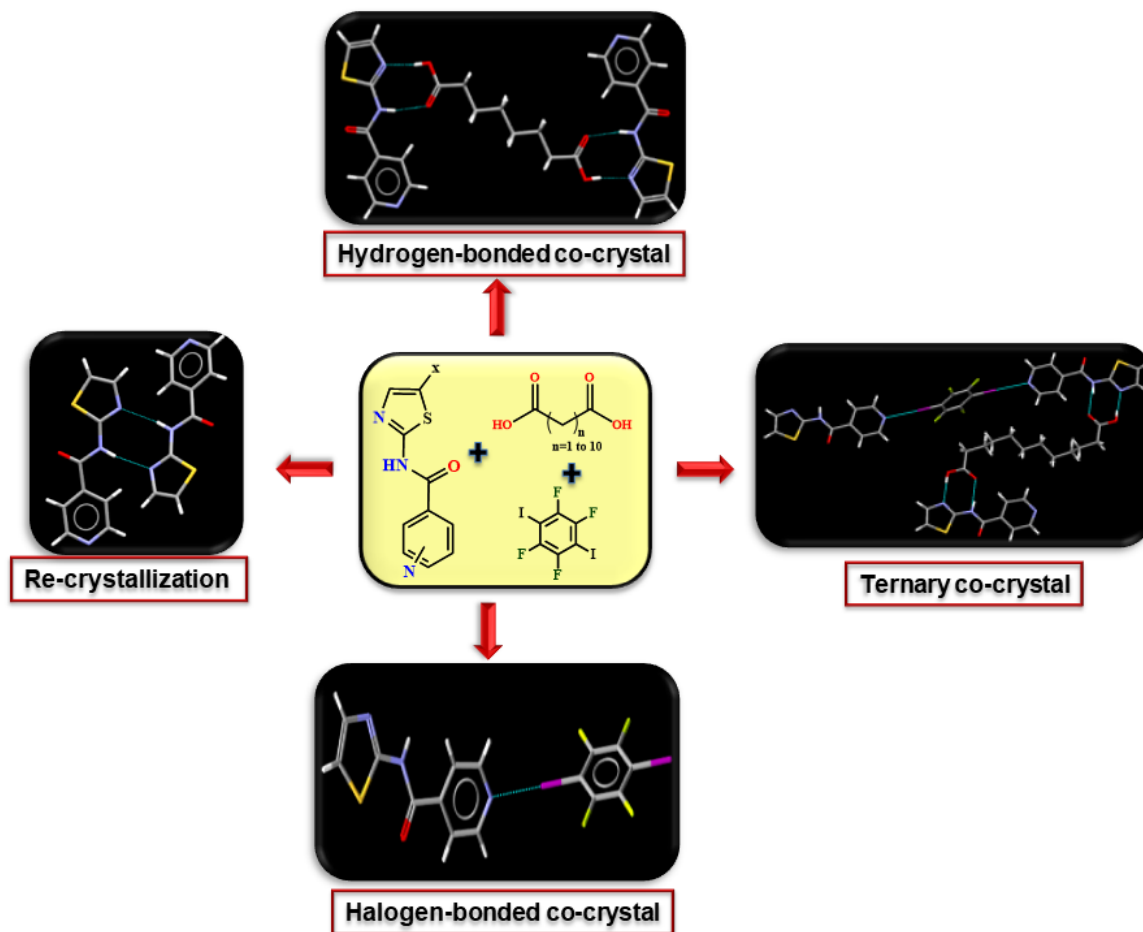
Figure 6.14 Summary of re-crystallization, binary and ternary co-crystals.

This system worked well in the solid-state, but it still raises questions about interaction strength and solubility issues while getting crystals from solution state. It is important to have interactions

of graded strength. If there are molecules **T1**, **H1** and **X1** where the interaction **T1-X1** (say halogen bond) is stronger than the interaction **T1-H1** (say hydrogen bond), then it is fair to expect that initial formation of **T1X1** in solution can lead to the association of **H1** to give an **T1-H1-X1** aggregate giving a ternary cocrystal. Yet, **T1-X1** should not be too strong and/or the resulting binary cocrystal too insoluble, because then, it will be preferentially isolated[36]. A fine balance of interactions and solubilities is therefore needed to get a ternary cocrystal, the design of which remains one of the big synthetic challenges in crystal engineering of molecular organic solids.

6.6 References

1.   Etter, M. C., *The Journal of Physical Chemistry,* **1991,** *95* (12), 4601-4610.

2.   Desiraju, G. R., *Angewandte Chemie International Edition in English,* **1995,** *34* (21), 2311-2327.

3.   Aakeroy, C. B.; Desper, J.; Urbina, J. F., *Chemical Communications,* **2005,** (22), 2820-2822.

4.   Desiraju, G. R., *Accounts of Chemical Research,* **2002,** *35* (7), 565-573.

5.   Bosch, E., *CrystEngComm,* **2007,** *9* (3), 191-198.

6.   Walsh, R. B.; Padgett, C. W.; Metrangolo, P.; Resnati, G.; Hanks, T. W.; Pennington, W. T., *Crystal Growth & Design,* **2001,** *1* (2), 165-175.

7.   Metrangolo, P.; Neukirch, H.; Pilati, T.; Resnati, G., *Accounts of Chemical Research,* **2005,** *38* (5), 386-395.

8.   Shirman, T.; Freeman, D.; Posner, Y. D.; Feldman, I.; Facchetti, A.; van der Boom, M. E., *Journal of the American Chemical Society,* **2008,** *130* (26), 8162-8163.

9.   Nguyen, H. L.; Horton, P. N.; Hursthouse, M. B.; Legon, A. C.; Bruce, D. W., *Journal of the American Chemical Society,* **2004,** *126* (1), 16-17.

10.   Bhogala, B. R.; Nangia, A., *Crystal Growth & Design,* **2003,** *3* (4), 547-554.

11.   Nguyen, T. L.; Fowler, F. W.; Lauher, J. W., *Journal of the American Chemical Society* **2001,** *123* (44), 11057-11064.

12.   Dale, S. H.; Elsegood, M. R. J.; Hemmings, M.; Wilkinson, A. L., *CrystEngComm* **2004,** *6* (36), 207-214.

13.   Aakeröy, C. B.; Salmon, D. J.; Smith, M. M.; Desper, J., *Crystal Growth & Design* **2006,** *6* (4), 1033-1042.

14.   Aakeröy, C. B.; Schultheiss, N. C.; Rajbanshi, A.; Desper, J.; Moore, C., *Crystal Growth & Design* **2009,** *9* (1), 432-441.

15.   Aakeroy, C. B.; Desper, J.; Helfrich, B. A., *CrystEngComm* **2004,** *6* (5), 19-24.

16.   Cavallo, G.; Metrangolo, P.; Milani, R.; Pilati, T.; Priimagi, A.; Resnati, G.; Terraneo, G., *Chemical Reviews* **2016,** *116* (4), 2478-2601.

17.   Desiraju, G. R., *Current Opinion in Solid State and Materials Science* **1997,** *2* (4), 451-454.

18.   Aakeröy, C. B.; Chopade, P. D.; Desper, J., *Crystal Growth & Design* **2011,** *11* (12), 5333-5336.

19.   Aakeröy, C. B.; Beatty, A. M.; Helfrich, B. A., *Angewandte Chemie International Edition* **2001,** *40* (17), 3240-3242.

20.   Almarsson, O.; Zaworotko, M. J., *Chemical Communications* **2004,** (17), 1889-1896.

21.   Schultheiss, N.; Newman, A., *Crystal Growth & Design* **2009,** *9* (6), 2950-2967.

22.   Aakeroy, C. B.; Desper, J.; Smith, M. M., *Chemical Communications* **2007,** (38), 3936-3938.

23.   Chakraborty, S.; Rajput, L.; Desiraju, G. R., *Crystal Growth & Design* **2014,** *14* (5), 2571-2577.

24.   Etter, M. C., *Accounts of Chemical Research* **1990,** *23* (4), 120-126.

25.   Tothadi, S.; Mukherjee, A.; Desiraju, G. R., *Chemical Communications* **2011,** *47* (44), 12080-12082.

26.   Dubey, R.; Mir, N. A.; Desiraju, G. R., *IUCrJ* **2016,** *3* (2), 102-107.

27.   Mir, N. A.; Dubey, R.; Desiraju, G. R., *IUCrJ,* **2016,** *3* (2), 96-101.

28.   Paul, M.; Chakraborty, S.; Desiraju, G. R., *Journal of the American Chemical Society* **2018,** *140* (6), 2309-2315.

29.   Topić, F.; Rissanen, K., *Journal of the American Chemical Society,* **2016,** *138* (20), 6610-6616.

30.   Aitipamula, S.; Wong, A. B. H.; Chow, P. S.; Tan, R. B. H., *CrystEngComm,* **2013,** *15* (29), 5877-5887.

31.   Bolla, G.; Nangia, A., *IUCrJ* **2016,** *3* (2), 152-160.

32.   Seaton, C. C.; Blagden, N.; Munshi, T.; Scowen, I. J., *Chemistry – A European Journal* **2013,** *19* (32), 10663-10671.

33.   Sandhu, B.; McLean, A.; Sinha, A. S.; Desper, J.; Sarjeant, A. A.; Vyas, S.; Reutzel-Edens, S. M.; Aakeröy, C. B., *Crystal Growth & Design,* **2018,** *18* (1), 466-478.

34.   Tothadi, S.; Sanphui, P.; Desiraju, G. R., *Crystal Growth & Design,* **2014,** *14* (10), 5293-5302.

35.   Bolla, G.; Nangia, A., *Chemical Communications,* **2015,** *51* (85), 15578-15581.

36.   Tothadi, S.; Desiraju, G. R., *Chemical Communications,* **2013,** *49* (71), 7791-7793.

Chapter 7 - Solid-State landscape of amide containing activated halogen bond donors

7.1 Introduction

Crystal packing is the result of the optimization of various possible intermolecular interactions between the molecules in the solid state.[1] Identification of supramolecular synthons between various functional groups simplifies the understanding and prediction of crystal structures to some extent.[2] Therefore, predicting supramolecular synthons[3-5] is less complex if the given molecule has a small number of functional groups which are self-complementary to form predictable robust synthons. As the number of functional groups on a molecule increases, the interference between supramolecular synthons becomes a major concern leading to synthon polymorphism[6-8] and synthon crossover,[9-10] Chapter 3-5.
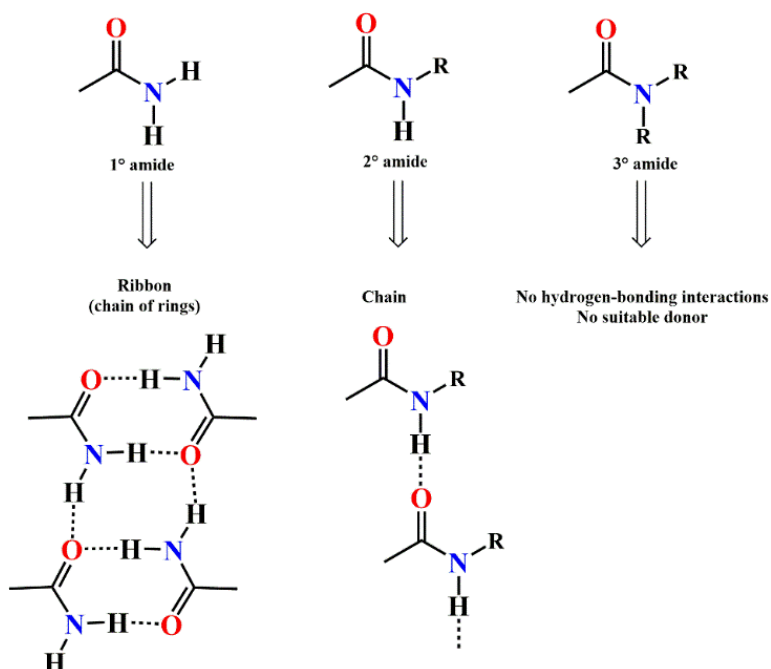
Scheme 7.1 Supramolecular synthons formed by 1°, 2° and 3° amides.[1]

One approach to minimize synthon crossover[11-13] and polymorphism[14-15] is to incorporate into the molecule a small number of functional groups that can interact intermolecularly and to use these interactions to limit the possible arrangements of the molecules in space with respect to one another. The amide functionality serves as a building block in this context from which a variety of

supramolecular structures can be assembled based on the number of substituents on the nitrogen atom, scheme 7.1. The type of structure formed is dependent on various factors such as geometry of amide group, the number and type of substituents attached to nitrogen atom and the number of amide groups present in a molecule as well steric and electronic effects of neighboring functional groups in a molecule. 2° Amides tend to form repeated chains via NH(amide)….C=O interactions. However, this interaction can be easily disrupted if a stronger acceptor, such as pyridine is present in the molecule or due to supramolecular chelating effects associated with NH amide (see chapter 3-5), Scheme 7.2.

(a)   (b)[16]
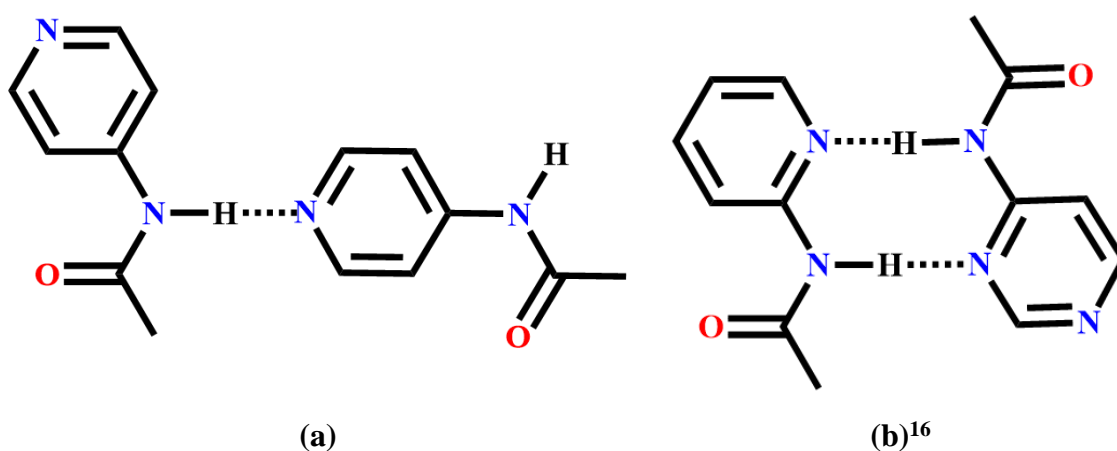
Scheme 7.2 Two possible ways the amide…amide chain interaction is at risk of synthon crossover.

In this context, a good synthetic strategy is needed to obtain desired synthons whilst keeping the amide…amide interaction intact. A good approach could be to introduce functional groups such as iodine and pyridine that can form halogen bonding, Scheme 7.3. Therefore, we want to design molecules in such a way to form both hydrogen and halogen bonding in a molecule without having synthon crossovers. Our design strategy includes two donors (hydrogen bond; NH amide and halogen bond donor, iodine) and two acceptors (pyridine and C=O) in a same molecule. There are two synthon possibilities; Synthon A includes amide disruption (preferred based on Etter's rule) and synthon B keeps amide…amide interaction intact.
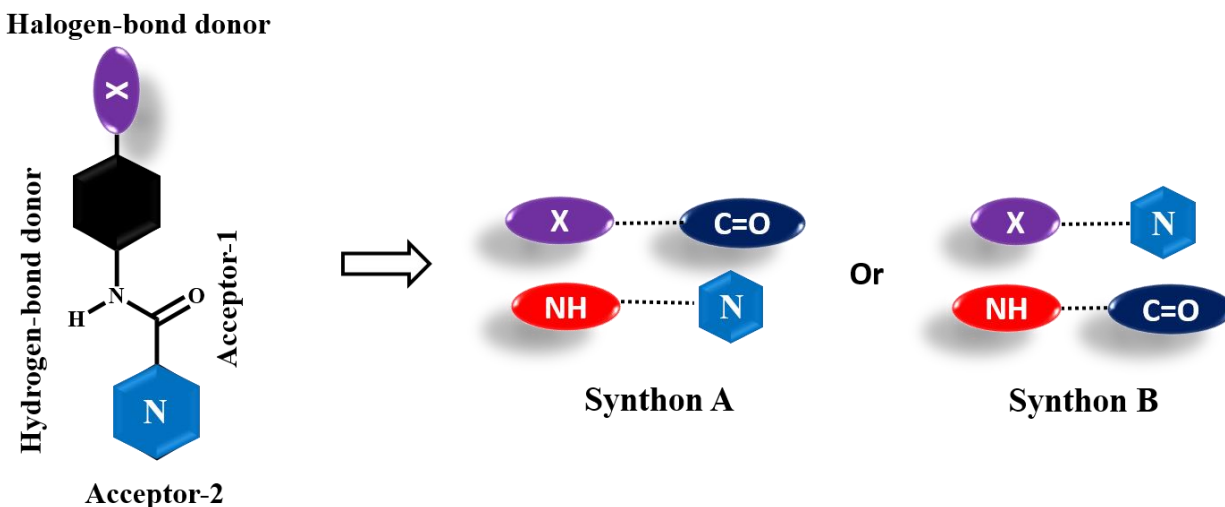
Scheme 7.3 Two types of synthons are possible in our design molecule.

In order to keep the amide functionality intact, synthon **B** is the desired interaction. The desired interaction can be achieved via two methods.

1. **Activation of halogen atom**: The first possibility is by activating the iodine atom on the molecule to make a strong halogen bond donor. Literature has shown that by introducing an *sp*-hybridized carbon atom[1] or by incorporating a fluorinated backbone[2] next to the halogen atom, the molecular electrostatic potential that determines the σ-hole is enhanced. In order to develop a new series of potent XB donors and keep amide functionality intact, we decided to activate the halogen bond donors by both methods; by incorporating fluorinated backbone next to iodine and by introducing an sp-hybridized carbon atom (Scheme 7.4).
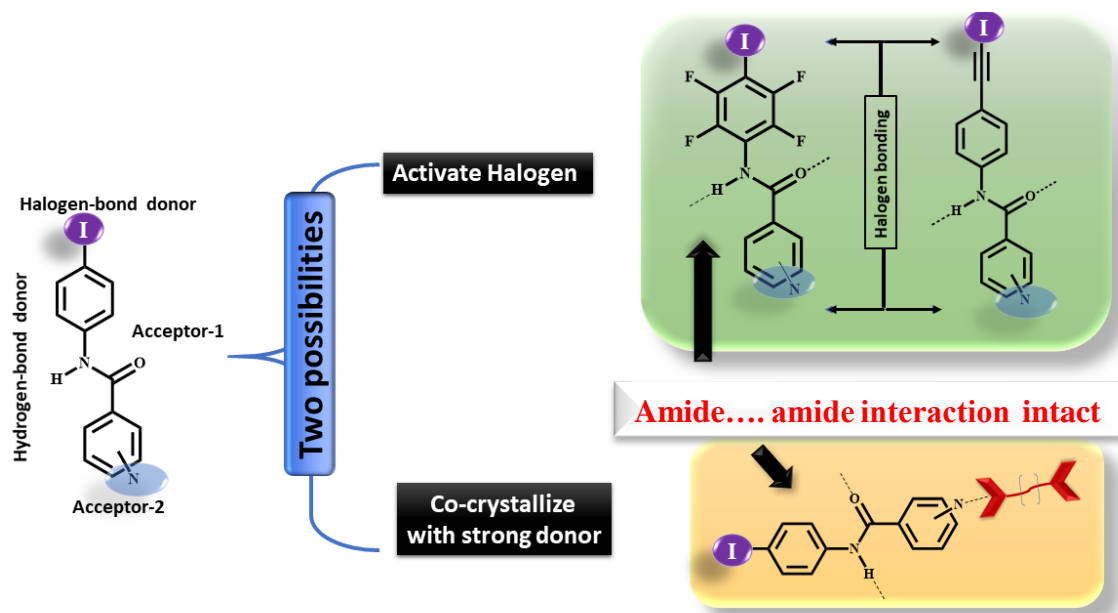
Scheme 7.4 Two possibilities to keep the amide functionality intact either by activating the halogen group or via hydrogen-bonded co-crystallization.

2. **Co-crystallization**: The second possibility to keep the amide functionality intact is by finding a suitable stronger hydrogen bond donor than amide NH such as carboxylic acid OH which can bind to pyridine N via hydrogen-bond co-crystallization and leave the amide functionality intact (Scheme 7.4). Our hypothesis is that hydrogen-bonding will be possible in unactivated molecules whereas it might be challenging in activated molecules because all donors and acceptors will be already satisfied, Scheme 7.5.

Hydrogen-bonded co-crystals
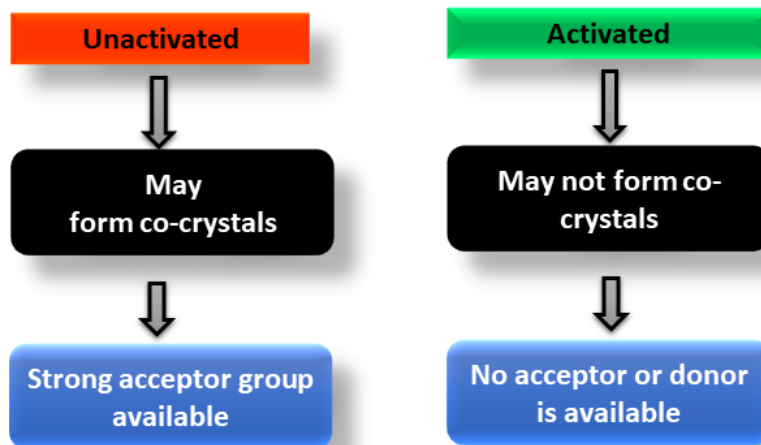
Scheme 7.5 Our hypothesis for hydrogen-bonded co-crystals with this group of molecules.

Eight different molecules are designed for this study; **3N-IB** and **4N-IB** are considered unactivated molecules, whereas **3N-ITFB**, 3N-EB, 3N-IEB, **4N-ITFB**, **4N-EB** and **4N-IEB** are considered activated molecules.

(a)   (b)
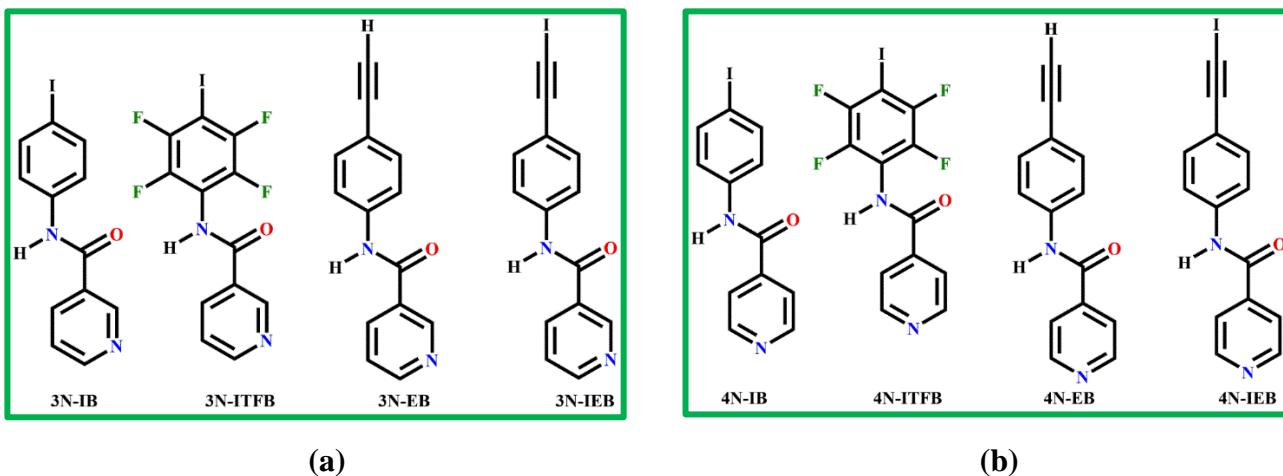
Scheme 7.6 (a) meta and (b) para substituted unactivated and activated target molecules.
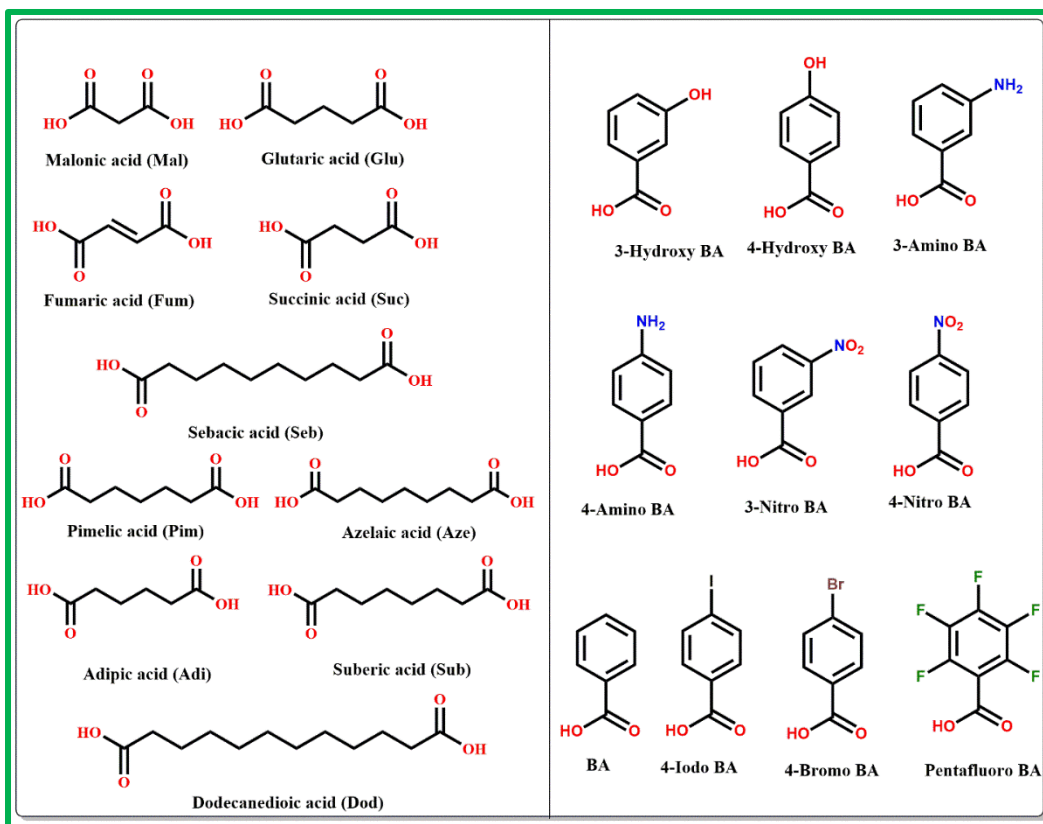
Scheme 7.7 Aliphatic and aromatic acids used in this study.

In this chapter we will examine/answer four different questions:

1. Can we keep the amide functionality intact by activating the halogen-bond donors in homomeric synthons?

2. Can we keep amide functionality intact via hydrogen-bonded co-crystallization?

3. What is the tipping point for heteromeric vs homomeric hydrogen bond interactions in this series of molecules?

4. Can ethynyl based H and iodoethynyl molecules based I; two atoms of radically different size and chemical characteristics display "synthon mimicry" in the solid state?
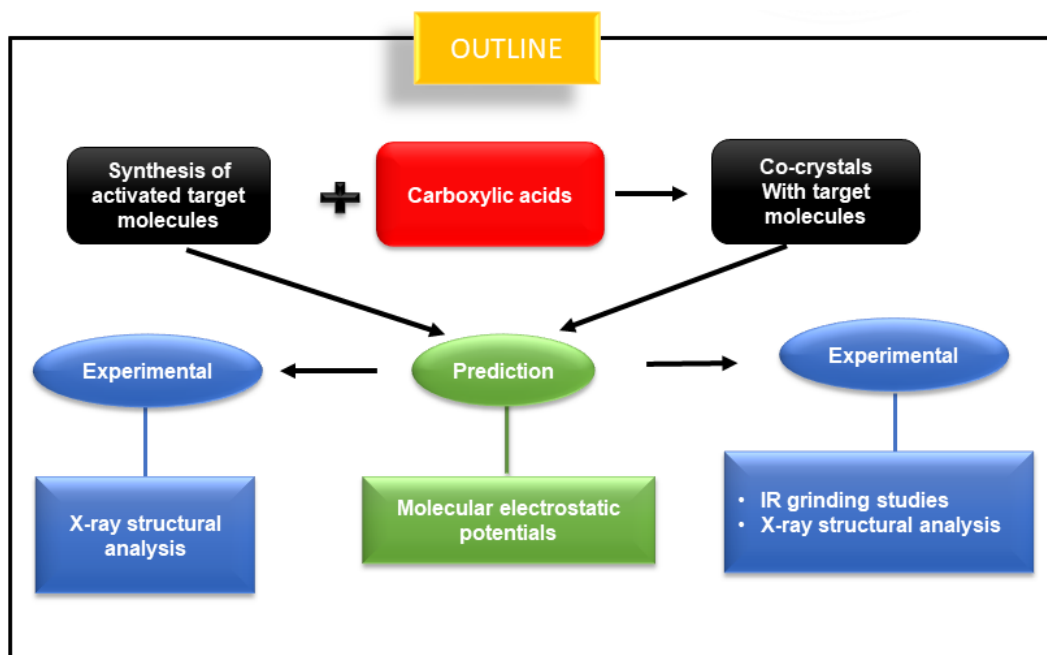
Scheme 7.8 An outline of this study

7.2 Experimental

7.2.1 General

All precursors, solvents and donors were purchased from commercial sources and used without further purification. ^{1}H NMR spectra and ^{13}C NMR spectra were recorded on Varian Unity plus 400 MHz spectrometer. Melting points of target molecules and co-crystals were determined using a Fischer–Johns Mel-Temp melting point apparatus and uncorrected.

7.2.2 Synthesis

7.2.2.1 Synthesis of N-(4-iodophenyl) nicotinamide, 3N-IB[17]
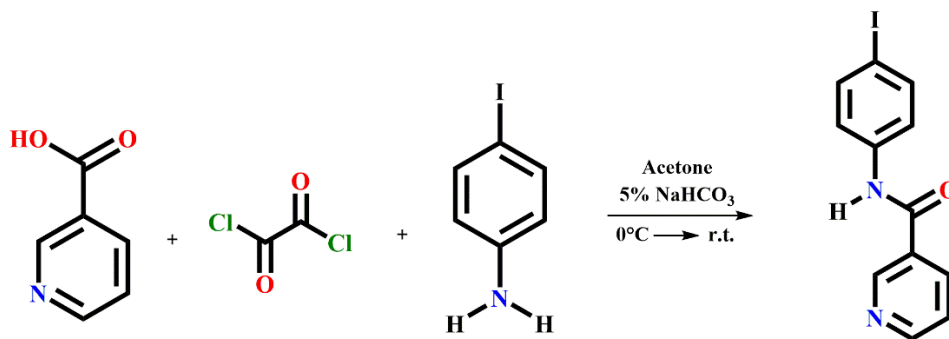
To stirred ice bathed neat oxalyl chloride, nicotinic acid was added to form thick slurry. The reaction mixture was left at room temperature for one hour during which excess oxalyl chloride was allowed to evaporate in fume hood to yield white powder. After one hour, 4-iodoaninline (1.5g, 6.85 mmol) in acetone was added under vigorous stirring to the resulting powder under ice bath conditions. The reaction was subsequently warmed to room temperature and stirred for 15 minutes. The reaction was terminated by quenching with 5% aqueous $NaHCO_3$ solution. The resulting crude product was washed with water and filtered to yield final compound as an off-white powder. Yield: 2.0g (90%); ^{1}H NMR (400 MHz, DMSO) δ ppm: 7.57 (d, 2H), 7.61(m, 2H), 7.73 (m, 1H), 8.27(d, 1H), 8.76 (d,1H), 9.09 (s, 1H), 10.52(m, 1H).

7.2.2.2 Synthesis of N-(4-iodophenyl) isonicotinamide, 4N-IB[18]
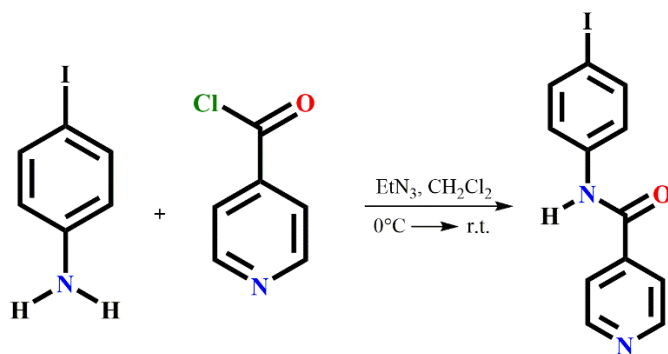
Isonicotinoyl chloride (2.00 g, 14.2 mmol) in 3 ml of triethylamine was added to 40ml of methylene chloride. The mixture was chilled to 0° C in an ice-bath for 5 minutes. Then, 4-iodoaniline (2.70g, 12 mmol) was added slowly to the cold solution over a period of 15 minutes. The reaction mixture was stirred at room temperature overnight. The resulting precipitate was collected on a frit, recrystallized with methanol to afford the final product as an off-white powder. Yield: 3.5g (87%); ^{1}H NMR (400 MHz, DMSO) δ ppm: 7.63 (d, 2H), 7.71(d, 2H), 7.84 (d, 2H), 8.78(d, 2H), 10.57(m, 1H).

7.2.2.3 Synthesis of 2,3,5,6-tetrafluoro-4-iodoaniline[19]
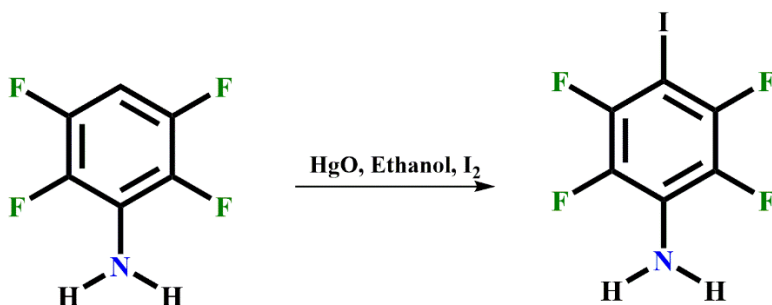
1.194g (5.5 mmol) of yellow/orange HgO was added to a solution of 1.35g (8.2 mmol) of 2,3,5,6-tetrafluoroaniline in 60 ml of ethanol. The mixture was stirred for 30 minutes and then 1.86 g (7.3g) of I_2 was added. The mixture was stirred overnight. The compound was rota evaporated to remove the solvent, then the product was re-dissolved in dichloromethane and washed several times with saturated $Na_2S_2O_3$ solution. The compound was dried over $MgSO_4$ solution. It was rota evaporated to get dark brown crystalline solid as the desired product. The product was confirmed using ^1H NMR. Yield: 2.2g (89%); ^1H NMR (400 MHz, DMSO) δ ppm: 4.11 (d, 2H).

7.2.2.4 Synthesis of N-(2,3,5,6-tetrafluoro-4-iodophenyl) nicotinamide, 3N-ITFB
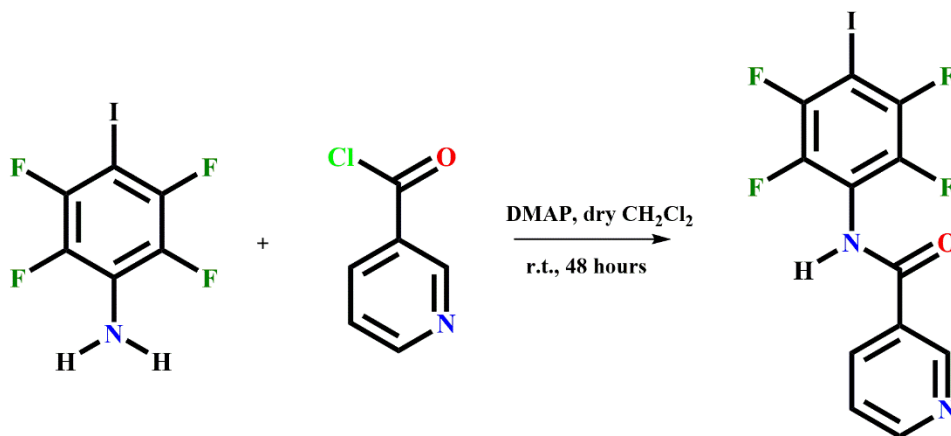
To an oven dried 50ml flask was added 3-nicotinic acid (2.46g, 20 mmol) under nitrogen atmosphere. An excess of thionyl chloride (2ml) was added via syringe to the reaction flask. The reaction mixture was refluxed at 90°C for 4-6 hours and excess thionyl chloride was removed under reduced pressure. The 3-nicotinoyl chloride (0.848g, 6 mmol) was used for next step without further purification. To another oven dried 100ml flask was added 4-iodotetrafluoroaniline (1.74g, 6 mmol)), 4-Dimethylaminopyridine (0.366g, 3 mmol) and dry methylene chloride (~20ml) under nitrogen atmosphere. The reaction mixture was stirred at room temperature for 48 hours. The white precipitate

was filtered to yield a white crystalline solid. yield: 2.12g (90%); m.p. 285-286 ℃, ¹H NMR (400 MHz, DMSO) δ ppm: 7.61 (m, 1H), 8.35(d, 1H), 8.82(d,1H), 9.15(s, 1H), 10.84(m, 1H).

7.2.2.5 Synthesis of N-(2,3,5,6-tetrafluoro-4-iodophenyl) isonicotinamide, 4N-ITFB
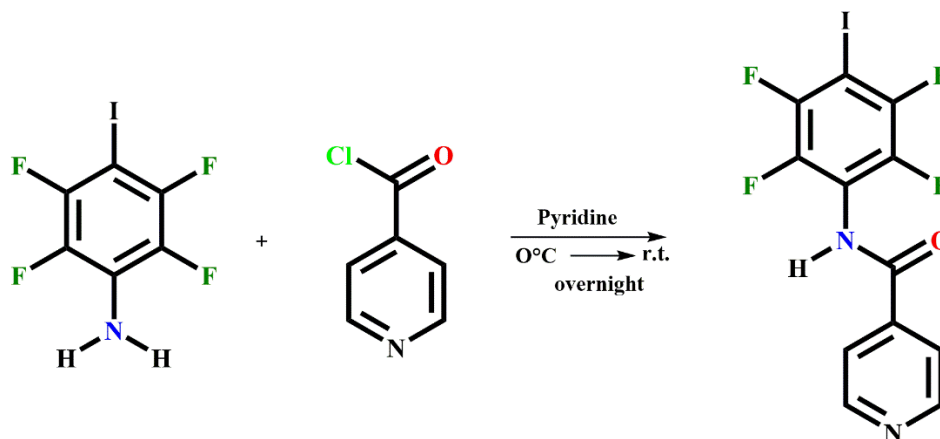
To an oven dried 50ml flask was added 4-nicotinic acid (2.46g, 20 mmol) under nitrogen atmosphere. An excess of thionyl chloride (20ml) was added via syringe to the reaction flask. The reaction mixture was refluxed at 90° C for 20 hours and excess thionyl chloride was removed under reduced pressure. The acid chloride was used without further purification. 4-nicotinoyl chloride (1.1375g, 8.06 mmol) was added into round bottom flask. Pyridine (10ml) was added to the mixture. After that, 4-iodotetrafluoroaniline (1.76g, 6.04 mmol) was added to the above mixture in an ice-water bath. The reaction mixture was stirred in an ice-water bath for 30 minutes and then stirred at room temperature under nitrogen for overnight. The crude product was washed with water and dried to get the final product as white crystalline solid. Yield: 2.01g (84%); ¹H NMR (400 MHz, DMSO) δ ppm: 7.80 (d, 2H), 8.77(d,2H), 13.65(m, 1H).

7.2.2.6 Synthesis of 4-((trimethylsilyl)ethynyl) aniline[20]
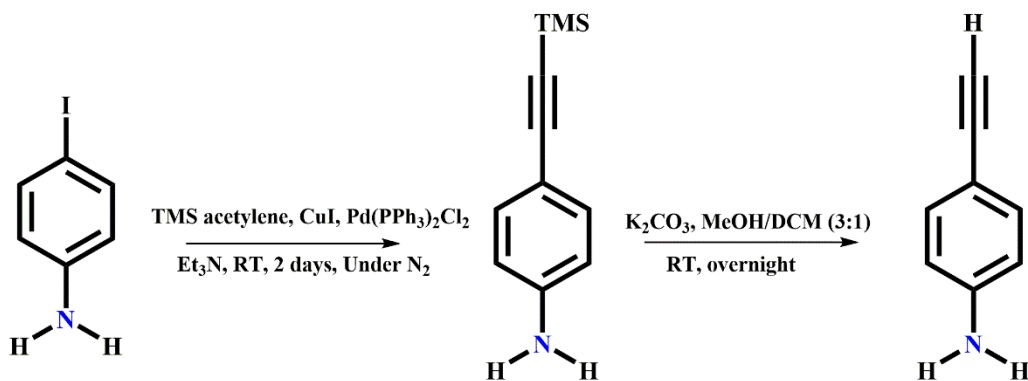
In a 250ml round bottom flask fitted with a stir bar, 20 ml of triethylamine was degassed and placed under nitrogen atmosphere. To this, 4-iodoaniline (2.4gm, 11mmol), copper(I) iodide (0.067gm, 0.109mmol), and dichlorobis(triphenylphosphine) palladium (II) (0.038gm, 0.055mmol) were added. The reaction mixture was degassed again for 20 minutes and kept under nitrogen atmosphere. To this, trimethylsilyl acetylene (1.67ml, 11.8mmol) was added and the reaction mixture was stirred at room temperature for 2 days. After this, the solvent was removed from the reaction mixture under vacuum. The slurry was dissolved in 50 ml chloroform and washed with 2M ammonium chloride solution and further with 50ml 1M sodium chloride solution. The solvent was removed from the organic layer to obtain the product. Yield: 1.89g (98%); ^{1}H NMR (400 MHz, DMSO) δ ppm: 0.24 (S, 9 H), 3.80 (br, 2 H), 6.57 (d, 2 H), 7.24 (d, 2 H).

The TMS protected alkyne (1g, 5.28mmol) was dissolved in methanol (15ml) and dichloromethane (5ml). Then, K_2CO_3 (2.07g, 15mmol) was added into the solution. The mixture was stirred at room temperature before being poured into water. The solution was extracted with ethyl acetate and washed with brine. After drying over magnesium sulfate ($MgSO_4$), the solvent was evaporated to afford the product in dark brown color which forms needles. Yield: 0.66g (94%); ^{1}H NMR (400 MHz, DMSO) δ ppm: 3.76 (s, 1 H), 5.48 (br, 2 H), 6.48 (d, 2 H), 7.09 (d, 2 H).

7.2.2.7 Synthesis of N-(4-ethynylphenyl) nicotinamide, 3N-EB
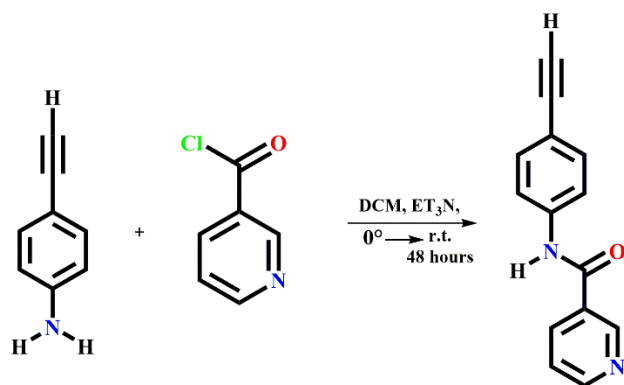
In a 2-neck round bottom 50 ml flask, 0.5 g (4.2 mmol) of 4-ethynylaniline and 2.07ml of triethylamine was dissolved in 50 ml of dichloromethane. The solution was kept in an ice-water bath at $0°$ C under N_2 atmosphere. Then 0.8g (5.6 mmol) of freshly prepared isonicotinoyl chloride was added into the reaction mixture. The mixture was allowed to stir in an ice-bath for 2 hours and then at room temperature. The reaction was monitored with TLC every 6-12 hours. After 24 hours, the 3-nicotinoyl chloride was added and the reaction was allowed to stir at room temperature. After 48 hours, the solution was rota-evaporated to get the crude product which was dissolved in ethanol. It was washed with saturated $NaHCO_3$ solution and water twice to get rid of any unreacted acid chloride and the solution was rota-evaporated to get the pure product. Yield: 0.82g (87%); ^{1}H NMR (400 MHz, DMSO) δ ppm: 4.12 (s, 1 H), 7.50 (d, 2 H), 7.59 (m, 1H), 7.82 (d, 2H), 8.28(d, 1H), 8.78(d, 1H),9.10(s, 1H), 10.58 (br, 1H).

7.2.2.8 Synthesis of N-(4-ethynylphenyl) isonicotinamide, 4N-EB
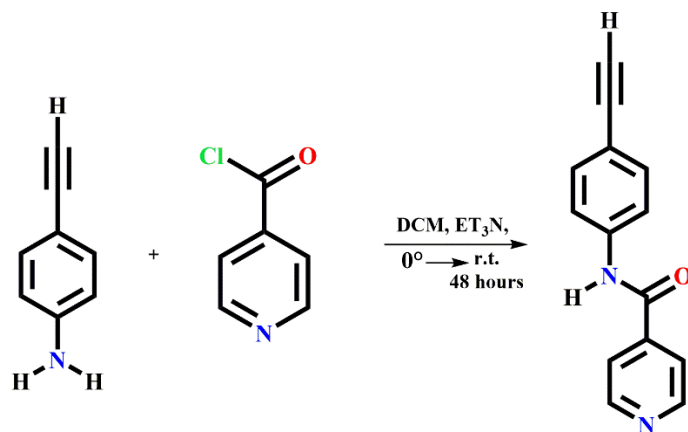
In a 2-neck round bottom 50 flask, 0.5 g (4.2 mmol) of 4-ethynylaniline and 2.07ml of triethylamine were dissolved in 50 ml of dichloromethane. The solution was kept in an ice-water bath at $0°$ C under

N^2 atmosphere. Then 0.8g (5.6 mmol) of freshly prepared isonicotinoyl chloride was added into reaction mixture. The mixture was allowed to stir in an ice-bath for 2 hours and then stir at room temperature. The reaction was monitored by TLC every 6-12 hours. After 24 hours, the 4-nicotinoyl chloride was added and the reaction was allowed to stir at room temperature. After 31 hours, the solution was rota-evaporated to get the crude product which was dissolved in ethanol. It was washed with saturated NaHCO$_3$ solution and water twice to get rid of any unreacted acid chloride and the solution was rota-evaporated to get the pure product. Yield: 0.70g (75%); ^{1}H NMR (400 MHz, DMSO) δ ppm: 4.13 (s, 1 H), 7.48 (d, 2 H), 7.80 (d, 2H), 7.85(d, 2H), 8.80(d, 2H), 10.65 (br, 1H).

7.2.2.9 Synthesis of N-(4-(iodoethynyl) phenyl) nicotinamide, 3N-IEB
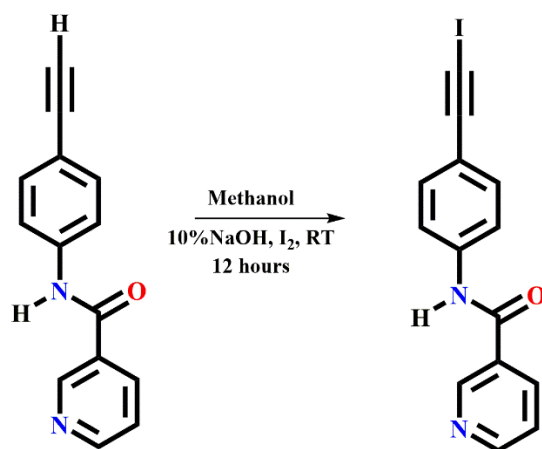
To a solution of N-(4-ethynylphenyl) isonicotinamide (0.25 g, 1.08 mmol) dissolved in THF (50 mL), were added dropwise simultaneously a concentrated solution of iodine in methanol (0.30 g, 1.19 mmol) and a 10% sodium hydroxide solution over 30 min, vigorously stirring. The mixture was stirred overnight and quenched with 100 mL water upon which a light-yellow color precipitate formed. The filtered solid was washed with sodium bisulfite solution and afforded pure pale-yellow color powder of **3N-IEB**. Yield: 0.36g (92%); ^{1}H NMR (400 MHz, DMSO) δ ppm: 7.43 (d, 2 H), 7.59 (m, 1H), 7.80 (d, 2H), 8.29(d, 1H), 8.78(d, 1H),9.09(s, 1H), 10.57 (br, 1H).

7.2.2.10 Synthesis of N-(4-(iodoethynyl) phenyl) isonicotinamide, 4N-IEB
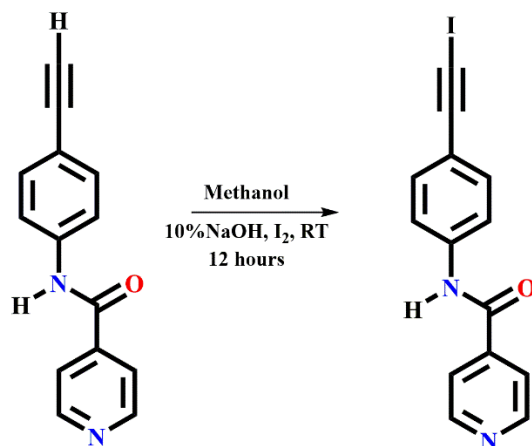
To a solution of N-(4-ethynylphenyl) isonicotinamide (0.2 g, 0.9 mmol) dissolved in THF (50 mL), added dropwise simultaneously a concentrated solution of iodine in methanol (0.30 g, 1.20 mmol) and a 10% sodium hydroxide solution over 30 min, vigorously stirring. The mixture was stirred overnight and quenched with 100 mL water upon which a light-yellow color precipitate forms. The filtered solid washed with sodium bisulfite solution afforded pure pale-yellow color powder of **4N-IEB**. Yield: 0.27g (86%); 1H NMR (400 MHz, DMSO) δ ppm: 7.44 (d, 2 H), 7.78 (d, 2H), 7.85(d, 2H), 8.80(d, 2H), 10.63 (br, 1H).

7.2.3 Growing crystals of target molecules
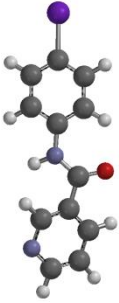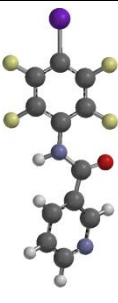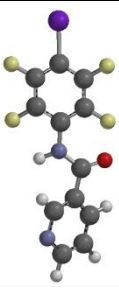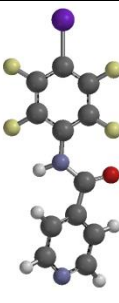
Crystals suitable for single crystal X-ray analysis for **3N-IB** and **4N-IB** were grown by slow evaporation in methanol. Crystals of **3N-ITFB** and **4N-ITFB** were obtained by slow evaporation from ethanol and nitromethane. Crystals of **3N-EB** and **4N-EB** were obtained from a mixture of methanol/water (1:1). Attempts to grow crystals of **3N-IEB** and **4N-IEB** were done in various solvents and binary solvents (methanol, ethanol, DCM, nitromethane, ethyl acetate, acetone, chloroform, methanol/water (1:1), methanol/DCM(1:1) and ethanol/DCM (1:1), but were unsuccessful.
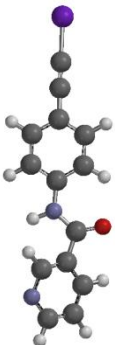
7.2.4 Molecular electrostatic potential calculations

As mentioned in chapter 3 and 5, only *trans* amide conformations are considered for this study based on a CSD analysis which shows that more than 90% of structures with an amide functionality exist in a *trans* conformation rather than *cis* conformation. Therefore, the calculations are restricted to finding the most stable conformation of the *trans* form (using MMFF) and that calculating the molecular

electrostatic potentials on the most stable conformation geometrically optimized using density functional B3LYP level of theory with 6-311++G** basis set in vacuum as described in section 2.2.4.

Table 7.1 Energies of *trans* amide conformation relative to the most stable *trans* conformation is for meta and para substituted molecules are shown below in kJ/mol.

	Conformation 1	Conformation 2		Conformation 1
3N-IB			**4N-IB**	
kJ/mol	0	+3		0
3N-ITFB			**4N-ITFB**	
kJ/mol	0	+4		0
3N-EB			**4N-EB**	
kJ/mol	0	+3		0

3N-IEB			4N-IEB	
kJ/mol	0	+3		0

7.2.5 Co-crystal screening and crystallography

Initial screening was carried out through solvent-assisted grinding using methanol. Target molecules were combined with each hydrogen-bond donor in stoichiometric ratios (2:1 for aliphatic and 1:1 for aromatic acids) leading to a total of 160 grinding experiments. In each experiment, 10 mg of the target molecule was used. Once the solvent had evaporated the ground mixtures were analyzed using IR spectroscopy to determine whether a co-crystal had formed or not. Successful interactions between the target molecules and donor were identified using the specific shifts of the peaks of the mixture compared to the starting compounds as well as well as the appearance of broad stretches around 2,300 and 1,800 cm^{-1} as a result of intermolecular O-H…N hydrogen bonds were taken as an indication of co-crystal formation (such hydrogen bonds would not be possible in either of the pure compounds, see appendix). For each successful reaction, the resulting solid was dissolved in a minimum amount (2 mL) of methanol/dichloromethane (1:1) or methanol/acetone (1:1) or methanol/ethyl acetate (1:1) mixture and then left in a vial for slow evaporation in order to obtain crystals suitable for single crystal X-ray diffraction. Most solvents resulted in fine needle like crystals which were unsuitable for single crystal diffraction analysis. Good quality crystals were obtained for only three combinations. Table 7.1 summarizes the experimental details. X-ray crystallographic data and all halogen-bond geometries are provided in the Appendix B.5.

Table 7.2 Experimental details for the ten co-crystals obtained

Compound ID	Solvent	Crystal stoichiometry	Color and morphology
3N-IB	Methanol	N/A	Colorless, blocks
3N-ITFB	Methanol	N/A	colorless, blocks
4N-ITFB	Ethanol	N/A	Colorless, blocks
3N-EB	Methanol/water (1:1)	N/A	Yellow, needles
4N-EB	Methanol/water (1:1)	N/A	Yellow, blocks
3N-IEB	Ethyl acetate	N/A	orange, plates
3N-IB: Seb	Methanol	2:1	Colorless, plate
3N-EB: Fum	Methanol/water (1:1)	2:1	Colorless, blocks
3N-IEB: Adi	Methanol	2:1	Yellow, plates

7.3 Results

7.3.1 Molecular electrostatic potentials

The maximum value of the potential, corresponding to the depth of the σ-hole on each halogen atom, was determined for the eight donors (Figure 7.1).
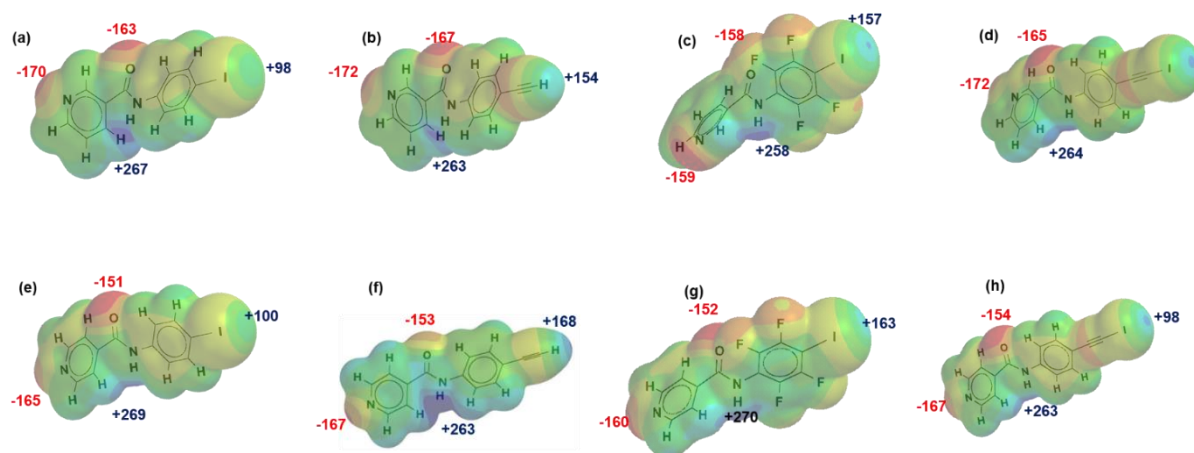
Figure 7.1 Molecular electrostatic potentials of unactivated and activated target molecules

7.3.2 Experimental structures of halogen bond donors

Once the compounds were successfully synthesized the next step was to analyze them with single crystal X-ray diffraction to determine the nature of their intermolecular interactions. Of the eight

target molecules that were synthesized, suitable crystals were obtained for five (**3N-IB**, **3N-ITFB**, **4N-ITFB**, **3N-EB** and **4N-EB**) molecules. Single crystal X-ray experimental data and the halogen bond geometries for the solved structures are provided in the Appendix D. For comparison purposes, the crystal structure of **3N-BB** is also presented from literature. [13]

The crystal structure of **3N-BB** is shown in Figure 7.2. The NH of the amide group binds to the pyridine nitrogen disrupting the amide-amide chain.
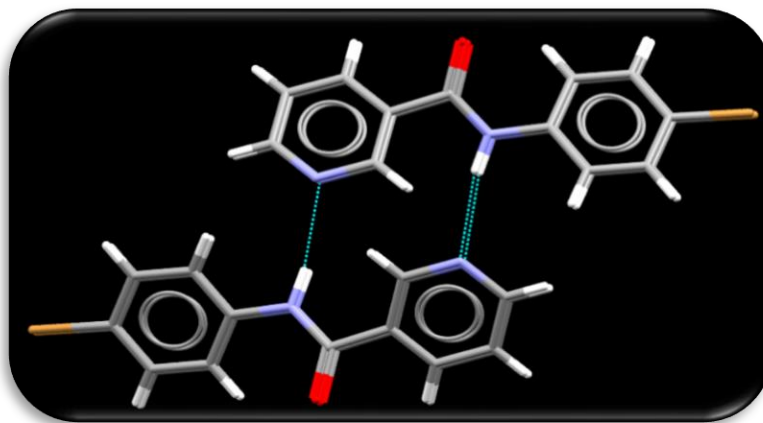
Figure 7.2 Main intermolecular interactions in the crystal structure 3N-BB (refcode:COFVUQ)[21]

The crystal structure of **3N-IB is** shown in Figure 7.3. The NH amide participates in the hydrogen bonding with the carbonyl oxygen of the neighboring molecule. The iodine participates in short contact interaction with pyridine N via halogen bonding.
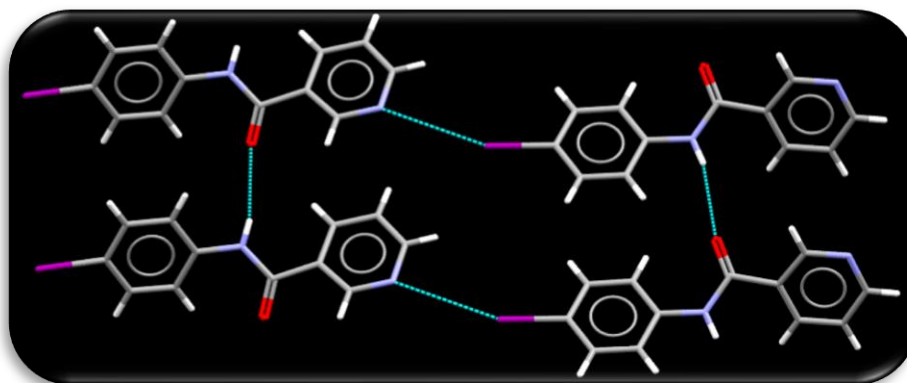
Figure 7.3 Main intermolecular interactions in the crystal structure 3N-IB

The crystal structures of **3N-ITFB** and **4N-ITFB** are shown in Figure 7.4 and Figure 7.5 respectively. The NH amide participates in the hydrogen bonding with the carbonyl oxygen of the

neighboring molecule. The activated iodine (via fluorinated aromatic ring) participates in halogen bonding with pyridine N.
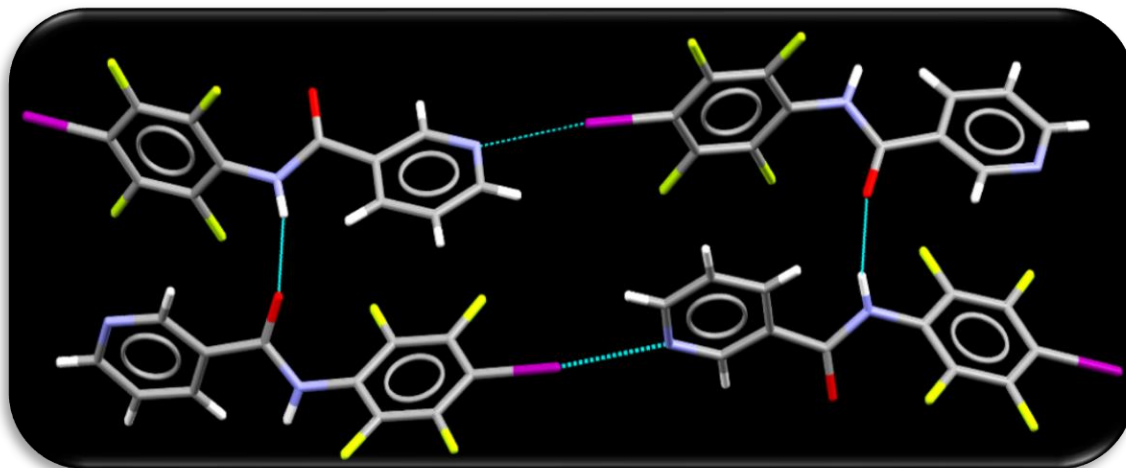
Figure 7.4 Main intermolecular interactions in the crystal structure 3N-ITFB
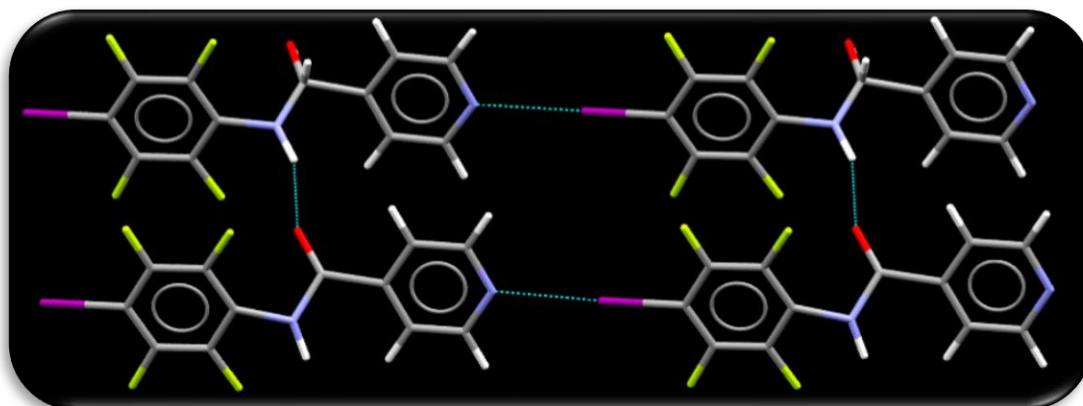
Figure 7.5 Main intermolecular interactions in the crystal structure 4N-ITFB

The crystal structure of **3N-EB and 4N-EB is** shown in Figure 7.6 and Figure 7.7 respectively. The NH amide participates in the hydrogen bonding with the carbonyl oxygen of the neighboring molecule. The activated ethynyl group participates in halogen bonding with pyridine N.
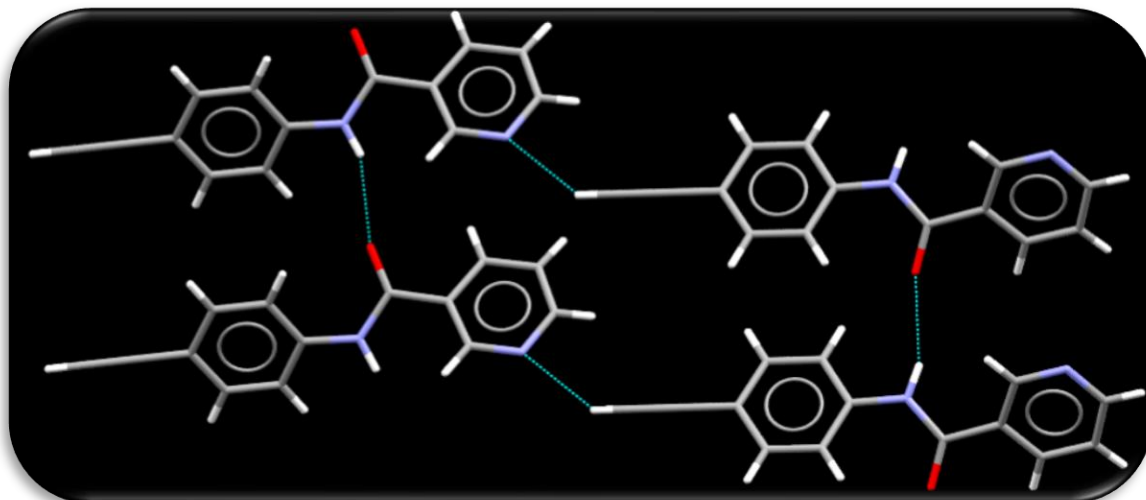
Figure 7.6 Main intermolecular interactions in the crystal structure 3N-EB
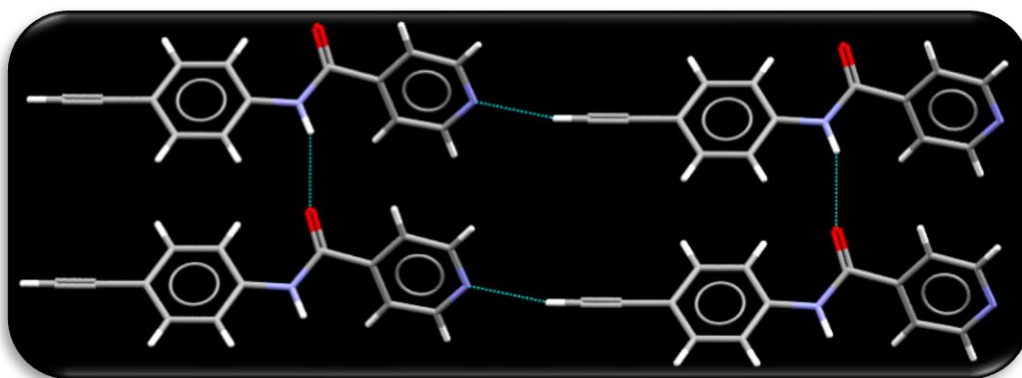
Figure 7.7 Main intermolecular interactions in the crystal structure 4N-EB

In the crystal structure of **3N-IEB**, amide…amide interaction is intact and activated iodine binds to pyridine.
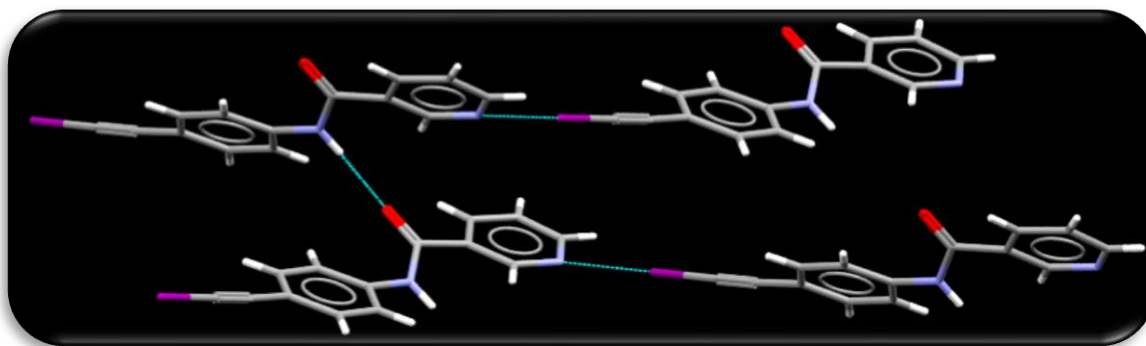
Figure 7.8 Main intermolecular interactions in the crystal structure 3N-IEB

7.3.3 Grinding experiments and characterization by IR spectroscopy

Solvent-assisted grinding experiments between each of the eight target molecules and 20 donors were analyzed using IR spectroscopy to identify notable interactions between the two potential co-formers, Figure 7.9-7.10 and Table 7.3 summarize the outcomes. Detailed IR analysis for grinding based experiments is provided in Appendix C.
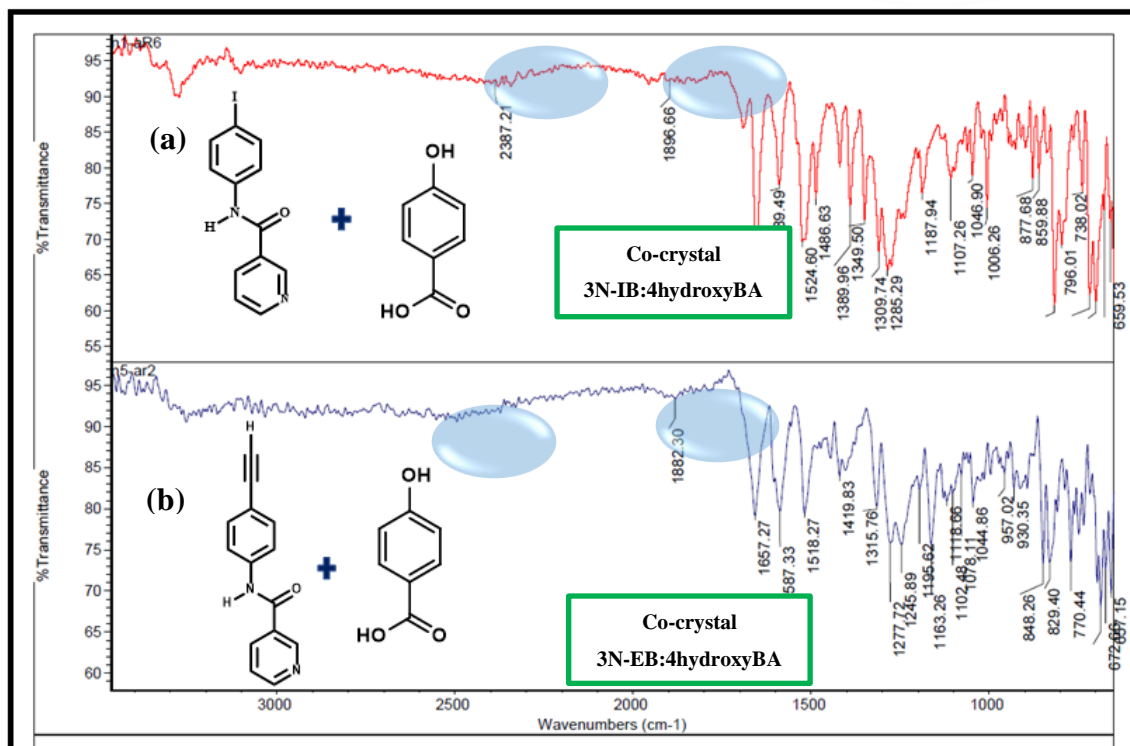
Figure 7.9 An example of hydrogen-bond co-crystal formation based on appearance of broad hydrogen-bond stretches for (a) 3N-IB:4hydroxyBA and (b) 3N-EB:4hydroxyBA.
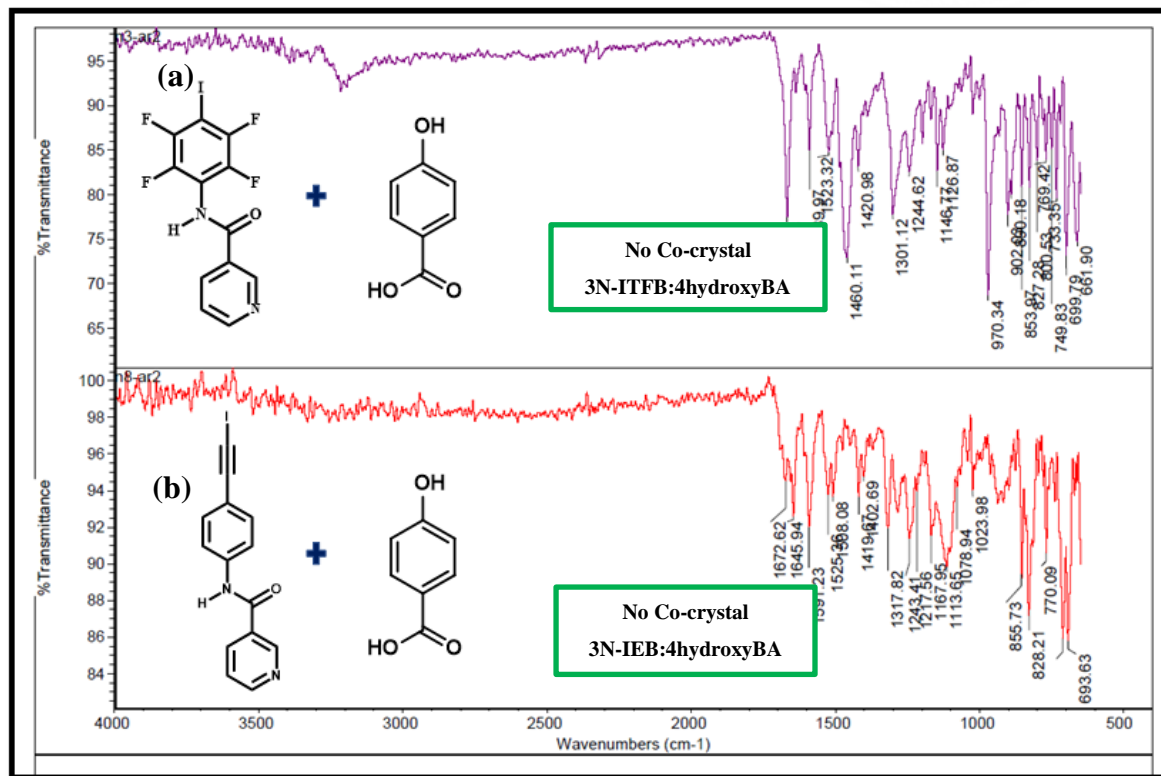
Figure 7.10 An example of no hydrogen-bond co-crystal formation for activated halogen bond donors based on no appearance of broad hydrogen-bond stretches; (a) 3N-ITFB:4hydroxyBA and (b) 3N-IEB:4hydroxyBA.

Table 7.3 Summary of grinding results. The green box indicates the positive co-crystal outcome and red box indicates negative co-crystal outcome.

		Target molecules							
		3N-IB	4N-IB	3N-EB	4N-EB	3N-ITFB	3N-ITFB	3N-IEB	4N-IEB
Hydrogen bond donors	Suc	green	green	green	green	red	red	red	red
	Adi	green	red	green	green	red	red	green	red
	Sub	green	red	green	red	red	red	red	red
	Seb	green	red	red	green	red	red	red	red
	Fum	green	green	green	green	red	red	red	red
	Mal	green	green	green	green	red	red	red	green
	Glu	green	red	green	green	red	red	red	red
	Pim	green	red	green	green	red	red	red	red
	Aze	green	red	green	green	red	red	red	red
	Dod	green	red	red	green	red	red	red	red
	3OHBA	green	green	green	green	red	red	red	red
	4OHBA	green	green	red	green	red	red	red	red
	3AminoBA	green	green	red	green	red	red	red	red
	4AminoBA	green	green	green	green	red	red	red	red

	3NitroBA								
	4NitroBA								
	BA								
	4BromoBA								
	4IodoBA								
	PentaFBA								
% Success rate		17/20	10/20	16/20	16/20	0/20	0/20	3/20	3/20
		85%	50%	80%	80%	0%	0%	15%	15%

7.3.4 Experimental structures of co-crystals

The structure determination of **3N-IB: seb** shows the expected trimer constructed from two symmetry related OH···N(py) hydrogen bonds, resulting in a trimer with a **2:1** stoichiometry of target molecule to the sebacic acid. The amide…amide interaction remains intact.
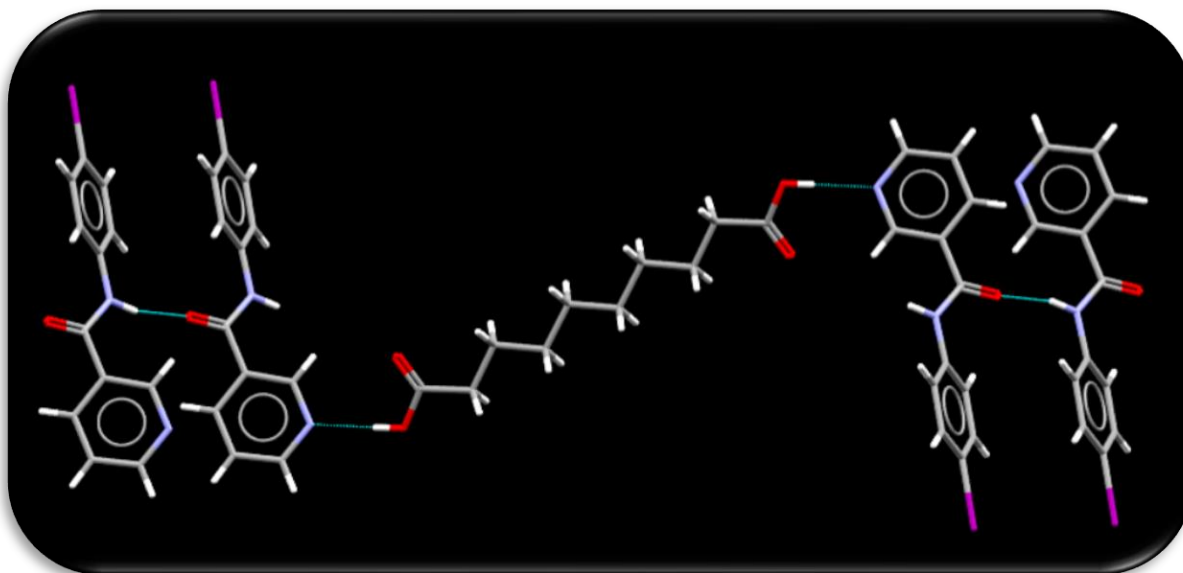
Figure 7.11 Main intermolecular interactions in the crystal structure 3N-IB: Seb

The structure determination of **3N-EB: fum** shows the formation of a hydrogen bond between pyridine N on the target molecules and the OH on the acid. However, the amide…amide interaction was disrupted by a water molecule. The OH group of water binds to a carbonyl oxygen of the target molecule. The carbonyl group of the acid binds to the amide NH.
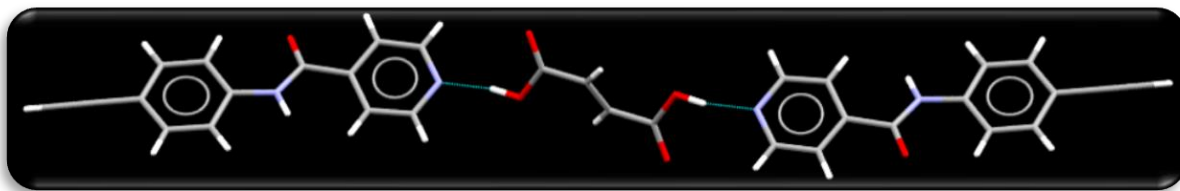
Figure 7.12 Main intermolecular interactions in the crystal structure 3N-EB: Fum
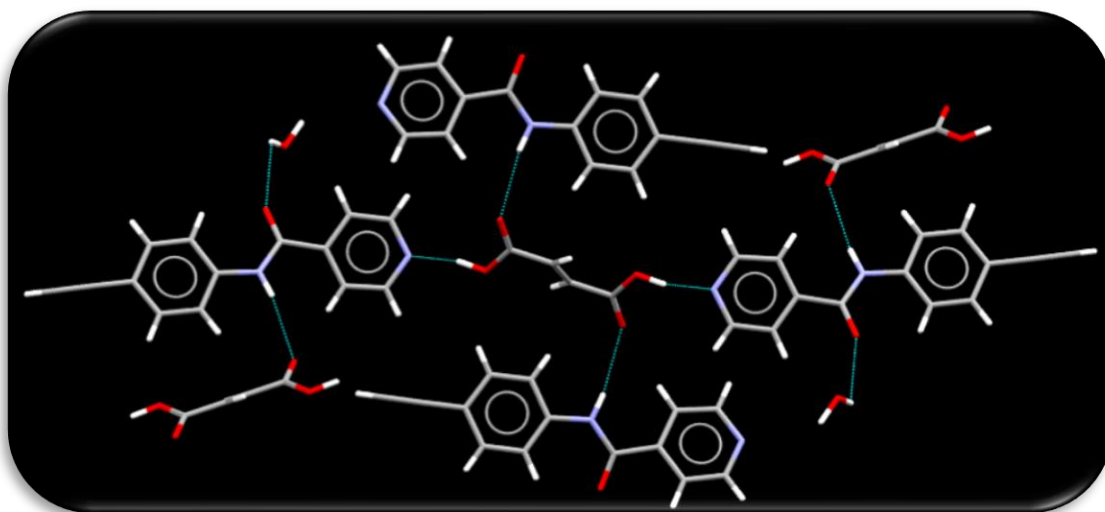
Figure 7.13 Main intermolecular interactions in the crystal structure 3N-EB: Fum

The crystal structure of **3N-IEB: Adi** involves hydrogen bonding between OH of the acid and the pyridine N as well as halogen bonding between the activated iodine and carbonyl oxygen of acid. The amide interaction however, remains intact.
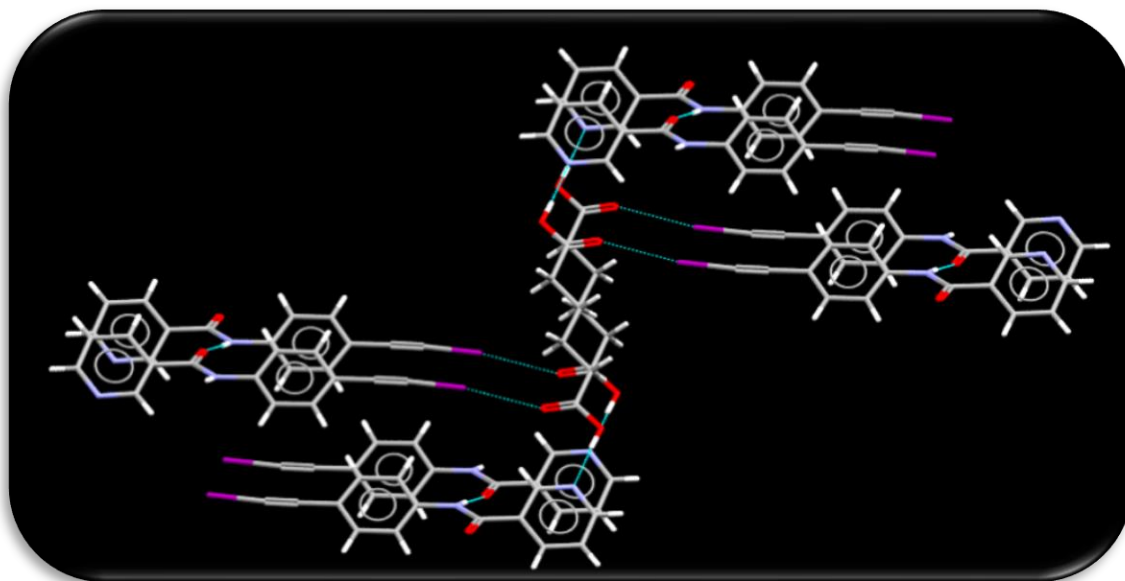
Figure 7.14 Main intermolecular interactions in the crystal structure 3N-IEB: Adi

7.4 Discussion

7.4.1 Molecular electrostatic potentials of target molecules

The premise for this study is the assumption that most conventional halogen bonds are dominated by electrostatics, a premise that means that a more pronounced σ-hole would produce a more effective XB donor. The unactivated halogen bond donor had the value for $V(r)$ of 98 kJ/mol. As we substituted the benzene ring with tetraflourinated ring, the halogen bond donor had the value for $V(r)$ of 157 kJ/mol.
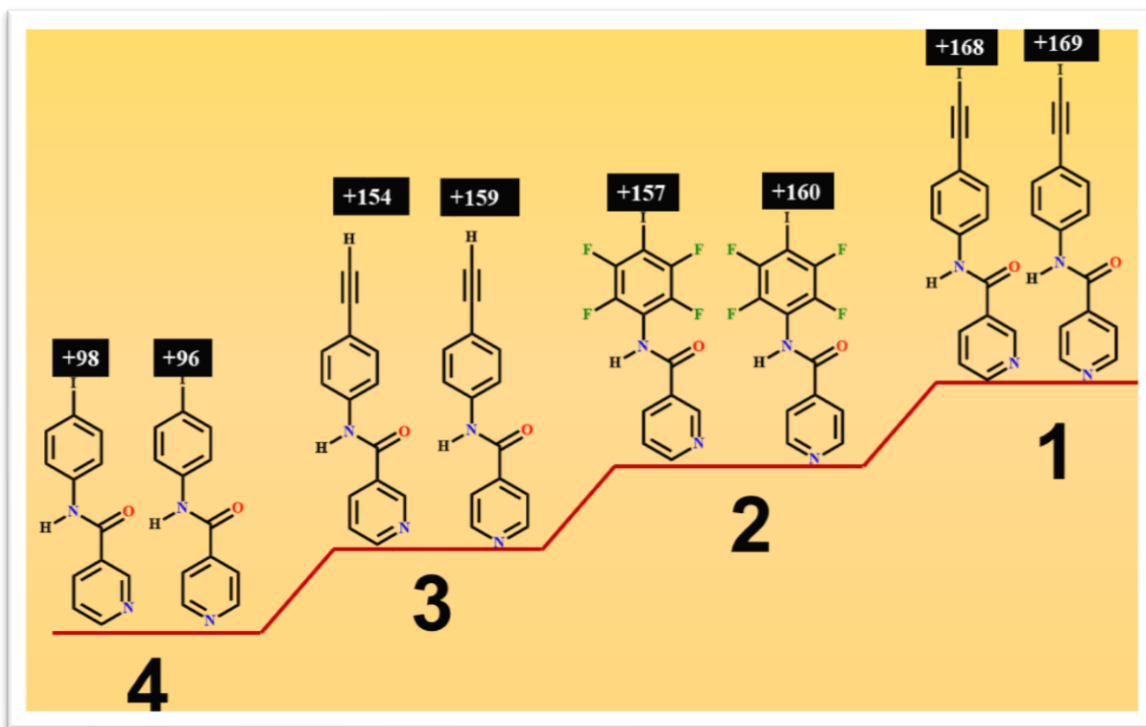
Figure 7.15 From left to right: The eight target molecules are presented in increasing order of molecular electrostatics potential associated with the most positive σ hole among the halogen atoms in the molecules. Values are presented in kJ/mol.

The second possibility is to add a sp-hybridized carbon next to halogen bond donor, the strong polarizing effect of the sp carbon atom allows the ethynyl-based hydrogen and iodine atom to display a σ-hole of higher magnitude of 154 kJ/mol and 168 kJ/mol respectively. The results show strong similarity between the σ hole on each donor and the resulting MEPs values. A less polarized iodine atom on **3N-IB** and **4N-IB** is more strongly affected by a *sp* triple bond iodine containing moiety than it is by the proximity of a *sp2*-perflourinated ring or *sp*-hybridized ethynyl group as shown by the increases in the electrostatic potential as we go from **3N/4N-IB**<**3N/4N-EB**<**3N/4N-ITFB**<**3N/4N-IEB**. Synthon B will be formed in all activated donors (**3N-ITFB, 4N-ITFB, 3N-EB, 4N-EB, 3N-IEB** and **4N-IEB**)

To summarize, the calculations have identified the iodoethynyl moiety in IEB as the top ranked halogen bond donor as shown in Figure 7.15. These results are in strong agreement with Aakeröy *et. al.* research work published in 2013[22] where supramolecular hierarchy was established by activating the iodine and bromine atoms.

7.4.2 Synthon prediction based on MEPs

NH amide is the best donor and pyridine is the best acceptor, therefore, based on Etter's rule NH amide will bind to N forming synthon A will be formed in the crystal structure of **3N-IB** and **4N-IB**. In activated target molecules (**3N-ITFB, 4N-ITFB, 3N-EB, 4N-EB, 3N-IEB** and **4N-IEB**), the iodine atom or activated ethynyl H are better donors compared to amide NH, therefore activated donor (I or H) will bind to pyridine and NH amide will bind to C=O forming synthon B.

7.4.3 Homomeric synthon analysis in target molecules

Synthon A was predicted in unactivated target molecules and synthon B was predicted in activated target molecules. Six crystal structures of target molecules (**3N-IB, 3N-ITFB, 4N-ITFB, 3N-EB, 4N-EB** and **3N-IEB**) and one literature structure **3N-BB** were analyzed and compared to predicted synthons based on electrostatics, Figure 7.16.
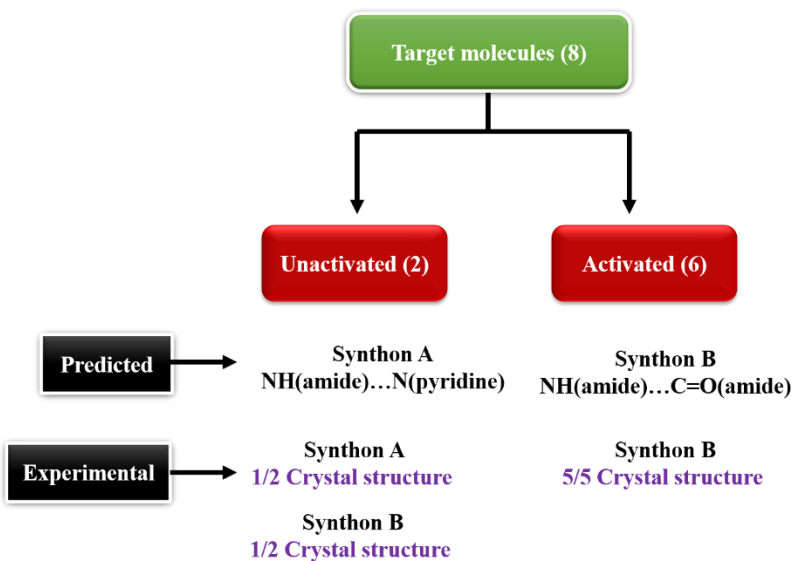
Figure 7.16 Predicted (based on MEPs) vs experimental synthons in the target molecules

In all six crystal structures obtained in this study, the amide functionality was intact irrespective of whether the halogen atom was activated or not. In **3N-IB**, the halogen bond was weak between I…N (more of a short contact than a halogen bond) and could be easily disrupted if iodine is replaced with bromine atom as was seen in the crystal structure of **3N-BB**, where based on Etter's rule; the best donor NH amide binds to best acceptor pyridine N. These results are also supported by electrostatic potential values on **3N-IB** and **3N-BB**, where these molecules have weak σ-hole

192

value and can only form short contacts with the stronger acceptor group. On the other hand, in all activated crystal structures, amide…amide interaction is observed and strong halogen bonding interaction between activated iodine and pyridine N is also observed in **3N-ITFB, 4N-ITFB, 3N-EB, 4N-EB and 3N-IEB** molecules. These results were complemented by electrostatic potentials.

7.4.4 Experimental co-crystal screening analysis

7.4.4.1 Unactivated vs activated halogen bond donors

Based on grinding experiment results, the unactivated halogen-bond donors (**3N-IB** and **4N-IB**) showed 67% success rate with the 20 hydrogen-bond donors whereas the activated halogen-bond donors (**3N-ITFB, 4N-ITFB, 3N-IEB,** and **4N-IEB**) showed only 7% success rate with hydrogen-bonded co-crystallization, Figure 7.17.
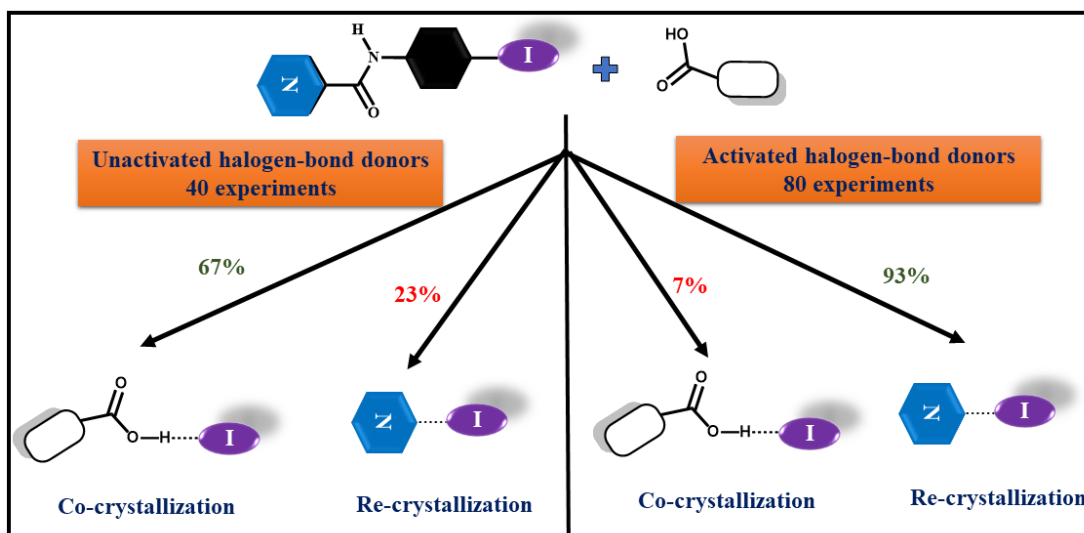
Figure 7.17 Co-crystallization vs re-crystallization summary of unactivated and activated class of halogen bond donors

These results indicated that the weak homomeric interactions between Iodine and nitrogen in unactivated molecules were easy to break via suitable donors whereas in unactivated molecules, all donors and acceptors were involved in strong homomeric interactions and therefore, were difficult to break via co-crystallization. These results agree with out hypothesis and are supported by molecular electrostatics potential values on the target molecules.

7.4.4.2 Ethynyl vs iodoethynyl based target molecules

Another comparison was done between ethynyl and iodoethynyl molecules, it was found that ethynyl molecules where only hydrogen bonding was involved in the molecules by themselves gave 80% positive co-crystal outcome whereas iodoethynyl molecules where strong halogen bonding is involved between activated iodine and nitrogen in homomeric system gave only 15% positive co-crystal outcome, Figure 7.18.
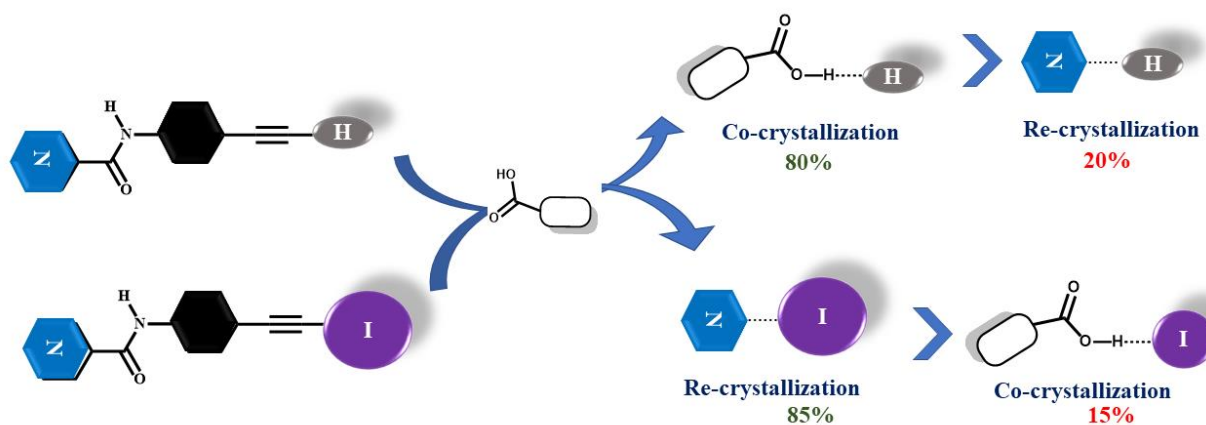
Figure 7.18 Co-crystallization vs re-crystallization summary of ethynyl vs iodoethynyl halogen bonds

Therefore, in this study, ethynyl based H and iodoethynyl molecules based I; two atoms of radically different size and chemical characteristics doesn't behave similarly in solid state towards aliphatic and aromatic hydrogen-bond donors. It could be due to presence of hydrogen-bonding interaction in the former and halogen bonding in the later at the pyridine site in the homomeric systems.
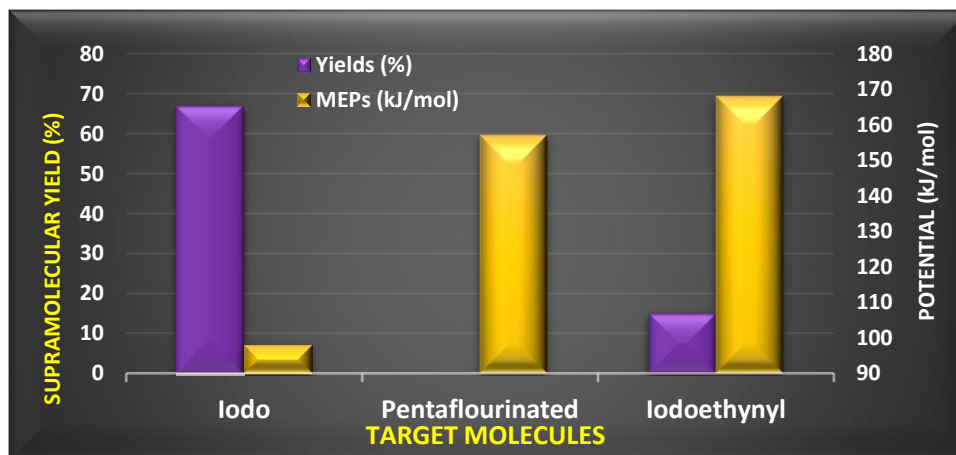
Figure 7.19 Comparison between supramolecular yield and electrostatic potentials of unactivated and activated molecules.

7.4.5 Experimental structures of co-crystals

Three crystal structures of co-crystals were obtained (**3N-IB: Seb, 3N-EB-Fum** and **3N-IEB: Adi**) and in 2 out 3 crystal structures (67%), the amide…amide interaction was intact. In 3N-EB: Fum, it was disrupted because of presence of water molecule in the lattice. The summary of three crystal structures is provided in Table 7.4.

Table 7.4 Interactions observed in the crystal structures of three co-crystals.

	Unactivated 3N-IB: Seb	Activated 3N-EB-Fum	Activated 3N-IEB: Adi
What type of cocrystal?	Neutral co-crystal	Hydrate co-crystal	Neutral cocrystal
Amide functionality Intact?	Yes	No (disrupted)	Yes
Which interaction is observed?	Only hydrogen bonding	Only hydrogen bonding	Hydrogen and halogen bonding
	OH(acid)…N(pyridine)	OH(acid)…N(pyridine) NH(amide)…O=C(acid)	OH(acid)…N(pyridine) I…O=C(acid)

7.5 Conclusions

Eight new compounds were synthesized in this study (six of them have halogen bond donor and 2 contains hydrogen bond donor). Two of the six are unactivated and four of the six are activated halogen bond donors. The goal of keeping the amide functionality intact by activating the halogen bond donor either by introducing fluorinated ring or by sp-hybridized carbon was achieved.

1. Both the hydrogen and halogen bonding was systematically controlled in the target molecules based on knowledge of stronger halogen bonding between I…. N and strong hydrogen bonding between neighboring amide functionality. In each crystal structure obtained for target molecules, the amide functionality was kept intact.

2. The second goal of keeping the amide functionality was also achieved via co-crystallization as three crystal structures were obtained, and each crystal structure has amide functionality intact.

3. The unactivated halogen bonds donors had 67% supramolecular yield whereas activated halogen bond donors had only 7% success with the 20 di and mono acids, Figure 7.20. The

results were readily rationalized against calculated molecular electrostatic surface potentials which serves to emphasize the importance of the electrostatic contributions to the strength of halogen bonds.

(a)   (b)
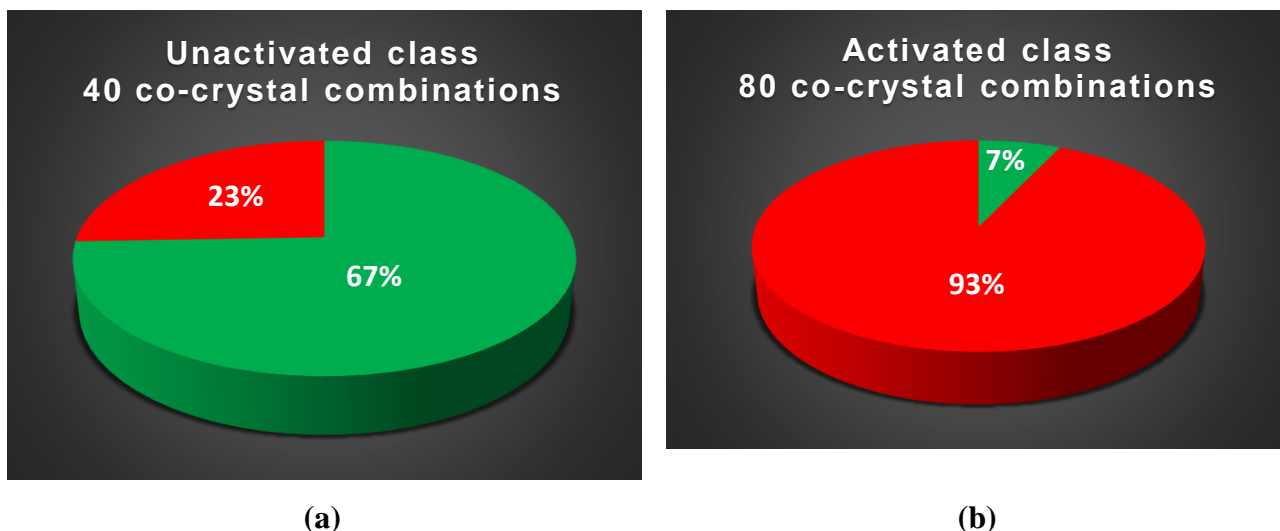
Figure 7.20 Success rates of co-crystallization experiments; green-positive results and red negative results for (a) activated and (b) unactivated molecules.

4. The ethynyl based 'hydrogen' and iodoethynyl based 'Iodine' do not display same supramolecular field in the solid state against 20 carboxylic acids studied as ethynyl based target molecules had 80% supramolecular yield and iodoethynyl based target molecules had 15% supramolecular yield, Figure 7.21.

(a)   (b)
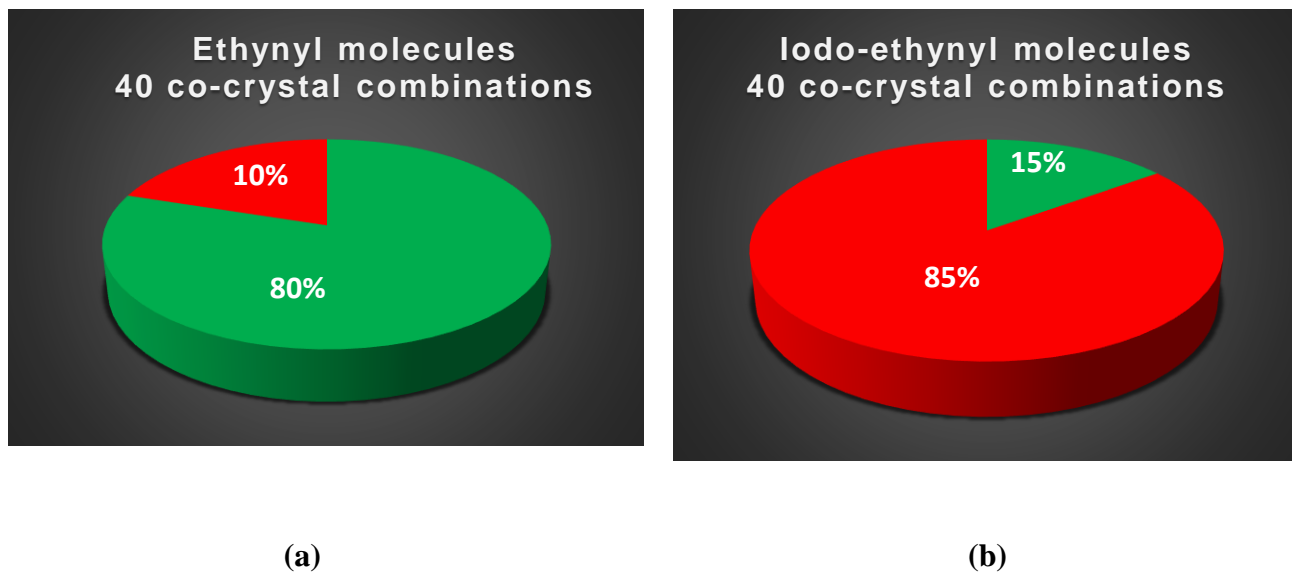
Figure 7.21 Success rates of co-crystallization experiments; green-positive results and red negative results for (a) ethynyl and (b) iodoethynyl molecules.

7.6 References

1. Greenberg, A.; Breneman, C. M.; Liebman, J. F., The Amide Linkage: Structural Significance in Chemistry, Biochemistry, and Materials Science. Wiley: 2002.

2. Sarma, J. A. R. P.; Desiraju, G. R., Crystal Growth & Design **2002,** 2 (2), 93-100.

3. R., D. G., Angewandte Chemie International Edition in English **1995,** 34 (21), 2311-2327.

4. Joel, B.; E., D. R.; Liat, S.; Ning-Leh, C., Angewandte Chemie International Edition in English **1995,** 34 (15), 1555-1573.

5. Mirzaei, M.; Eshghi, H.; Akhlaghi Bagherjeri, F.; Mirzaei, M.; Farhadipour, A., Journal of Molecular Structure **2018,** 1163, 316-326.

6. Pogoda, D.; Janczak, J.; Videnova-Adrabinska, V., Acta Crystallographica Section B **2016,** 72 (2), 263-273.

7. Mukherjee, A.; Desiraju, G. R., Chemical Communications **2011,** 47 (14), 4090-4092.

8. Mukherjee, A.; Tothadi, S.; Chakraborty, S.; Ganguly, S.; Desiraju, G. R., CrystEngComm **2013,** 15 (23), 4640-4654.

9. B. Aakeröy, C.; D. Chopade, P.; Desper, J., Avoiding "Synthon Crossover" in Crystal Engineering with Halogen Bonds and Hydrogen Bonds. 2011; Vol. 11.

10. Aakeroy, C. B.; Chopade, P. D.; Ganser, C.; Rajbanshi, A.; Desper, J., CrystEngComm **2012,** 14 (18), 5845-5853.

11. Aakeröy, C. B.; Chopade, P. D.; Desper, J., Crystal Growth & Design **2011,** 11 (12), 5333-5336.

12. Khavasi, H. R.; Esmaeili, M., CrystEngComm **2014,** 16 (36), 8479-8485.

13. Khavasi, H. R.; Tehrani, A. A., CrystEngComm **2013,** 15 (29), 5813-5820.

14. Moulton, B.; Zaworotko, M. J., Chemical Reviews **2001,** 101 (6), 1629-1658.

15. Rodríguez-Spong, B.; Price, C. P.; Jayasankar, A.; Matzger, A. J.; Rodríguez-Hornedo, N. r., Advanced Drug Delivery Reviews **2004,** 56 (3), 241-274.

16. Sandhu, B.; McLean, A.; Sinha, A. S.; Desper, J.; Sarjeant, A. A.; Vyas, S.; Reutzel-Edens, S. M.; Aakeröy, C. B., Crystal Growth & Design **2018,** 18 (1), 466-478.

17. Al-Sha'er, M. A., Pharma Chem. 2014 **2014,** 6, 261-291.

18. Mishra, A.; Jung, H.; Park, J. W.; Kim, H. K.; Kim, H.; Stang, P. J.; Chi, K.-W., Organometallics **2012,** 31 (9), 3519-3526.

19. Bushuyev, O. S.; Tomberg, A.; Friščić, T.; Barrett, C. J., Journal of the American Chemical Society **2013,** 135 (34), 12556-12559.

20. Chadwick, R. C.; Khan, U.; Coleman, J. N.; Adronov, A., Small **2013,** 9 (4), 552-560.

21. Taylor, R. G. D.; Yeo, B. R.; Hallett, A. J.; Kariuki, B. M.; Pope, S. J. A., CrystEngComm **2014,** 16 (21), 4641-4652.

22. Aakeröy, C. B.; Baldrighi, M.; Desper, J.; Metrangolo, P.; Resnati, G., Chemistry – A European Journal **2013,** 19 (48), 16240-16247.

Chapter 8 - Predicting the crystallization propensity of drug like molecules with diverse functionalities

8.1 Introduction

Predicting the crystallization propensity of drug-like molecules is one of the biggest challenges facing pharmaceutical scientists today.[1] Specifically, understanding which molecules in a series of similar compounds will be the most difficult to crystallize would be extremely useful, as would any indication of the experimental conditions (such as temperature, solvent polarity, molar concentration, etc.) that might make crystallization take place more readily. When it comes to predictability, it is somewhat easier to predict synthons in molecules with similar functionalities but as complexity of molecules grows, synthon prediction becomes more difficult as we have seen in chapters 2-5. Crystallizability of a molecule is dependent on many factors such as molecular weight, rotatable bond, # of conformers, enthalpy, entropy, glass transition temperature; experimental, environmental conditions and hydrogen bond donors and acceptors, Figure 8.1.
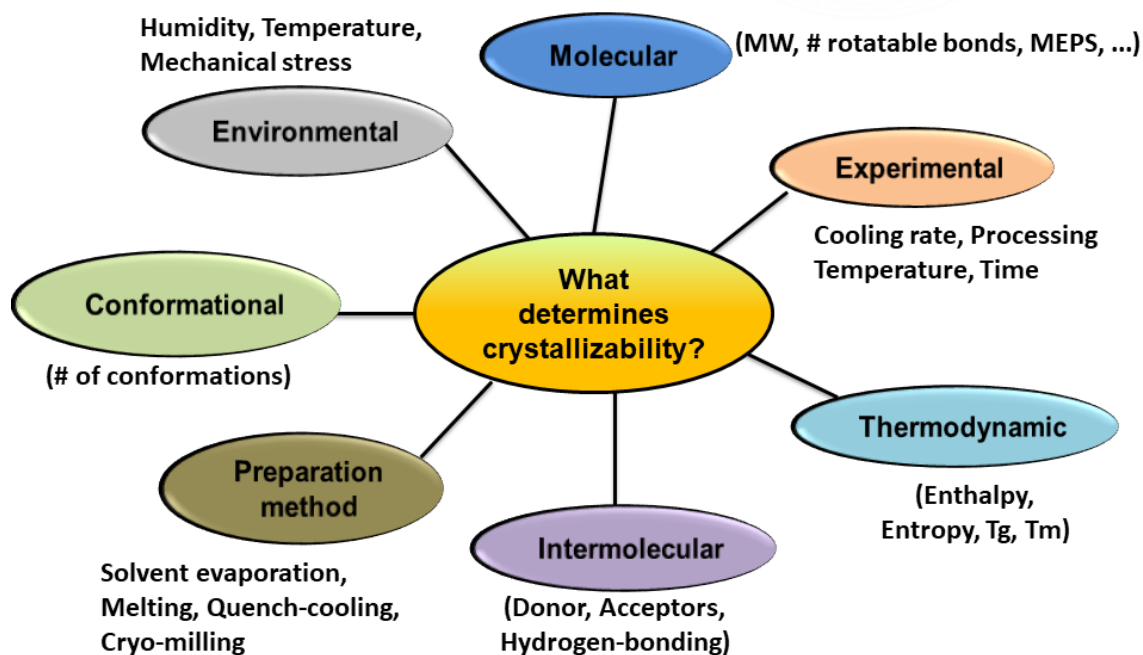
Scheme 8.1 Various experimental and chem-informatics variables affecting crystallizability.

8.1.1 Current methodologies

Despite the importance of being able to understand which structural features of a molecule (such as molecular size, polarity, etc.) and which experimental conditions (such as temperature, concentration, etc.) permit a molecule to readily crystallize, there has been very little work published in the literature focused on this topic. Different approaches have been taken in recent years to model and predict crystallization propensity. Machine learning approaches to understanding molecular crystallizability in terms of forming 'good crystals' have found key attributes the number of rotatable bonds and a 'molecular connectivity index' effectively describing molecular size.[2] Statistical (Random Forest) modeling has also been used to model and predict crystallizability in terms of giving single crystals suitable for X-ray diffraction.[3] In this study, descriptors of the relative energy, atom connectivity, conformation and flexibility (number of rotatable bonds) contributed to the 70% accuracy of the crystallizability prediction. Taylor, et al. have also investigated the crystallization tendency of organic molecules from both undercooled melts and supersaturated solutions, having shown that small, rigid molecules of low molecular weight were more likely to crystallize, although viscosity was also a factor.[4-5]

# of Rotatable Bonds	MW <300	MW >300
0-3	Easy	Moderate
4-6	Easy to moderate	Moderate to difficult
>6	Moderate to difficult	Difficult

(a) (b)
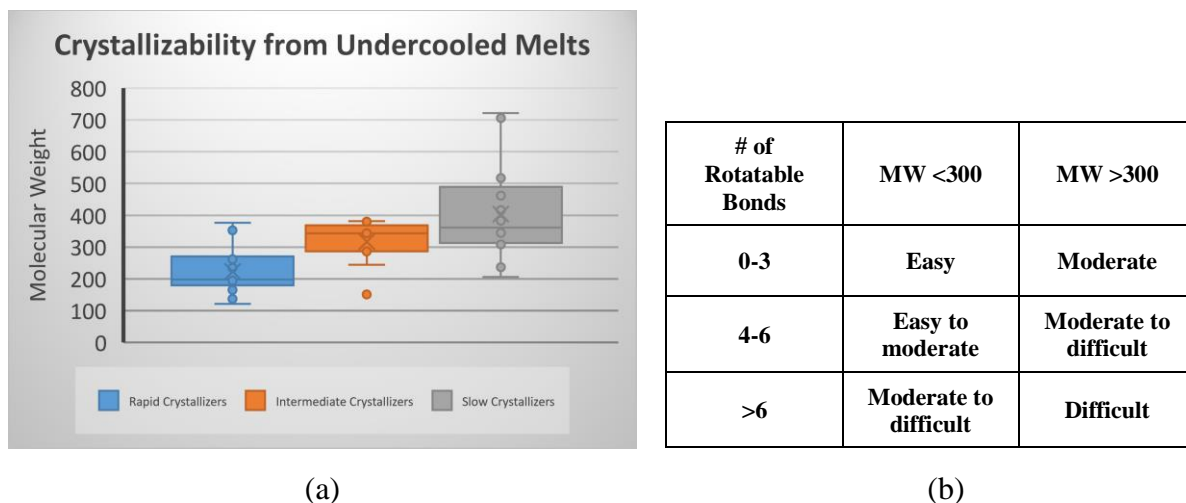
Figure 8.1 (a and b) Literature studies showing the correlation between molecular weight, rotatable bonds and crystallizing ability of molecules.

A recent study found that amorphous stability is strongly correlated with multiple factors, including molecular weight, thermodynamic parameters (T_m, H, G) and molecular descriptors (# of H bond donors, carbon-to-heteroatom ratio).[6] These varied approaches may not give a consistent picture of crystallizability, as the difficulty in nucleation does not necessarily correlate to slow

crystal growth rates or poor crystal quality.[7] More importantly, these approaches do not provide a path forward for overcoming inherently poor crystallizability, i.e., crystallizing a molecule for the first time.

In this study, cheminformatics tools such as hydrogen-bond propensity[8-9] will be used to predict the polymorphs of drug molecules and then use this approach along with molecular descriptors to find a systematic way to predict crystallizability. We want to be able to predict if a given molecule will be easy to crystallize or not, i.e. the molecule's "crystallizability", Scheme 8.2.

8.1.2 Outline

The two main goals of this chapter are listed below:

1. Apply solid-state informatics tools on drug molecules with diverse functionalities and compare/analyze the predicate propensies with the experimentally observed structures or polymorphs.

2. Use solid state and chem informatics to determine a logistic regression model that can be useful in predicting the crystallizability of drug molecules.

The second goal of chapter involve six different steps as shown in scheme 8.5.

Step 1: Get a published dataset (training dataset) where the crystallizability is already established for a set of diverse molecules

Step 2: Choose descriptors and extract information

Step 3: Find a correlation between crystallizability and a descriptor if any.

Step 4: Develop a logistic regression model to predict crystallizability

Step 6: Calculate probability of crystallization on training dataset and test dataset

Scheme 8.2 Outline of second goal of the chapter.

8.1.3 Defining crystallizability

There is no simple way of defining crystallizability, and it is even questionable that it is a measurable physical quantity. If we ignore the ambiguity in the concept of crystallizability as a physically meaningful concept and instead limit ourselves to attempting to predict a probability of crystallization, it may be possible to re-phrase the problem in a meaningful way and to find a mathematical model for it. A probability is a real number in the interval [0; 1], and we are looking for an equation for calculating it, given a molecule and any external experimental factors as input. A crystallization experiment either yields crystals or not, it is a dichotomous outcome. We describe the outcome as a binary random variable that takes the values 1 (crystals formed) or 0 (no crystals). Experiments with a dichotomous outcome that (presumably) depends on several factors can often be modelled with a binomial logistic regression model.
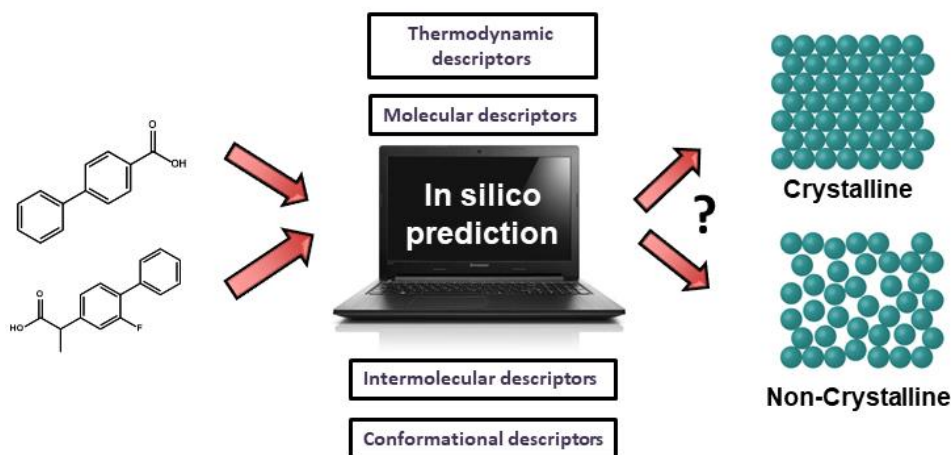
Scheme 8.2 Computational approach for prediction of crystallizability in drug-like molecules

In order to determine crystallizability, large flexible drug molecules will be used as the training dataset. The chemical structure of list of drug molecules used in each classification is shown in scheme 8.3 and 8.4. Each molecule was classified into Class I as crystallizable or Class III as non-crystallizable based on experimental classification established already in literature By Taylor et. al.[4] The crystallization classification was done using differential scanning calorimetry (DSC) where each sample was heated to above the melting temperature, held isothermally for 3 min, cooled at a rate of 20°C min^{-1} to -75°C, and reheated at 108°C min to just above the melting

temperature. The molecules were classified into three categories based on this classification, Figure 8.2, Table 8.1.

The classification was done as follows:

1. Class I (Rapid crystallizers): If a compound shows exotherm peak from the undercooled melt state prior to the Tg event, it was classified as class I molecule.

2. Class II (intermediate crystallizers): If the compound shows no crystallization upon cooling from the undercooled melt state to below Tg, however, crystallization was observed during reheating above Tg, it is classified as class II molecule.

3. Class III (Slow crystallizers): If no crystallization was observed upon either cooling to below Tg or upon subsequent reheating up to the melting point, it is classified as class III molecule.

| (a) Class I | (b) Class II | (c) Class III |
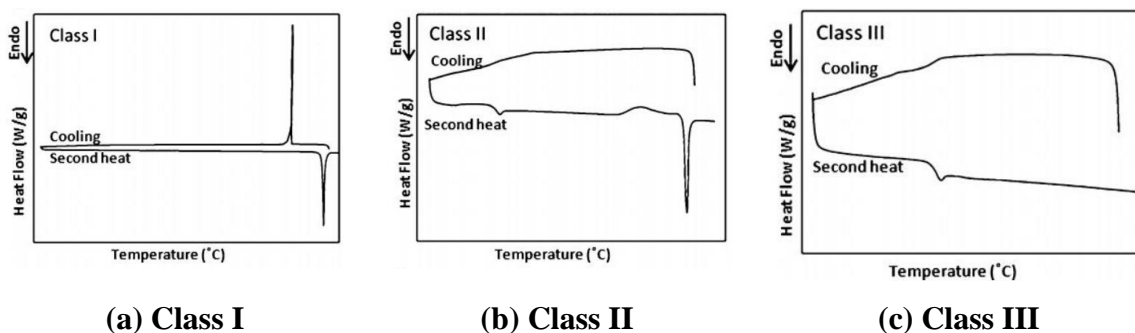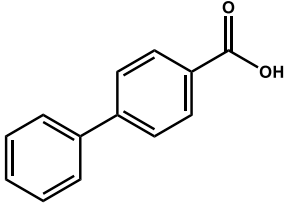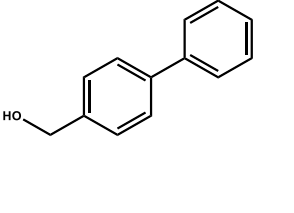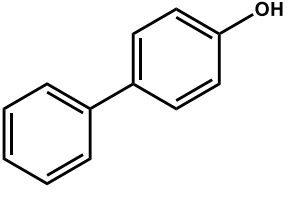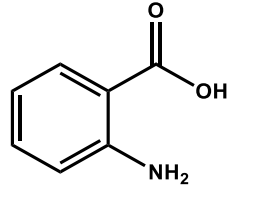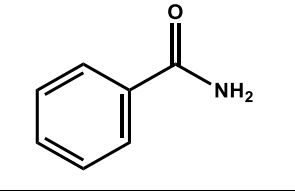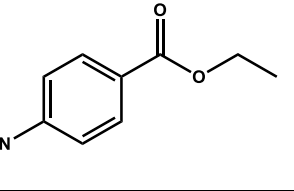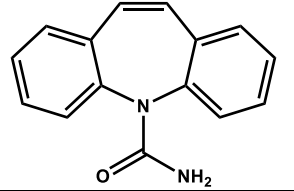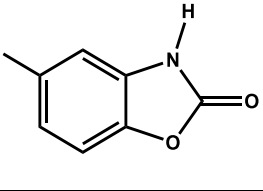
Figure 8.2 Classification of molecules based on DSC experiments as class I, class II and class III

Class I			
Biphenyl carboxylic acid	4-Biphenylmethanol	4-Phenyl-phenol	Anthranilic acid
Benzamide	Benzocaine	Carbamazepine	Chlorzoxazone

Felbinac	**Flufenamic acid**	**Haloperidol**	**Indoprofen**
Phenacetin	**Theophylline**	**Lidocaine**	**Tolfenamic acid**
Chlorpropamide	**Tolbutamide**		

Scheme 8.3 Chemical structures of class I molecules included in this study

Class III molecules			
Acetaminophen	**Celecoxib**	**Dibucaine**	**Nifedipine**
Droperidol	**Flurbiprofen**	**Tolazamide**	**Salicin**
Clofoctol	**Aceclofenac**	**Felodipine**	**Ibuprofen**
Nimusalide	**Pimozide**	**Indomethacin**	**Ketoprofen**

Procaine	Probucol	Ritonavir	
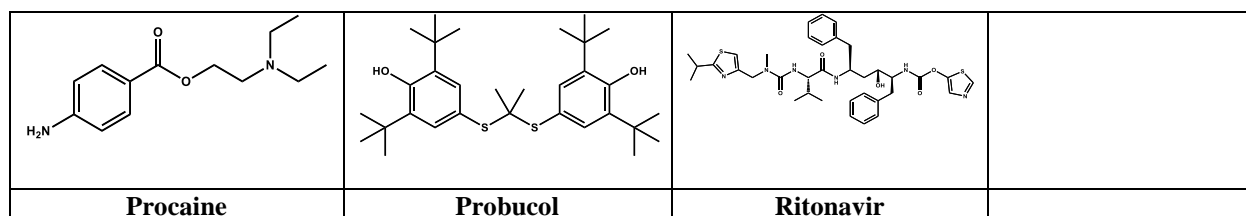
Scheme 8.4 Chemical structures of class III molecules included in this study

8.1.4 Logistic regression model

The logistic regression model[10] is a statistical model used to determine correlation between different variables and the response variable is categorical. The model involves three different methods to find the best correlation; backward elimination, forward selection and stepwise selection. The backward elimination method involves starting with all predictors in the model and removing the predictor with highest p-value greater than $\alpha_{critical}$, repeating the steps to get the model that has all descriptors p-value less than $\alpha_{critical}$. The $\alpha_{critical}$ here simply means "p-to-remove" and will be used as 0.10 cut off for this study. The forward selection method is reverse of backward elimination method. In this method, we start with no variables in the model and for all predictors not in the model, their p-value is checked if they are added to the model. The process is continued until descriptors with lowest p-value less than αcritical is chosen and no new descriptors can be added. Stepwise regression is a combination of backward elimination and forward selection.[11] It is similar to forward selection except that variables are removed from the model if they become nonsignificant as other predictors are added. Stepwise procedures are relatively cheap computationally, but they do have some drawbacks.

The model does tend to overfit data and gives an equation based on input file. Few ways to check the quality of model includes:

1. P-value= (correlation b/w descriptor and crystallizability), Less than 10% considered good
2. R^2 Value = Statistical measure of how close the data is to the fitted regression line (Above 50% considered good)
3. GOF= Tells us how good the model is (Above 75% considered good).

The regression model is the logit function:

$$P(crystallizability) = \frac{exp(Y)}{1+exp(Y)}$$ ....Equation 8.1

Here, P is the predicted outcome, a number between 0 and 1, which can be interpreted as a probability, and t is called the 'linear predictor' and depends on the experimental factors.

Let the factor levels that affect the experiment be denoted xi, i.e. the ith factor has level x. The importance of each factor (their regression coefficients) are denoted β_n. We can then write $t = \beta_0 + \beta_1 x_1$. The first β_0 is a constant, corresponding to the y-axis intercept in linear regression. If there are several factors, we write the 'linear predictor' as a scalar product

$$t = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \text{Equation 8.2}$$

8.2 Experimental

8.2.1 Hydrogen-bond propensity study

Hydrogen-bond propensity[8-9, 12] calculations (CSD Version 5.38 and Mercury 3.9) were used as a way of predicting the most likely interactions in the structures of the class I and class III molecules Each compound was sketched and auto-edited, functional groups were selected as suggested by Mercury, a training dataset (350-600 structures per functional group) was made and the propensities were calculated with an ROC curve higher than 0.800 ("good discrimination").

8.2.2 Logistic regression model

The logistic regression model is used because the response variable in this study is categorical i.e. the output can take only two values, I and III and there are one or more independent variables that determine an outcome. The MINITAB[13-14] software will be used to determine the regression model. Different methods of determining the model were used such as stepwise, forward selection and backward elimination, Figure 8.3.
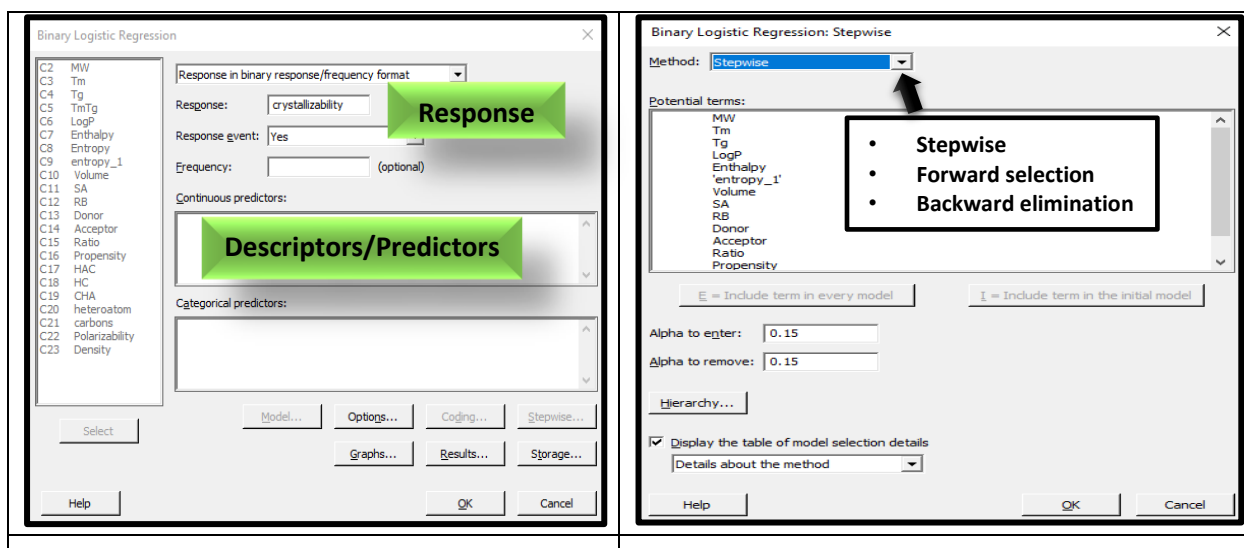
Figure 8.3 (a) Representation of prediction and response variables, (b) different prediction methods.

The logistic regression models were obtained using 36 molecules (half were classified as crystallizable and half were classified as non-crystallizable). The methods used to determine the model is stepwise selection and forward selection. Four different models were obtained (Figure 8.4). The details of model 1 and model 2 is provided in Table 8.1, Figure 8.3 and Table 8.2, Figure 8.4 respectively.

Table 8.1 Summary of the model 1 obtained in this study

Class I molecules	18
Class II molecules	18
Methods used	Stepwise selection, Forward selection
Significant descriptors	MW, Tg, SA, RB, heteroatom
GOF	.99, .99, .95
Model equation	33.5 -0.1508 MW+0.197Tg -0.282SA+1.84RB+3.25heteroatom
R^2 Value:	73.31%
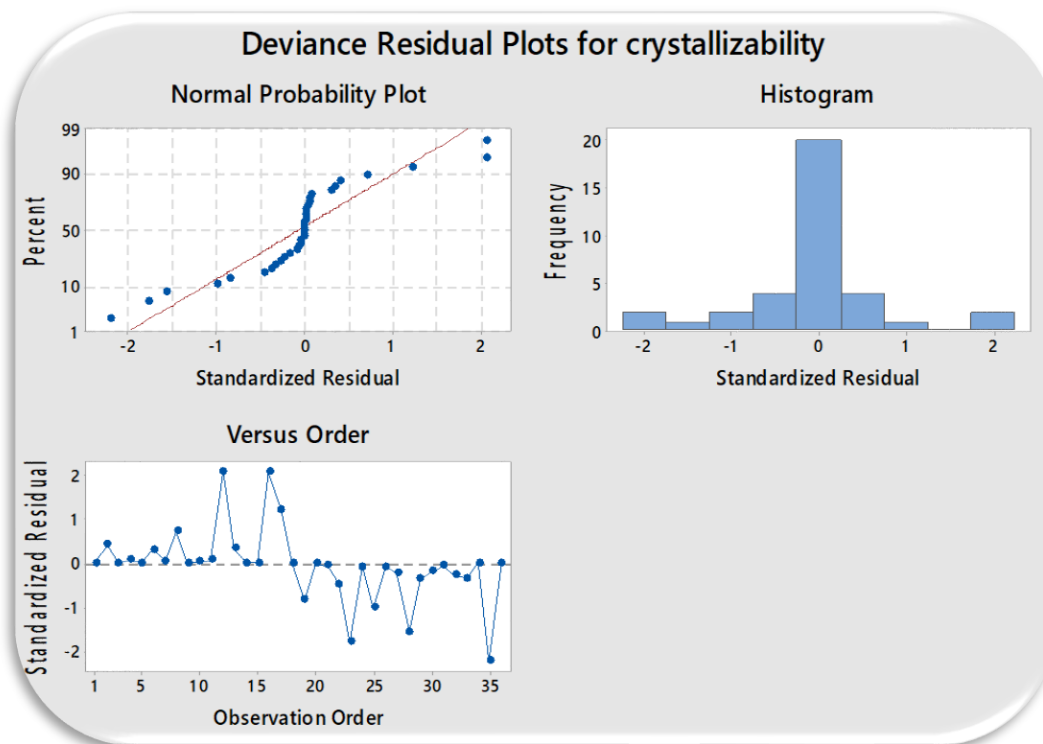
Figure 8.4 Normal probability plot and histogram for the model 1

The probability equation for crystallizability obtained from this model 2 is given below:

$$P = \frac{exp(33.5 - 0.1508MW + 0.971Tg - 0.282SA + 1.84RB + 3.25heteroatom)}{1 + exp(33.5 - 0.1508MW + 0.971Tg - 0.282SA + 1.84RB + 3.25heteroatom)} \dots\dots\dots\dots\dots Equation\ 8.3$$

Table 8.2 Summary of the model 2 obtained in this study

Class I molecules	18
Class II molecules	18
Methods used	Forward selection
Significant descriptors	MW, Tm, SA, RB, heteroatom
GOF	.99, .87, .97
Model equation	15.53 - 0.1227 MW + 0.0891 Tm - 0.276 SA + 1.039 RB + 3.90 heteroatom
R² Value:	68.9%
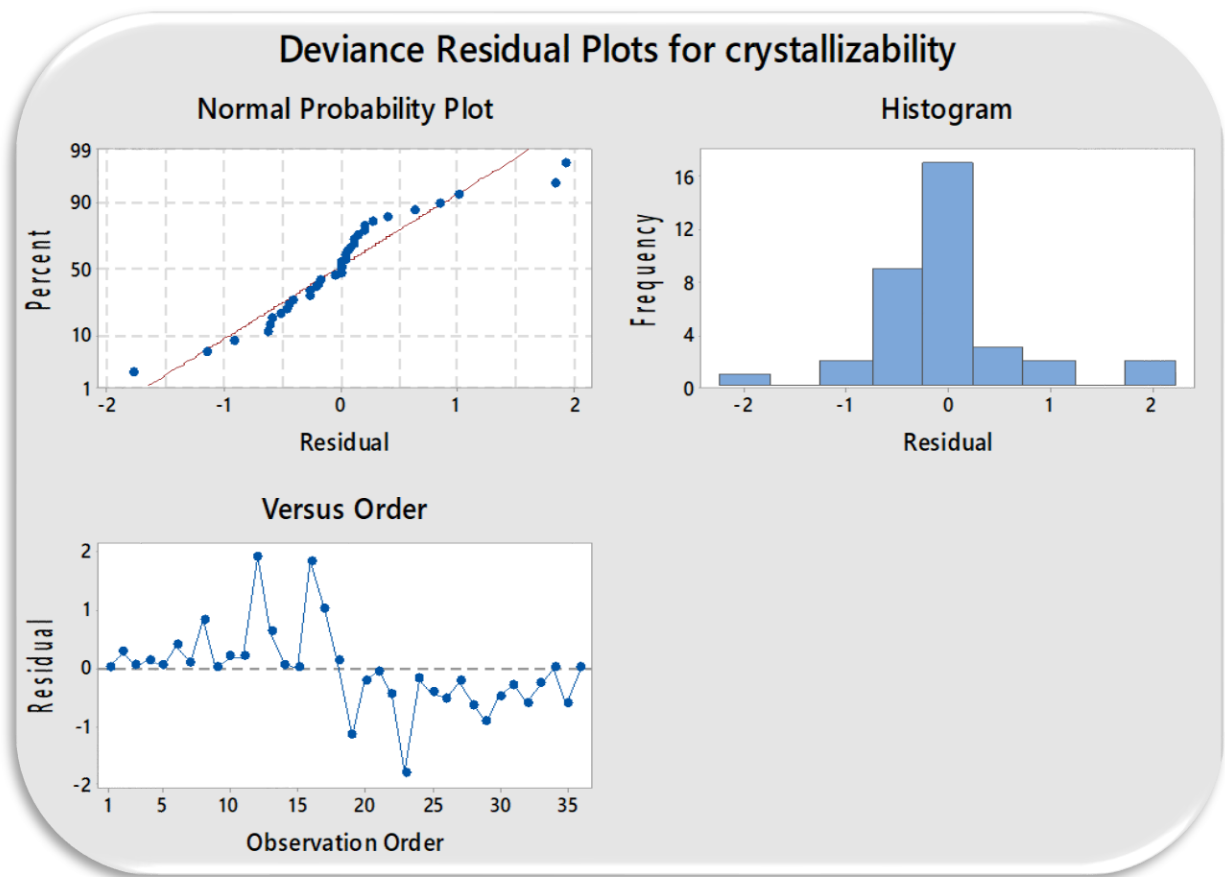
Figure 8.5 Normal probability plot and histogram for the model 2

The probability (crystallizability) equation obtained from this model 2 is given below:

$$P = \frac{exp(15.53 - 0.1227 \text{ MW} + 0.0891 \text{ Tm} - 0.276 \text{ SA} + 1.039 \text{ RB} + 3.90 \text{ heteroatom})}{1 + exp(15.53 - 0.1227 \text{ MW} + 0.0891 \text{ Tm} - 0.276 \text{ SA} + 1.039 \text{ RB} + 3.90 \text{ heteroatom})} \dots\dots\dots\dots \text{Equation 8.4}$$

Model 1 — Descriptors: MW, Tg, SA, RB, heteroatom
$Y' = 33.5 - 0.1508\,MW + 0.197\,Tg - 0.282\,SA + 1.84\,RB + 3.25\,heteroatom$

Model 2 — Descriptors: MW, Tm (experimental), SA, RB, heteroatom
$Y' = 15.53 - 0.1227\,MW + 0.0891\,Tm - 0.276\,SA + 1.039\,RB + 3.90\,heteroatom$

Model 2 (modified) — Descriptors: MW, Tm (online), SA, RB, heteroatom
$Y' = 15.53 - 0.1227\,MW + 0.0891\,Tm - 0.276\,SA + 1.039\,RB + 3.90\,heteroatom$

Model 3 — Descriptors: MW, Tg, TPSA(Tot), S/L axis ratio
$Y' = 30.0 - 0.0618\,MW + 0.0680\,Tg - 0.1186\,TPSA(Tot) - 14.4\,S/L\ axis\ ratio$

Model 4 — Descriptors: MW, Tm (experimental), TPSA(Tot), S axis (Å)
$Y' = 25.7 - 0.0399\,MW + 0.0482\,Tm - 0.1209\,TPSA(Tot) - 2.15\,S\ axis\ (Å)$

Figure 8.6 Four different models and the significance of each descriptor.

8.2.3 Training dataset

Taylor's approach based on DSC melt experiments was used to determine the final classification, Table 8.1. Taylor et. al classified molecules as class I, class II and class III but we can only have response variable as two outputs. We combined class II and class III molecules together because based on initial co-relation studies between descriptors and crystallizability, we found that class II and class III were indistinguishable.

Table 8.3 Classification of each molecule based on melt study and classification used for logistic regression model.

	ID	Based on melt	Final classification
1	4-Biphenylcarboxylic acid	Class I	Class I
2	4-Biphenylmethanol	Class I	Class I
3	4-biphenylcarboxaldehyde	Class I	Class I
4	4-Phenylphenol	Class I	Class I
5	Anthranilic acid	Class I	Class I
6	Antipyrin	Class I	Class I
7	Atenolol	Class I	Not included in the study
8	Benzamide	Class I	Class I
9	Benzocaine	Class I	Class I
10	caffeine	Class I	Class I
11	Carbamazepine	Class I	Class I
12	Chlorpropamide	Class I	Class I

13	Chlorzoxazone	Class I	Class I
14	Felbinac	Class I	Class I
15	Flufenamic acid	Class I	Class I
16	Griseofulvin	Class I	Class I
17	Haloperidol	Class I	Class I
18	Indoprofen	Class I	Class I
19	Lidocaine	class I	Class I
20	Phenacetin	Class I	Class I
21	Theophylline	Class I	Class I
22	Tolbutamide	Class I	Class I
23	Tolfenamic acid	Class I	Class I
24	Acetaminophen	Class II	Class III
25	Bifonazole	Class II	Class III
26	Celecoxib	Class II	Class III
27	cinnarizine	Class II	Class III
28	Clofoctol	Class II	Class III
29	Dibucaine	Class II	Class III
30	Droperidol	Class II	Class III
31	Flurbiprofen	Class II	Class III
32	Nifedipine	class II	Class III
33	Salicin	Class II	Class III
34	Tolazamide	Class II	Class III
35	Aceclofenac	Class III	Class III
36	clotrimazole	Class III	Class III
37	Felodipine	Class III	Class III
38	Fenofibrate	Class III	Class III
39	Ibuprofen	Class III	Class III
40	Indomethacin	Class III	Class III
41	itraconazole	Class III	Class III
42	ketoconazole	Class III	Class III
43	Ketoprofen	Class III	Class III
44	loratadine	Class III	Class III
45	miconazole	Class III	Class III
46	Nilutamide	Class III	Class III
47	Nimesulide	Class III	Class III
48	Pimozide	Class III	Class III
49	Probucol	Class III	Class III
50	Procaine	Class III	Class III
51	Ritonavir	Class III	Class III

8.2.4 Choice of descriptors

Number of descriptors including physical parameters such as molecular weight, donors, acceptors, ratio of D/A, rotatable bond, polarizability, enthalpy etc were chosen for this study. The list of descriptors is shown in Figure 8.8. Descriptor Information for each molecule is provided in Appendix F.
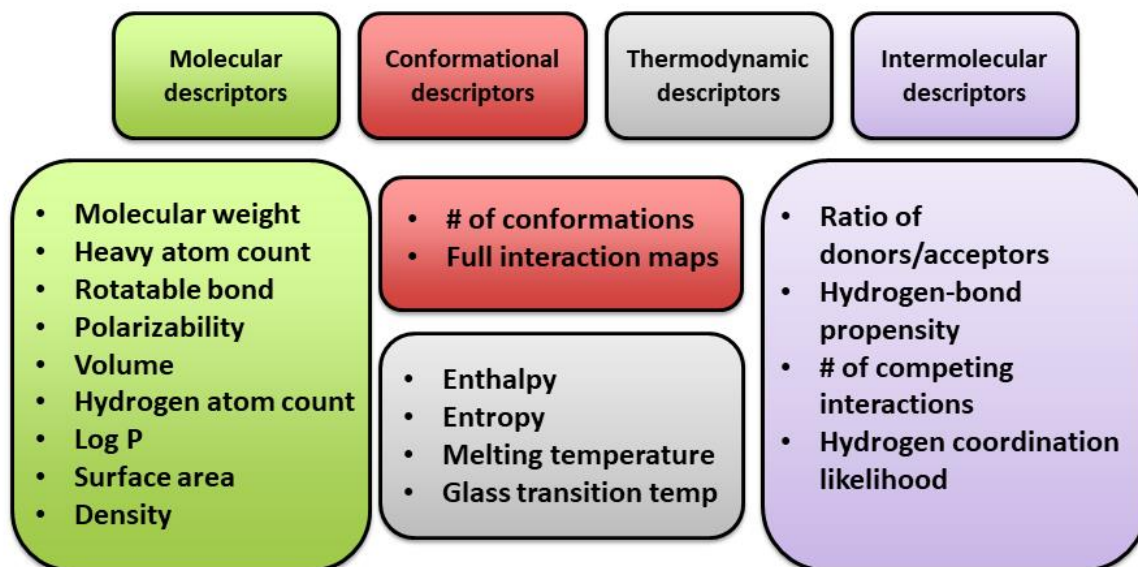
Figure 8.7 Various molecular, conformational, thermodynamic and intermolecular descriptors chosen for this study are shown here.

8.3 Results

8.3.1 Propensity comparison study

Table 8.4 Representation of highest propensity interaction observed in each crystal structure and comparison with the known forms. White box indicates no more polymorphs, green box indicates experimental results match with prediction, red indicates experimental does not match with prediction and yellow indicates that the crystal structure was not available.

		Best propensity	# polymorphs	Form I	Form II	Form III	Form IV	Form V
1	4-Biphenylcarboxylic acid	0.42	1					
2	4-Biphenylmethanol	0.27	1					
3	4-Phenylphoneol	0.17	2					
4	Anthranilic acid	0.92	3	Intra	Intra	Intra		
5	Atenolol	0.71	2					
6	Benzamide	0.73	3					
7	Benzocaine	0.56	3					
8	Carbamazepine	0.62	5	Dimer	Dimer	Dimer	Dimer	Catemer
9	Chlorpropamide	0.67	5					
10	Chlorzoxazone	0.71	1					
11	Felbnac	0.39	1					
12	Flufenamic acid	0.91	8					
13	Haloperidol	0.38	1					
14	Indoprofen	0.3	1					

15	Phenacetin	0.59	1					
16	Theophylline	0.38	3	C=O		C=O		
17	Tolbutamide	0.64	4					
18	Lidocaine	0.31	1					
19	Tolfenamic acid	0.91	2					
20	Acetaminophen	0.56	6					
21	Celecoxib	0.7	1					
22	Clofoctol	0.07	NO CS					
23	Dibucaine	0.25	1					
24	Droperidol	0.26	2					
25	Flurbiprofen	0.32	2					
26	Nifedipine	0.26	3					
27	Salicin	0.59	1					
28	Tolazamide	0.59	2					
29	Aceclofenac	0.94	1					
30	Felodipine	0.32	5					
31	Ibuprofen	0.37	2					
32	Indomethacin	0.36	2					
33	Nimesulide	0.68	2					
34	Ketoprofen	0.19	1					
35	Pimozide	0.27	No CS					
36	Probucol	0.02	1					
37	Procaine	0.59	2					
38	Ritonavir	0.4	2					

8.3.2 Influence of molecular descriptors on the crystallizability prediction

Based on molecular weight, molecules with molecular weight less than 300 g/mol falls in Class I (21 out of 26 molecules, 81%) and above 300 g/mol falls in Class III (14 out of 23 molecules, 61%), Figure 8.8a.

Based on rotatable bond, 22 out of 26 (85%) of class I molecules have # of rotatable bonds less than 4 and 19 out of 23 molecules (82%) of class III molecules # of rotatable bonds ≥ 4, Figure 8.8b.

Based on heavy +hydrogen atom: 18 out of 26 (69%) of class I molecules have Heavy atom+ hydrogen atom < 35 and 18 out of 23 (78%) of class III molecules have Heavy atom+ hydrogen atom > 35, Figure 8.9a.

Class III molecules have larger surface area compared to class I molecules, Figure 8.9b.

213
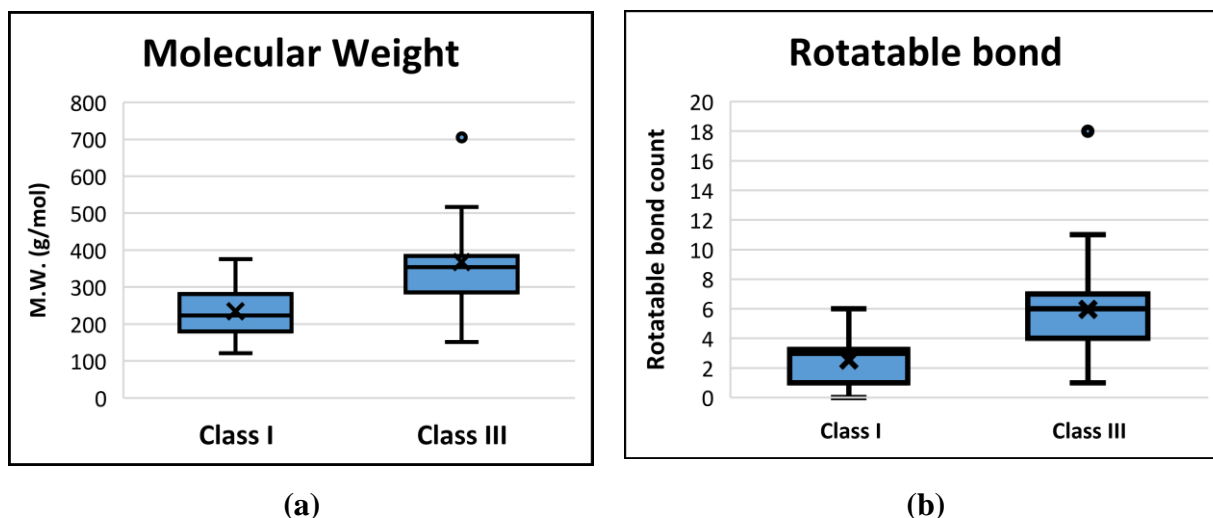
Figure 8.8 (a) Molecular weight vs crystallizability, (b) Rotatable bond vs crystallizability.
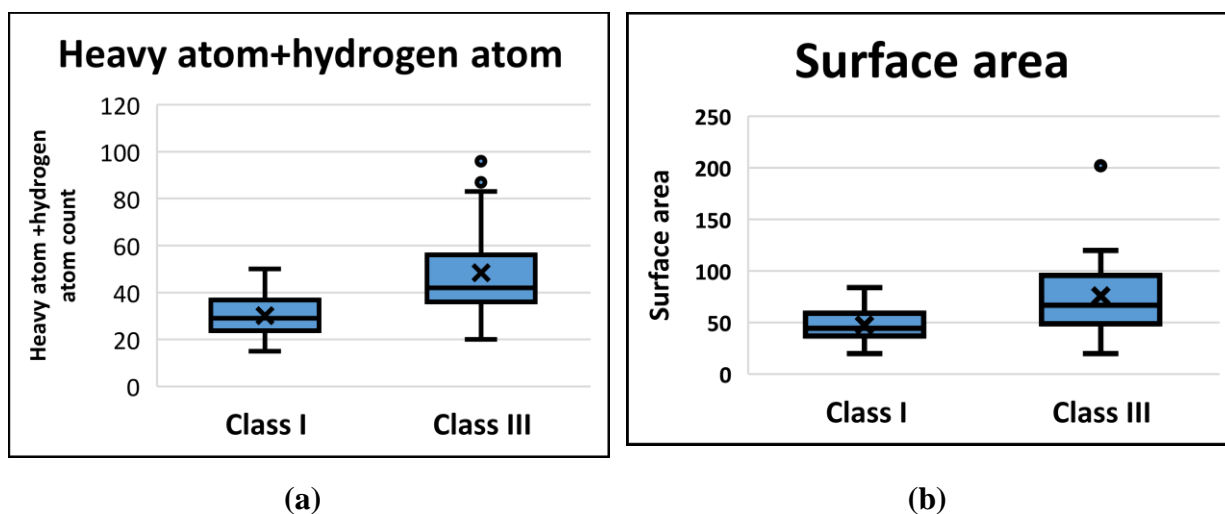
Figure 8.9 (a) Heavy atom + hydrogen atom count and (b) surface area vs crystallizability classification of molecules.

8.3.3 Influence of conformational descriptors on crystallizability prediction

The number of conformations increases as we go from class I to class III molecules (Table 8.5).

Table 8.5 # of conformers for each molecule in class I and class III

ID	Crystallizability	conformers	ID	Crystallizability	conformers
4-Biphenylcarboxylic acid	Class I	3	Tolazamide	Class III	2000
4-Biphenylmethanol	Class I	2	Flurbiprofen	Class III	38
4-Phenylphenol	Class I	2	Clofoctol	Class III	384
Anthranilic acid	Class I	2	Celecoxib	Class III	72

214

Benzamide	Class I	1	Dibucaine	Class III	2000
Benzocaine	Class I	19	Droperidol	Class III	2000
Carbamazepine	Class I	4	Nifedipine	Class III	223
Chlorpropamide	Class I	1872	Salicin	Class III	258
Chlorzoxazone	Class I	0	Aceclofenac	Class III	2000
Felbinac	Class I	24	Felodipine	Class III	546
Flufenamic acid	Class I	95	Ibuprofen	Class III	46
Haloperidol	Class I	2000	Indomethacin	Class III	384
Indoprofen	Class I	48	Nimesulide	Class III	71
Phenacetin	Class I	40	Ketoprofen	Class III	153
Theophylline	Class I	0	Pimozide	Class III	2000
Tolbutamide	Class I	2000	Probucol	Class III	174
Lidocaine	Class I	2000	Procaine	Class III	2000
Tolfenamic acid	Class I	46	Ritonavir	Class III	2000

8.3.4 Influence of intermolecular descriptors on crystallizability prediction

Ratio of donor/acceptor: 14 out of 20 (70%) of class I molecules have ratio of donor/acceptor $\geq$ 0.5 and 9 out of 12 (75%) of class III molecules have ratio of donor/acceptor < 0.5 (Figure 8.10).
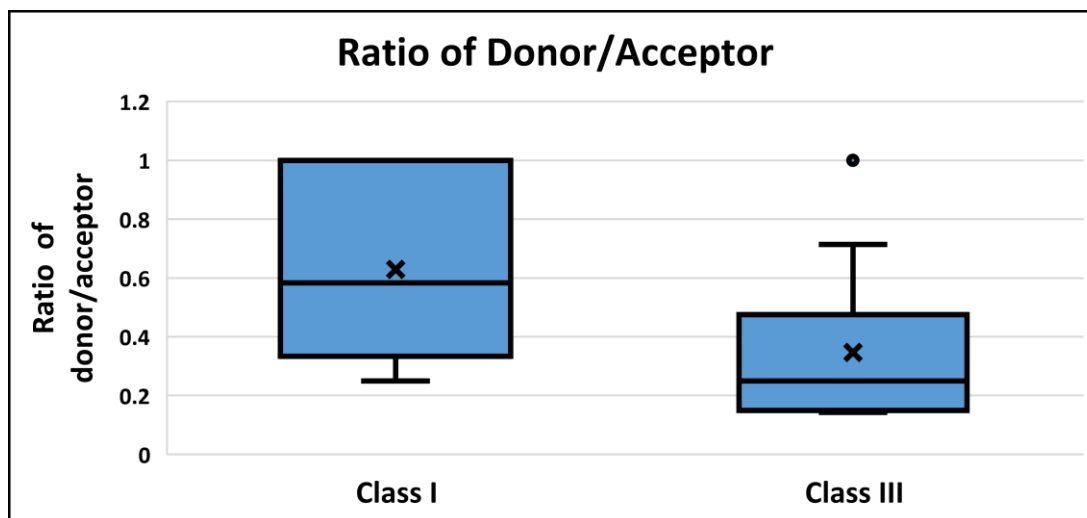
Figure 8.10 Ratio of donor/acceptor vs crystallizability

Propensity: In class I: 11 out of 21molecules, 52% molecules have propensity higher than 0.5 and In Class III, 6 out of 9 molecules, 67% molecules have propensity lower than 0.5 (Figure 8.11).
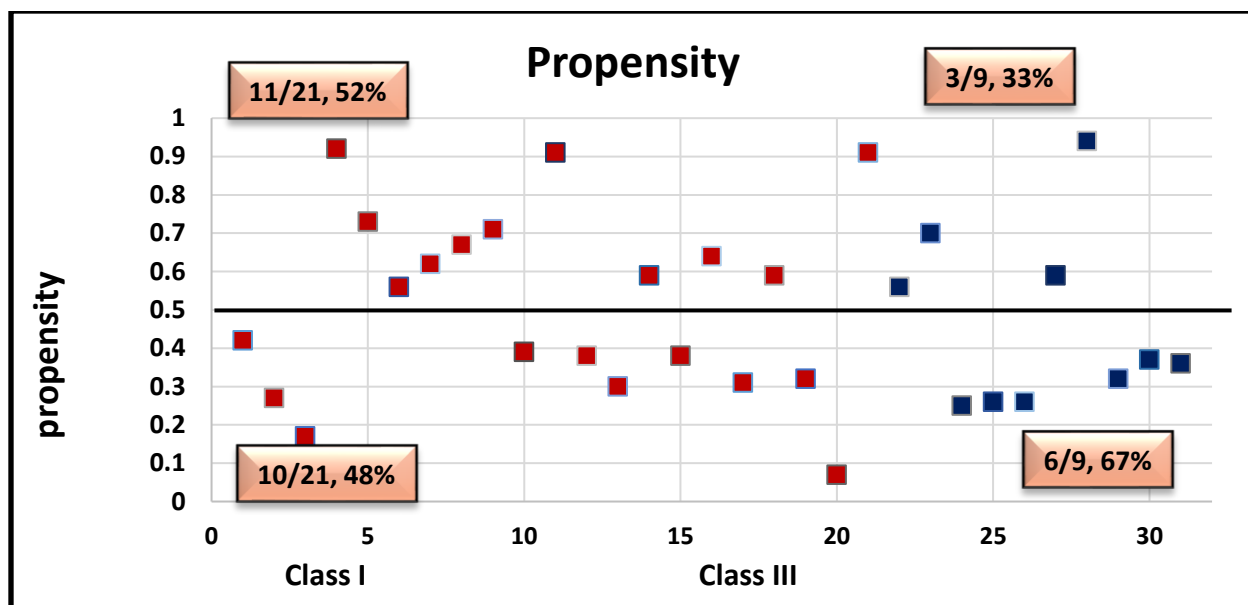
Figure 8.11 Propensity vs crystallizability

# of competing interactions: In class I, 10 out of 22 molecules, 45% have 0 competing interactions and in class III, 3 of 10 molecules, 30% have 0 competing interactions (Figure 8.12).
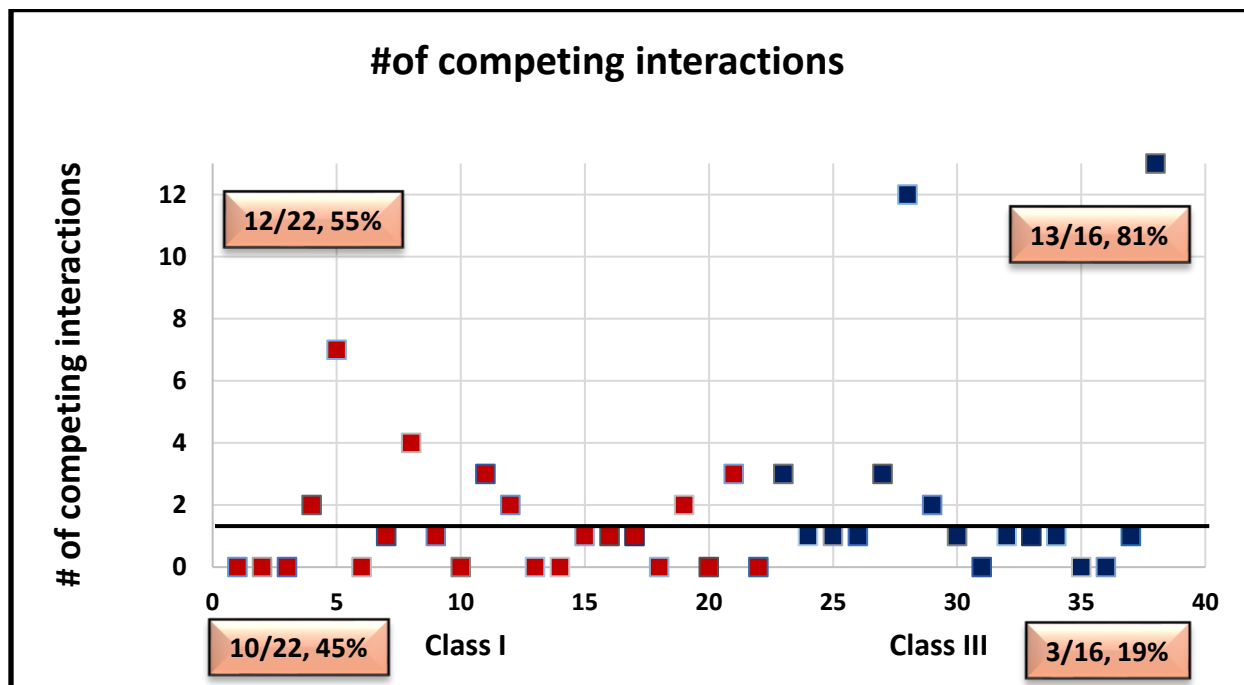
Figure 8.12 # of competing interactions vs crystallizability

Risk of synthon polymorphism: The risk of synthon polymorphism is determined by # of competing interaction in the molecules using propensity chart. Anything below 0.10 was not considered a valid interaction for this comparison. In class I, 11 out of 19 molecules, 58% have 0 competing interactions, Table 8.6 and in class III, 6 of 19 molecules, 32% have 0 competing interactions, Table 8.7.

Table 8.6 Risk of synthon polymorphism in class I molecules.

	Molecules	# of structures predicted	# of structures after cif. Insertion	Is the best putative structure experimentally observed?	Risk of synthon polymorphism?
1	4-Biphenylcarboxylic acid	1	1	yes	No
2	4-biphenylmethanol	1	1	yes	No
3	4-phenylphoneol	1	1	yes	No
4	Anthranilic acid	2	7	no	Yes
5	Atenolol	41	4	yes	Yes
6	Benzamide	1	2	yes	No
7	Benzocaine	2	3	yes	No
8	Carbamazepine	2	3	yes	No
9	Chlorpropamide	2	4	no	Yes
10	Chlorzoxazone	1	1	yes	No
11	Felbnac	1	1	yes	No
12	Flufenamic acid	2	1	no	Yes
13	Haloperidol	1	2	no	Yes
14	Indoprofen	2	2	no	Yes
15	Phenacetin	2	1	yes	No
16	Theophylline	6	6	yes	Yes
17	Tolbutamide	2	7	no	Yes
18	Lidocaine	1	4	no	No
19	Tolfenamic acid	0	1	yes	No
			Success rate	12/19 (63%)	11/19 (58%)

Table 8.7 Risk of synthon polymorphism in class III molecules.

	Molecules	# of structures predicted	# of structures aftr cif. Insertion	Is the best putative structure experimentally observed ?	Risk of synthon polymorphism?
1	Acetaminophen	2	2	no	Yes
2	Celecoxib	3	3	no	Yes
3	Clofoctol			N/A	No
4	Dibucaine	1	2	yes	Yes
5	Droperidol	2	3	yes	Yes
6	Flurbiprofen	1	3	no	No

7	Nifedipine	0	1	Yes	Yes
8	Salicin	>40	>20	no	Yes
9	Tolazamide	1	17	yes	Yes
10	Aceclofenac	1	4	yes	Yes
11	felodipine	2	2	yes	Yes
12	Ibuprofen	1	2	yes	No
13	Indomethacin	3	3	No	Yes
14	Nimusalide	2	3	no	Yes
15	Ketoprofen	2	2	yes	Yes
16	Pimozide			N/A	No
17	Probucol	1	3	no	No
18	Procaine	2	3	No	No
19	Ritonavir	>50	>50	no	Yes
			Success rate	8/17 (47%)	6/19 (32%)

8.3.5 Logistic regression model results

The model 1 was very good with 73% chances of predicting the unknown correctly. The five most significant descriptors were molecular weight (MW), glass transition temperature (Tg), surface area (SA), rotatable bond (RB) and heteroatom. The model 2 was based on molecular weight (MW), melting temperature (Tm), surface area (SA), rotatable bond (RB) and heteroatom with R2 value of 69%. All the descriptors in model 2 can be easily determined except melting point, therefore an online software (https://www.epa.gov/tsca-screening-tools/epi-suitetm-estimation-program-interface) was used to predict the melting temperature. The model 3 was based on molecular weight (MW), glass transition temperature (Tg), topological surface area (TPSA), small/large (S/L axis ratio) with R2 value of 76%. The model 4 was based on molecular weight (MW), melting temperature (Tm), topological surface area (TPSA), S axis (Å) with R2 value of 76%.

8.3.6 Regression Output Analysis (training dataset)

All four models were tested on the training dataset and the comparison between prediction and experimental crystallizability was done and the success rate with each model is shown in Figure 8.13.
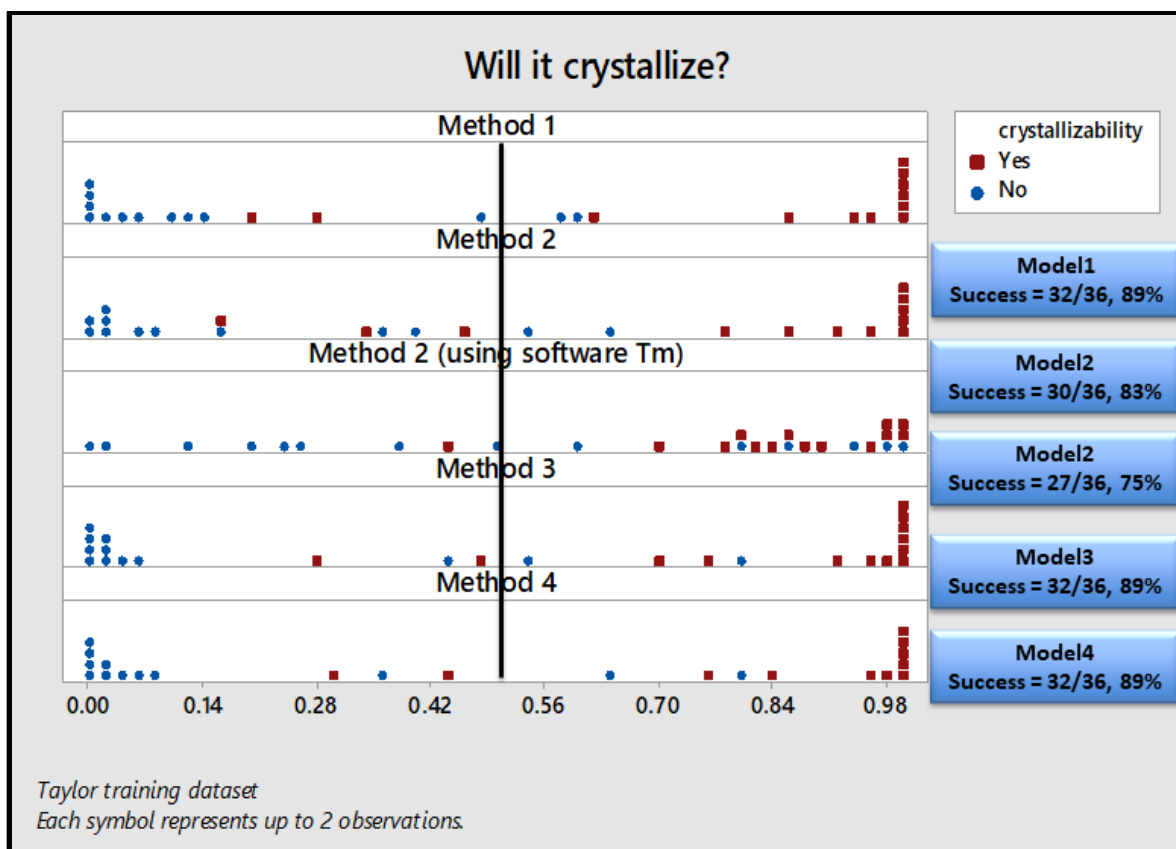
Figure 8.13 The probability of crystallizability prediction of training dataset using four different models and the resulting success rate with each model.

8.3.7 External validation test datasets

The model was then tested on the external unknown molecules where experimental crystallizability is already established. The prediction results were compared with the experimental results. Three different test datasets were studied, and the results are shown in Figure 8.14-8.16.

8.3.7.1 Test datatset-1 (Bergstrom et.al. study (22 molecules)
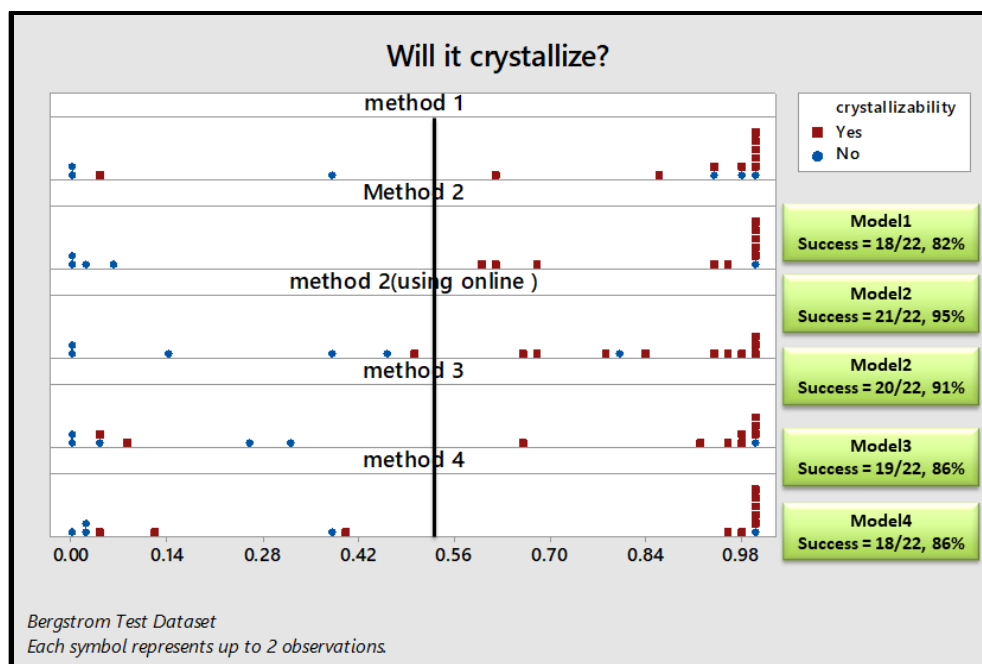
Figure 8.14 The probability of crystallizability prediction of bergstrom *et.al.* study[15] using four different models and the resulting success rate with each model.

8.3.7.2 Test datatset-2 (Kohrenon et.al. study (12 molecules)
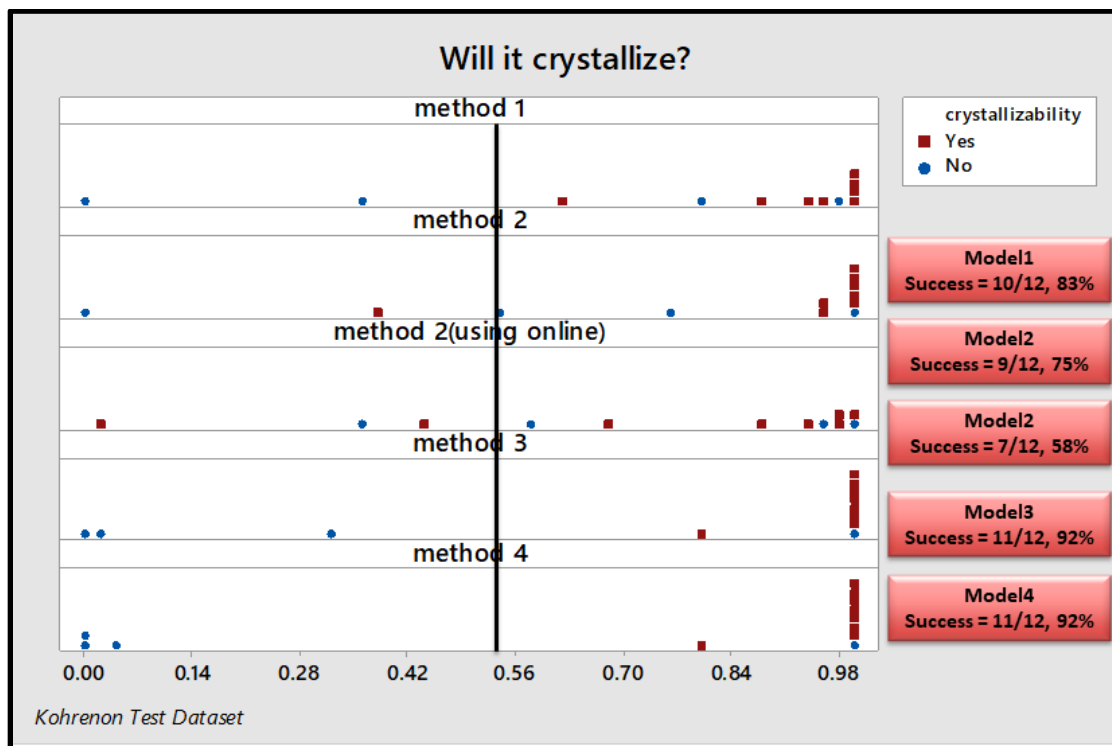
Figure 8.15 The probability of crystallizability prediction of Kohrenon *et.al.* study[16] using four different models and the resulting success rate with each model.

8.3.7.3 Test datatset-3 (Taylor et.al. study (13 molecules)
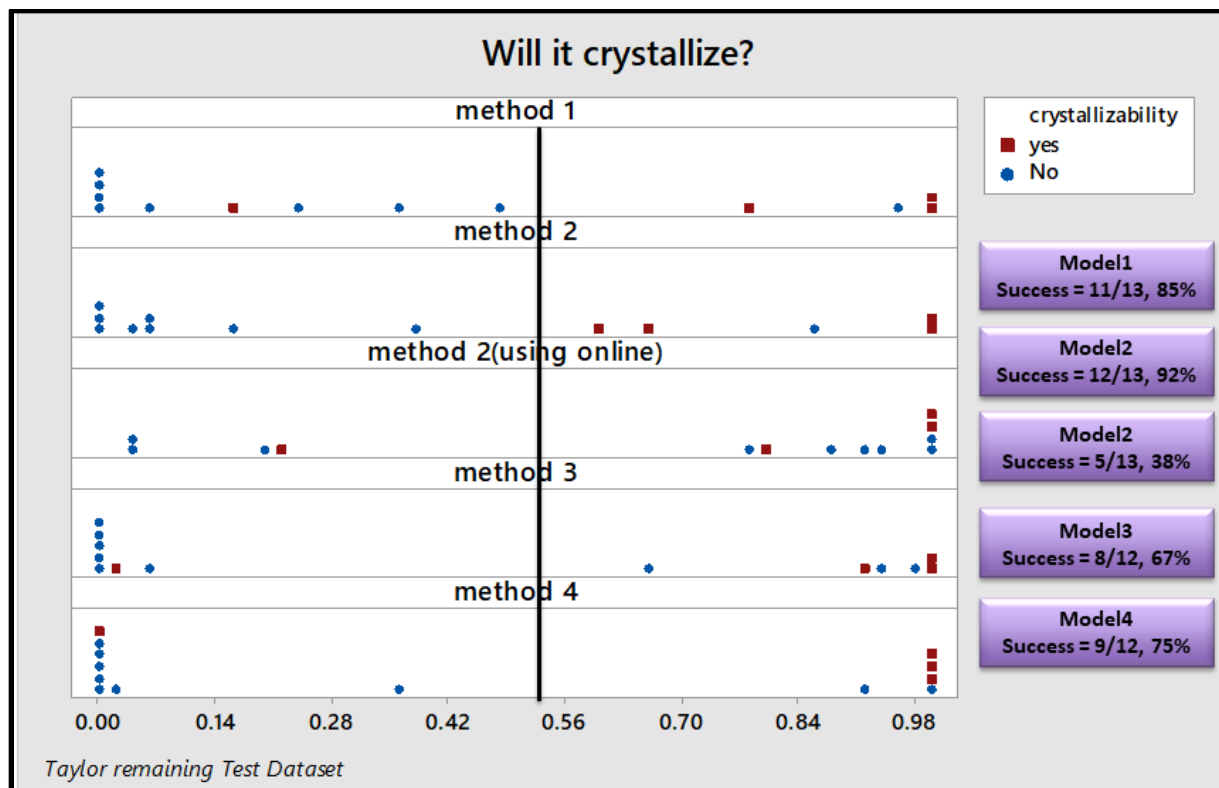
Figure 8.16 The probability of crystallizability prediction of Taylor *et.al.* study[4] using four different models and the resulting success rate with each model.

8.4 Discussion

8.4.1 Influence of various descriptors on the crystallizability

Based on comparison between various descriptors and crystallizability classification, it was found that molecular weight, rotatable bond, # of competing interactions, risk of synthon polymorphism, ratio of donor/acceptor affects whether a molecule crystallizes or not, Table 8.8. A low molecular weight molecule with less number of rotatable bonds and lower risk of synthon polymorphism is more prone to fall in class I (crystallizable) whereas high molecular weight compounds with higher number of rotatable bonds and high risk of synthon polymorphism are more prone to fall in class III (non-crystallizable).

Table 8.8 Summary of correlation between descriptors and each classification

	Class I	Class III
Molecular weight	Low (<300g/mol)	High (>300g/mol)
Rotatable bond	Low (<4)	High (≥4)
Heavy +hydrogen atom	low	High
Ratio of donor/acceptor	High (>0.5)	Low (<0.5)
Propensity	Medium	Low
# of competing interactions	Medium	High
Risk of synthon polymorphism	Low	High

8.4.2 Selection and validation of training and test datasets

A logistic regression model was derived to predict drug crystallizability. Acetaminophen was considered as an outlier on the model as it affected the model negatively when present in the dataset therefore, it was removed from the dataset. The chosen iterative procedure to select the optimal model algorithm and analysis criterion is illustrated in section 8.1.2. For each selection, the GOF, R^2 value, graphs were carefully analyzed to choose the best fit model. The model predicts the crystallizability outcome as 0 for non-crystallizable and 1 for crystallizable. The 0.5 cut off was chosen to analyze the results and in 32 out of 36 molecules (89% success rate), there was agreement between prediction and experimental results using model 1. Out of 36 molecules, 16 were true positive,2 were false positives, 2 were false negatives and 16 were true negatives. The model suggested that five descriptors play an important role in the crystallizability; molecular weight, rotatable bond, surface area, glass transition temperature and number of heteroatoms present in the molecule. The model 2 was same model as model 1 but the glass transition temperature descriptor was replaced with melting point temperature. The complete analysis of four models and the resulting success rate is shown in Table 8.9.

Table 8.9 Summary of success rate of each model for different datasets.

	Model 1 (%)	Model 2 (%)	Model 2 (online) (%)	Model 3 (%)	Model 4 (%)
Taylor	89	83	75	89	89

Bergstrom	82	95	95	86	86
Kohrenon	83	75	58	92	92
Tay-remaining	85	92	38	67	75
Overall	85%	88%	68%	84%	86%

In all validation studies, all four models predicted the experimental crystallization above 75% indicating the applicability of these simple model equations outside the scope of the sample set due to its moderately high predictive power and its structural simplicity. In a drug discovery setting, this degree of accuracy would be of high utility.
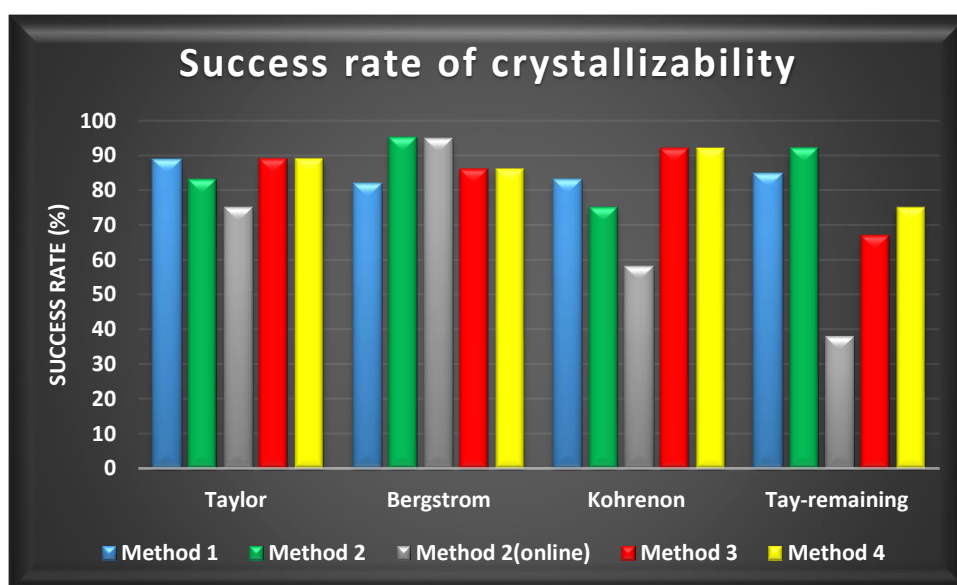
Figure 8.17 Summary of validation test results

Combining both training and validation dataset, we have 85% agreement (70 out of 83 molecules) between predicted and experimental crystallizability outcome using model 1 and the confusion matrix showing TP (true positive), FN(false negative), FP(false positive) and TN(true negative) values are shown in Figure 8.18.

Probability of crystallizability		Predicted outcome	
		Yes	No
Experiment	Yes	True positive	False negative
	No	False positive	True negative

(a)

Probability of crystallizability		Predicted outcome	
		Yes	No
Experiment	Yes	41	4
	No	9	29

(b)

Figure 8.18 Confusion matrix of predicted vs experimental results for 83 molecules involved in this study is shown using model 1

8.5 Conclusion

This work demonstrates that crystallizability of drug molecules can be predicted with a good degree of confidence using a combination of easily calculated, predicted, or measured parameters. The correlations of various descriptors such as molecular weight, RB, surface area, heteroatom, melting temperature, glass transition temperature, molecular shape/volume with the crystallizability provide insight into the key factors that influence mechanisms which drive crystallization. Four different models were used to predict the crystallizability of 83 different molecules and the predictions were made successfully with above 85% success rate.[6] This success rate is higher than what is reported in the literature with other predictive methods. Nevertheless, the predictive power of the selected models should be further validated on a larger external data set. Once successfully validated, such models could assist in faster and more cost effective decision making, especially in preformulation phases of future direction where crystallization strategies are under consideration.
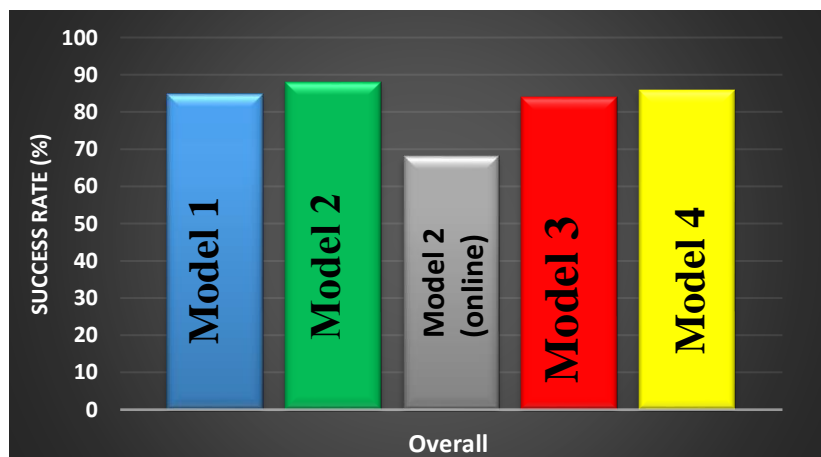
8.6 References

1. Hancock, B. C., *Journal of Pharmaceutical Sciences 106* (1), 28-30.

2. Wicker, J. G. P.; Cooper, R. I., *CrystEngComm* **2015,** *17* (9), 1927-1934.

3. Bhardwaj, R. M.; Johnston, A.; Johnston, B. F.; Florence, A. J., *CrystEngComm* **2015,** *17* (23), 4272-4275.

4. Baird, J. A.; Van Eerdenbrugh, B.; Taylor, L. S., *Journal of Pharmaceutical Sciences 99* (9), 3787-3806.

5. Van Eerdenbrugh, B.; Baird, J. A.; Taylor, L. S., *Journal of Pharmaceutical Sciences 99* (9), 3826-3838.

6. Nurzyńska, K.; Booth, J.; Roberts, C. J.; McCabe, J.; Dryden, I.; Fischer, P. M., *Molecular Pharmaceutics* **2015,** *12* (9), 3389-3398.

7. Hursthouse, M. B.; Huth, L. S.; Threlfall, T. L., *Organic Process Research & Development* **2009,** *13* (6), 1231-1240.

8. Galek, P. T. A.; Allen, F. H.; Fabian, L.; Feeder, N., *CrystEngComm* **2009,** *11* (12), 2634-2639.

9. Wood, P. A.; Feeder, N.; Furlow, M.; Galek, P. T. A.; Groom, C. R.; Pidcock, E., *CrystEngComm* **2014,** *16* (26), 5839-5848.

10. 128., D. A. F. S. M. T. a. P. C. U. P. p.

11. http://www.jerrydallal.com/lhsp/simplify.htm.

12. Galek, P. T. A.; Chisholm, J. A.; Pidcock, E.; Wood, P. A., *Acta Crystallographica Section B* **2014,** *70* (1), 91-105.

13. Arend, D. N. C. V. C. s. C., IL: U.S. Army Corps of Engineers Research Laboratory. (CERL Report No.CH7-22510).

14. Minitab 17 Statistical Software (2010). [Computer software]. State College, P. M., Inc. (www.minitab.com).

15. Pajula, K.; Taskinen, M.; Lehto, V.-P.; Ketolainen, J.; Korhonen, O., *Molecular Pharmaceutics* **2010,** *7* (3), 795-804.

16. Pajula, K.; Lehto, V.-P.; Ketolainen, J.; Korhonen, O., *Molecular Pharmaceutics* **2012,** *9* (10), 2844-2855.

Chapter 9 - Modulating physical properties of solid forms of urea using co-crystallization technology

9.1 Introduction

Urea is an inexpensive fertilizer with the highest nitrogen content among commonly used solid fertilizers. The solubility of urea in water is approx. 110-170 g/100 ml, in the 20-40 ˚C range, and it is rapidly hydrolyzed after application.[1] The high solubility of urea in water and rapid moisture intake after application makes it susceptible to mobility and/or runoff from treated areas where it kills aquatic plants and animals causing environmental damage. Because of the mobility, higher concentrations of urea are oftentimes applied to seeds and plants to ensure that a sufficient amount of active ingredient is available for utilization by the plant. Excessive application results in inefficiencies and large loss of excess nitrogen to the environment which can impact air and water quality, biodiversity and human health.[2] The high moisture content of urea at relative humidity of 40-50% causes stickiness, leading to storage and stability issues.
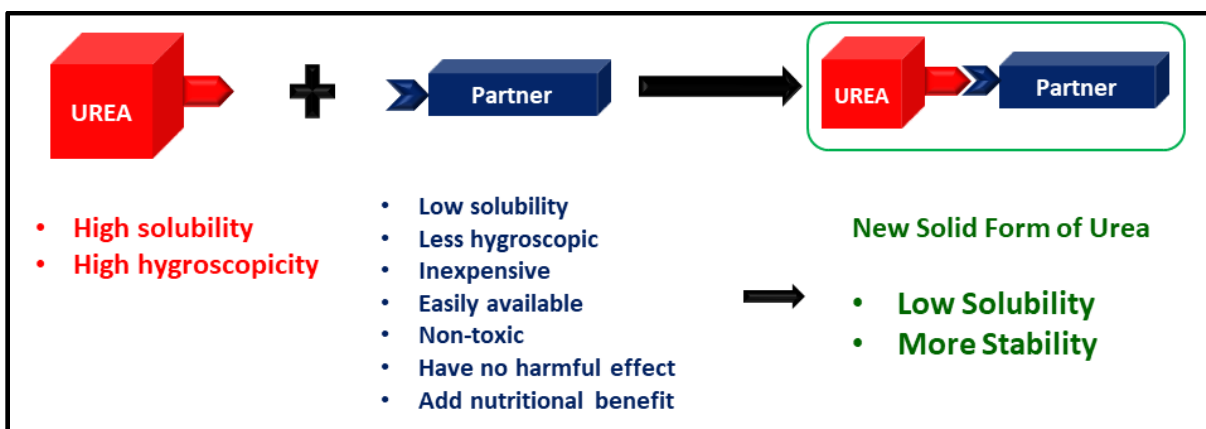
Scheme 9.1 An outline of co-crystallization event in the urea.

Various slow-release fertilizer (SRF) technologies have been used, thus preventing problems related to leaching, volatization, denitrification and run-off.[3] The uncoated slow release fertilizer (urea:formaldehyde,[4] isobutylidene-diurea[5] and inorganic salts[6]), coated slow-release fertilizers

(sulfur coated,[7] polymer coated[8] and hybrid coated urea[9]) as well as urease and nitrification bioinhibitors[7, 10] have been used to tackle the issues.[11] These methods have disadvantages related to manufacture cost of production, and irregular nutrient release.
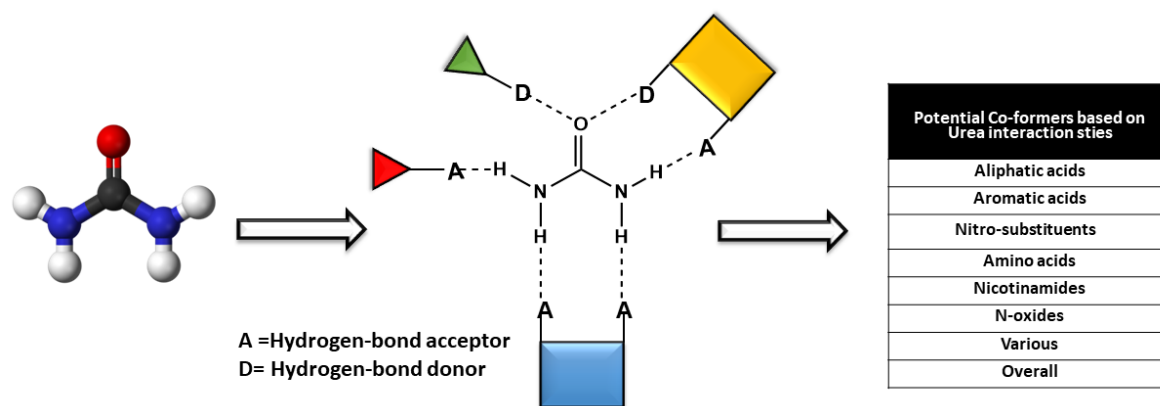
Scheme 9.2 Four potential avenues for intermolecular 'intervention' on urea (**A** and **D** = hydrogen-bond acceptor and donor, respectively) and resulting choice of potential co-formers.

Co-crystals has proven to be a powerful tool to alter the physiochemical properties of molecules of interest without modifying their biological properties. Therefore, by fine-tuning the crystalline environment of a compound without altering its molecular structure, we could potentially "dial-in" desirable physical properties, which would be highly significant to manufacturers/consumers of organic specialty chemicals.[12] It is able to do so using the strength of non-covalent intermolecular interactions, such as hydrogen bonding,[13-14] halogen bonding,[15] $\pi \cdots \pi$,[16] and other non-covalent interactions.[17] Co-crystals have proven to show solubility advantage over various pharmaceutical compounds with diverse chemical and pharmacological nature which includes myricetin (anticancer flavonol),[18] itraconazole (antifungal),[19] adefovir (antiviral),[20] aceclofenac (NSAID),[21] gabapentine (anticonvulsant),[22] aspirin(NSAID),[23-24] norfloxacin (antibiotic),[25] stanolone(API),[26] caffeine(stimulant),[27] efavirenz(antiviral)[28] and mirtazapine (antidepressant)[29] displaying potential to improve solubility and hence bioavailability. The technique has also been used for creating less-sensitive explosives, as well as modifying and optimizing other properties of energetic materials, a class of materials including explosives, propellants and pyrotechnics. Cocrystals of TNT, HMX and CL-20 have clearly demonstrated the capacity for modifying materials properties through co-crystallization.[30] Co-crystals have been used to stabilize volatile

liquid chemicals,[31] make non-linear optic materials[32-35] and are gaining interest in agro-chemicals.[36-37]

The study is done to address the following questions,

1. Can co-crystallization technology be utilized for producing new solid forms of urea (Scheme 9.1)?

2. Can the physiochemical properties of new solid forms of urea be altered without affecting the biological properties?

   2.1 Can we decrease solubility in a predetermined manner?

   2.2 Can we increase stability and hygroscopicity sensitivity?

We hypothesize that if we can incorporate a suitable co-former within a series of crystalline solids characterized by considerable structural consistency, we may be able to fine-tune properties such as aqueous solubility and moisture sensitivity. Changes to the physical properties will be achieved by varying the co-crystallizing agents in a systematic fashion without altering the precise nature of the molecular recognition events that drive the supramolecular assembly. This will allow us to modify physical properties and alter aqueous solubility, stability and hygroscopicity in order to optimize the performance and benefits of urea as a plant nutrient. It would be highly advantageous if it will be possible to alter/control the aqueous solubility in a predetermined manner.

In this study, we present the results from a systemic co-crystallization study comprising 60 experiments with selected co-formers intended to provide more details of new co-crystals of urea that can add value to the resulting solid form (Scheme 9.3). The co-formers chosen for this study are relatively inexpensive, easily available, non-toxic, have no harmful effect on the soil, or can add nutritional benefit to the soil. The choice of co-formers was also based on finding chemical functionalities that can bind in a complementary and effective manner with the molecular recognition sites that urea present to its surroundings (Scheme 9.2).

Aliphatic acids

Citric acid (CA) **Citramalic acid (CmA)** **Malonic acid (MA)** **Succinic acid (SA)** **Fumaric acid (FA)**

Glutaric acid (GA) **Adipic acid (AA)** **Pimelic acid (PA)** **Suberic (SbA)**

Tartaric acid (TA) **Maleic acid (MeA)** **Azelaic acid (AzA)** **Sebacic acid (SeA)**

Aromatic acids

3-Hydroxy benzoic acid (3HBA) **4-Hydroxy benzoic acid (4HBA)** **3-Amino benzoic acid (3ABA)** **4-Amino benzoic acid (4ABA)** **Phthalic acid (PtA)**

Pyrazine-2-carboxylic acid (Py2_CA) **Pyrazine2,3–dicarboxylic acid (Py2, 3DCA)** **Pyridine2,5-dicarboxylic acid (P2, 5DCA)** **Pyridine 2,6-dicarboxylic acid (P2, 6DCA)** **Pyrazine2,6-dicarboxylic acid (Py3,5DCA)**

Hippuric acid (HipA) **2,5-dihydroxybenzoic acid (2,5DHBA)** **3,4-dihydroxybenzoic acid (3,4DHBA)** **3,5-dihydroxybenzoic acid (3,5DHBA)**

229

Amino acids

Lysine (Lys)

Histidine (His)

Arginine (Arg)

Aspartic acid (Asp)

Alanine (Ala)

Nicotinamide's

Glutamic acid (Glu)

Nicotinamide (N)

N-methyl-nicotinamide (Me_N)

2-chloro-nicotinamide (2Cl_N)

6-chloro-nicotinamide (6Cl_N)

N-oxides

Pyrazine N-oxide (Py_NO)

Tetramethyl pyrazine N-oxide (TmPy_NO)

Pyrazine-NN-oxide (Py _NNO)

Tetramethylpyrazine-NNoxide (TmPy_NNO)

Bipyridine-N-oxide (BiP_NO)

Nitro-subtituents

o-nitrophenol (2NP)

m-nitrophenol (3NP)

p-nitrophenol (4NP)

1,2-dinitrobenzene 1,2DNB

1,3-dinitrobenzene 1,3DNB

1,4-dinitrobenzene 1,4DNB

o-nitrobenzoic acid 2NBA

m-nitrobenzoic acid 3NBA

p-nitrobenzoic acid 4NBA

230

Various

Glucose (Gl)   Biuret (Biu)   Imidazolidone (Im)   Maleimide (Ma)   Resorcinol (Re)

2,6-Lutidine (Lut)   Quinoline (Qui)   Uric acid (Ur)   1,10-Phenanthrolin (Phen)

Scheme 9.3 List of co-formers used in this study

9.2 Experimental

9.2.1 Materials

Urea and all co-formers were purchased from Aldrich and utilized without further purification.
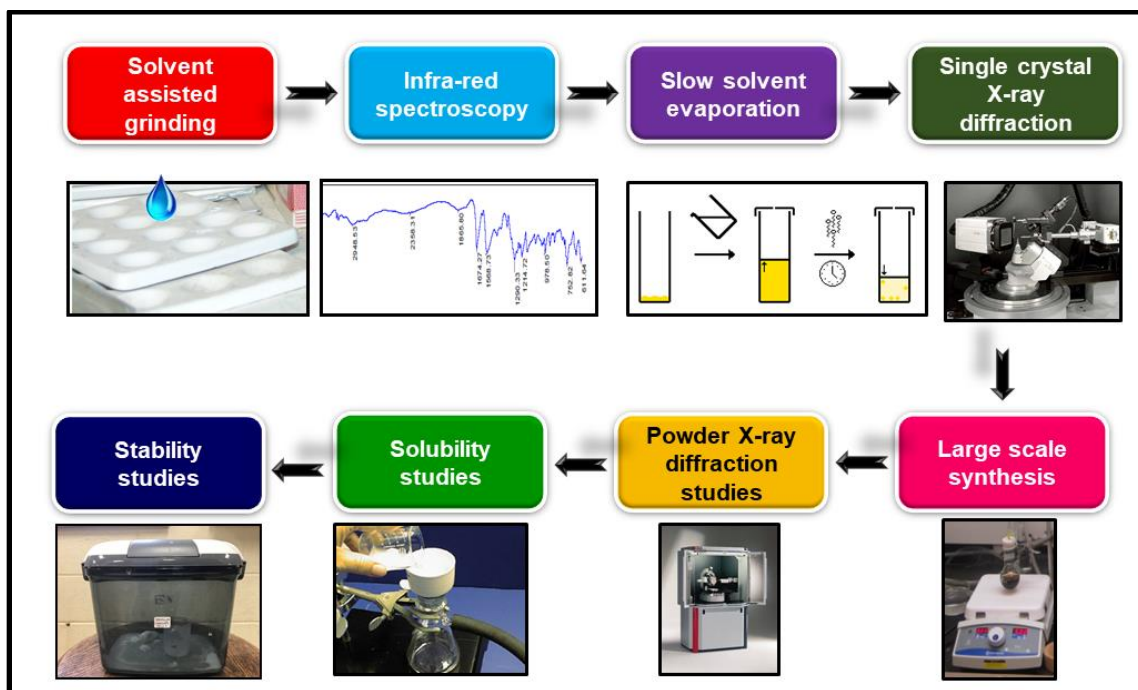
Scheme 9.4 Flow chart of experimental strategy employed in this study

Melting points were measured using Fisher-Johns melting point apparatus. Solution ^{1}H NMR data were collected on a Varian Unity plus 400 MHz spectrometer in DMSO. IR spectra were recorded with a Nicolet 380 FT-IR. An outline of experimental steps followed in this study is displayed in scheme 9.4

9.2.2 Isostar and full interaction maps (FIMs) search

The first phase of this study involves the use of several complementary approaches pursued in parallel. In order to develop robust protocols for co-crystal synthesis of urea we combine a knowledge-based approach for mapping out the intermolecular interactions that are most likely to take place with different parts of the urea functionality. To this end, a careful analysis of all relevant structural data from the Cambridge Crystallographic Data Centre will furnish a map of the most likely short contacts with the C=O and the N-H moieties. The data is analyzed using Isostar which provides 3-D maps of intermolecular interactions based upon existing experimental structural data. In this work, an intermolecular contact was defined as any contact shorter than the sum of the van der Waals radii of the atoms involved. Two terminal urea compounds with *cis* and *trans* confirmation were studied. The FIMs were done using Mercury CSD database 5.38 (Nov 2016) with updates from Feb and May 2017 with 2, 4, 6 and contour levels and using uncharged NH nitrogen, carbonyl oxygen and aromatic CH carbon.
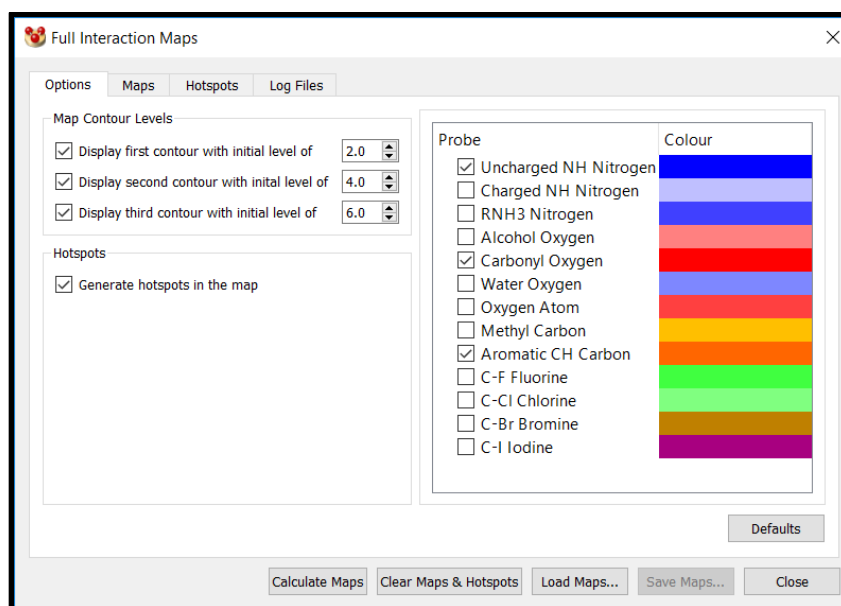
Figure 9.1 Flow chart of experimental strategy employed in this study

9.2.3 Co-crystal screening

The co-crystal screening experiments were based on appropriate stoichiometric ratios of target:co-former and all reactant combinations were put through slow solvent evaporation in different solvents to grow crystals suitable for single-crystal x-ray diffraction studies, (Table 9.1).
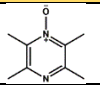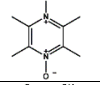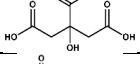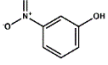
Table 9.1 Stoichiometric ratio, solvents used in the solution co-crystallization

	Conformers	M.W. (g mol^{-1})	Mass (g)	mmole	Solvent	Ratio Co-former: urea
			Aliphatic acids			
1	Citric acid (CA)	192	0.0192	0.10	Methanol, Water	1:2
2	Citramalic acid (CmA)	148.12	0.0148	0.10	Methanol, Ethanol	1:2
3	Malonic acid (MA)	52.00	0.0052	0.05	Methanol, Ethanol	1:2
4	Succinic acid (SA)	118.09	0.0059	0.05	Methanol, Ethanol	1:2
5	Fumaric acid (FA)	116.07	0.0058	0.05	Methanol, Ethanol	1:2
6	Glutaric acid (GA)	132.12	0.0066	0.05	Methanol, Ethanol	1:2
7	Adipic acid (AA)	146.14	0.0073	0.05	Methanol, Ethanol	1:2
8	Pimelic aid (PA)	160.17	0.0084	0.05	Methanol, Ethanol	1:2
9	Suberic acid (SbA)	174.20	0.0087	0.05	Methanol, Ethanol	1:2
10	Sebacic acid (SeA)	202.25	0.0101	0.05	Methanol, Ethanol	1:2
11	Tartaric acid (TA)	150	0.0150	0.10	Methanol, Ethanol	1:2
12	Maleic acid (MeA)	120	0.0120	0.10	Methanol, Ethanol	1:2
13	Azelaic acid (AzA)	188.22	0.0094	0.05	Methanol, Ethanol	1:2
			Aromatic acids			
14	3-Hydroxy benzoic acid (3HBA)	138.12	0.0138	0.10	Methanol, Ethanol	1:1
15	4-Hydroxy benzoic acid (4HBA)	138.12	0.0138	0.10	Methanol, Ethanol	1:1
16	3-Amino benzoic acid (3ABA)	137.13	0.0137	0.10	Methanol, Ethanol	1:1
17	4-Amino benzoic acid (4ABA)	137.13	0.0137	0.10	Methanol, Ethanol	1:1
18	Phthalic acid (PtA)	166	0.0166	0.10	Methanol, Ethanol	1:2
19	Pyrazine-2-carboxylic acid (Py2_CA)	124	0.0124	0.10	Methanol, Ethanol	1:2
20	Pyrazine2,3-dicarboxylic acid (Py2,3_CA)	168	0.0168	0.10	Methanol, Ethanol	1:2
21	Pyridine2,5-dicarboxylic acid (Py2,5_CA)	168	0.0168	0.10	Methanol, Ethanol	1:2
22	Pyridine 2,6-dicarboxylic acid (Py2,6_CA)	168	0.0168	0.10	Methanol, Ethanol	1:2
23	Pyrazine3,5-dicarboxylic acid (Py3,5_CA)	168	0.0168	0.10	Methanol, Ethanol	1:2
24	Hippuric acid (HipA)	179	0.0179	0.10	Methanol, Ethanol	1:2
25	2,5-Dihydroxybenzoic acid (2,5DHBA)	154	0.0154	0.10	Methanol, Ethanol	1:1
26	3,4-Dihydroxybenzoic acid (3,4DHBA)	154	0.0154	0.10	Methanol, Ethanol	1:1
27	3,5-Dihydroxybenzoic acid (3,5DHBA)	154	0.0154	0.10	Methanol, Ethanol	1:1
28	o-Nitrophenol (2NP)	140	0.028	0.20	DMF, Methanol	1:1
29	m-Nitrophenol (3NP)	140	0.028	0.20	DMF, Methanol	1:1
30	p-Nitrophenol (4NP)	140	0.028	0.20	DMF, Methanol	1:1
31	1,2-Dinitrobenzene (1,2DNB)	168	0.0168	0.10	Methanol, Ethanol	1:2
32	1,3-Dinitrobenzene (1,3DNB)	168	0.0168	0.10	Methanol, Ethanol	1:2
33	1,4-Dinitrobenzene (1,4DNB)	168	0.0168	0.10	Methanol, Ethanol	1:2
34	o-Nitrobenzoic acid (2NBA)	167	0.0167	0.10	Methanol, Ethanol	1:2
35	m-Nitrobenzoic acid (3NBA)	167	0.0167	0.10	Methanol, Ethanol	1:2
36	p-Nitrobenzoic acid (4NBA)	167	0.0167	0.10	Methanol, Ethanol	1:2
			Amino acids			
37	Lysine (Lys)	146	0.0146	0.10	Methanol/Water (1:1)	1:1
38	Histidine (His)	155	0.0155	0.10	Methanol/Water (1:1)	1:1

39	Arginine (Arg)	170	0.017	0.10	Methanol/Water (1:1)	1:1
40	Aspartic acid (Asp)	133	0.0133	0.10	Methanol/Water (1:1)	1:1
41	Glutamic acid (Glu)	147	0.0147	0.10	Methanol/Water (1:1)	1:1
42	Alanine (Ala)	89	0.0089	0.10	Methanol/Water (1:1)	1:1
	Nicotinamides					
43	Nicotinamide (N)	122	0.0122	0.10	Methanol, Ethanol	1:2
44	N-methyl-nicotinamide (Me_N)	136	0.0136	0.10	Methanol, Ethanol	1:2
45	2-chloro-nicotinamide (2Cl_N)	156	0.0156	0.10	Methanol, Ethanol	1:2
46	6-chloro-nicotinamide (6Cl_N)	156	0.0156	0.10	Methanol, Ethanol	1:2
	N-Oxides					
47	Pyrazine N-oxide (Py_NO)	96	0.0096	0.10	Methanol, Ethanol	1:2
48	Tetramethyl pyrazine N-oxide (TmPy_NO)	110	0.011	0.10	Methanol, Ethanol	1:2
49	Pyrazine-NN-oxide (Py_NNO)	112	0.0112	0.10	Methanol, Ethanol	1:2
50	Tetramethylpyrazine-NNoxide (TmPy_NNO)	168	0.0168	0.10	Methanol, Ethanol	1:2
	Bipyridine-N-oxide (BiP_NO)	172	0.0172	0.10	Methanol, Ethanol	1:2
	Various					
52	Glucose (Gl)	180	0.018	0.10	Water	1:1
53	Biuret (Biu)	103	0.0103	0.10	Methanol, Ethanol	1:2
54	Imidazolidone (Im)	86	0.0086	0.10	Methanol, Ethanol	1:1
55	Maleimide (Ma)	97	0.0097	0.10	Methanol, Ethanol	1:2
56	Resorcinol (Re)	110	0.011	0.10	Acetone/Methanol (1:1)	1:1
57	2,6-Lutidine (Lut)	107	0.010	0.10	Methanol, Ethanol	1:1
58	Quinoline (Qui)	129	0.013	0.10	Methanol, Ethanol	1:1
59	Uric acid (Ur)	168	0.0168	0.10	Methanol, Ethanol	1:2
60	1,10-Phenanthroline (Phen)	180	0.0180	0.10	Methanol, Ethanol	1:1

60 different co-formers were chosen based on the Iso-star and FIMs search and urea was put through a co-crystal screen using liquid-assisted grinding (LAG) in a drop of methanol against 60 co-formers in stoichiometric amounts. The solid resulting from each reaction was characterized using IR spectroscopy to determine if a co-crystal had formed (See Appendix C). A positive co-crystal outcome was considered when peaks from both urea and co-former were present in the grinded mixture and shifted more than 7 wavenumbers as well as the appearance of broad stretches around 2,300 and 1,800 cm^{-1} as a result of intermolecular O-H…N hydrogen bonds were taken as an indication of co-crystal formation (such hydrogen bonds would not be possible in either of the pure compounds). Four new crystal structures of co-crystals were obtained, and the experimental details are provided in Table 9.2

Table 9.2 Experimental details for new crystal structures of co-crystals obtained in this study

Co-crystals	Co-former	Solvent	Stoic (urea:co-former)	M.P. (℃)	Morphology
U1		Ethanol	2:1	155-164 ℃	Plate, colorless
U2		Ethanol	2:1	205-208 ℃	Plate, colorless
U3		Ethanol, water	2:1	97-99 ℃	Prism, colorless
U4		DMF, methanol	1:1	70-72 ℃	Plate, colorless

9.2.4 Large scale synthesis of co-crystals

Supersaturated solutions of co-crystals of urea with phthalic acid, pimelic acid, and p-nitrophenol were prepared via a solvothermal method (Figure 9.2).[1] After 10-20 min of heating at 40-50 °C, the resulting solution was cooled to 5° C. The precipitate was filtered, dried and analyzed using IR spectroscopy to confirm co-crystal formation. The PXRD pattern of bulk co-crystal and simulated single crystal pattern were compared to confirm the homogeneity of the bulk sample.

Step1 ➡ **Step 2** ➡ **Step 3** ➡ **Step 4** ➡ **Step 5**

Heating at 40~50º C	Cooling at 0~5ºC	Suction filtration	Solid co-crystal	XRPD
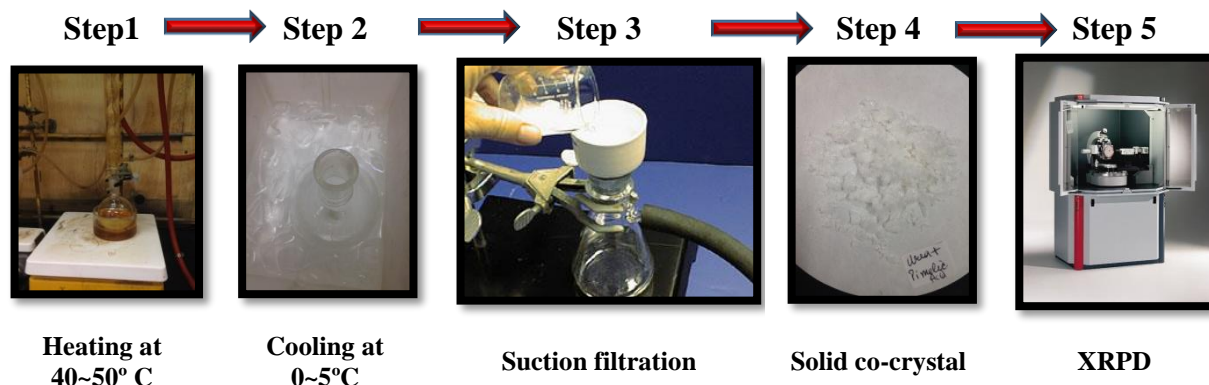
Figure 9.2 Stepwise synthesis and characterization of large scale co-crystallization

9.2.4.1 Urea phthalic acid (2:1)

Urea (1.212g, 20mmol) and phthalic acid (1.660 g, 10mmol) were dissolved in 20 ml of 95% ethanol in a 50 ml round bottom flask. The mixture was refluxed at 95ºC for 10 minutes. After 10 minutes, the resulting solution was left to slowly evaporate at room temperature. The solid phase

was harvested by vacuum filtration and dried at room temperature on a Fisherbrand filter paper for 1 hour to remove loosely bound solvent (Melting point 85-89 °C). The PXRD of the experimental co-crystal was compared with simulated pattern (Figure 9.3).

(a)   (b)
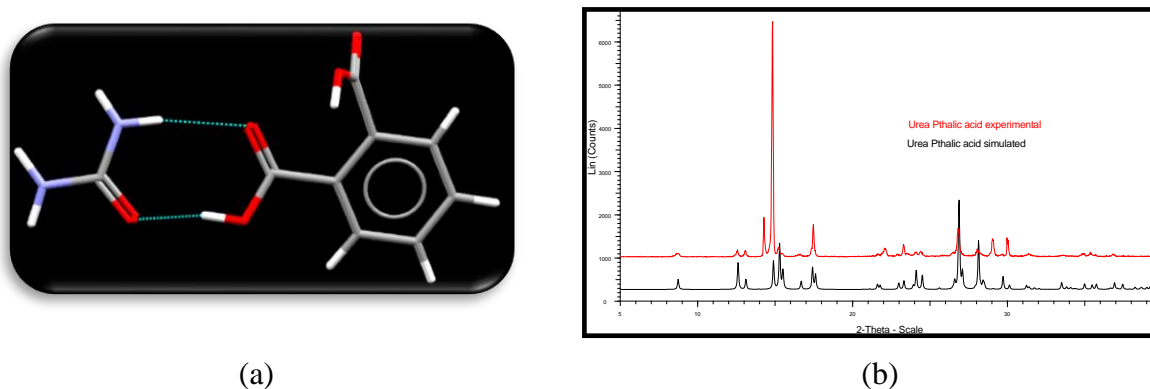
Figure 9.3 (a) Supramolecular synthon observed in the crystal structure of U: PhA, (b) Simulated and experimental XRD pattern comparison for U:PhA co-crystal

9.2.4.2 Urea Pimelic acid (2:1)

Urea (12.1g, 200 mmole) and pimelic acid (16.20 g, 100 mmole) were dissolved in 250 ml of ethanol in a 500ml round bottom flask. The mixture was refluxed at 40°C for 30 minutes, after which the resulting solution was cooled to 5° C using an ice bath. The solid was harvested by vacuum filtration and dried at room temperature on a Fisherbrand filter paper for 12 hours to remove loosely bound solvent. (Melting point 118-121 °C). The PXRD of the experimental co-crystal was compared with simulated pattern (Figure 9.4).

(a)   (b)
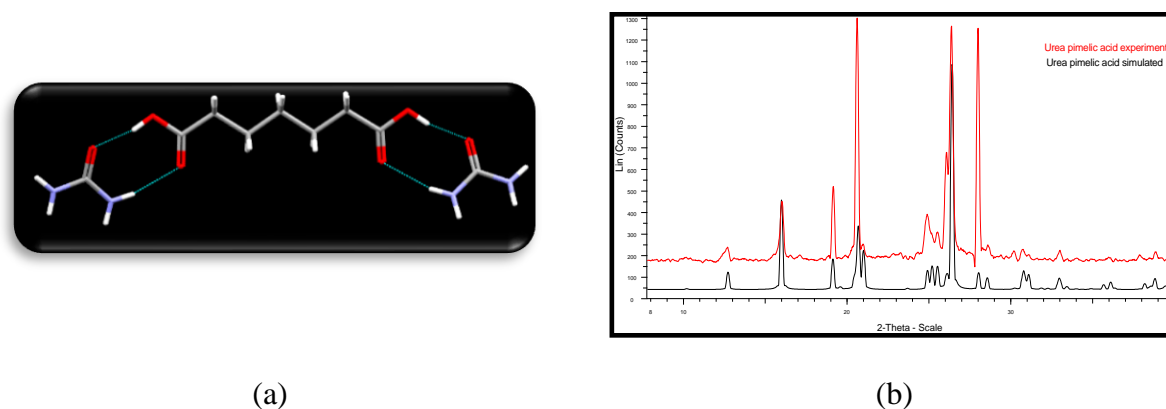
Figure 9.4 (a) Supramolecular synthon observed in the crystal structure of U:PA, (b) Simulated and experimental XRD pattern comparison for U:PA co-crystal.

9.2.4.3 Urea 4-nitrophenol (1:1)

Urea (12.10g, 200 mmole) and 4-nitrophenol (28.52 g, 200 mmole) were dissolved in 250 ml of methanol in a 500 ml round bottom flask. The mixture was refluxed at 40°C for 45 minutes, after which the resulting solution was cooled to 5° C using ice bath. The solid (yellow crystals) was harvested by vacuum filtration and dried at room temperature on a fisherbrand filter paper for 12 hours to remove loosely bound solvent. (Melting point 98-100 °C). The PXRD of the experimental co-crystal was compared with a simulated pattern (Figure 9.5).

(a)   (b)
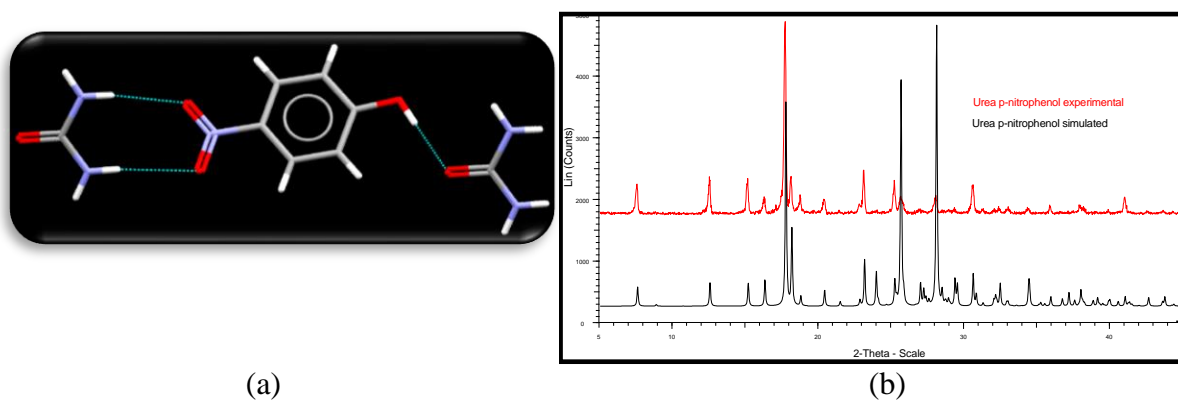
Figure 9.5 (a) Supramolecular synthon observed in the crystal structure of U:4NP, (b) Simulated and experimental XRD pattern comparison for U:4NP co-crystal.

9.2.5 Solubility studies

9.2.5.1 Qualitative analysis

The qualitative analysis was performed as follows: Control experiment: ~1.00 gm of pure urea was added to the vial and dissolved in approximately 1ml of water. Test experiment: An amount containing ~1.00 gm of urea in each co-crystal was added into each vial and the water was added until the co-crystal was completely dissolved. The vials where more than 1 ml was needed to dissolve the solid were chosen as good candidates for initial solubility studies.

9.2.5.2 Quantitative analysis

The solubility studies of pure urea and **U:PA**, **U:4NP** were done using a gravimetric method. Pure urea: A saturated solution of urea was prepared by dissolving a known amount of urea in 10ml of distilled water in a 25ml beaker in trial 1. The beaker was sealed with para-film and stirred in a water bath to maintain constant temperature (22-23°C) for 4 hours, 24 hours, 10 days and 1 month. After stirring for a given time, the remaining solid was filtered from the solution and dried overnight. The weight of the leftover solid was recorded. The average concentration of the filtrate was determined by the mass difference of urea in 10ml of the water and leftover solid after stirring. Three trials were performed at each 4 hours, 24 hours, 10 days and 1 month respectively. The same procedure was repeated for **U:PA** and **U:4NP** co-crystals.

9.2.6 Hygroscopicity studies

Hygroscopicity studies were performed at 43% (K_2CO_3) and 85% (KCl) humidity conditions using humidity chamber (Figure 9.6). A supersaturated solution was made for each solution and was maintained saturated for a period of 1 month. Urea and three new solid forms were kept inside the humidity chamber and the samples were analyzed every week using TGA and DSC analysis and the photographs were taken after every week to qualitatively assess the hygroscopicity effects.
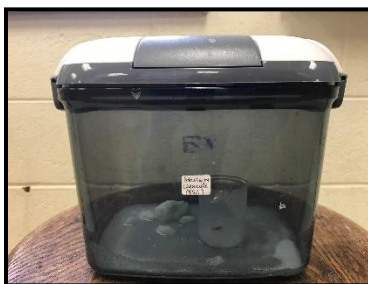
Figure 9.6 Humidity chamber used for hygroscopicity studies

9.2.7 Powder X-ray diffraction (PXRD)

The PXRD spectra were collected on a Bruker X-ray powder diffractometer using CuK (= 1.54059A)° radiation obtained at 30 kV and 15 mA. The scans were run from 5.0° to 30.0° 2, increasing at a step size of 0.05° 2 with a counting time of 2 s for each step.

9.2.8 Differential scanning calorimetry (DSC)

The melting points were measured with a Mettler Toledo DSC 822e differential scanning calorimeter (Greifensee, Switzerland). Accurately weighed samples (∼3 mg) were prepared in a covered aluminum crucible having pierced lids to allow escape of volatiles. The sensors and samples were under nitrogen purge during the experiments. The temperature calibration was carried out using the melting point of highly pure indium in the medium temperature range. Heating rate of 5 ◦C/min was selected.

9.2.9 Thermogravimetric analysis (TGA)

TGA were performed on a Mettler-Toledo TGA/SDTA 851e instrument. Approximately 2 mg sample was heated from 25 to 300˚ C at 10˚ C/min under nitrogen purge.

9.2.10 Qualitative plant effect test

Three tomato plants were treated with exactly the same amount of **Urea, U:PA** and **U:4NP** along with potassium and phosphorus at three different time intervals.

9.3 Results

9.3.1 Iso-star search and FIMs results

The FIMs results for urea are shown in Figure 9.7 displaying all potential hydrogen-bond donors and acceptors.
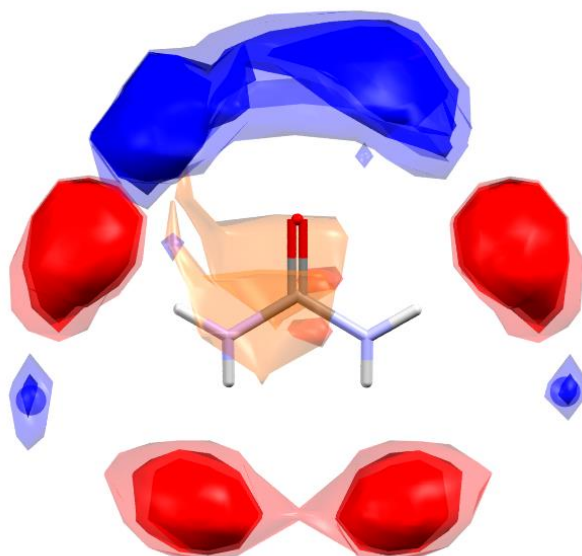
Figure 9.7 FIMs for urea showing all conventional donors and acceptors fully available for hydrogen-bond interactions

9.3.2 Grinding experiment results

The IR analysis from the grinding experiments is shown in Table 9.3. The detailed description of how the ground mixture was compared to the pure components is shown in Figure 9.8.

Table 9.3 Experimental outcome of grinding experiments for co-crystallizations of urea with co-formers. Co-formers with subscripts are already reported in the literature.

	Urea co-crystal	Outcome		Urea co-crystal	Outcome		Urea co-crystal	Outcome
	Aliphatic acids			**Aromatic acids**			**Nitro substituents**	
1	U:CA	✓	**14**	U:3HBA	✓	**28**	U:2NP	✗
2	U:CmA	✓	**15**	U:4HBA	✓[38]	**29**	U:3NP	✓
3	U:MA	✓[39]	**16**	U:3ABA	✗	**30**	U:4NP	✓[40-41]
4	U:SA	✓[42,43, 44]	**17**	U:4ABA	✓[45]	**31**	U:1,2DNB	✓
5	U:FA	✓[46-47, 43]	**18**	U:PtA	✓[45]	**32**	U:1,3DNB	✓
6	U:GA	✓[48, 49]	**19**	U:Py2_CA	✓	**33**	U:1,4DNB	✓
7	U:AA	✓[50]	**20**	U:Py2,3_DCA	✓[45]	**34**	U:2NBA	✓
8	U:PA	✓[51]	**21**	U:Py2,5_DCA	✗	**35**	U:3NBA	✓[52]
9	U:SbA	✓	**22**	U:Py2,6_DCA	✓[45]	**36**	U:4NBA	✓
10	U:TA	✓[53-54]	**23**	U:Py3,5_DCA	✗			
11	U:MeA	✓[44]	**24**	U:HipA	✗			
12	U:AzA	✓	**25**	U:2,5DHBA	✓[38]			
13	U:SeA	✓	**26**	U:3,4DHBA	✓			
			27	U:3,5DHBA	✓[38]			
	Amino acids			**Nicotinamides**			**N-oxides**	
37	U:Lys	✓	**43**	U:N	✓[55]	**47**	U:Py_NO	✓
38	U:His	✗	**44**	U:Me_N	✓	**48**	U:TmPy_NO	✓
39	U:Arg	✓	**45**	U:2Cl_N	✗	**49**	U:Py_NNO	✓

40	U:Asp	✓	46	U:6Cl_N	✗	50	U:TmPy_NNO	✓
41	U:Ala	✗				51	U:BiP_NO	✓
42	U:Glu	✓						
Various								
52	U:Gl	✓[56]	57	U:Lut	✓[57]			
53	U:Biu	✗	58	U:Qui	✓[58]			
54	U:Im	✓[59]	59	U:Ur	✗			
55	U:Ma	✓	60	U:Phen	✓[60]			
56	U:Re	✓[61]						
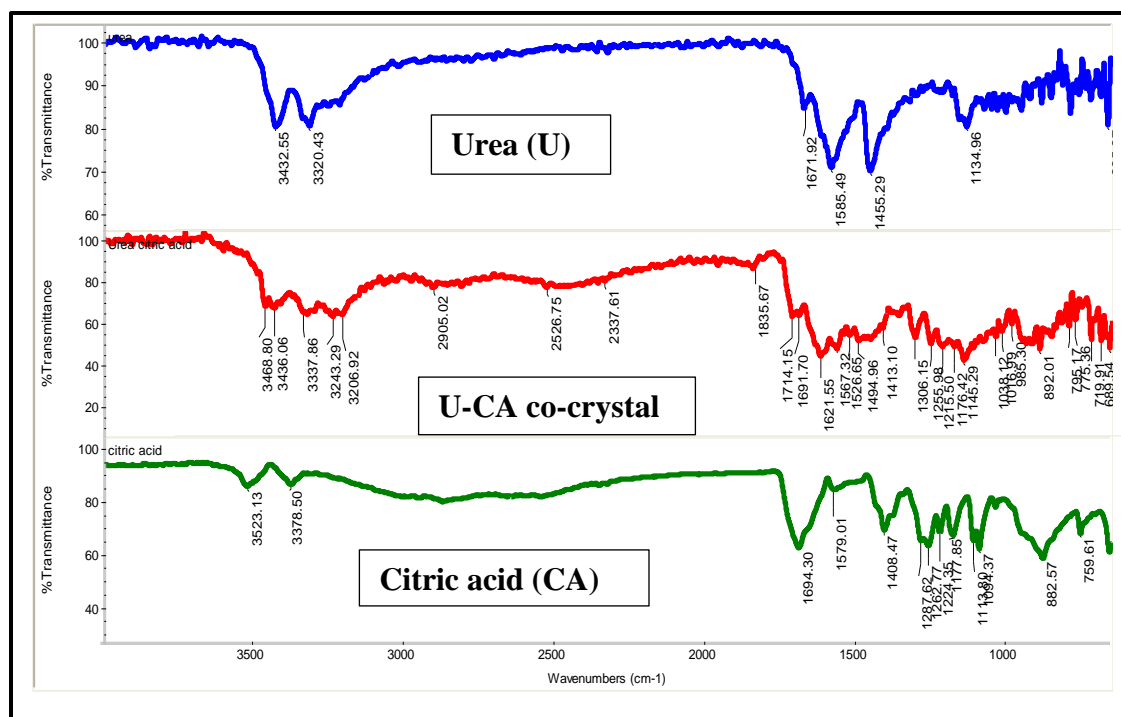
Figure 9.8 Examples of IR spectra for a successful co-crystallization: U(blue), U:CA (red) and CA (green)

9.3.3 Crystal structures of co-crystals

Four new crystal structures (**U1, U2, U3** and **U4**) were obtained and different synthons were observed in each co-crystal (Figure 9.9). The expanded crystal packing for each crystal structure is shown in Figure 9.10.

(a)   **(b)**   **(c)**   **(d)**
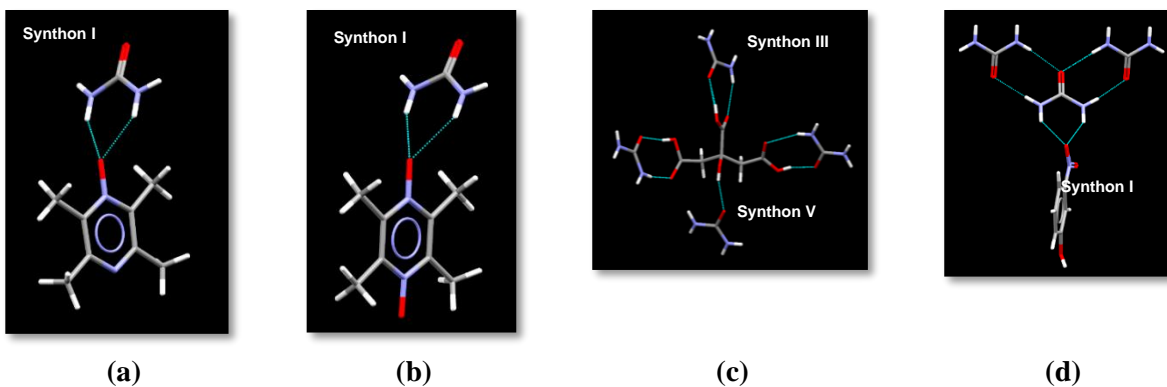
Figure 9.9 Synthons observed in the crystal structures of (a) **U1**, (b) **U2**, (c) **U3** and (d) **U4**

(a)   **(b)**

(c)   **(d)**
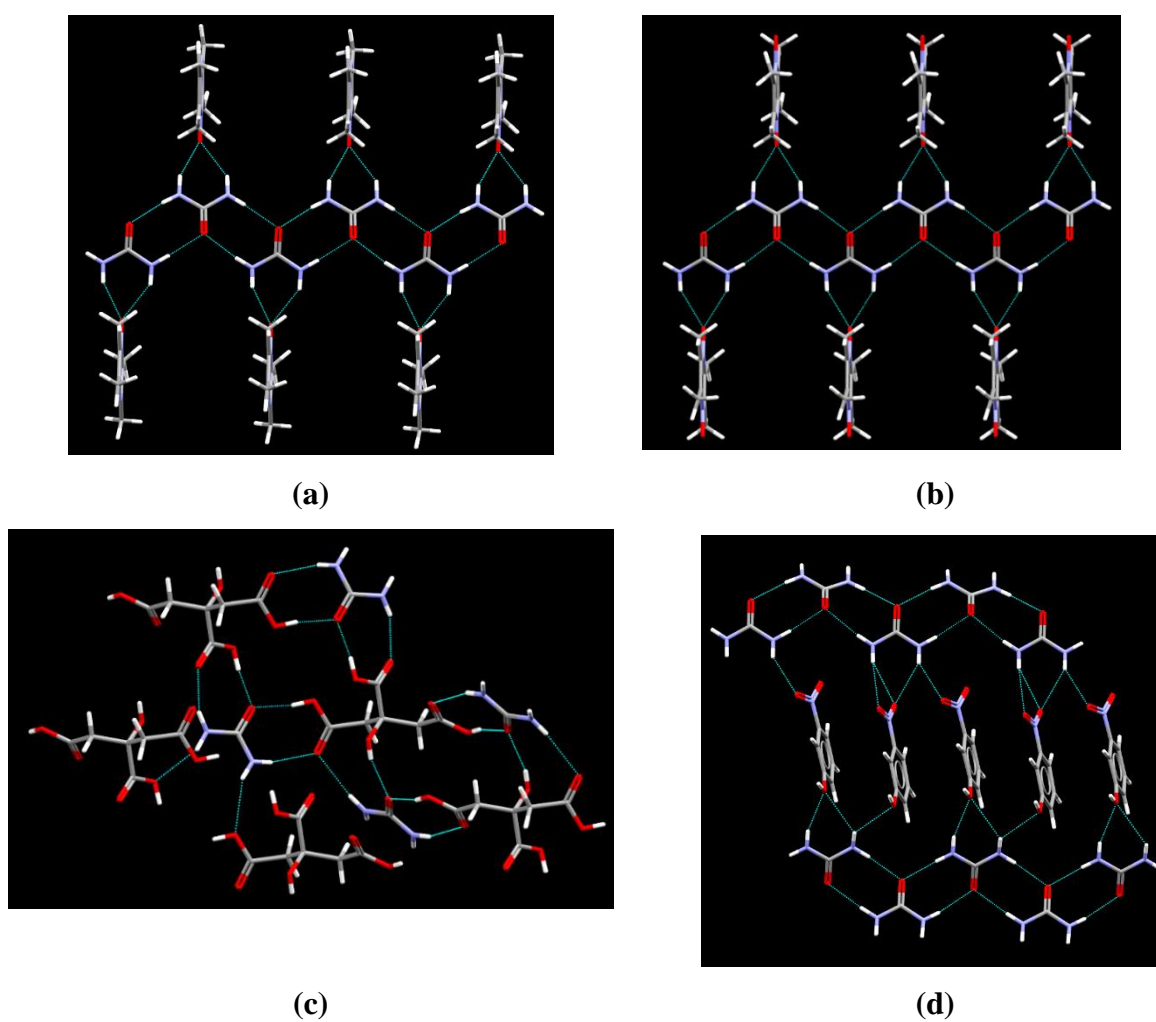
Figure 9.10 Packing in the crystal structures of **(a) U1, (b) U2, (c) U3** and **(d) U4** in the unit cell

9.3.4 Solubility studies

9.3.4.1 Qualitative analysis

The results of qualitative solubility test are displayed in Figure 9.11.

(a)

Figure 9.11 Qualitative analysis of urea and urea-co-former co-crystals in distilled water

9.3.4.2 Quantitative analysis

The quantitative solubility analysis of urea and three co-crystals of urea were performed using a gravimetric method. The homogeneity test was performed using PXRD and the result of each co-crystal is shown in Figure 9.12, 9.13 and 9.14.
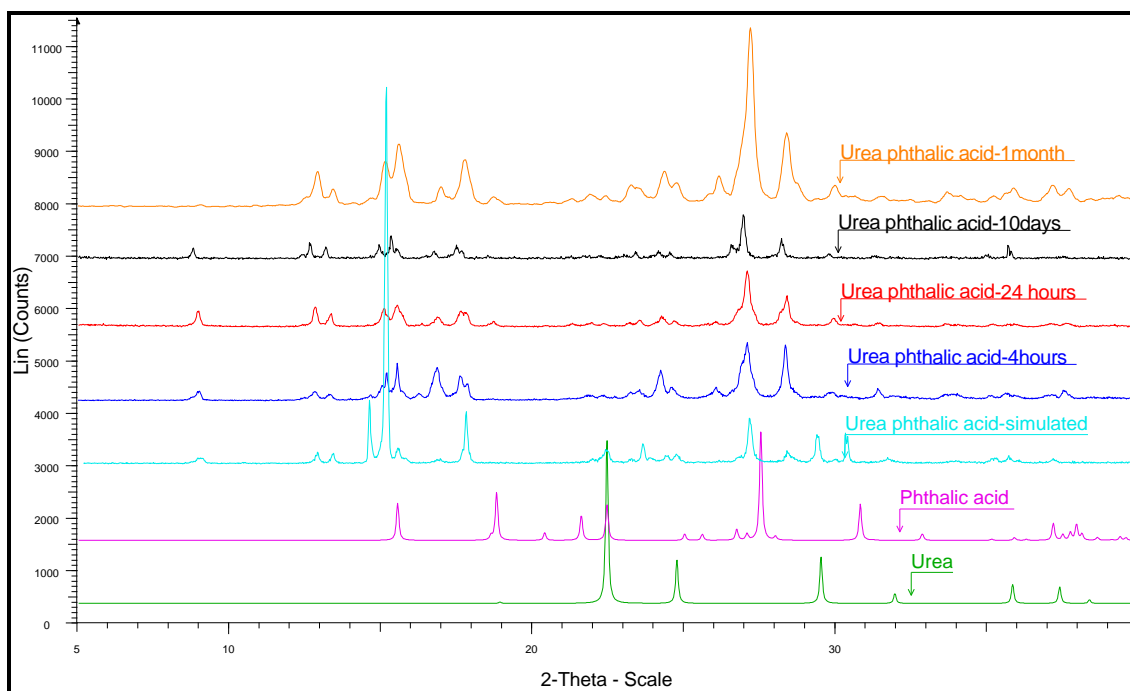
Figure 9.12 PXRD analysis of U:PhA after 4 hours, 24 hours 10 days and 1month
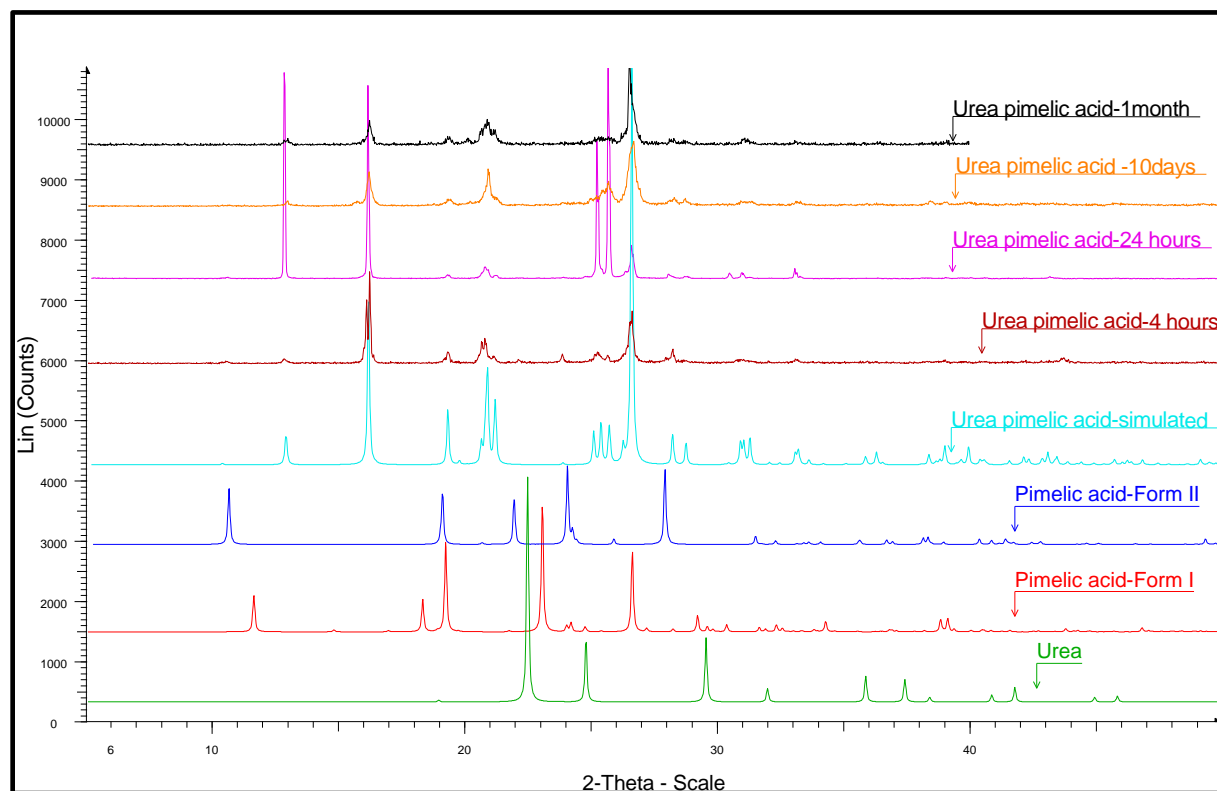
Figure 9.13 PXRD analysis of U:PA after 4 hours, 24 hours 10 days and 1month
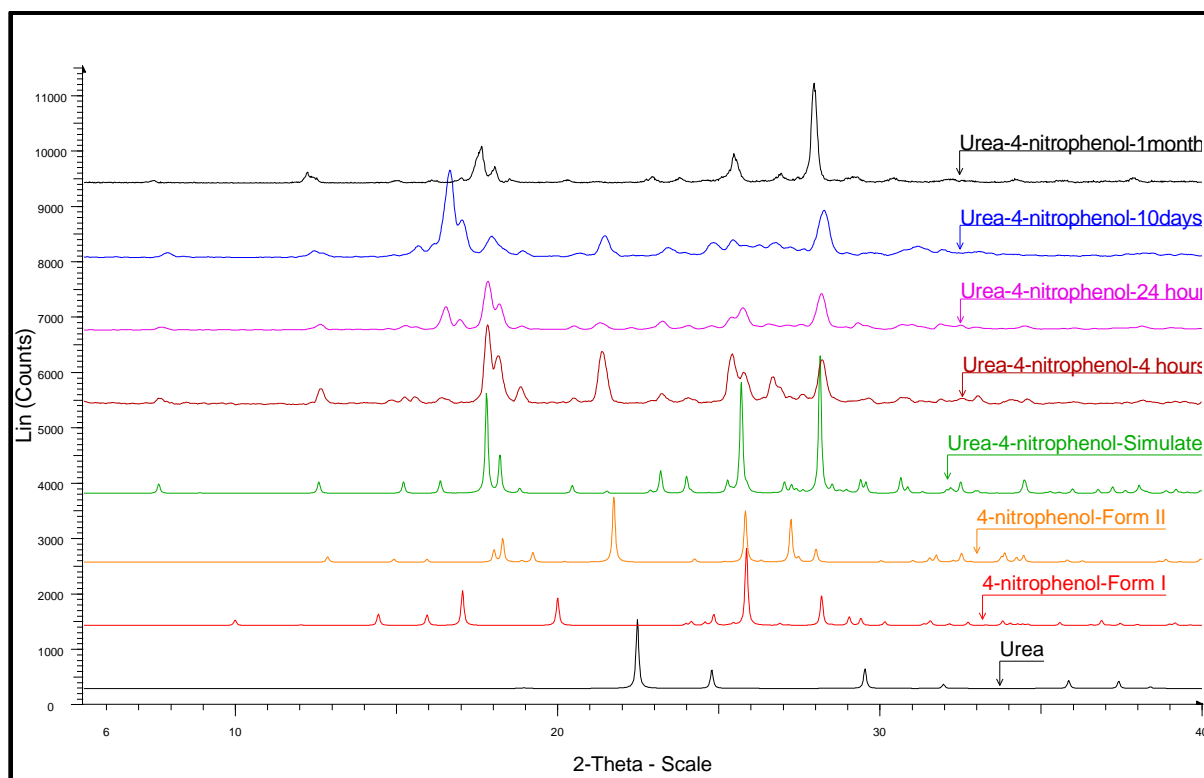
Figure 9.14 PXRD analysis of U:4NP after 4 hours, 24 hours 10 days and 1month

The solubility was measured after 4 hours, 24 hours, 10 days and 1 month and results are shown in Table 9.4 and Figure 9.15.

Table 9.4 Solubility of pure urea, **U:PhA**, **U:PA** and **U:4NP** after 4 hours, 24 hours, 10days and 1 month.

	4 hours	24 hours	10 days	1month
Pure urea	17.79±0.02	17.25±0.20	17.22±0.24	17.18±0.3
Concentration of urea in **U:PhA** co-crystal	2.87± 0.08	3.28 ± 0.6	3.21±0.4	3.54±0.8
Concentration of urea in **U:PA** co-crystal	0.84± 0.26	0.71±0.04	1.01±0.3	0.76±0.1
Concentration of urea in **U:4NP** cocrystal	0.28±0.04	0.2±0.09	0.30±0.03	0.27±0.07

245
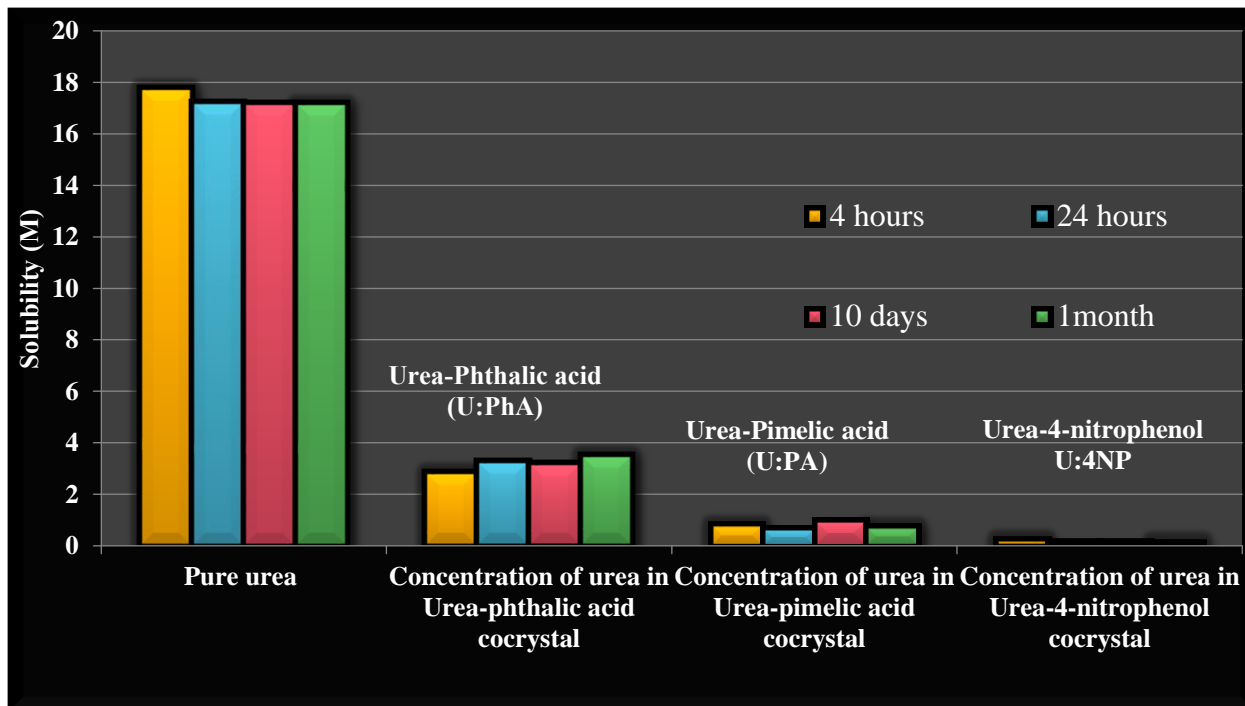
Figure 9.15 Solubility profile of urea co-crystals for 4 hours, 24 hours, 10 days and 1 month

9.3.4.3 Qualitative hygroscopicity studies

The qualitative hygroscopicity test was performed in a humidity chamber at 43% and 85% relative humidity on urea and two co-crystals (**U:PA** and **U:4NP**) (Figure 9.16). The pictures were taken after each time interval (0 day, 1 day and 1 month).
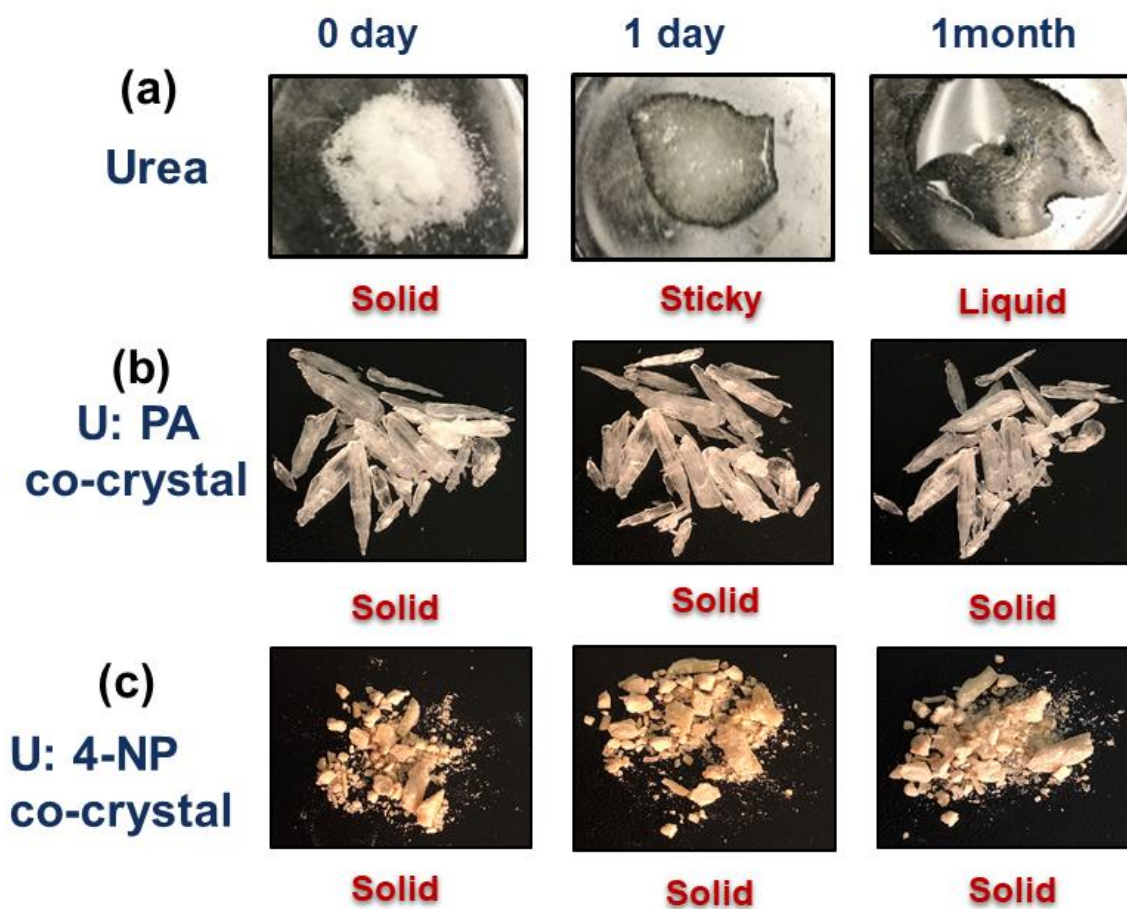
Figure 9.16 Response to 85% humidity of (a) urea, (b) **U:PA** co-crystal and (c) **U:4NP** co-crystal

9.3.4.4 Quantitative stability studies

In order to quantitively confirm that each co-crystal remained solid after exposure to high humidity, the TGA and DSC analysis were performed on urea (Figure 9.17a) and two new co-crystals (Figure 9.17b).
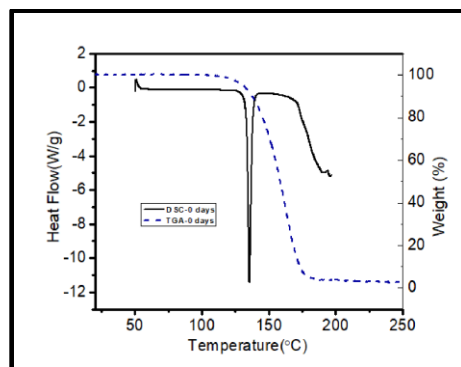
(a)
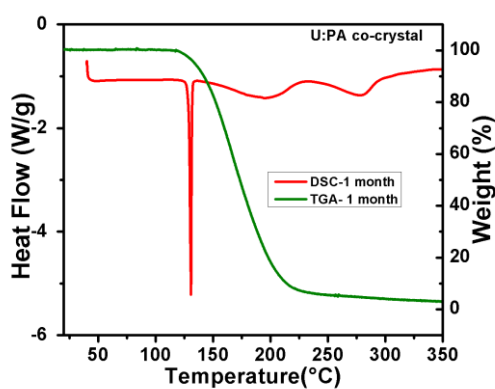
(b)
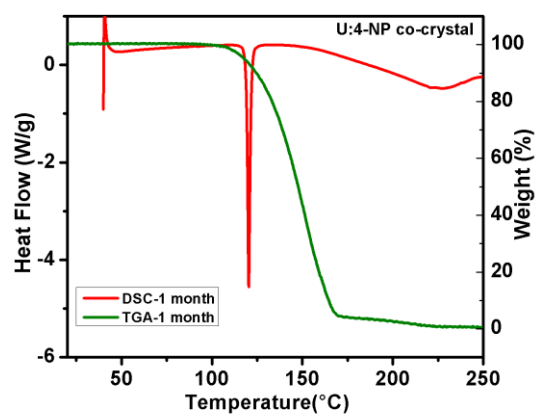
(c)

Figure 9.17 DSC-TGA profile of (a) **Urea**, (b) **U:PA** and (c) **U:4NP** co-crystal after 1month exposure to 80% humidity conditions.

9.3.4.5 Plant study test

The results of applying the pure urea and two co-crystals on tomato plants are shown in Figure 9.18.

Figure 9.18 Urea and the two new co-crystals tested on the tomato plant

9.4 Discussion

9.4.1 FIMs

The structural landscape occupied by urea was explored using full interaction maps (FIMs). As expected, a large number of short-contacts were found in close proximity of (i) both lone pairs on the carbonyl oxygen atom and (ii) the four N-H protons, leading to four avenues for synthesizing co-crystals of urea, Scheme 9.2.

9.4.2 Synthon search in CSD

The FIM provides an overview of short contacts to urea in the solid state, but to add more granularity to the map, an analysis of relevant data from the CSD[30-32] was conducted. The type of interactions that urea tends to engage in can be described by five different synthons, Scheme 9.5

(a) Synthon I   (b) Synthon II   (c) Synthon III

(d) Synthon IV   (e) Synthon V
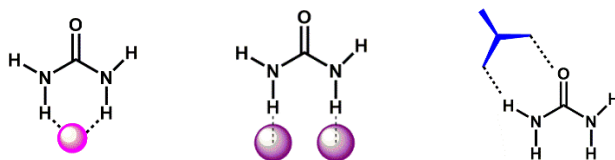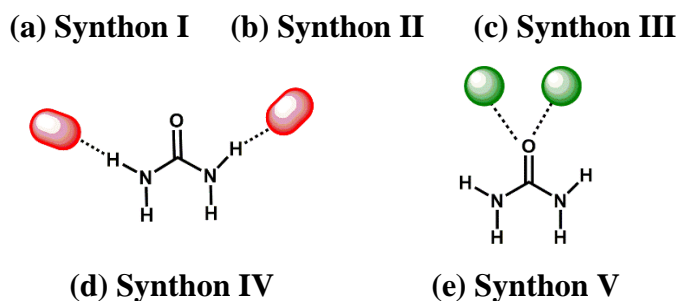
Scheme 9.5 Five likely synthons in co-crystals of urea

Synthon I: In the first category of co-crystals, urea forms robust $R_2^2(8)$ self-complementary homosynthons and the suitable co-former (mostly hydrogen bond acceptors) binds to two N-H groups of urea via bifurcated hydrogen bond interactions forming synthon **I**. The following co-formers forms synthon I with urea-2,2-bipyridine (refcode AMILEN)[62] urea-1,10-phenanthroline (refcode AMILUD)[62], urea-2,9-dimethyl-1,10-phenanthroline (refcode AMIMAK)[62], urea-picoline N-oxide (refcode BAQZOI),[63] urea-4,4'-dicyanobiphenyl (refcode BAFSOQ),[64] and urea-4,4'-bipyridine N,N-dioxide (refcode BAFTEH)[64], urea-2,6- or 3,5-lutidine (refcodes JOLCIY, JOLCOE, and LUTDUR)[65], urea-dihydroxyalkane (refcodes OTIYUM, OTIZAT, OTIZEX, OTIZIB, OTIZOH, and OTIZUN)[66].

Synthon II: In the crystal structures of urea of with 4,4'-Dinitobiphenyl, and 18-Crown-6 pentakis (CRWNUR), the cocrystals forms synthon **II**. The crystal structure of 2,7-dimethyl-1,8-naphthyridine and urea (IDELIO)[67] and 4-aminobenzoic acid with urea (NUHYEU) [68] forms synthon **II**.

Synthon III: In the third category of co-crystals, a suitable ditopic coformer with a hydrogen bond donor and acceptor site binds to urea via robust $R_2^2(8)$ heterosynthon breaking the self-complementary hydrogen bond interaction observed in category I. Synthon **III** is observed in the crystal structures of urea-adipic acid (refcode ERIWUY), urea-4,6-dimethyIpyrimidin-2-one (refcode JELSEY)[69], Urea-tartaric acid (refcode NEHPIZ), Urea-5-nitrosalicyclic acid (refcode NUHXUJ), Urea-pyrazine-2,3-dicarboxylic acid (refcode NUHYOE)[68],Urea-o-phthalic acid (refcodes NUHYIY, NUHYUK)[68], urea-2-Methyl-5-nitraminotetrazole (refcode OTAZUF), urea-fumaric acid (refcode TIPWIY), urea-glutaric acid (refcode TONGUS), urea-succinic acid (refcode UNIRIT)[42] , urea-oxalic acid (refcode UROXAL), urea-lenalidomide (XODPOX).[71]

Synthon IV: In the fourth category of co-crystals, urea forms self-complementary bifurcated hydrogen bond with itself and the suitable coformer (symmetric or asymmetric acceptor) binds to the NH group of urea leading to 1D or 2D architectures. Synthon IV is observed in the crystal structures of urea- dinitrile of the formula $(CH2)_nCN$ when n = 3-5, 6,8,10, and 12 (refcodes HEZSIU, HEZSOU, HEZTAH, and HEZTEL), urea:4,4′-bipyridine (refcode KOPJUW)[72], urea-1,4-dioxane (refcode SIDJEW)and urea-morpholine (refcode SIDJIA).[73]

Synthon V: In the fifth category of co-crystals, no self-complementary hydrogen bond homosynthons are observed. The NH, hydrogen bond donor group of urea binds to suitable acceptor whereas the carbonyl group of urea binds to suitable co-former with hydrogen bond donor groups. Synthon **V** is observed in the crystal structures of urea-estradiol, urea-2-Amino-4-methylthio-5-cyano-6(1H)-pyrimidinethione (refcode GIGWOI), urea-12-tridecanedione (refcode MISNOR)[74], and urea-2,6-bis(2-benzimidazolyl)pyridine (PERTOW, PERTUC).[75]

9.4.3 New crystal structures

In our study, 49 of the 60 experiments (82%), the IR analysis showed prominent new features or significant changes to the spectra of one or both reactants, indicative of the formation of a co-crystal, Table 9.3 and Table 9.5.

Table 9.5 Co-crystallization summary and synthons expected in the co-former used (summary of each synthon is shown below)

Co-formers	# of co-formers studied	SAG success rate	Synthons expected
Aliphatic acids	13	13/13	III, V
Aromatic acids	14	10/14	II, III, V
Nitro-substituents	9	8/9	I, III
Amino acids	6	4/6	III, V
Nicotinamides	4	2/4	II, III
N-oxides	5	5/5	I
Various	9	7/9	I-V
Overall	60	49/60	
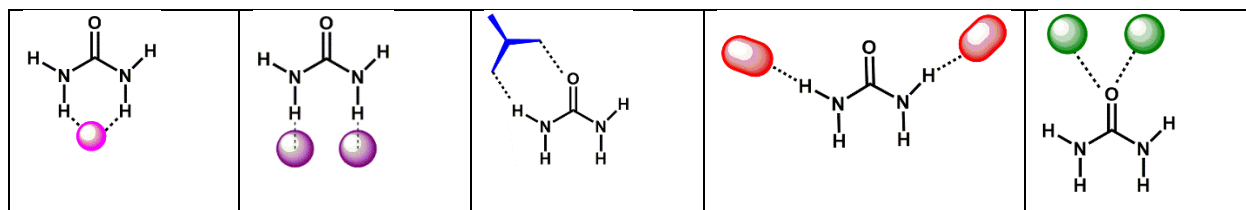
(a) Synthon I	(b) Synthon II	(c) Synthon III	(d) Synthon IV	(e) Synthon V

In our study, four crystals suitable for single-crystal XRD were obtained and analyzed using single crystal X-ray diffraction in order to study the details of binding interactions between urea and various co-formers, Figure 9.9 and Figure 9.10.

Two crystal structures were obtained for urea with n-oxides (**U1** and **U2**). In both cases, urea forms bifurcated hydrogen bonds with the oxygen of nitrogen-oxides (Figure 9.8a and b). The urea molecules form a linear ribbon that is constituted with dimer synthons arranged in a zig-zag manner with syn N-H···O hydrogen bonds as described in synthon I. The Iso-star prediction of two N-H groups of *trans*-urea tend to form bifurcated hydrogen bond with terminal oxygen atom of the co-former agrees with the obtained results in case of N-oxides. In both cases, the bifurcated hydrogen bonding was observed (Figure 9.9).

In the crystal structure of urea-citric acid (**U3**), each asymmetric unit contains one molecule of citric acid and two crystallographically unique molecules of urea. The primary supramolecular synthons in this structure comprises three $R_2^2(8)$ hydrogen-bonded synthon III formed using O-H (acid)···O(urea) and N-H(urea) ···O(acid) hydrogen bonds resulting from the interactions between dicarboxylic acid and urea molecules (Figure 9.8c). All of the acidic hydrogens are engaged in hydrogen bonding. The Iso-star prediction agrees well with the urea-citric acid structure obtained. Both lone pairs on the carbonyl oxygen participate in hydrogen bond with O-H functional groups of citric acid. Also, the carbonyl oxygen of urea forms strong hydrogen-bonding with the O-H hydroxyl group of citric acid forming synthon V (Figure 9.9c).

In the crystal structure of urea:m-nitrophenol, synthon I was observed (Figure 9.9d). The nitro group of m-nitrophenol forms bifurcated hydrogen-bond with the NH groups of urea. The urea molecules form a linear ribbon composed of dimer synthons arranged in a zig-zag manner with syn N-H···O hydrogen bonds as described in synthon I.

9.4.4 Solubility studies

Three of the co-crystals identified in the qualitative solubility screen test were selected for further solubility studies; urea:phthalic acid,[76] (**U:PhA**) urea: pimelic acid,[51] (**U:PA**) and urea:4-nitrophenol[41, 77-78] (**U:4NP**). There was 6-fold decrease in the concentration of urea in the urea:

phthalic acid co-crystal as compared to the pure urea in water. On the other hand, the concentration of phthalic acid in the urea: phthalic acid cocrystal is considerably increased as compared to pure phthalic acid. The solubility of urea in **U:PA** decreased to 0.76 (±0.10) M and in **U:4NP** decreased to 0.27 (±0.07) M after 1 month. There is 23-fold decrease in **U:PA** and 64-fold decrease **U:4NP** co-crystal, Table 9.4.

9.4.5 Hygroscopicity studies

Another major concern with urea is its sensitivity to moisture, and in order to examine if these new solid formulations of urea, **U:PhA**, **U:PA** and **U:4NP**, are better equipped to tolerate moisture, humidity studies were performed. Urea becomes a transparent liquid in less than 24 hours at 85% humidity, whereas **U:PhA** remained solid for 10 days and then gained moisture and became sticky. Both **U:PA** and **U:4NP** remain solid for over one month, Figure 9.16. An additional thermal analysis (DSC and TGA) of the two co-crystals (**U:PA** and **U:4NP**) after having been exposed to 85% humidity for a month demonstrate that there are no thermodynamic events in either case prior to their respective melting points (Figure 9.17), indicating that the two co-crystals are very tolerant of prolonged exposure to high humid conditions.

9.4.6 Qualitative plant testing

Three tomato plants were treated with exactly same amount of each co-crystal; urea**, U:PA** and **U:4NP** along with potassium and phosphorus at three different time intervals (Figure 9.18). All three plants were treated each with different nutrient, **U:4NP** for one, **U:PA** for second and **U:4NP** and the it was found that all plants had same growth and no visible changes were observed to the plants comparing the pure urea and the co-crystals indicating that the new solid forms doesn't have any harmful effect on the plant shelf life or growth.

9.4.7 Conclusions

In conclusion, FIMs combined with the CSD analysis, guided the choice of sixty co-formers for co-crystallization resulting in 49 out of 60 positive results and four new co-crystals. Furthermore, solubility and stability studies of **U:PA** and **U:4NP** indicates that the two co-crystals were able to

reduce the solubility of urea to 23-fold (in **U:PA**) and 64-fold (in **U:4NP**) respectively as well as making the new solid forms of urea more tolerant toward 85% humidity conditions as confirmed using PXRD, DSC and TGA. These results indicate that systematic changes to the molecular structure of urea can be made by altering the way it packs in the crystal lattice via co-crystals. The goal of reducing solubility and lowering hygroscopicity of the parent compound was achieved, which, in turn, offers new opportunities for a highly-efficient slow-release fertilizer with limited hygroscopicity thereby reducing problems of transport, handling, and storage of urea. And the physicochemical properties such as solubility, and stability of urea can be tuned in a predictable and desired manner to optimize the performance and benefits of urea as a plant nutrient. This technique can also be used to alter solubility problems associated with agrochemicals such as herbicides, pesticides, insecticides and hence, positively impact the environment.
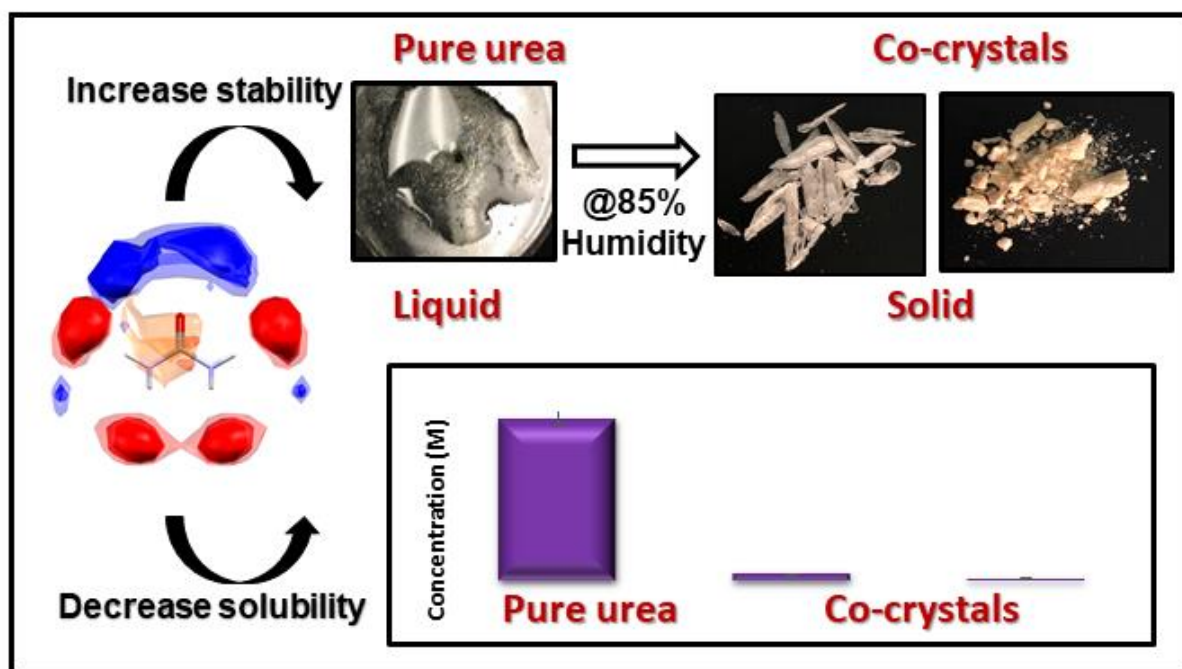
9.5 Conclusions

1.   Lee, F.-M.; Lahti, L. E., *Journal of Chemical & Engineering Data* **1972,** *17* (3), 304-306.

2.   Goulding, K.; Jarvis, S.; Whitmore, A., *Philosophical Transactions of the Royal Society B: Biological Sciences* **2008,** *363* (1491), 667-680.

3.   Ruark, M., **2012**, 1-33.

4.   Shi, J.; Li, J.; Zhou, W.; Zhang, D., *Frontiers of Forestry in China* **2007,** *2* (1), 104.

5.   Benzian, B.; Freeman, S. C. R.; Mitchell, J. D. D., *Plant and Soil* **1971,** *35* (1), 517-532.

6.   Peng, P. H.; Ernst, W. R.; Bridger, G. L.; Hartley, E. M., *Industrial & Engineering Chemistry Process Design and Development* **1979,** *18* (3), 453-458.

7.   Zhao B, D. S., Zhang J, Liu P; 2013; PLoS ONE 8(8): e70569. .

8.   Landis, T. D. D., R. Kasten., **2009**, 5-12.

9.   Kottegoda, N.; Sandaruwan, C.; Priyadarshana, G.; Siriwardhana, A.; Rathnayake, U. A.; Berugoda Arachchige, D. M.; Kumarasinghe, A. R.; Dahanayake, D.; Karunaratne, V.; Amaratunga, G. A. J., *ACS Nano* **2017,** *11* (2), 1214-1221.

10.   Zaman, M.; Saggar, S.; Blennerhassett, J. D.; Singh, J., *Soil Biology and Biochemistry* **2009,** *41* (6), 1270-1280.

11.   A.D. Blaylock, J. K., R.D. Dowbenko *West. Nutr. Manag.* **(2005**, 8-13.

12.   Aakeröy, C. B.; Chopade, P. D., Cocrystals: Synthesis, Structure, and Applications. In *Supramolecular Chemistry*, John Wiley & Sons, Ltd: 2012.

13.   Basavoju, S.; Boström, D.; Velaga, S. P., *Pharmaceutical Research* **2008,** *25* (3), 530-541.

14.   Braga, D.; Dichiarante, E.; Palladino, G.; Grepioni, F.; Chierotti, M. R.; Gobetto, R.; Pellegrino, L., *CrystEngComm* **2010,** *12* (11), 3534-3536.

15.   Liantonio, R.; Metrangolo, P.; Pilati, T.; Resnati, G.; Stevenazzi, A., *Crystal Growth & Design* **2003,** *3* (5), 799-803.

16.   Orola, L.; Veidis, M. V.; Mutikainen, I.; Sarcevica, I., *Crystal Growth & Design* **2011,** *11* (9), 4009-4016.

17.   Olenik, B.; Boese, R.; Sustmann, R., *Crystal Growth & Design* **2003,** *3* (2), 175-181.

18.   Zhang, Y.-N.; Yin, H.-M.; Zhang, Y.; Zhang, D.-J.; Su, X.; Kuang, H.-X., *Journal of Molecular Structure* **2017,** *1130*, 199-207.

19.   Shevchenko, A.; Miroshnyk, I.; Pietilä, L.-O.; Haarala, J.; Salmia, J.; Sinervo, K.; Mirza, S.; van Veen, B.; Kolehmainen, E.; Nonappa; Yliruusi, J., *Crystal Growth & Design* **2013,** *13* (11), 4877-4884.

20.   Jung, S.; Lee, J.; Kim, I. W., *Journal of Crystal Growth* **2013,** *373*, 59-63.

21.   *Nanomedicine* **2017,** *12* (6), 615-638.

22. Reddy, L. S.; Bethune, S. J.; Kampf, J. W.; Rodríguez-Hornedo, N., *Crystal Growth & Design* **2009,** *9* (1), 378-385.

23. Trask, A. V., *Molecular Pharmaceutics* **2007,** *4* (3), 301-309.

24. Schultheiss, N.; Newman, A., *Crystal Growth & Design* **2009,** *9* (6), 2950-2967.

25. Basavoju, S.; Boström, D.; Velaga, S. P., *Crystal Growth & Design* **2006,** *6* (12), 2699-2708.

26. Takata, N.; Shiraki, K.; Takano, R.; Hayashi, Y.; Terada, K., *Crystal Growth & Design* **2008,** *8* (8), 3032-3037.

27. Bučar, D.-K.; Henry, R. F.; Lou, X.; Duerst, R. W.; MacGillivray, L. R.; Zhang, G. G. Z., *Crystal Growth & Design* **2009,** *9* (4), 1932-1943.

28. de Melo, A. C. C.; de Amorim, I. F.; Cirqueira, M. d. L.; Martins, F. T., *Crystal Growth & Design* **2013,** *13* (4), 1558-1569.

29. Sarkar, A.; Rohani, S., *Journal of Pharmaceutical and Biomedical Analysis* **2015,** *110*, 93-99.

30. Landenberger, K. B.; Bolton, O.; Matzger, A. J., *Journal of the American Chemical Society* **2015,** *137* (15), 5074-5079.

31. Aakeroy, C. B.; Wijethunga, T. K.; Benton, J.; Desper, J., *Chemical Communications* **2015,** *51* (12), 2425-2428.

32. D'silva, E. D.; Podagatlapalli, G. K.; Rao, S. V.; Rao, D. N.; Dharmaprakash, S. M., *Crystal Growth & Design* **2011,** *11* (12), 5362-5369.

33. Bosshard, C.; Pan, F.; Wong, M. S.; Manetta, S.; Spreiter, R.; Cai, C.; Günter, P.; Gramlich, V., *Chemical Physics* **1999,** *245* (1), 377-394.

34. Vijayalakshmi, S.; Kalyanaraman, S., *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **2014,** *120* (Supplement C), 14-18.

35. Gryl, M.; Seidler, T.; Stadnicka, K.; Matulkova, I.; Nemec, I.; Tesarova, N.; Nemec, P., *CrystEngComm* **2014,** *16* (26), 5765-5768.

36. Nauha, E.; Nissinen, M., *Journal of Molecular Structure* **2011,** *1006* (1–3), 566-569.

37. Nauha, E.; Kolehmainen, E.; Nissinen, M., *CrystEngComm* **2011,** *13* (21), 6531-6537.

38. Przybyłek, M.; Ziółkowska, D.; Kobierski, M.; Mroczyńska, K.; Cysewski, P., *Journal of Crystal Growth* **2016,** *433* (Supplement C), 128-138.

39. Bandoli, G.; Clemente, D. A.; Brustolon, M.; Corvaja, C.; Pinzino, C.; Colligiani, A., *Molecular Physics* **1980,** *39* (5), 1145-1152.

40. Muthuraja, P.; Sethuram, M.; Sethu Raman, M.; Dhandapani, M.; Amirthaganesan, G., *Journal of Molecular Structure* **2013,** *1053*, 5-14.

41. Rathika, A.; Ganapathi Raman, R., *Optik - International Journal for Light and Electron Optics* **2014,** *125* (12), 2978-2982.

42. Alhalaweh, A.; George, S.; Boström, D.; Velaga, S. P., *Crystal Growth & Design* **2010,** *10* (11), 4847-4855.

43. Jones, A. O. F.; Leech, C. K.; McIntyre, G. J.; Wilson, C. C.; Thomas, L. H., *CrystEngComm* **2014,** *16* (35), 8177-8184.

44. Videnova-Adrabińska, V., *Journal of Molecular Structure* **1996,** *374* (1), 199-222.

45. Smith, G.; E. Baldry, K.; Byriel, K.; Kennard, C., *Molecular Cocrystals of Carboxylic Acids. XXV The Utility of Urea in Structure Making with Carboxylic Acids and the Crystal Structures of a Set of Six Adducts with Aromatic Acids*. 1997; Vol. 50, p 727-736.

46. Videnova-Adrabinska, V., *Acta Crystallographica Section B* **1996,** *52* (6), 1048-1056.

47. Smith, G.; Kennard, C.; Byriel, K., *The Preparation and Crystal Structures of a Series of Urea Adducts: with Fumaric Acid (2 : 1), with Itaconic Acid (1 : 1) and with Cyanuric Acid (1 : 1)*. 1997; Vol. 50, p 1021-1025.

48. Videnova-Adrabińska, V.; Etter, M. C., *Journal of Chemical Crystallography* **1995,** *25* (12), 823-829.

49. Chadwick, K.; Davey, R.; Sadiq, G.; Cross, W.; Pritchard, R., *CrystEngComm* **2009,** *11* (3), 412-414.

50. Chang, H.-S.; Lin, J.-L., *Acta Crystallographica Section E: Structure Reports Online* **2011,** *67* (Pt 6), o1317-o1317.

51. Xu, W.; Huang, W.-x.; Chen, H.-y., *Acta Crystallographica Section E: Structure Reports Online* **2011,** *67* (Pt 7), o1795-o1795.

52. Rai, R. N.; Reddi, R. S. B.; Rai, U. S., *Progress in Crystal Growth and Characterization of Materials* **2013,** *59* (2), 73-111.

53. Kostyanovsky, R. G.; Kostyanovsky, V. R.; Kadorkina, G. K., *Mendeleev Communications* **2009,** *19* (1), 17-18.

54. Lü, M. K.; Meng, F. Q.; Yang, Z. H.; Yu, W. T.; Zeng, H., *Crystal Research and Technology* **1996,** *31* (7), 833-836.

55. Y.Inoue, *CSD Communication* **2016**.

56. Snyder, R. L.; Rosenstein, R. D., *Acta Crystallographica Section B* **1971,** *27* (10), 1969-1975.

57.   Lee, J. D.; Wallwork, S. C., *Acta Crystallographica* **1965,** *19* (3), 311-313.

58.   Mahmoud, M. M.; Wallwork, S. C., *Acta Crystallographica Section B* **1975,** *31* (2), 338-342.

59.   Deutsch, Z.; Bernstein, J., *Crystal Growth & Design* **2008,** *8* (10), 3537-3542.

60.   Mahapatra, A. K.; Sahoo, P.; Hazra, G.; Goswami, S.; Fun, H.-K., *Journal of Luminescence* **2010,** *130* (8), 1475-1480.

61.   Pickering, M.; Small, R. W. H., *Acta Crystallographica Section B* **1982,** *38* (12), 3161-3163.

62.   Donnelly, P. S.; Skelton, B. W.; White, A. H., *Australian Journal of Chemistry* **2003,** *56* (12), 1249-1253.

63.   Videnova-Adrabinska, V.; Janeczko, E., *Chemical Communications* **1999,** (16), 1527-1528.

64.   Thaimattam, R.; Shekhar Reddy, D.; Xue, F.; C. W. Mak, T.; Nangia, A.; R. Desiraju, G., *Journal of the Chemical Society, Perkin Transactions 2* **1998,** (8), 1783-1790.

65.   Taouss, C.; Jones, P. G., *CrystEngComm* **2014,** *16* (25), 5695-5704.

66.   Marti-Rujas, J.; Kariuki, B. M.; Hughes, C. E.; Morte-Rodenas, A.; Guo, F.; Glavcheva-Laleva, Z.; Tastemur, K.; Ooi, L.-l.; Yeo, L.; Harris, K. D. M., *New Journal of Chemistry* **2011,** *35* (7), 1515-1521.

67.   Seth, S. K.; Das, N. K.; Aich, K.; Sen, D.; Fun, H.-K.; Goswami, S., *Journal of Molecular Structure* **2013,** *1048*, 157-165.

68.   Graham Smith, K. E. B., Karl A. Byriel, Colin H.L. Kennard, *Aust. J. Chem.* **1997,** *50*, 727-736.

69.   Emsley, J.; Ma, L. Y. Y.; Hursthouse, M. B.; Karaulov, S. A.; Motevalli, M., *Journal of the Chemical Society, Perkin Transactions 2* **1990,** (7), 1077-1080.

70.   Krawczuk, A.; Gryl, M.; Pitak, M. B.; Stadnicka, K., *Crystal Growth & Design* **2015,** *15* (11), 5578-5592.

71.   Song, J.-X.; Yan, Y.; Yao, J.; Chen, J.-M.; Lu, T.-B., *Crystal Growth & Design* **2014,** *14* (6), 3069-3077.

72.   Tothadi, S., *CrystEngComm* **2014,** *16* (32), 7587-7597.

73.   Taouss, C.; Thomas, L.; Jones, P. G., *CrystEngComm* **2013,** *15* (34), 6829-6836.

74.   Hollingsworth, M. D.; Brown, M. E.; Dudley, M.; Chung, H.; Peterson, M. L.; Hillier, A. C., *Angewandte Chemie International Edition* **2002,** *41* (6), 965-969.

75.   Chetia, B.; Iyer, P. K., *Tetrahedron Letters* **2006,** *47* (46), 8115-8117.

76.   Smith, G.; Baldry, K. E.; Byriel, K. A.; Kennard, C. H. L., *Australian Journal of Chemistry* **1997,** *50* (7), 727-736.

77.   Li, -. J. Z. a. X.-H., *Acta Crystallogr., Sect. E: Struct. Rep. Online* **2005,** *61*, 3366.

78.   Selvakumar, S.; Leo Rajesh, A., *Journal of Materials Science: Materials in Electronics* **2016,** *27* (7), 7509-7517.

79.   Shnidman, L.; Sunier, A. A., *The Journal of Physical Chemistry* **1931,** *36* (4), 1232-1240.

Chapter 10 - Conclusions

This dissertation involved fundamental studies of crystal engineering of two main categories of small rigid molecules (12 pyrazole based and 12 thiazole based molecules) with multiple functional groups. This fundamental study was done to validate methods such as hydrogen-bond propensity (HBP), hydrogen-bond energies (HBE) and electrostatics for synthon prediction in these molecules. Based on our study, we found that in 14 out of 18 target molecules (78% success rate) where crystal structures were obtained, the synthon predictions correct using a combination of methods indicating that both HBP and HBE are suitable for synthons predictions (Figure 10.1).
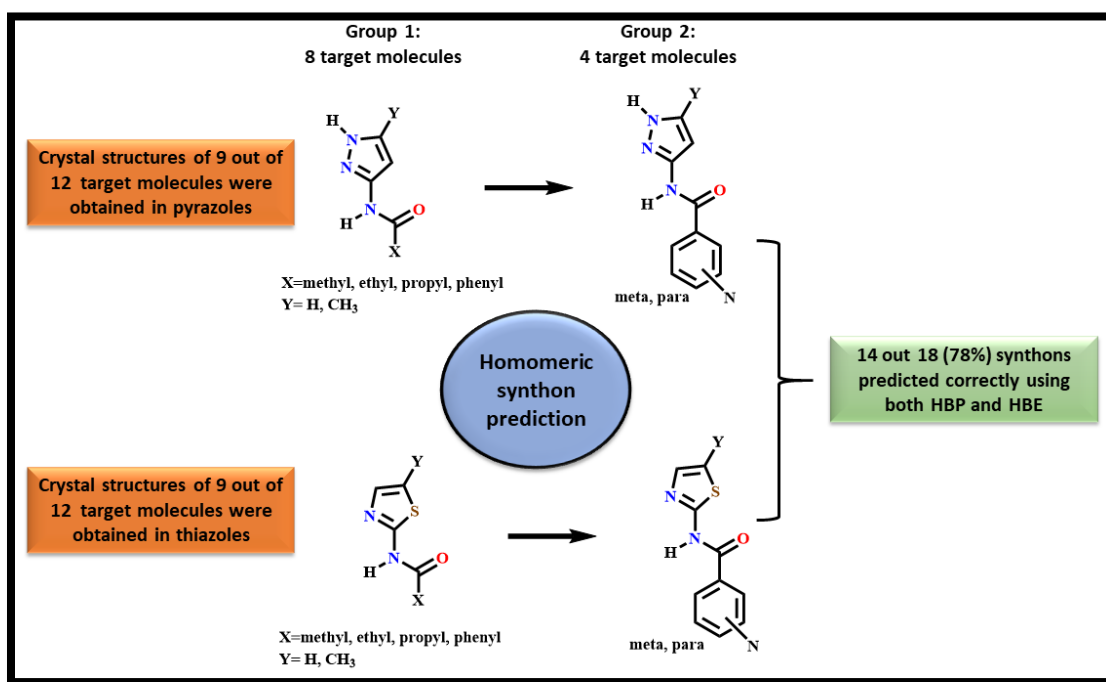
Figure 10.1 Summary of synthon prediction in 12 pyrazole based and 12 thiazole based target molecules

In order to study the robustness of homomeric interactions and to determine whether it was possible to break these interactions, the experimental co-crystallization was performed on (12 thiazole and 12 pyrazole based molecules) with 20 carboxylic acids leading to 480 co-crystallization experiments. In each category, the molecules were sub-divided into two groups based on the number of functional groups present. Following information was achieved from this study (Figure 10.2),

1. When an additional strong acceptor group such as pyridine was introduced, the experimental co-crystallization success rate increased from 64% to 83% in pyrazole based molecules and from 49% to 91% in thiazole based molecules.

2. In group1, changing pyrazole to thiazole moiety decreased the success rate from 64% to 49% and in group 2, changing pyrazole to thiazole increases the success rate from 83% to 91%.

3. Overall, pyrazole showed a slightly higher success rate (69%) than thiazole based molecules (58%).
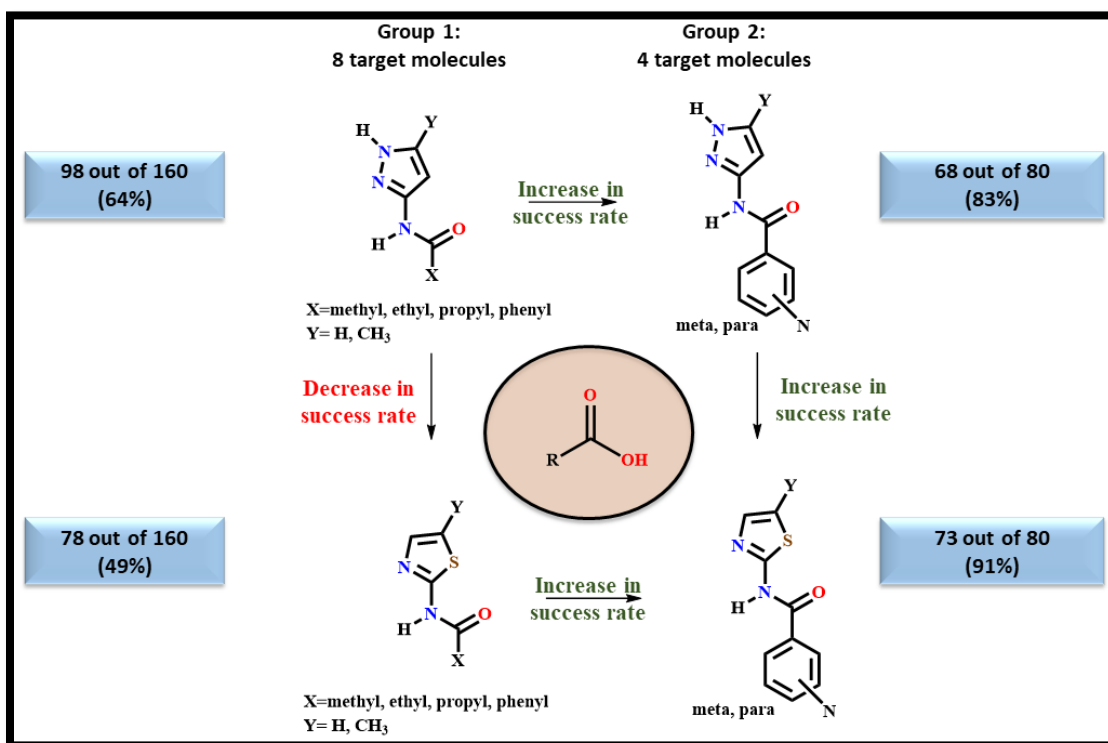
Figure 10.2 Summary of experimental co-crystal screen

Different prediction methods were used to analyze and predict the experimentally observed co-crystal screen. In the thiazole based molecules, hydrogen-bond propensity was used as the prediction method and 70% of group1 molecules and 68% of group 2 molecules were predicted correctly. In the pyrazole molecules, hydrogen-bond propensity and hydrogen bond energies were used as prediction methods. Based on HBP, 44% of group1 molecules and 15% of group 2 molecules were predicted correctly. Because of poor prediction using HBP, HBE was used as an alternative method and it was found that 70% of group 1 and 85% of group 2 molecules were

predicted correctly using this method. In conclusion, a combination of both HBP and HBE are suggested for prediction studies. The synthon prediction were done on each co-crystal combination and combinations where co-crystal was obtained, 82% and 100% success was obtained in predicting the correct synthon observed in the crystal structures of pyrazole and thiazole based molecules (Figure 10.3).
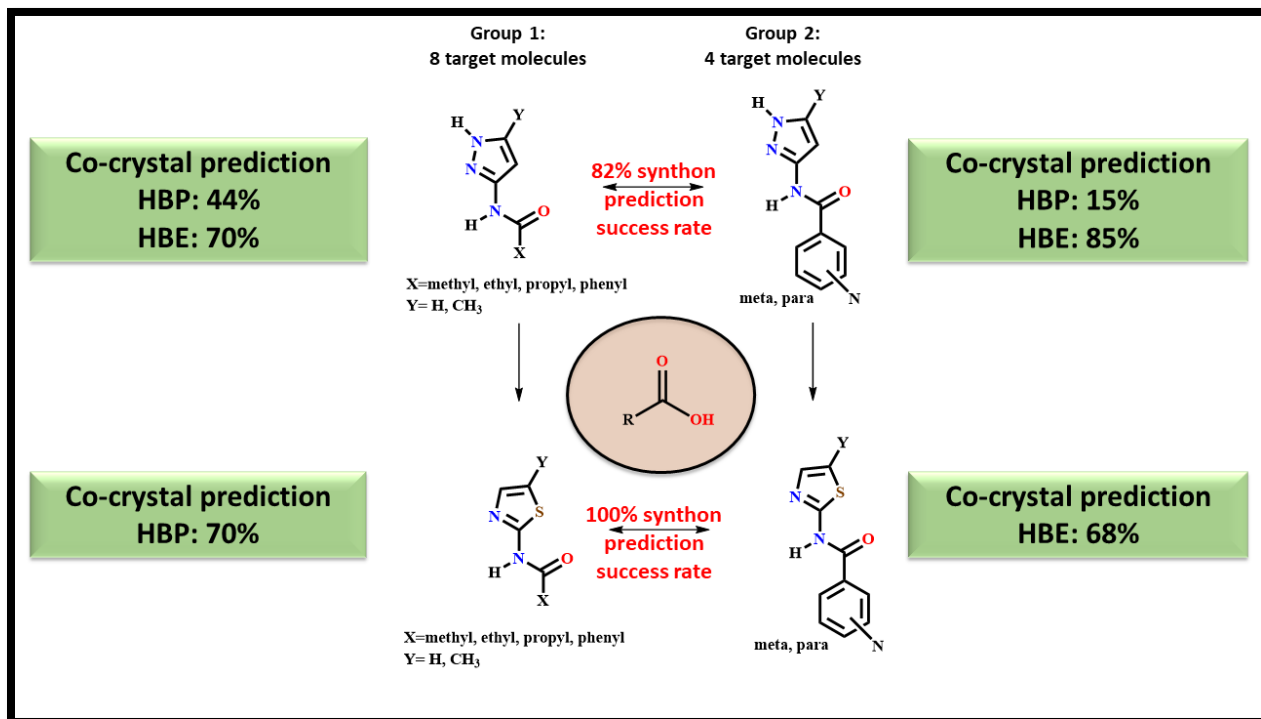
Figure 10.3 Summary of experimental vs predicted comparison using HBE and HBP for co-crystal screening

The ability of pyridine-based molecules to form binary hydrogen-bonding was successfully studied before. The pyridine functionality was also used to study binary halogen-bonding. Because of 100% supramolecular yield with both binary hydrogen and halogen bonding, we were successfully able to design ternary complexes with high supramolecular yield by carefully designing the target molecules with two-point interaction for hydrogen-bonding and one-point interaction for halogen bonding. In the end, we were able to obtain two crystal structures of ternary system showing both hydrogen and halogen bonding (Figure 10.4).
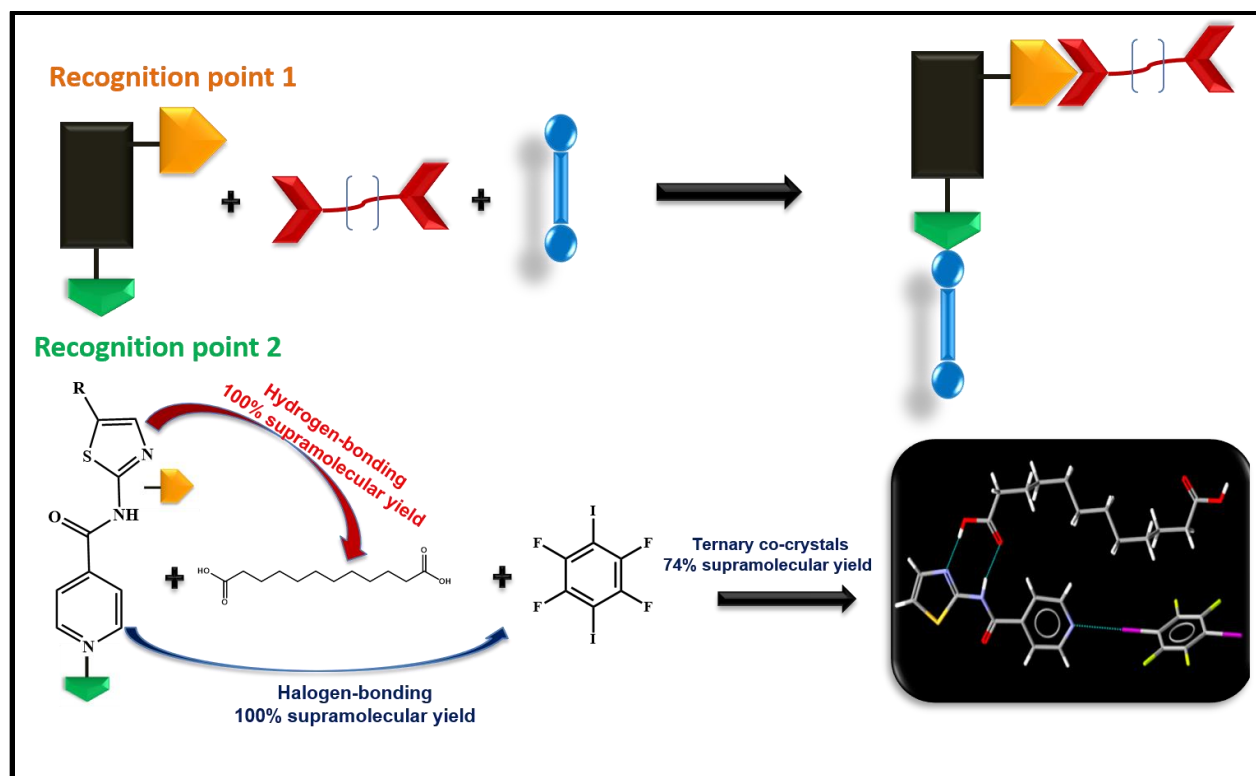
Figure 10.4 Summary of binary and ternary co-crystals of thiazole-pyridine based molecules

These small rigid molecules were mostly acceptors either for hydrogen-bonding or halogen bonding. In order to include halogen bond donors on the target molecules while keeping amide functionality intact, the supramolecular hierarchy of activated and unactivated halogen bond donors was explored. Eight new halogen bond donors were synthesized, and it was found that iodoethynyl based target molecules were the best halogen bond donors followed by tetraflourinated iodine and ethynyl based molecules (Figure 10.5). Moreover, the activated molecules didn't form co-crystals with the carboxylic acids whereas unactivated halogen did because of the availability of strong pyridine nitrogen group in the latter.
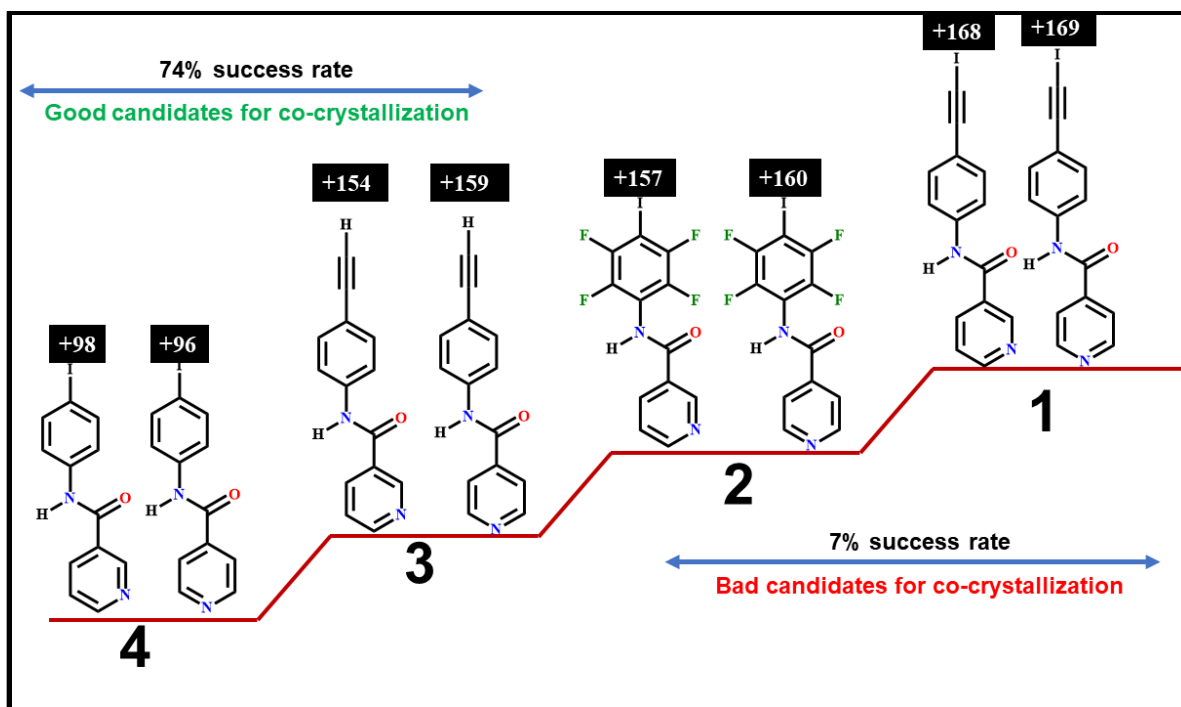
Figure 10.5 Summary of supramolecular hierarchy and co-crystallization results of target molecules

The next part of thesis involved moving from small rigid molecules to large flexible molecules with diverse functionalities. The major concern with large diverse molecules was the ability to predict their crystallizability. A statistical logistic regression model was developed using easily defined, measured and experimental parameters to predict crystallizability of drug-like molecules. The model was tested on molecules where experimental crystallizability is already established and a success rate >85% was obtained as validation between prediction and experimental results (Figure 10.6).
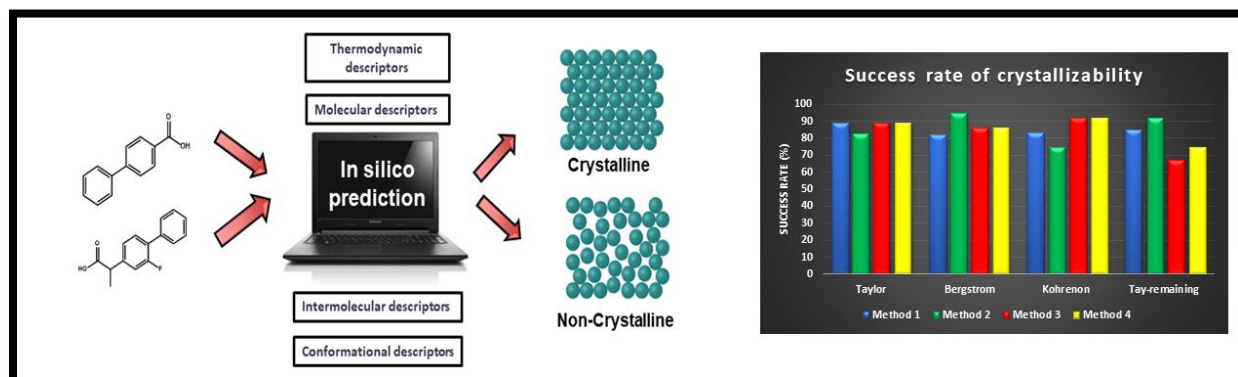
Figure 10.6 Summary of successful prediction of crystallizability using statistical logistic regression model

One of the most exciting aspects of this dissertation is that we were able to make a transition from basic and very fundamental research on structural chemistry, to new materials that may have wide-ranging practical applications coupled to considerable environmental benefits.
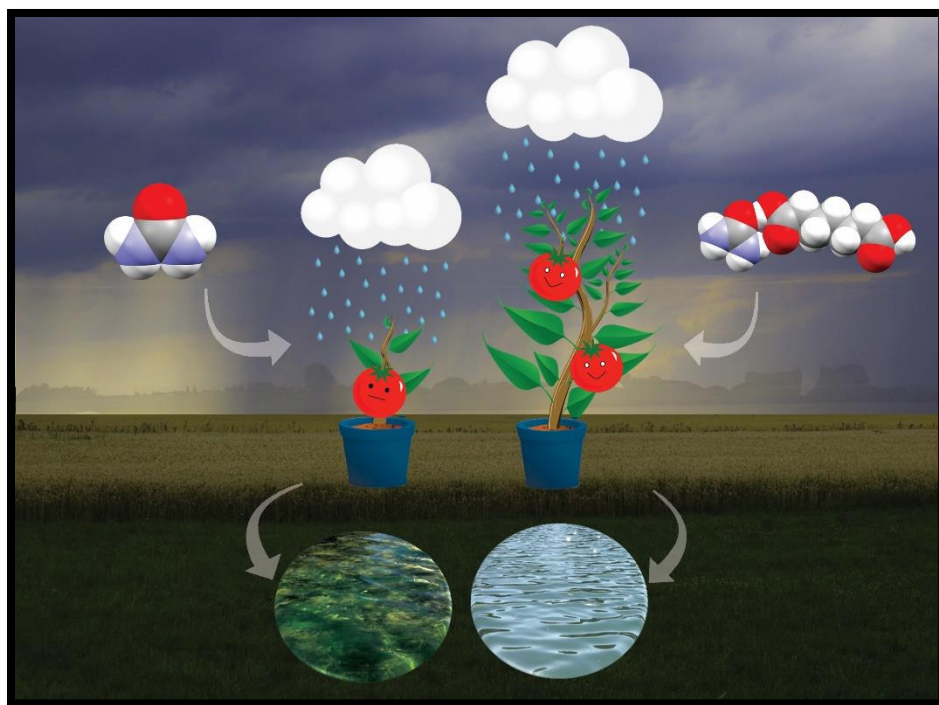
Figure 10.7 Summary of decreased solubility and increased stability of urea via co-crystallization technology

Urea has gained worldwide importance as a nitrogen rich fertilizer and is used globally on a multi-ton scale each year. However, due to its high solubility in water it is also prone to mobility and run-off problems. Farmers therefore tend to use more than required to meet the needs of plants and crops which, in turn, increases the amount of nitrogen in the environment thereby creating numerous unwanted environmental consequences such as eutrophication of lakes and rivers. A simple and easily scalable protocol based on knowledge-based method was used to design new solid forms of urea with more appropriate solubility and much enhanced stability with respect to humidity (Figure 10.7).

Appendix A - Hydrogen-bond propensity for solid state risk assessment

A.1 Introduction

The identification of the most stable form of a drug candidate is a crucial step in the development of new active pharmaceutical ingredients. Experimental solid form screening is a protocol followed by industry to determine the most stable form and to identify any polymorphic forms[1]. This process is time consuming, non-exhaustive and at same time involves risk ("have we done enough experiments?") and opportunity ("can we design a better material?"). Therefore, it is of considerable interest to acquire as much detailed knowledge as possible on the crystal form landscape of a compound in the course of research and development from a lead compound to a marketed product.[2]

A new approach to modeling and predicting stable polymorphs and for finding the most stable form utilizes solid form informatics. The Cambridge Crystallographic Data Centre (CCDC), in partnership with the Solid Form Design Centre[3], its predecessor the Pfizer Institute for Material Science, and more recently members of the Crystal Form Consortium (CfC), has developed hydrogen-bond propensity[4-7] and hydrogen-bond coordination likelihood tools to augment solid form design and risk assessment (Figure A.1).

(a)	**(b)**
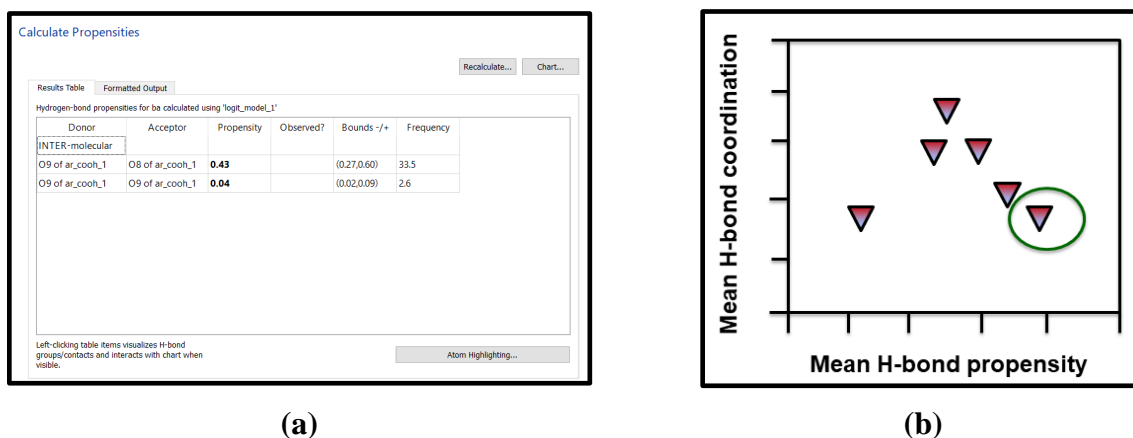
Figure A.1 (a) Hydrogen-bond propensity and (b) hydrogen bond coordination likelihood charts.

The knowledge-based tool allows information on molecular geometries and intermolecular interactions encapsulated in the Cambridge Structural Database[8] (CSD), the world's repository of

all published and patented small molecule crystal structures (900,000+), to be efficiently and thoroughly mined. The tools search for similar structures from CSD based on functional groups present in the molecule of interest, this is followed by a statistical survey to find the most likely hydrogen bonds considering functional groups, competition, aromaticity and steric density and calculates propensity values for the hydrogen bonds in the range 0 to 1 for the specific molecule. A form with high propensity values indicates optimal intermolecular bonding, whereas low propensity hydrogen bonds can indicate that alternative ways of packing might be preferable, suggesting likely polymorphism.

The hydrogen-bond coordination likelihood tool produces a landscape of possible structures based on hydrogen-bonds derived by propensity tool. When there is only one combination with high propensity bonds and all donors and acceptors being used (denoted as "high coordination"), the risk of polymorphism is ranked as low. On the other hand, when there are many combinations with similar propensity and coordination the risk of polymorphism is considered high. The analysis can be performed on a single molecule to assess the "risk" of polymorphs or on a combination of molecules to determine the likelihood of the formation of multicomponent crystal forms.

In this study, we will cover basics of hydrogen-bond propensity and hydrogen-bond coordination likelihood, interpretation/comparison of different datasets, sensitivity of the model and limitations. Five main sections include following:

- Sensitivity analysis- comparing sample sizes
- Role of HBP in identifying stable polymorphs
- Using HBP to understand H-bond competition and their strength
- Importance of functional group definition when making comparisons
- Limitations

A.2 Experimental

A.2.1 Hydrogen-bond propensity calculations

Hydrogen-bond propensity[4-5, 9] calculations were employed using CSD Version 5.38 and Mercury 3.9, each compound was sketched, and auto-edited, functional groups were selected as suggested by Mercury (Table A.1 represents functional groups used for each molecule), a training dataset

(350-600 structures per functional group) was made and the propensities were calculated with an ROC curve higher than 0.800 ("good discrimination") unless otherwise indicated. The propensity calculations are done using outline shown in Figure A.2 to A.6.

Table A.1 Functional groups used to determine the hydrogen-bond propensities of small rigid molecules. The labels in the figures can be explained as follows: T_n = atom makes n bonds, c = atom is cyclic, ⓑ = bond is acyclic, and H_n = n bonded hydrogen atoms.

Isonicotinamide	4-Hydroxybenzamide	Benzoic acid	Isonicotinic acid

Step1: A default H-bond definition as shown in Figure A.2 is used but sometimes if a short contact is observed in the crystal structure, that can be avoided by changing the minimum H-bond distance range.
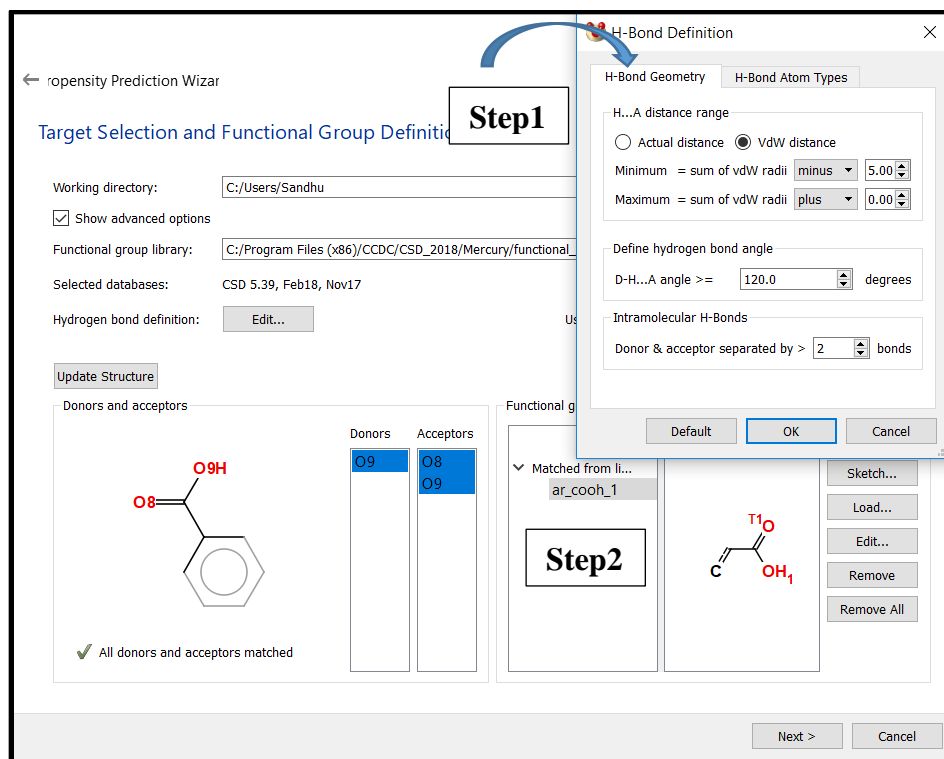
Figure A.2 Hydrogen-bond propensity wizard showing hydrogen-bond definition and functional groups.

Step 2: By default, the functional groups are chosen in this wizard, however if a different functional group is needed, click on the functional group chosen and either use the editing tool or the sketch tool to make the new functional group.
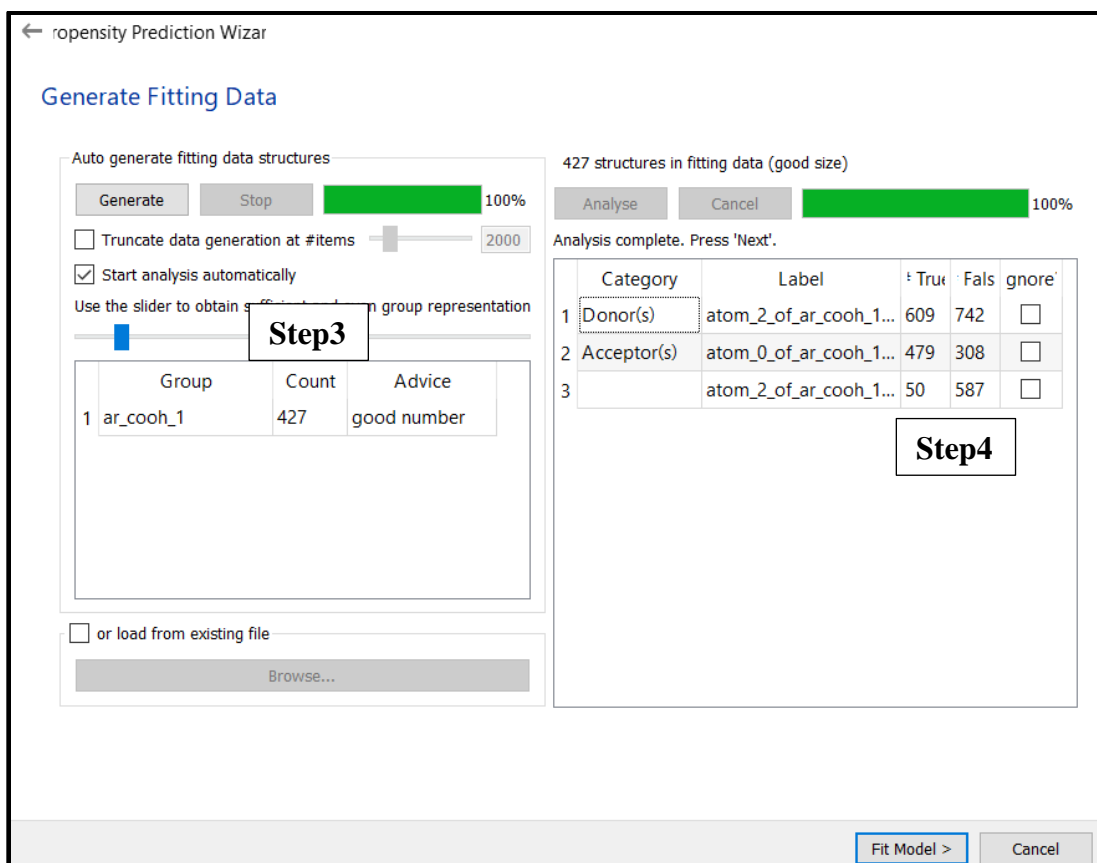
Figure A.3 Hydrogen-bond propensity wizard showing fitting data information.

Step 3: The fitting data is generated in this step. The number of structures in the fitting data can be changed by moving the blue slider. Alternatively, if there is specific data already generated that you wish to use, click on the load from existing file option.

Step 4: The true and false data are generated, and any value shown in red should be ignored.

Step 5: If the model fails this step, use the refine model option to delete descriptors that have significance code less than 2 one by one and select accept and continue. The area under ROC curve should be above 0.70 to be considered as good model.
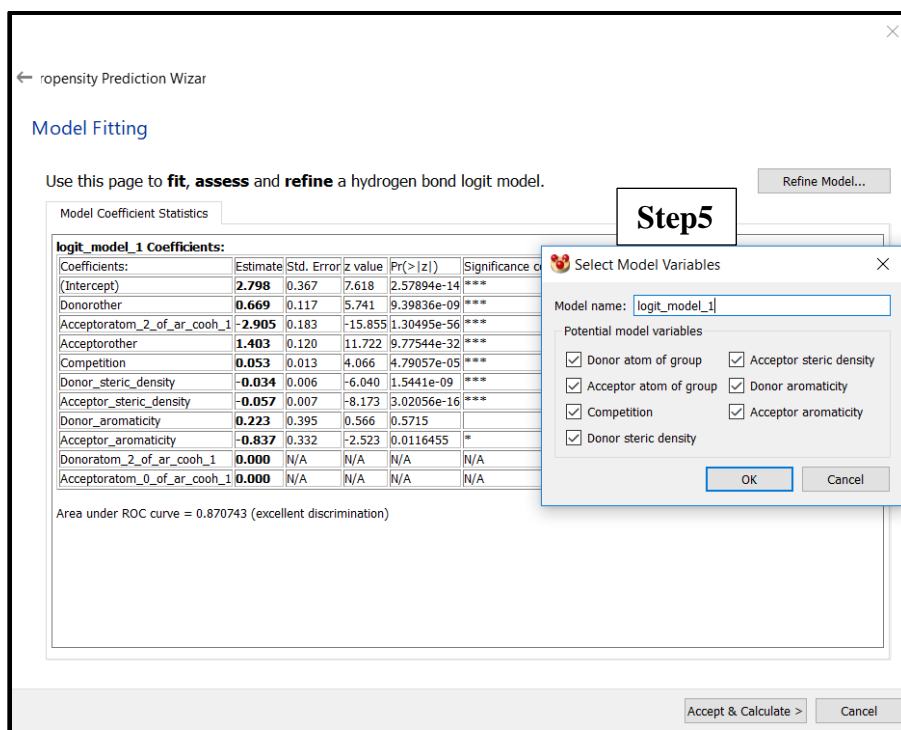
Figure A.4 Hydrogen-bond propensity wizard showing how to refine the model.

Step 6: Calculate the propensities and click on the chart to get the hydrogen-bond coordination likelihood table.
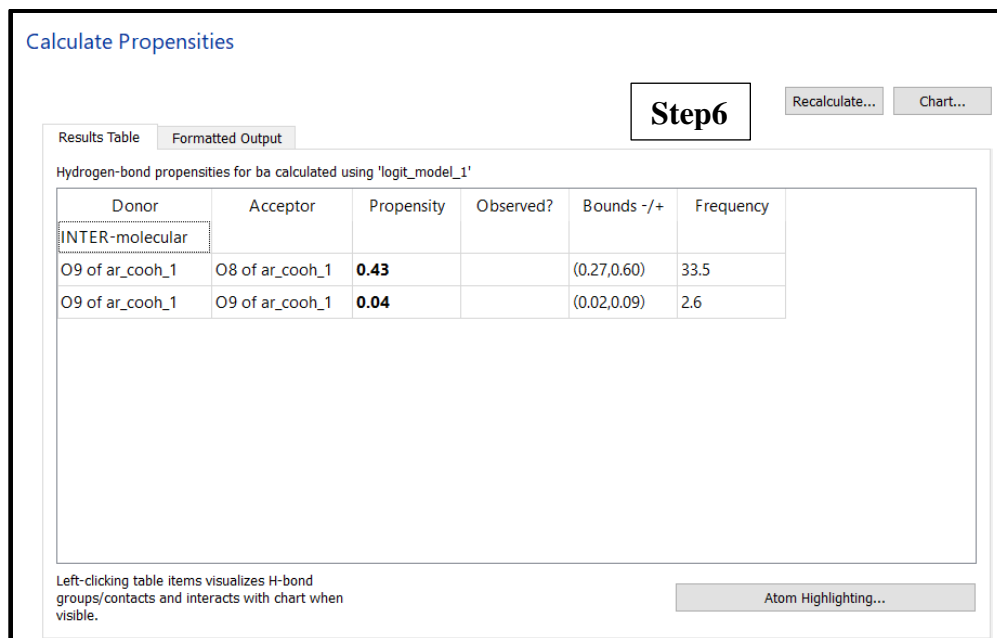
Figure A.5 Hydrogen-bond propensity wizard showing propensity calculations.
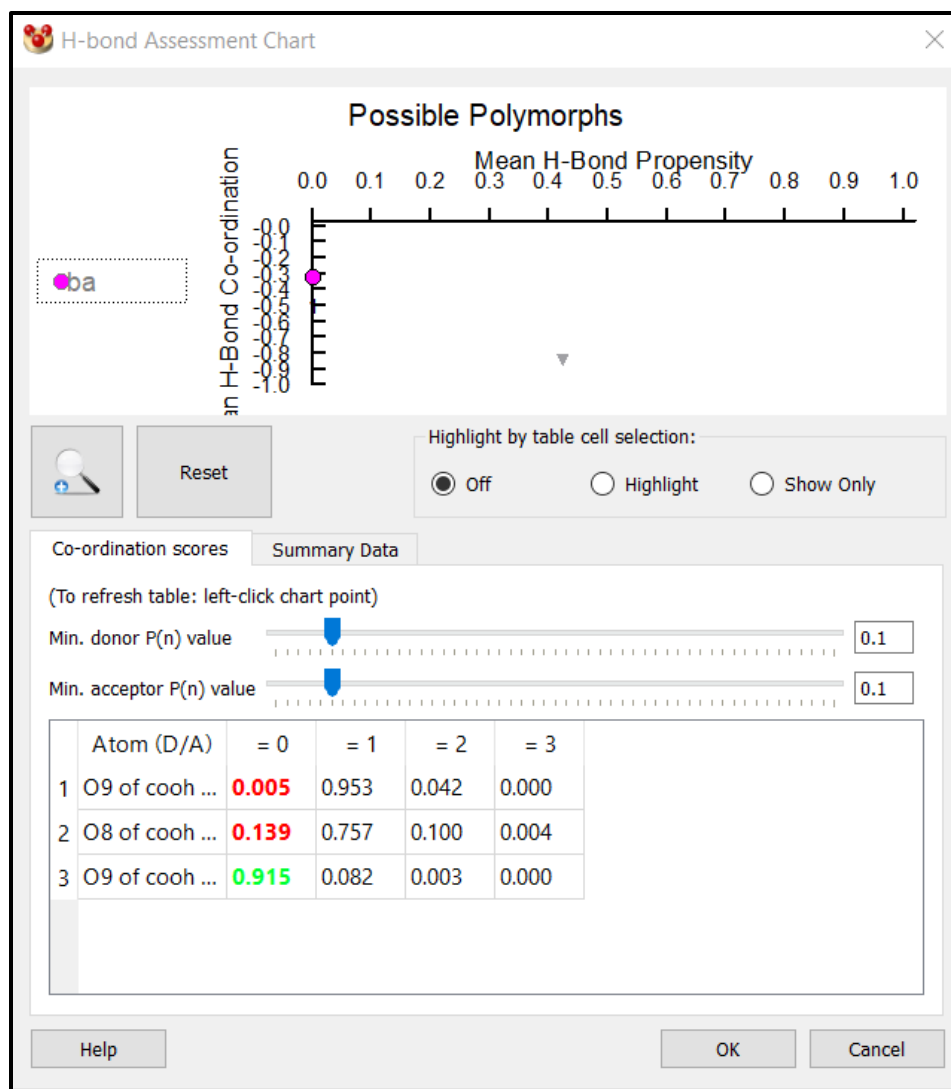
Figure A.6 Hydrogen-bond coordination likelihood chart.

A.2.2 Molecular electrostatic potentials

Molecular electrostatic potential surfaces (MEPS) of molecules studied were generated with DFT B3LYP level of theory using 6-311++G** basis set in vacuum. All calculations were carried out using Spartan'08 software. All molecules were geometry optimized with the maxima and minima in the electrostatic potential surface (0.002 e/au isosurface) determined using a positive point charge in vacuum as a probe. The numbers indicate the interaction energy (kJ/mol) between the probe and the surface of the molecule at that particular point.

A.3 Results

A.3.1 Sensitivity analysis

A.3.1.1 Comparing sample of Isonicotinamide

A "sensitivity" analysis was performed on isonicotinamide as it has two hydrogen-bond donors (amine NH_2 group) and two hydrogen-bond acceptors (pyridine N and C=O). There are two possible hydrogen-bond synthons; NH(amine) can either bind to pyridine N or C=O (Scheme A.1). To find out which is more likely to happen, and whether choosing a certain dataset affects the resulting propensity order/ranking of these two interactions, hydrogen-bond propensity analysis was performed.

Synthon A   **Synthon B**
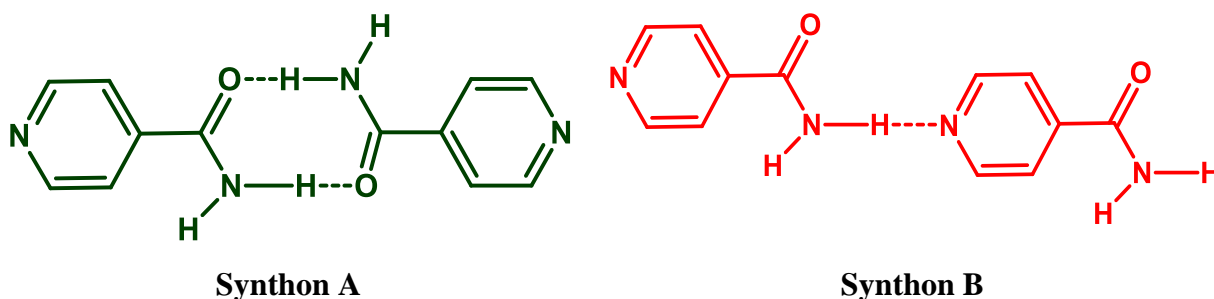
Scheme A.1 Two possible hydrogen-bond interactions in the isonicotinamide.

The results of hydrogen-bond propensity are shown in Table A.1.

Table A.1 ROC and hydrogen-bond propensity of each interactions obtained for each dataset for isonicotinamide.
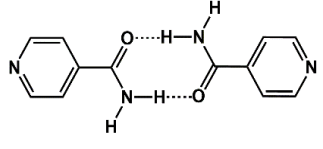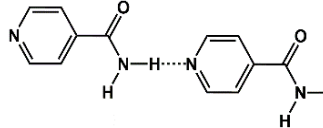
Total Dataset	ROC	NH(amine)...C=O	NH(amine)...N
351	0.66	0.62 (0.45, 0.77)	0.48(0.32, 0.64)
551	0.64	0.62 (0.50, 0.73)	0.56 (0.45,0.66)
800	0.64	0.64 (0.53, 0.73)	0.62 (0.53, 0.71)
1045	0.63	0.65(0.55, 0.73)	0.67 (0.59, 0.74)
3589	0.67	0.66 (0.60, 0.72)	0.78 (0.75, 0.81)
5727	0.69	0.66 (0.62, 0.71)	0.82 (0.80, 0.84)

A.3.2 Role of HBP in identifying stable polymorphs

A.3.2.1 Isonicotinamide[10]

Isonicotinamide has two hydrogen-bond donors (amine NH_2) and two acceptors (pyridine N and carbonyll C=O). The NH donor group can either bind to carbamoyl oxygen atom forming $R_2^2(8)$ type NH(amine)…O=C(carbamoyl) hydrogen-bond interaction[10] or one NH of amine binds to pyridine N forming NH(amine)…N(pyridine) and second NH amine binds to C=O forming NH(amine)…C=O(carbamoyl) hydrogen-bond interaction[11]. HBP was carried out to determine whether the model can predict the both experimental forms and determine which form is more favorable, Table A.2.

Table A.2 Results of hydrogen bond propensity calculations on isonicotinamide and the observed hydrogen bonds in each known polymorph.

Donor…Acceptor	Propensity	Refcode[10]	Refcode[11]
	0.62 (0.45, 0.77)	Yes	Yes
	0.48 (0.32, 0.64)	-	Yes

The hydrogen-bond coordination chart was compared to predicted hypothetical structures and experimentally observed synthons (Figure A.7 and table A.3).

273
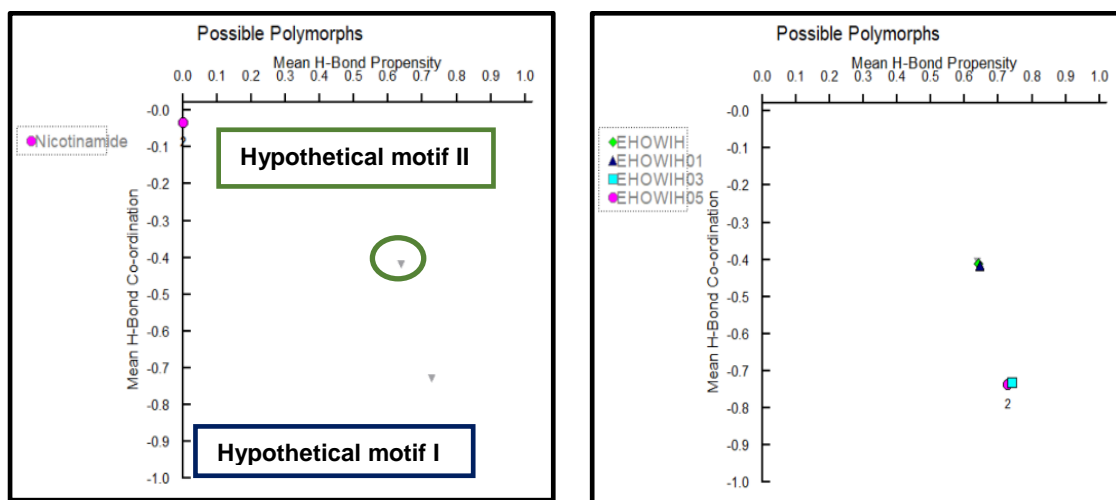
Figure A.7 Propensity coordination chart of isonicotinamide showing (a) two predicted motifs and (b) experimental observed motifs overlapping with predicted motifs.

Table A.3 Coordination chart for each motif predicted by propensity-coordination chart for isonicotinamide.
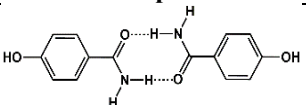
Structure #	Atom (D/A)	0	1	2	3
Hypothetical motif I	NH2 of carbamoyl (D)	0.00	0.05	**0.87**	0.08
	N(aromatic) (A)	0.12	**0.79**	0.09	0.00
	O=C of carbamoyl (A)	0.04	**0.62**	0.31	0.02
Hypothetical motif II	NH2 of carbamoyl (D)	0.00	0.05	**0.87**	0.08
	N(aromatic) (A)	**0.12**	0.79	0.09	0.00
	O=C of carbamoyl (A)	0.04	0.62	**0.31**	0.02
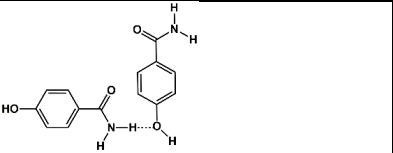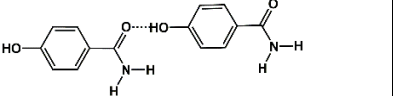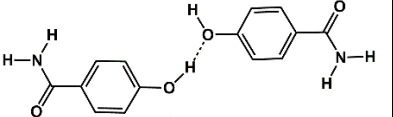
A.3.2.2 4-Hydroxybenzamide[12]

4-Hydroxybenzamide has three hydrogen-bond donors (amine NH_2 group and hydroxyl OH group) and two hydrogen-bond acceptors (C=O and hydroxyl O). Four different types of interactions are possible in 4-hydroxybenzamide, Table A.4

The HBP calculations were performed to determine which is the most optimal interaction and if there is competition between interactions. To perform HBP calculations, carbamoyl and aromatic OH functionalities were chosen, the hits for training dataset was 485 with 368 structures containing both functionalities Table A.4. The area under the ROC curve was 0.79 signifying good discrimination.

Table A.4 Results of Hydrogen bond propensity calculations on hydroxybenzamide and the observed hydrogen bonds in each experimental known polymorph.

Rank	Donor…Acceptor	Propensity	Refcode[12]	Refcode[13]
1		0.64 (0.51, 0.76	Yes	Yes

	Structure			
2		0.39 (0.28, 0.51)	Yes	-
3		0.31 (0.22, 0.43)	Yes	-
4		0.14 (0.10, 0.20)	-	Yes

The hydrogen-bond coordination chart was obtained for 4-hydroxybenzamide, Figure A.8 and Table A.5.

(a)   (b)
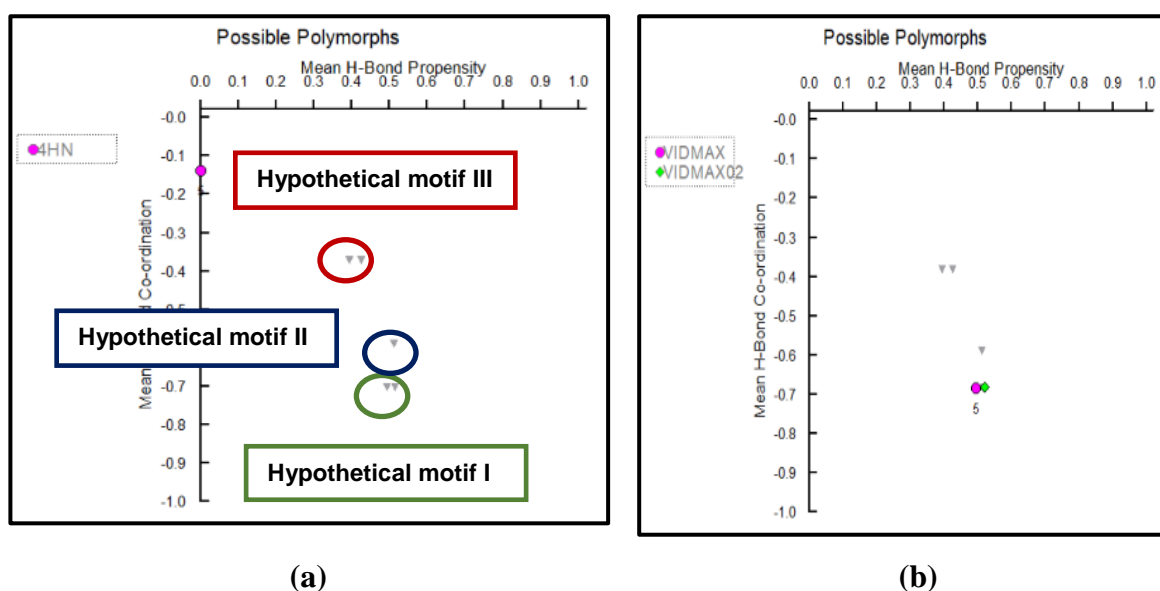
Figure A.8 Propensity coordination chart of hydroxybenzamide showing (**a**) five predicted motifs and (**b**) experimental observed motifs overlapping with two most optimal predicted motifs.

Table A.5 Coordination table for each hypothetical structure for 4Hydroxybenzamide.

Structure	Atom (D/A)	0	1	2	3
Hypothetical structure I and I'	NH$_2$ of carbamoyl (D)	0.01	0.12	**0.81**	0.07
	OH of hydroxyl (D)	0.02	**0.93**	0.05	0.00
	O of hydroxyl (A)	0.68	**0.31**	0.01	0.00
	O=C of carbamoyl (A)	0.00	0.14	**0.79**	0.07
Hypothetical structure II	NH$_2$ of carbamoyl (D)	0.01	**0.12**	0.81	0.07
	OH of hydroxyl (D)	0.02	**0.93**	0.05	0.00
	O of hydroxyl (A)	**0.68**	0.31	0.01	0.00
	O=C of carbamoyl (A)	0.00	0.14	**0.79**	0.07
Hypothetical structure III and III'	NH$_2$ of carbamoyl (D)	0.01	**0.12**	0.81	0.07
	OH of hydroxyl (D)	0.02	**0.93**	0.05	0.00

O of hydroxyl (A)	0.68	**0.31**	0.01	0.00
O=C of carbamoyl (A)	0.00	**0.14**	0.79	0.07

A.3.3 Using HBP to understand H-bond competition

Benzoic acid is a simple molecule with one hydrogen-bond donor (carboxylic OH) and two hydrogen-bond acceptors (carbonyl C=O and O(acid). When another possible strong acceptor group is introduced such as pyridine nitrogen as in isonicotinic acid, it increased the competition between two strong acceptors (C=O and Pyridine N) to bind to single OH group. The hydrogen-bond propensity value are shown in scheme A.2 and the molecular electrostatic potential values are shown in figure A.9.
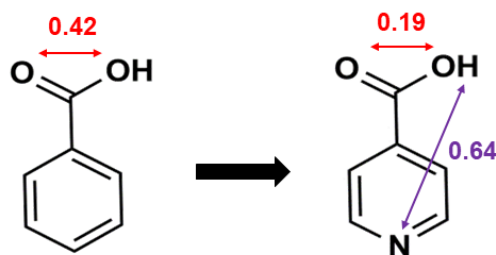
Scheme A.2 Chemical diagram of benzoic acid and isonicotinic acid showing the propensity. (the propensity score OH(acid)…C=O shown in red, and propensity score of OH(acid)…N(aromatic) is shown in purple.
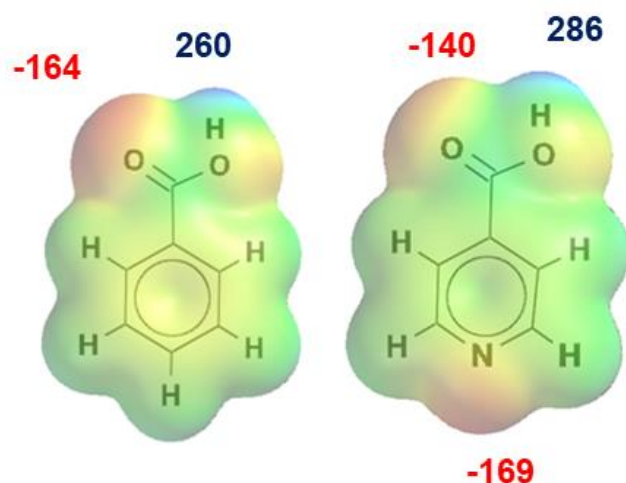
Figure A.9 Molecular electrostatic potentials of benzoic acid, and isonicotinic acid.

276

A.3.4 Importance of functional group definition when making comparisons

A.3.4.1 Replacing a functional group (isonicotinic acid to thioisonicotinic acid)
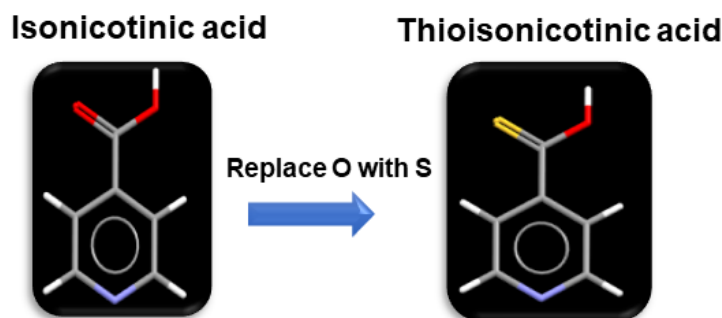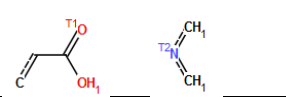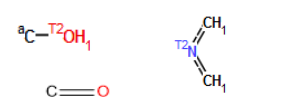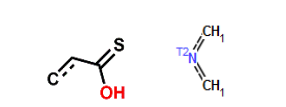
Scheme A.3 Chemical diagram of isonicotinic and thioisonicotinic acid.

To make comparisons of hydrogen bond propensity between two molecules, a careful attention must be paid to the functional group chosen to define each molecule.

We illustrate this with an example of isonicotinic acid by replacing carbonyl (C=O) of acid to thionyl (C=S) group. The functional groups and corresponding hydrogen-bond propensity values are shown in Table A.6.

1. When the isonicotinic acid is chosen for HBP, two functional groups are chosen, aromatic COOH and aromatic nitrogen.
2. HBP was performed on isonicotinic acid by fragmenting the COOH group into C=O and OH groups along with aromatic nitrogen.
3. Replacing C=O with C=S functional group and choosing aromatic C=SOH and aromatic N for thioisonicotinic acid, the regression failed because the aromatic C=SOH functionality does not have enough data in the CSD.
4. For thioisonicotinic acid, aromatic nitrogen, acrylic hydroxyl and C=S were chosen as the functional groups.

Table A.6 Functional groups chosen and resulting propensities for isonicotinic acid and thioisonicotinic acid.

	Functional group definition	Dataset	Propensity O-H…O=C/O-H…S=C	Propensity O-H…N(py)
Isonicotinic acid		375	0.19 (0.08, 0.37)	0.64(0.42, 0.81)
		375	0.34 (0.21, 0.48)	0.59 (0.44, 0.73)
Thioisonicotinic acid			Regression failed	
		364	0.25 (0.17, 0.36)	0.63 (0.51, 0.73)
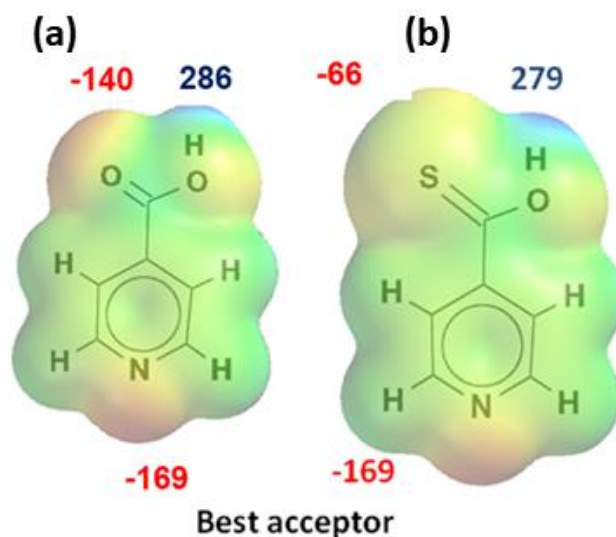
Figure A.10 Molecular electrostatic potentials of (a) isonicotinic acid and (b)thioisonicotinic acid.

A.3.4.2 Role of choosing correct functional group

In order to determine the right functional group choice for molecules where the fragmentation is possible such as in bis-urea substituents, we performed an analysis of *bis*-urea analogues by fragmenting each functional group into three different ways (Scheme A.4-Scheme A.6). The propensity for each molecule was calculated using each fragmented functional group and

compared to the experimental data to make conclusion regarding which fragmentation is the correct choice of functional group for such molecules (Table A.7 and Table A.8). Four molecules were chosen for this study.
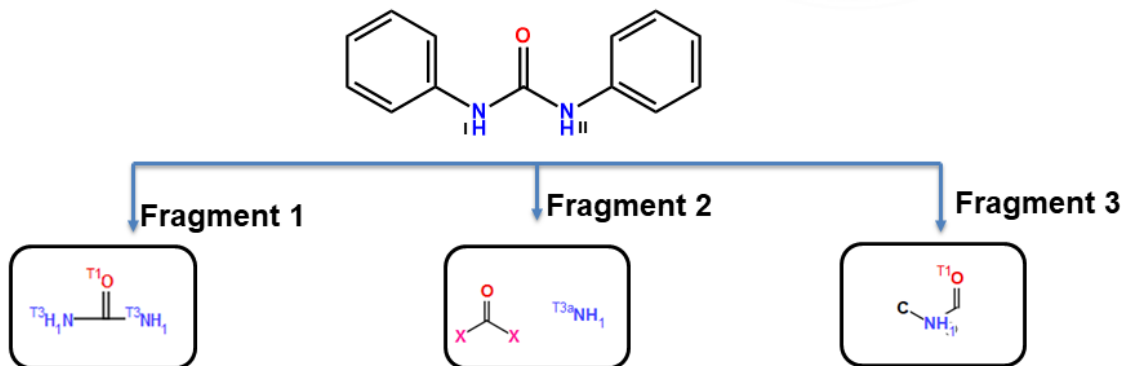
Scheme A.4 Three different fragmentation of functional group on 1,3-bisphenyl urea.
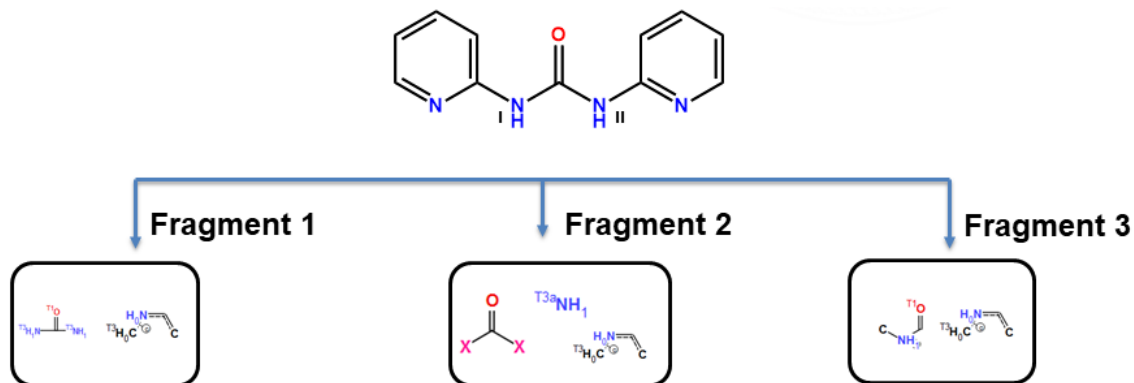
Scheme A.5 Three different fragmentations of functional group on 1,2-bis(pyridin-2-yl) urea.
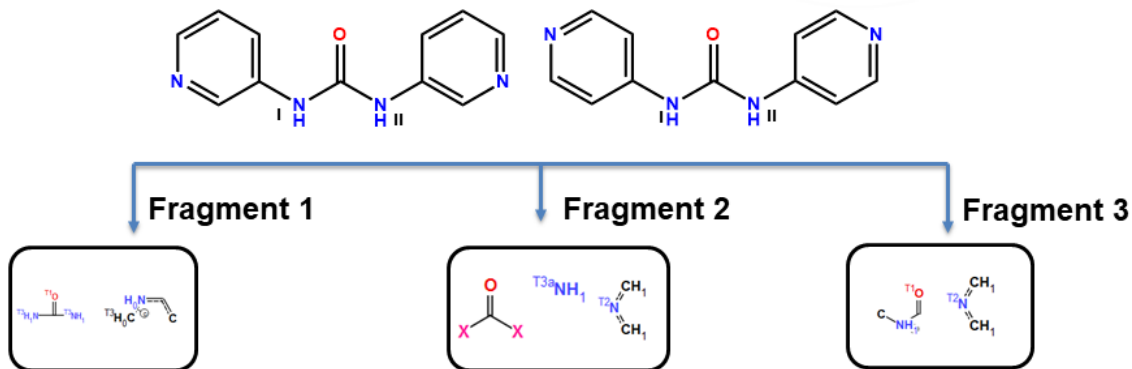
Scheme A.6 Three different fragmentations of functional group on 1,3-bis(pyridin-2-yl) urea and 1,4-di(pyridin-2-yl) urea.

Table A.7 Hydrogen-bond propensity score for each fragment in 1,3-bisphenylurea.

Donor…. Acceptor	Fragment 1	Fragment 2	Fragment 3
NH(1)…O=C	0.60 (0.42, 0.56)	0.49(0.39, 0.59)	0.40 (0.26, 0.56)
NH(2)…O=C	0.60 (0.42, 0.56)	0.49 (0.39, 0.59)	0.40 (0.26, 0.56)
ROC	0.731	0.773	0.720

Table A.8 Hydrogen-bond propensity score for each fragment in *bis*-urea analogues.

1,2-bis(pyridin-2-yl) urea			
Donor…. Acceptor	Fragment 1	Fragment 2	Fragment 3
NH(I)…N(pyridine) intra	0.86 (0.86, 0.86)	0.86 (0.86, 0.86)	0.86 (0.86, 0.86)
NH(II)…O=C	0.66 (0.48, 0.80)	0.29 (0.18, 0.44)	0.22 (0.11, 0.38)
NH(I)…N(pyridine) inter	0.58 (0.43, 0.73)	0.25 (0.15, 0.40)	0.31 (0.18, 0.49)
ROC	0.737	0.784	0.771
1,3-bis(pyridin-2-yl) urea			
Donor…. Acceptor	Fragment 1	Fragment 2	Fragment 3
NH(II)…O=C	0.55 (0.35, 0.73)	0.39 (0.26, 0.54)	0.43 (0.31, 0.57)
NH(I)…N(pyridine) inter	0.71 (0.56, 0.83)	0.34 (0.22, 0.49)	0.47 (0.36, 0.59)
ROC	0.711	0.727	0.653
1,4-bis(pyridin-2-yl) urea			
Donor…. Acceptor	Fragment 1	Fragment 2	Fragment 3
NH(II)…O=C	0.54 (0.35, 0.71)	0.38 (0.25, 0.53)	0.47 (0.30, 0.63)
NH(I)…N(pyridine) inter	0.76 (0.63, 0.85)	0.36 (0.23, 0.51)	0.53 (0.38, 0.68)
ROC	0.719	0.727	0.642

A.3.5 Limitations

A.3.5.1 Differentiation of equivalent synthon polymorphs

The polymorph prediction of benzamide is studied here (Figure A.11).
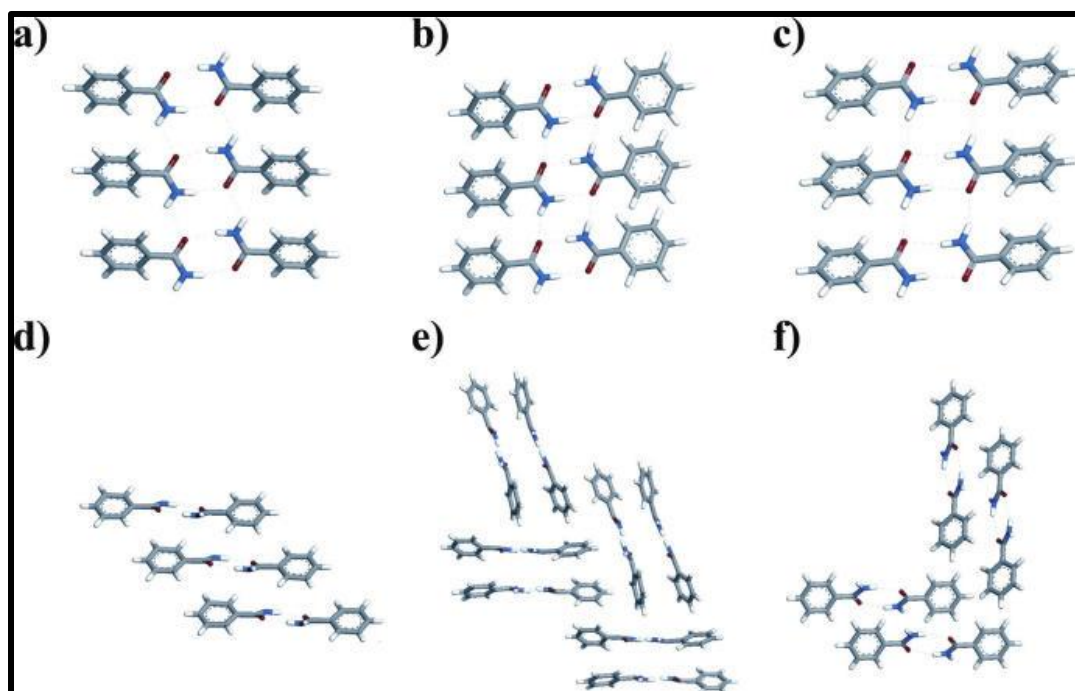
Figure A.11 Hydrogen bonded ribbons of (**a**) Form I, (**b**) Form II, and (**c**) Form III. The π-π interaction motifs are shown for (**d**) Form I along the b-axis, (**e**) Form II along the c-axis, and (**f**) Form III along a-axis.

The hydrogen-bond coordination chart was obtained for figure A.12.
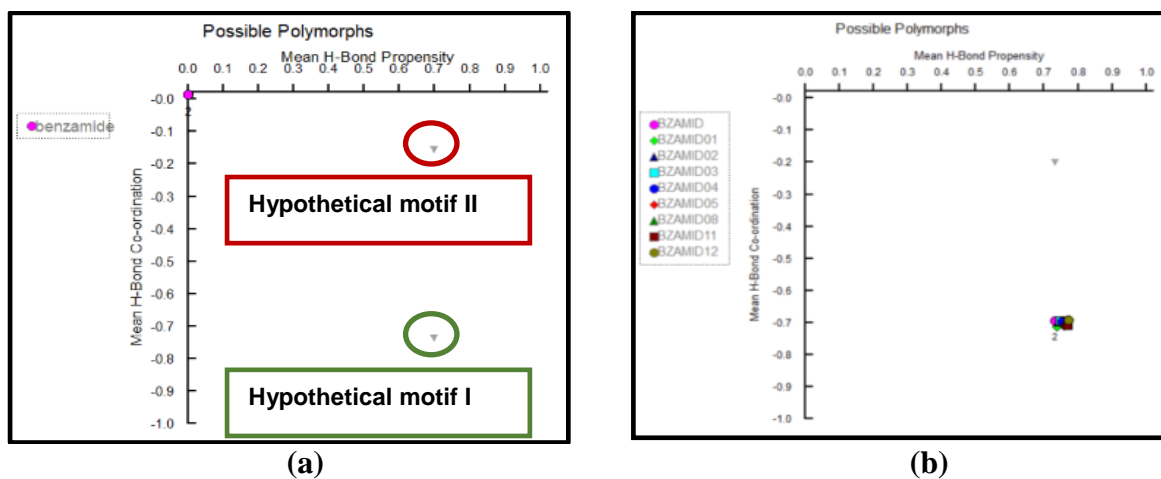
Figure A.12 Propensity coordination chart of benzamide showing (a) two predicted motifs and (b) all experimental observed motifs overlapping with the most optimal predicted motifs.

A.4 Discussion

A.4.1 Sensitivity analysis- comparing sample sizes

A.4.1.1 Isonicotinamide

We chose six different datasets; 351, 551, 800, 1045, 3589 and 5727 hits for training dataset. The hydrogen-bond propensity was calculated for each dataset and the ROC was obtained (Table A.1).
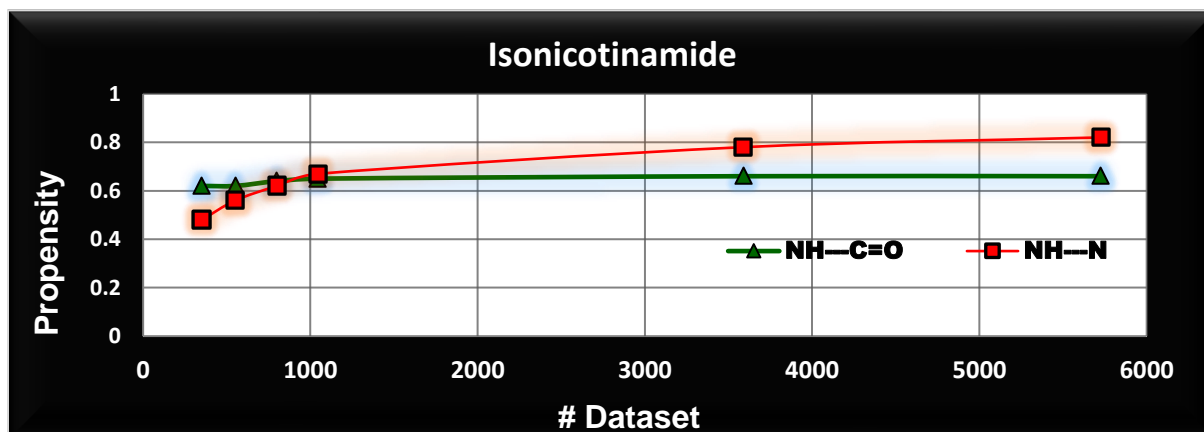
Figure A.13 Graph showing the propensity ranking for two interactions changes as we increases the dataset.

As the number of hits for each training dataset is increased, more structures are added into the model. As the dataset gets bigger, the training dataset deviates from the molecule of interest in terms of functionalities present on them. The larger dataset includes molecules with either aromatic N or carbamoyl functionalities as well as having them on two separate molecules in a crystal (co-crystals). In the case of isonicotinamide, more molecules with NH(amine)…N interactions are included in the dataset as we increase the training set. The propensity of NH(amine)…N increased, and it became rank 1 in the propensity order after 1000 dataset. The propensity of NH(amine)….C=O hydrogen-bond interaction also increased by 0.04 going from 351 to 5727 datasets.

A.4.2 Role of HBP in identifying stable polymorphs

A.4.2.1 Isonicotinamide[10]

The hydrogen-bond coordination table suggested that isonicotinamide has two hypothetical structures, both with similar propensity value but different coordination likelihood. Interestingly, the predicted structures match very well with the experimentally observed structures. Hypothetical

structure I resembles Form II and hypothetical structure II resembles form 1and the former is more likely structure compared to latter based on the satisfaction of each donor and acceptor group satisfaction (Figure A.14).

(a)	(b)	(c)
	Form 1[10]	Form II[11]
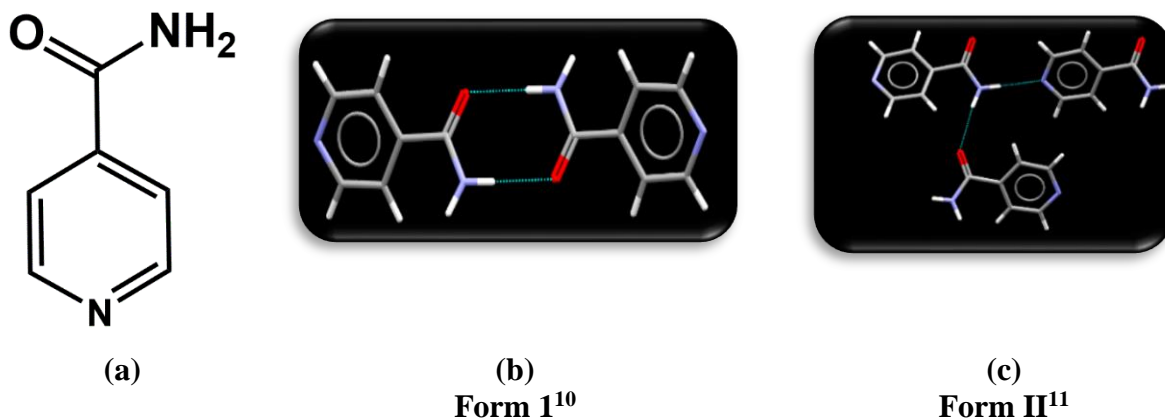
Figure A.14 Chemical diagram of isonicotinamide and two experimental known polymorphs.

A.4.2.2 4-Hydroxybenzamide

Experimentally, two synthon polymorphs of 4-hydroxybenzamide are known. In form 1, synthons 1, 2, and 3 ranked based on propensity were observed[12]. In form 2, synthon 1 and 4 were observed experimentally (Figure A.15)[13]. All conventional donors and acceptors are involved in the hydrogen-bonding and each donor and acceptor has same coordination in the both crystal structures. Based on hydrogen-coordination table, five hypothetical structures with different donor/acceptor coordination and propensity were predicted. The two most optimal structures indicated as hypothetical structure I and II have the same coordination for each donor and acceptor (primary NH amine can donate two, OH hydroxyl can donate one, carbamoyl C=O can accept two, and O(hydroxyl) can accept one) but propensities are different. The HBP tool is able to distinguish and predict both experimental forms correctly.

(a)	(b)	(c)
	Form 1[12]	Form 2[13]
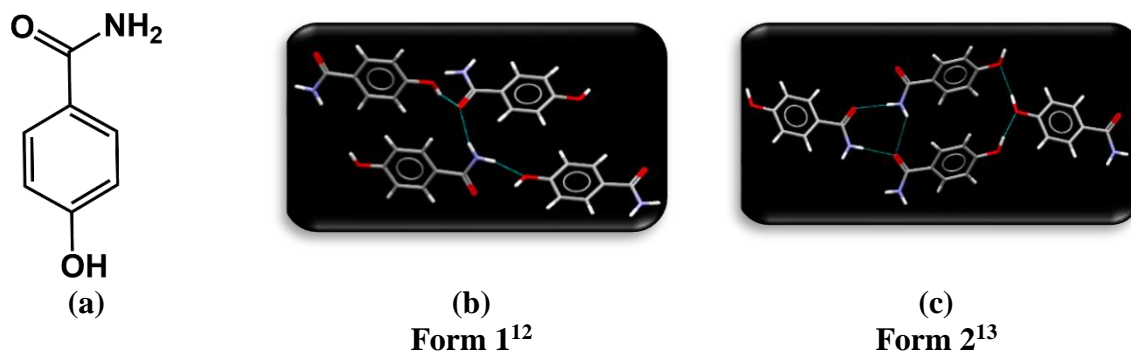
Figure A.15 Chemical diagram of hydroxybenzamide and two experimental known polymorphs.

Table A.9 Summary of two case studies for predicting stable polymorphs

Molecules	Predicted polymorphs	Experimental polymorphs	Does model predict the experimental polymorphs?
Isonicotinamide	2	2	Yes (both were predicted)
4-Hydroxybenzamide	5	2	Yes (both were predicted)

A.4.3 Using HBP to understand H-bond competition and their strength

The HBP of OH(acid)….C=O is 0.42 (0.25, 0.61) in benzoic acid. Experimentally, the interaction is observed in its structure[14] where the OH(acid) forms head to head dimer with the carbonyl oxygen of acid via R22(8) OH(acid)….O=C(acid) synthon. When a strong acceptor group is introduced, the HBP of OH(acid)…C=O decreases to 0.19 (0.08, 0.37) due to the presence of H-bond competition by a relatively stronger acceptor group. The maximum HBP value is 0.64 (0.42, 0.81) for OH(acid)…N(pyridine) which agrees with the experimentally observed interaction[15](Figure A.16).

(a)   (b)

Figure A.16 Hydrogen-bonding synthon observed in the crystal structure of (a) benzoic acid and (b) isonicotinic acid.

These HBP results were supported by electrostatic potential charges (Figure A.9). According to molecular electrostatic potentials, Pyridine N has the highest negative charge in isonicotinic acid making it better acceptor compared to carbonyl oxygen and therefore has higher chance of forming interaction with the OH of acid. Whereas in benzoic acid, there is no competition for the C=O, therefore, it forms hydrogen bond with the OH of acid. The HBP results were supported by MEPs and experimentally observed structures.

A.4.4 Importance of functional group definition when making comparisons

A.4.4.1 Replacing a functional group (isonicotinic acid to thioisonicotinic acid)

In order to make comparisons of HBP between two molecules, both molecules should have same definition of functional groups, same # of dataset. An example shown here is for isonicotinic and thioisonicotinic acid and we want to study how replacing C=O with C=S affects the resulting propensity. The propensity results are shown in table A.6. The HBP comparison of isonicotinic acid and thioisonicotinic acid indicates that when a strong acceptor such as C=O is replaced with a weaker acceptor group such as C=S, the propensity of OH(acid)…C=S (thionyl) decreases comparing to OH(acid)…C=O because S is less electronegative compared to oxygen as shown by MEPs calculations and the propensity of OH(acid)…N(aromatic) increases because C=S group provides less competition to aromatic N compared to C=O.

A.4.4.2 Importance of choosing correct functional group

The importance of choosing a correct functional group is discussed in this section.

A.4.4.2.1 Hydrogen bond propensity in 1,3-bisphenyl urea[16]

1,3-bisphenyl urea was chosen as a control experiment with two hydrogen-bond donors (NH) and one hydrogen-bond acceptor (C=O). The NH groups have only one choice for the acceptor group and it was established using propensity tool that both NH groups binds to C=O via bifurcated hydrogen-bond interactions. The prediction results match the experimental crystal structure. The propensity score decreases as we go from fragment 1 to fragment 3 indicating fragment 1 to be the best choice.

A.4.4.2.2 Hydrogen bond propensity 1,2-bis(pyridin-2-yl) urea[17]

Based on propensity all three fragments, 1,2-bis(pyridin-2-yl) urea is expected to have intramolecular hydrogen bonding between NH(I)…N(pyridine) intra which matches with the experimental observed interaction in the crystal structure. Fragment 1 and 2 predicts NH(II)…O=C to be the second-best interaction and it was also observed in the crystal structure as well. Fragment 1 is preferred over fragment 2 because it clearly distinguish between two possible intermolecular interactions (NH(II) binding to N(pyridine) or NH(II) binding to C=O).

(a)   (b)
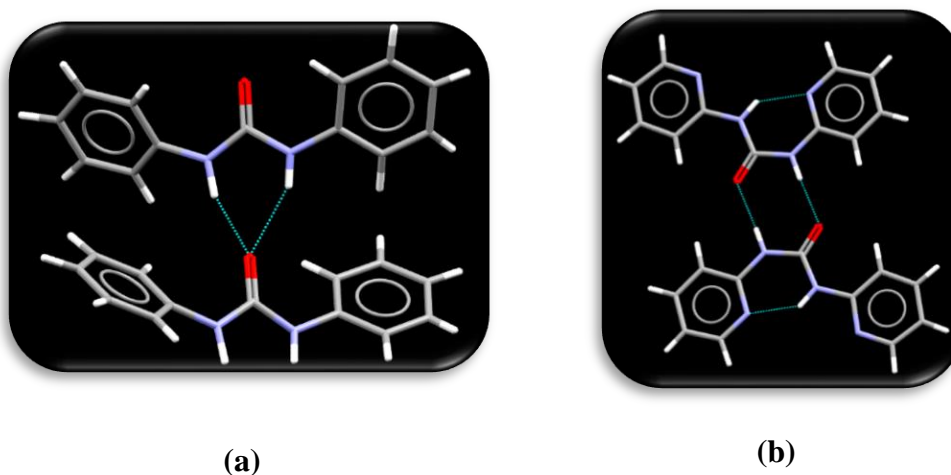
Figure A.17 Synthons observed in the crystal structures of (**a**) 1,3-bisphenyl urea, and (**b**) 1,2-bis(pyridin-2-yl) urea[17].

2.4.4.2.3 Hydrogen bond propensity 1,3-bis(pyridin-2-yl) urea[18] and 1,4-bis(pyridin-2-yl) urea[19]

Based on hydrogen-bond propensity, NH(I)…N(pyridine) intermolecular interaction was predicted to be the best interaction in both 1,3-bis(pyridin-2-yl) urea and 1,4-bis(pyridin-2-yl) urea via fragment 1 and 3. Experimentally, NH(I)…N(pyridine) is observed in both experimental structures. Fragment 2 predicted NH(II)…O=C as the best interaction.

(a)   (b)
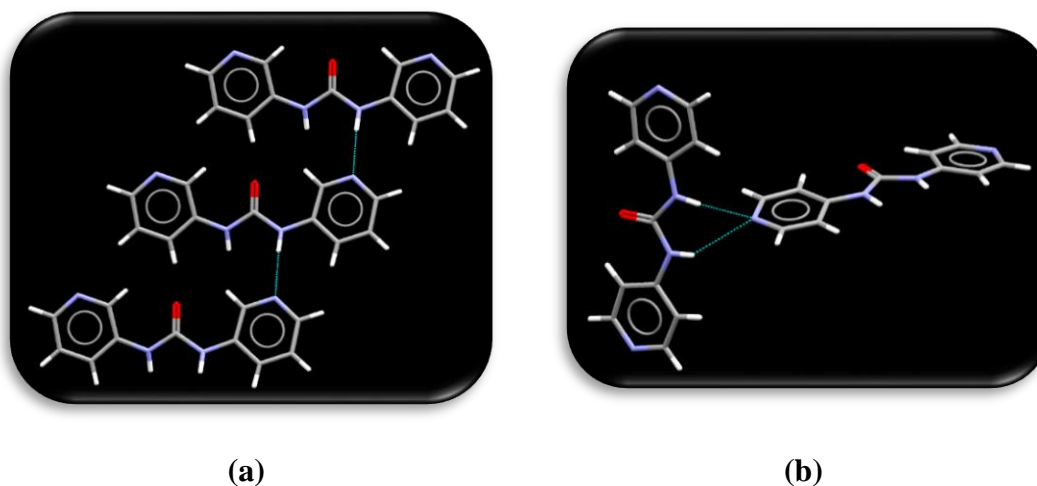
Figure A.18 Synthons observed in the crystal structures of (**a**) 1,3-bis(pyridin-2-yl) urea[18] and (**b**) 1,4-bis(pyridin-2-yl) urea[19].

Therefore, fragment 1 is the best fragment for this study as in 4 out of 4 molecules, it predicted the experimentally observed synthon correctly. These results also suggest that choosing functional group is a key step in the predicting the propensity and ranking as well as the magnitude of propensity can vary drastically based on choice of functional groups.

Table A.9 Summary of functional fragmentations in bis urea molecules

Molecules	Experimentally observed synthon	Which fragment predicted correctly?
1,3-bisphenyl urea	Bifurcated H-bond	Fragment 1-3
1,2-bis(pyridin-2-yl) urea	Intra (NH…N) Inter (NH…O=C)	Fragment 1
1,3-bis(pyridin-2-yl) urea	Inter (NH…N)	Fragment 1
1,4-bis(pyridin-2-yl) urea	Inter (NH…N)	Fragment 1

A.4.5 Limitation

A.4.5.1 Differentiation of equivalent synthon polymorphs

Benzamide has three known experimental forms. All three structures are built up by the same primary building unit, a dimer of benzamide. These ribbons are packed in three distinct patterns for each crystal structure. These significant differences in the observed packing motifs are triggered by the second-ranking intermolecular interactions, the π-π interactions of the benzyl moieties. Form I is dominated by shifted π stacks while in Form II and Form III a herringbone pattern like arrangement is observed.

We used hydrogen-bond propensity model to determine if all three forms or the distinct differences in the second-ranked intermolecular interaction will be predicted by the HBP tool. Since the propensity tool is solely based on hydrogen-bond interactions between conventional hydrogen-bond donors and acceptors and same type of synthon is observed in all three crystal structures of benzamide. The NH(carbamoyl)….C=O hydrogen-bond interaction has same propensity score of 0.73 (0.62, 0.82) and same coordination score (two for amine donor and two for C=O group acceptor) for all three structures. The software predicted all three forms to lie at same spot in the propensity-coordination chart. Therefore, it fails to predict the secondary interactions.

A.5 Conclusion

The object of this research was a proof of concept to test hydrogen bond propensity tool of Cambridge crystallographic database center. At the outset of this discussion, it is important to note that the propensity tool is not a means for predicting polymorphism.[1] The choice of the term "propensity" is meant to give an indication of the possibility of multiple crystal forms; in that sense it can provide some guidance on the amount of "time and money"[20] that might be expended in experimentally searching for multiple crystal forms. Nevertheless, it is a statistical tool, and there is no guarantee that any particular compound will adhere to the statistics.

The tool is sensitive to the number of datasets chosen for analysis. In order to have chemical similarity in the training model as well have decent training model, the dataset between 300-500 will be considered optimal. The role of HBP in identifying stable polymorphs was done (isonicotinamide and 4-hydroxybenxamide). In both molecules, the HBP model predicts the experimentally observed synthon polymorphs correctly. The tool is good to understand the H-bond competition between conventional donors and acceptors and determine the most optimal synthon and its strength. It is sensitive to functional group definition and plays an important role in the final outcome of propensity ranking. Therefore, careful analysis of functional group must be done to make comparisons between different molecules. The fragment 1 functional group was the right choice of functional groups in bis urea molecules as the predictions made by HBP matches 100% with the experimentally observed synthons in four molecules studied.

A.6 References

1.   Nauha, E.; Bernstein, J., *Journal of Pharmaceutical Sciences* **2015,** *104* (6), 2056-2061.

2.   Price, S. L.; Reutzel-Edens, S. M., *Drug Discovery Today* **2016,** *21* (6), 912-923.

3.   Feeder, N.; Pidcock, E.; Reilly, A. M.; Sadiq, G.; Doherty, C. L.; Back, K. R.; Meenan, P.; Docherty, R., *Journal of Pharmacy and Pharmacology* **2015,** *67* (6), 857-868.

4.   Galek, P. T. A.; Allen, F. H.; Fabian, L.; Feeder, N., *CrystEngComm* **2009,** *11* (12), 2634-2639.

5.   Wood, P. A.; Feeder, N.; Furlow, M.; Galek, P. T. A.; Groom, C. R.; Pidcock, E., *CrystEngComm* **2014,** *16* (26), 5839-5848.

6.   Galek, P. T. A.; Fabian, L.; Motherwell, W. D. S.; Allen, F. H.; Feeder, N., *Acta Crystallographica Section B* **2007,** *63* (5), 768-782.

7.   Galek, P. T. A.; Fabian, L.; Allen, F. H., *CrystEngComm* **2010,** *12* (7), 2091-2099.

8.   Groom, C. R.; Allen, F. H., *Angewandte Chemie International Edition* **2014,** *53* (3), 662-671.

9.   Galek, P. T. A.; Chisholm, J. A.; Pidcock, E.; Wood, P. A., *Acta Crystallographica Section B* **2014,** *70* (1), 91-105.

10.   Aakeröy, C. B.; Beatty, A. M.; Helfrich, B. A.; Nieuwenhuyzen, M., *Crystal Growth & Design* **2003,** *3* (2), 159-165.

11.   Eccles, K. S.; Deasy, R. E.; Fabian, L.; Braun, D. E.; Maguire, A. R.; Lawrence, S. E., *CrystEngComm* **2011,** *13* (23), 6923-6925.

12.   Hansen, L. K.; Perlovich, G. L.; Bauer-Brandl, A., *Acta Crystallographica Section E* **2007,** *63* (5), o2362.

13.   Tothadi, S.; Desiraju, G. R., *Crystal Growth & Design* **2012,** *12* (12), 6188-6198.

14.   Sim, G. A.; Robertson, J. M.; Goodwin, T. H., *Acta Crystallographica* **1955,** *8* (3), 157-164.

15.   Takusagawa, F.; Shimada, A., *Acta Crystallographica Section B* **1976,** *32* (6), 1925-1927.

16.   Deshapande, S. V.; Meredith, C. C.; Pasternak, R. A., *Acta Crystallographica Section B* **1968,** *24* (10), 1396-1397.

17.   Corbin, P. S.; Zimmerman, S. C., *Journal of the American Chemical Society* **2000,** *122* (15), 3779-3780.

18.   Kumar, D. K.; Jose, D. A.; Das, A.; Dastidar, P., *Chemical Communications* **2005,** (32), 4059-4061.

19.   Chandran, S. K.; Nath, N. K.; Cherukuvada, S.; Nangia, A., *Journal of Molecular Structure* **2010,** *968* (1), 99-107.

20.   Dutton, D., *Journal of the American Chemical Society* **1964,** *86* (8), 1654-1654.

Appendix B - NMR of target molecules

B.1 ¹H NMR for pyrazole based target molecules
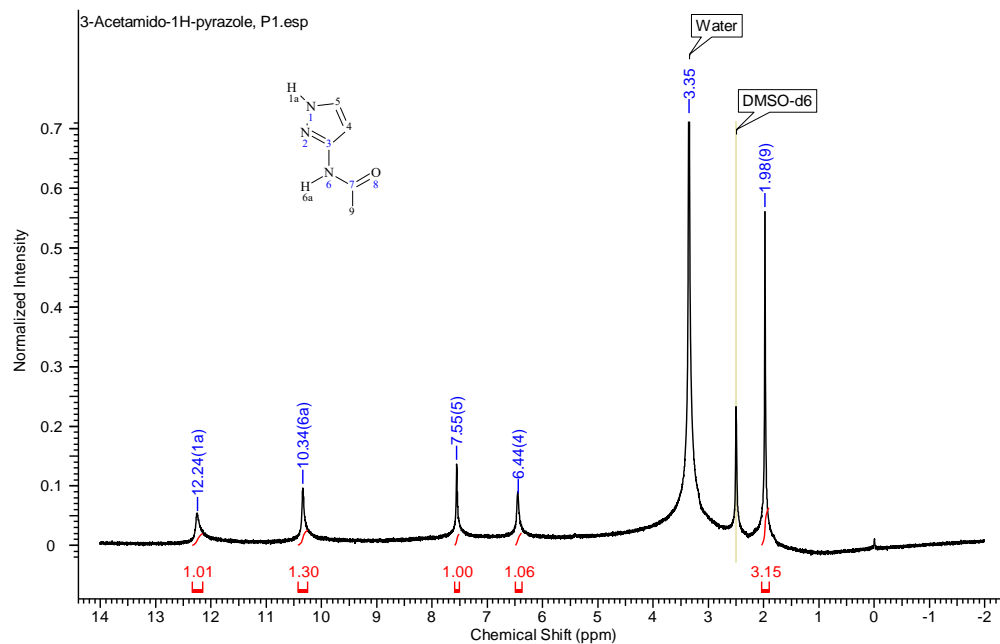
Figure B.1.1 ¹H NMR of 3-acetamido-1H-pyrazole, **P1**.
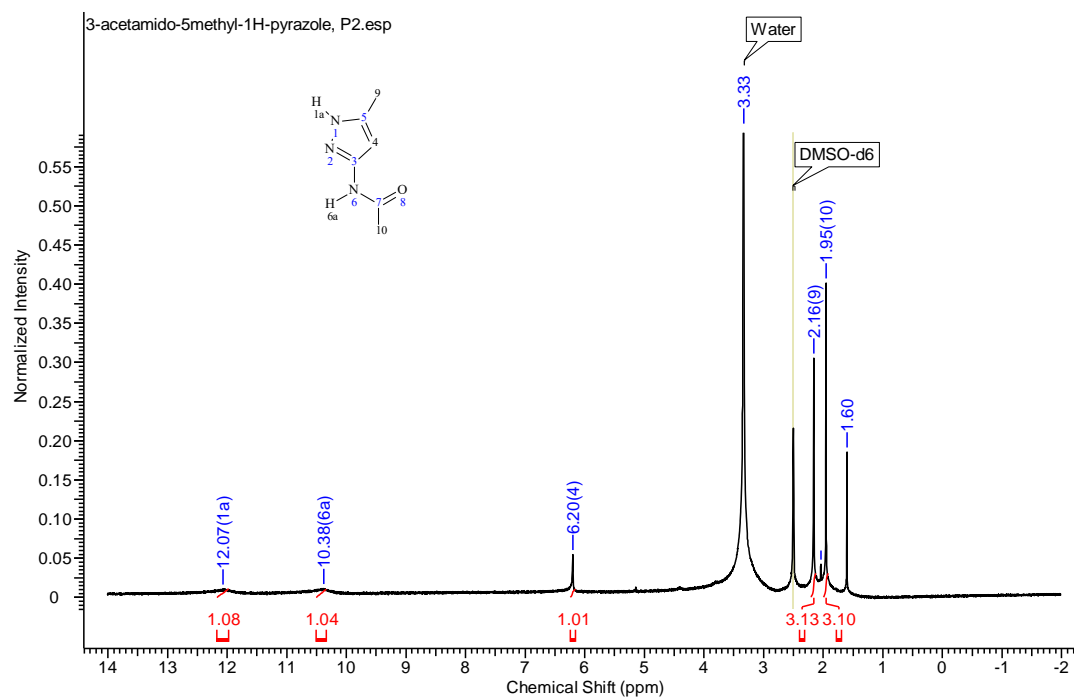
Figure B.1.2 ¹H NMR of 3-acetamido-5methyl-1H-pyrazole, **P2**.
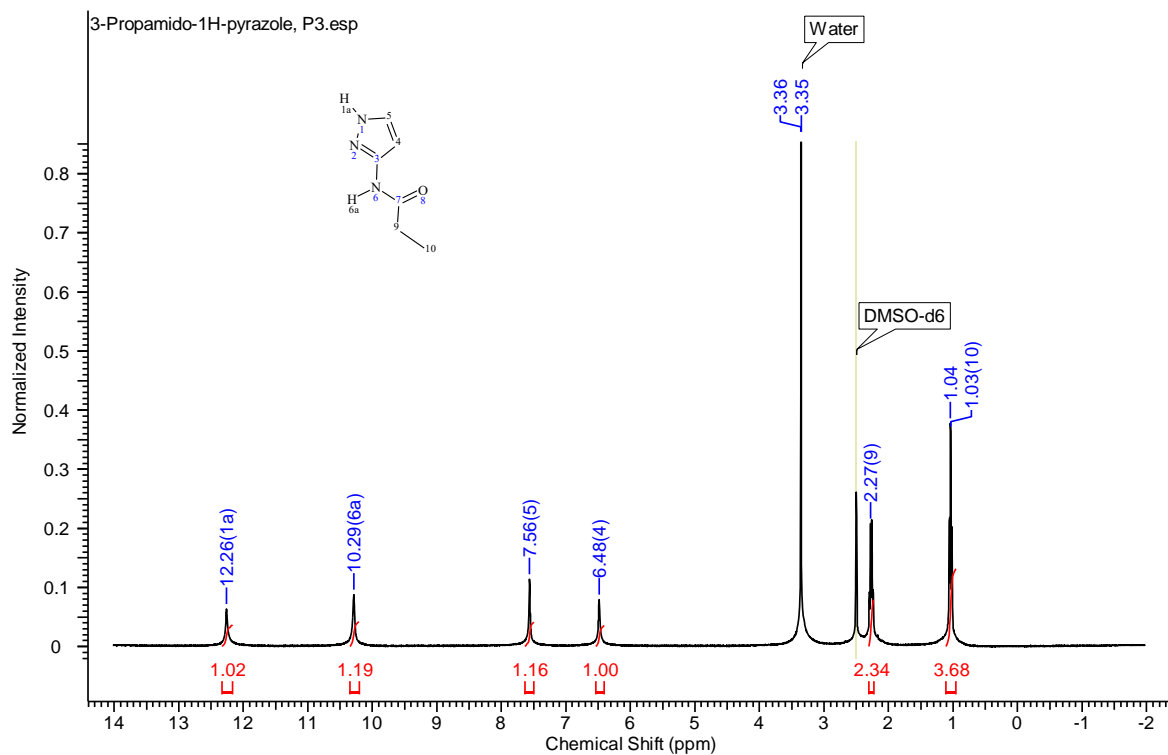
Figure B.1.3 ¹H NMR of 3-propamido-1H-pyrazole, **P3**.
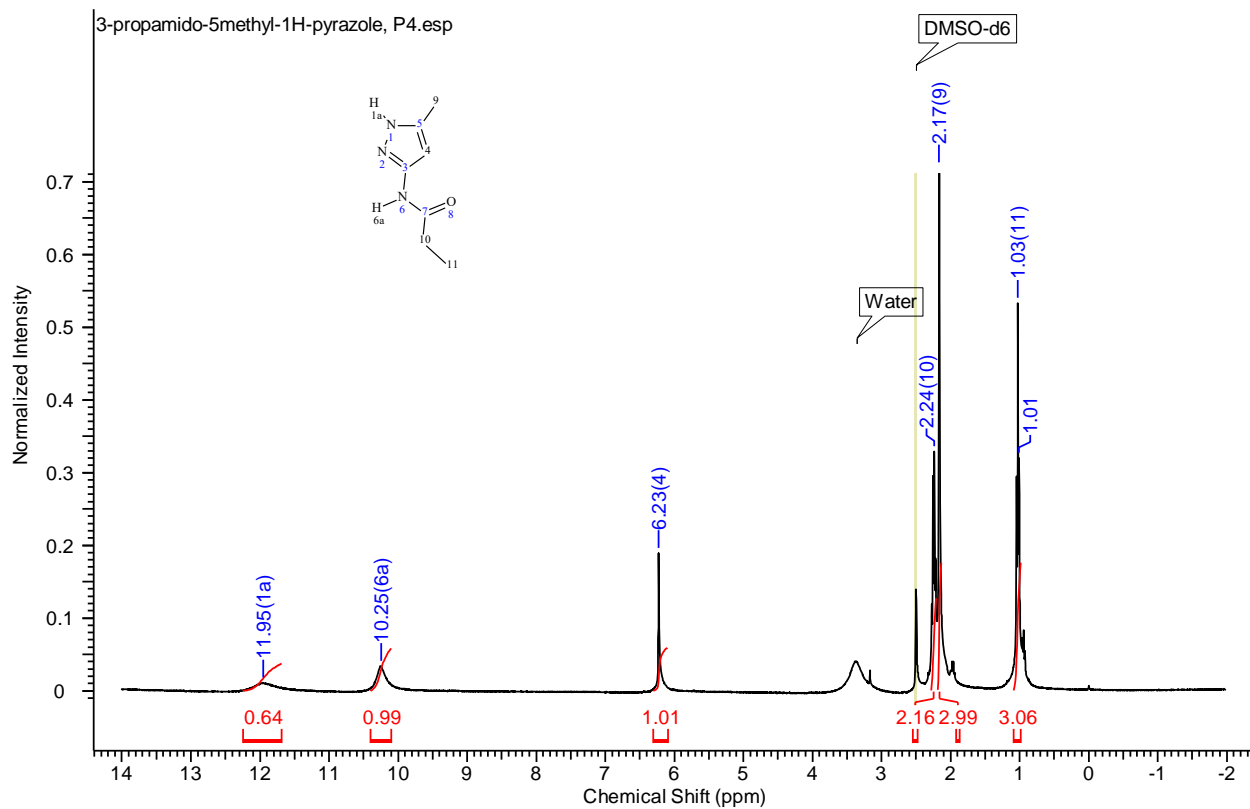
Figure B.1.4 ^{1}H NMR of 3-propamid

o-5methyl-1H-pyrazole, **P4**.
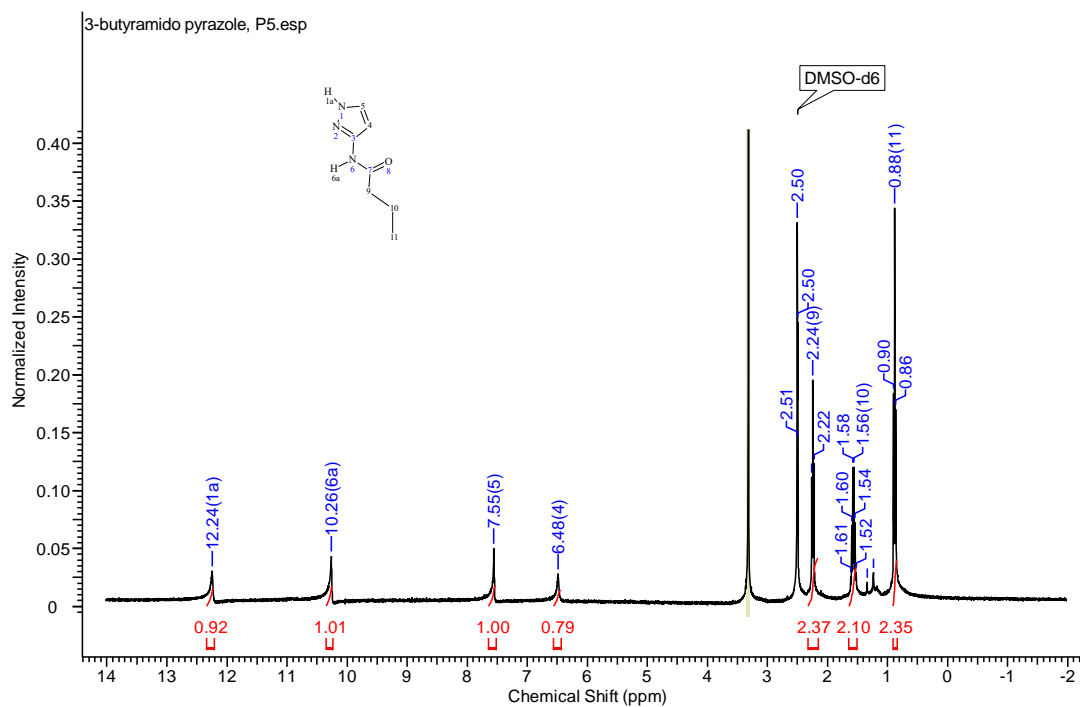
Figure B.1.5 NMR spectrum of 3-butyramido pyrazole, **P5**
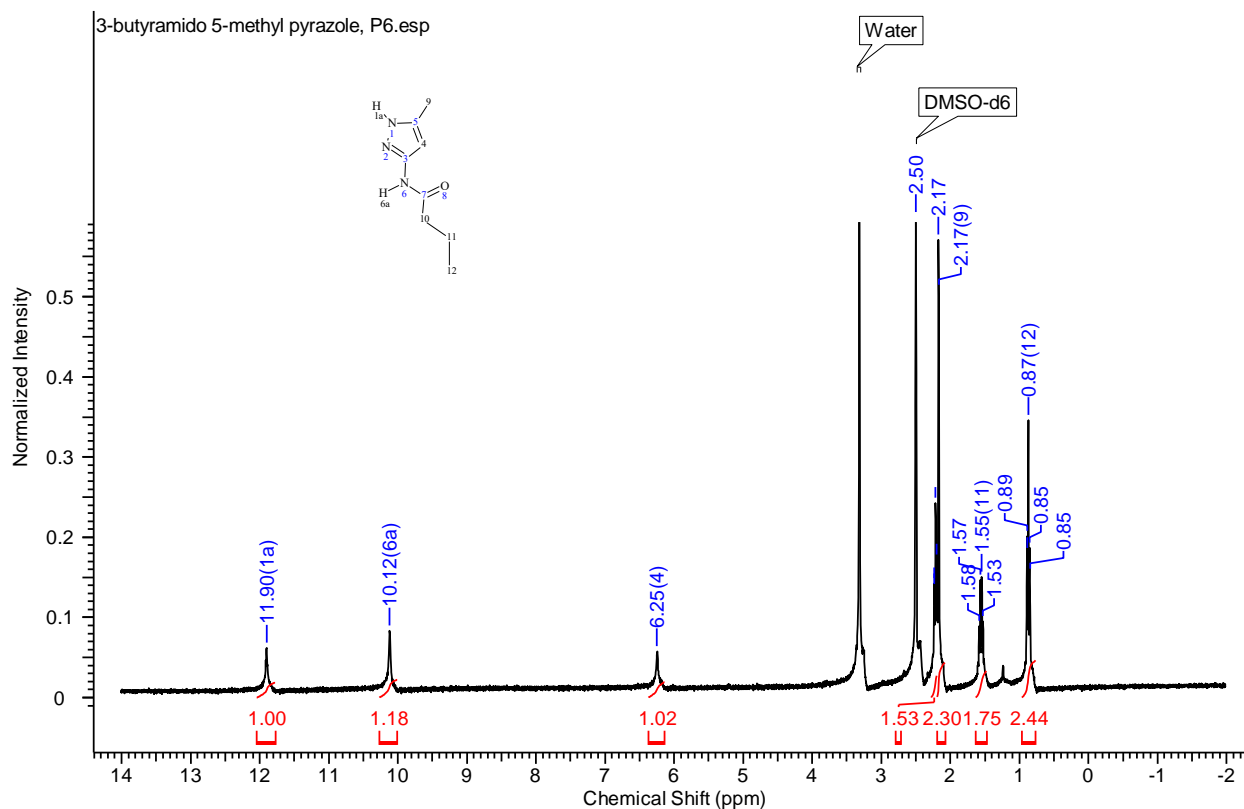
Figure B.1.6 NMR spectrum of 3-butyramido 5-methyl pyrazole, **P6**
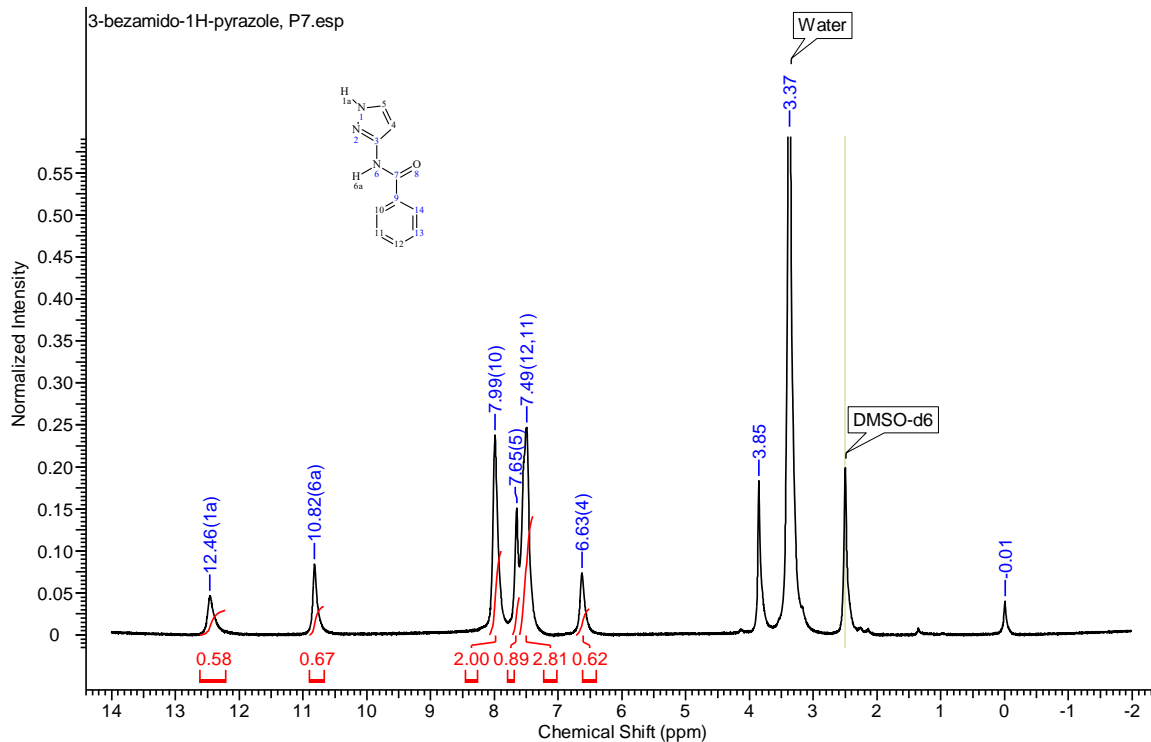
Figure B.1.7 ^{1}H NMR of 3-bezamido-1H-pyrazole, **P7**.
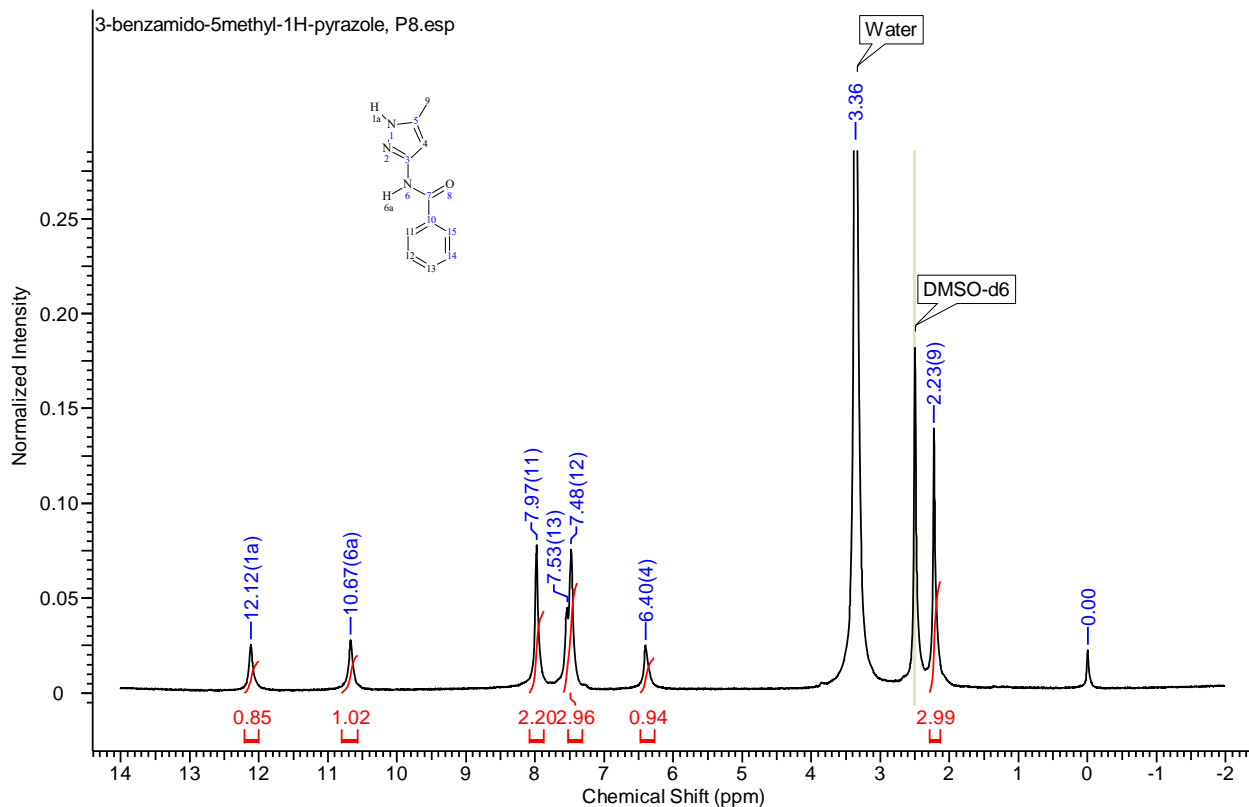
Figure B.1.8 ^{1}H NMR of 3-benzamido-5methyl-1H-pyrazole, **P8**.
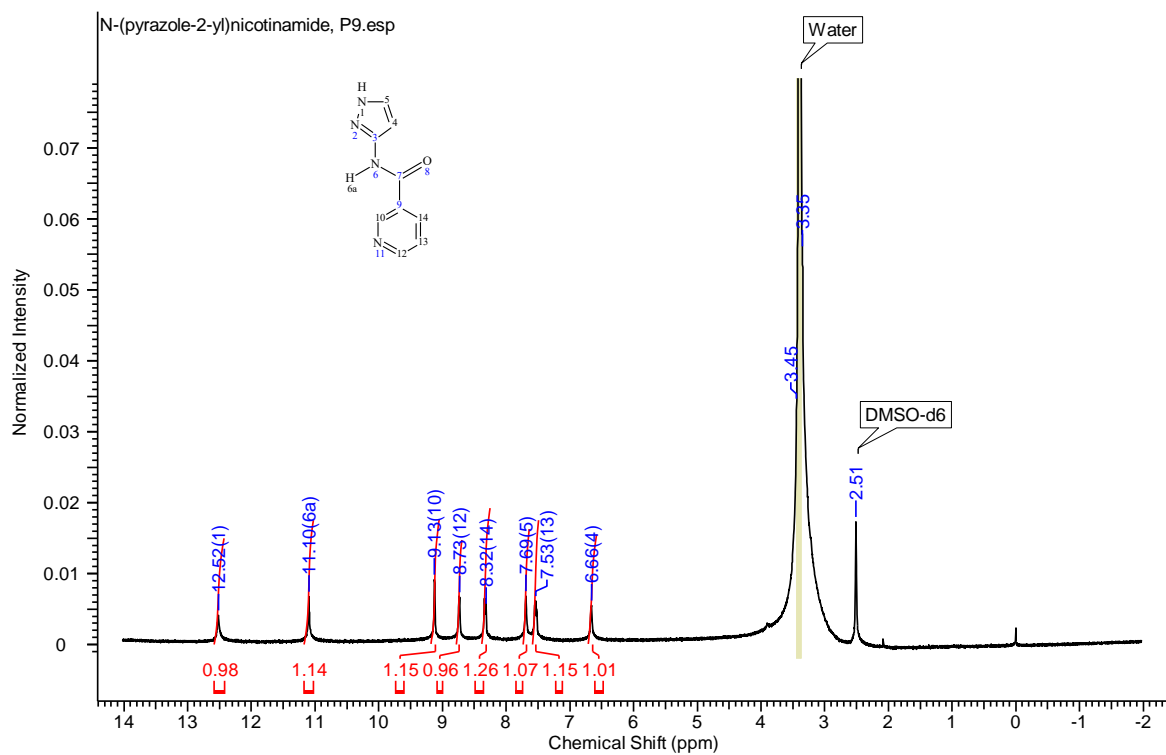
Figure B.1.9 ^{1}H NMR of N-(pyrazole-2-yl)nicotinamide, **P9**.
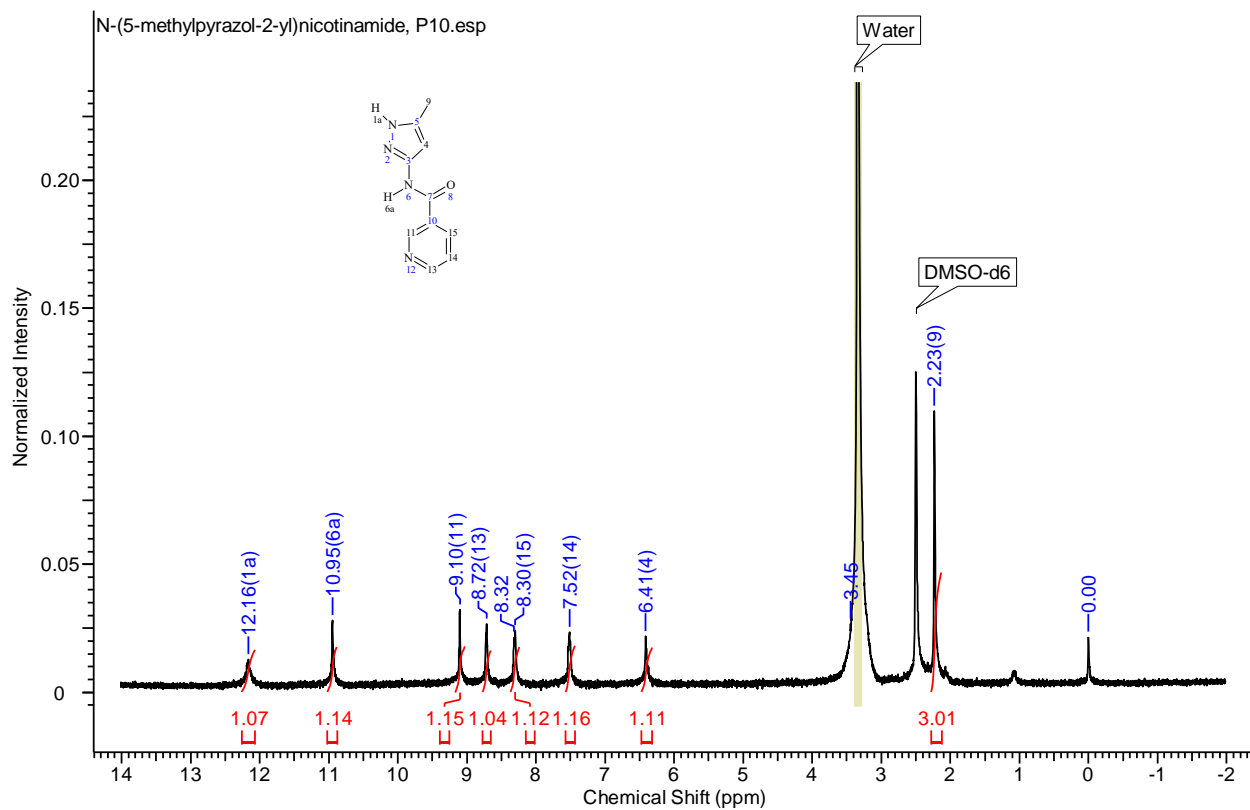
Figure B.1.10 NMR spectrum of N-(5-methylpyrazol-2-yl)nicotinamide, **P10**
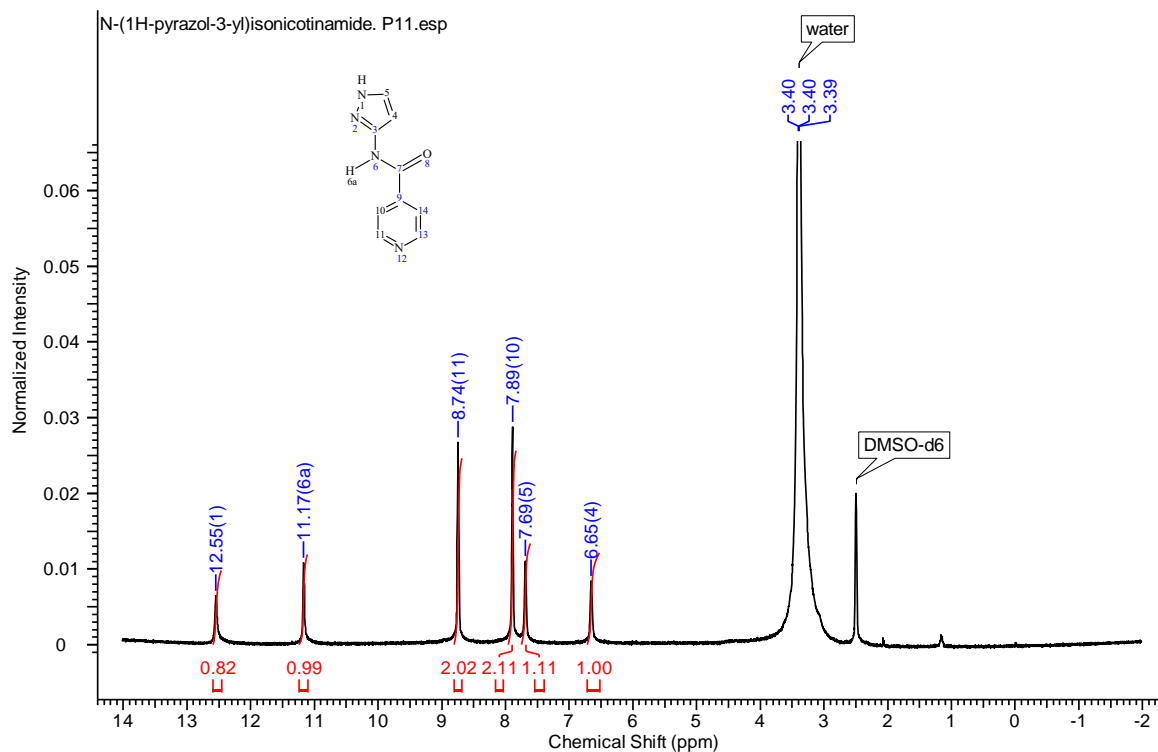
Figure B.1.11 NMR spectrum of N-(5-pyrazol-2-yl)isonicotinamide, **P11**
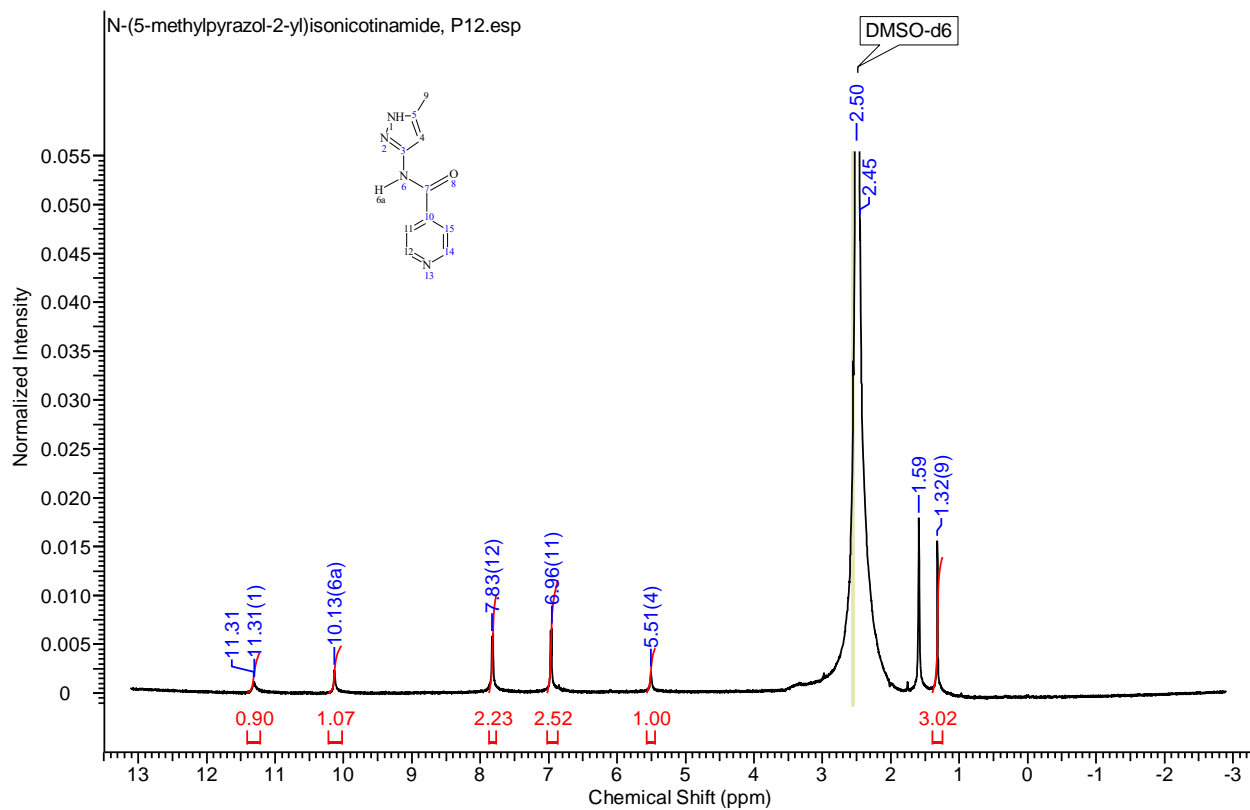
Figure B.1.12 NMR spectrum of N-(5-methylpyrazol-2-yl)isonicotinamide, **P12**

B.2 ^{1}H NMR for thiazole based target molecules
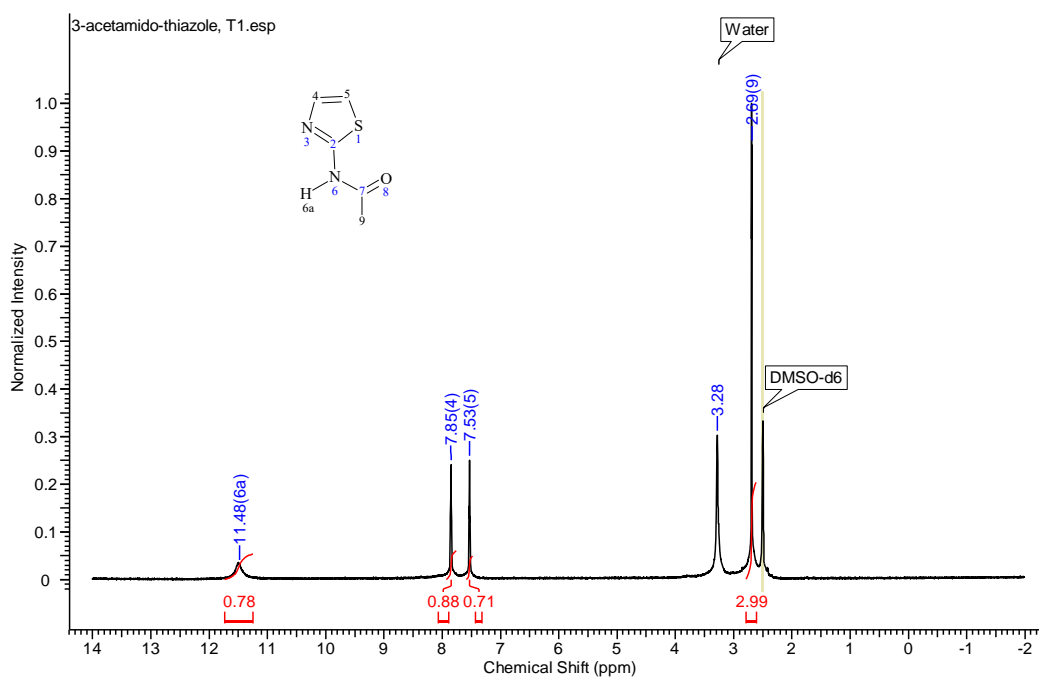
Figure B.2.1 NMR spectrum of 3-acetamido-thiazole, **T1**
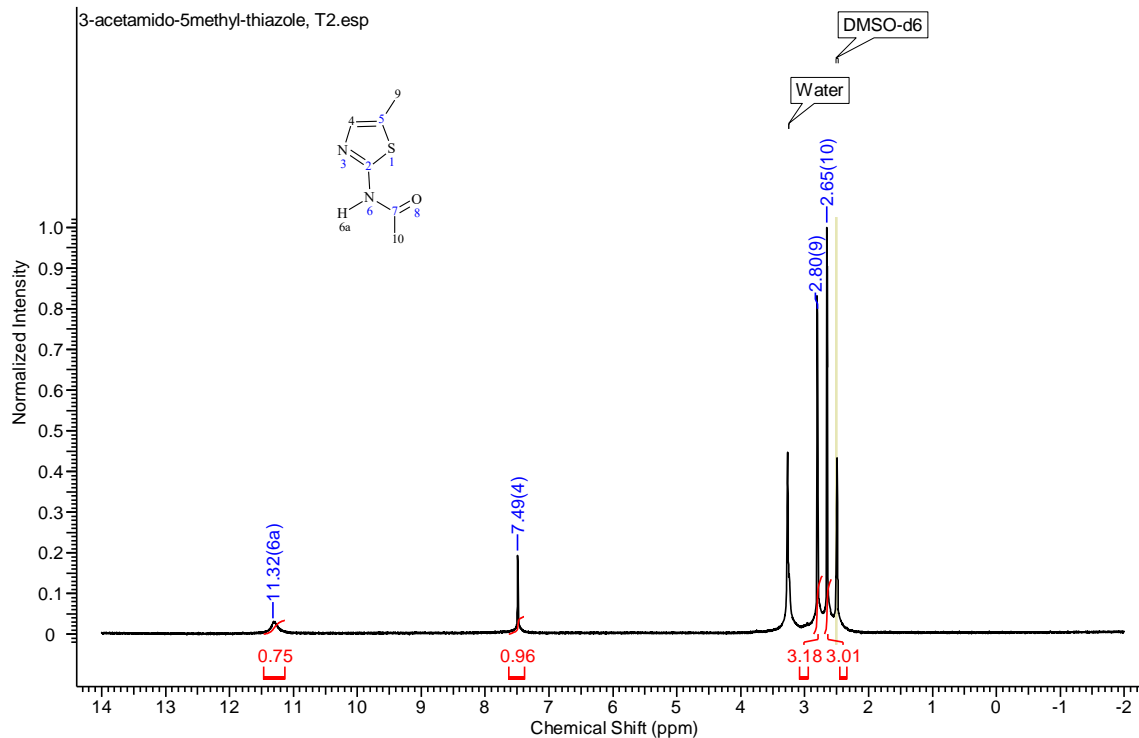
Figure B.2.2 NMR spectrum of 3-acetamido-5methyl-thiazole, **T2**
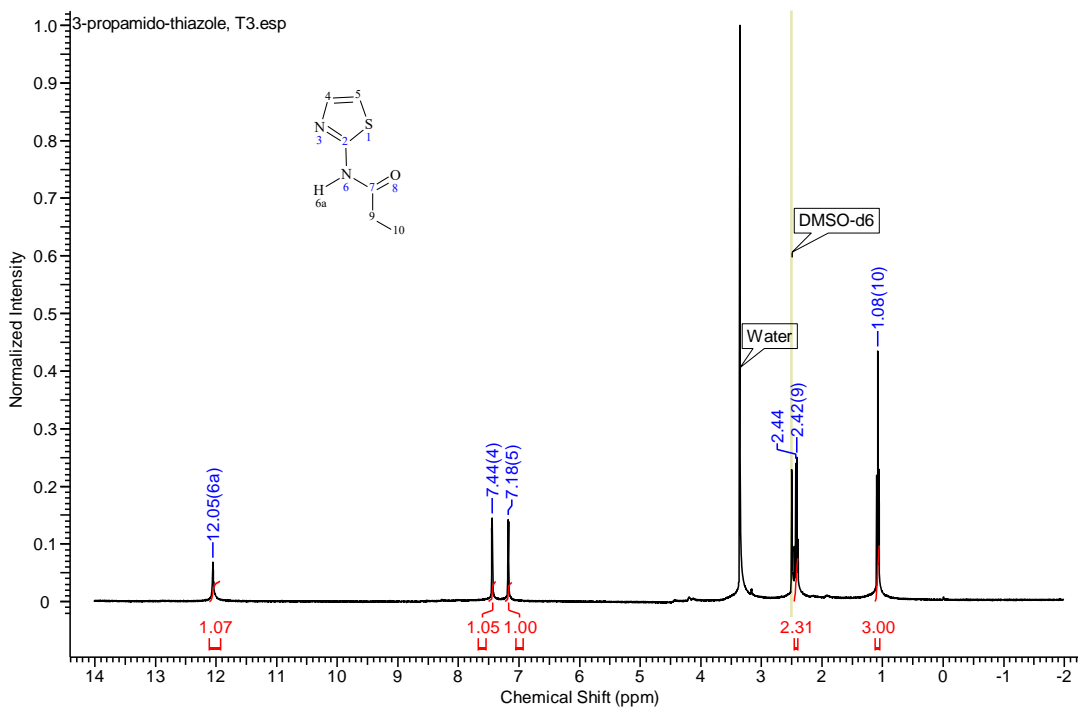
Figure B.2.3 NMR spectrum of 3-propamido-thiazole, **T3**
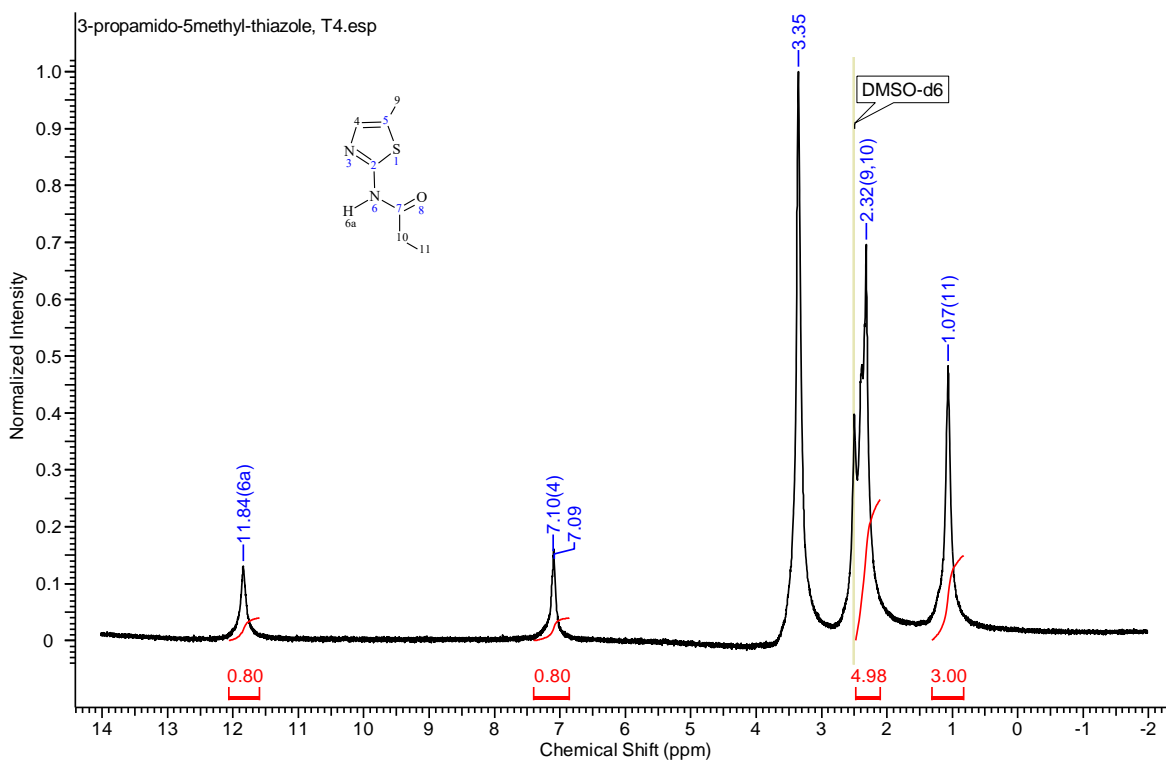
Figure B.2.4 NMR spectrum of 3-propamido-5methyl-thiazole, **T4**
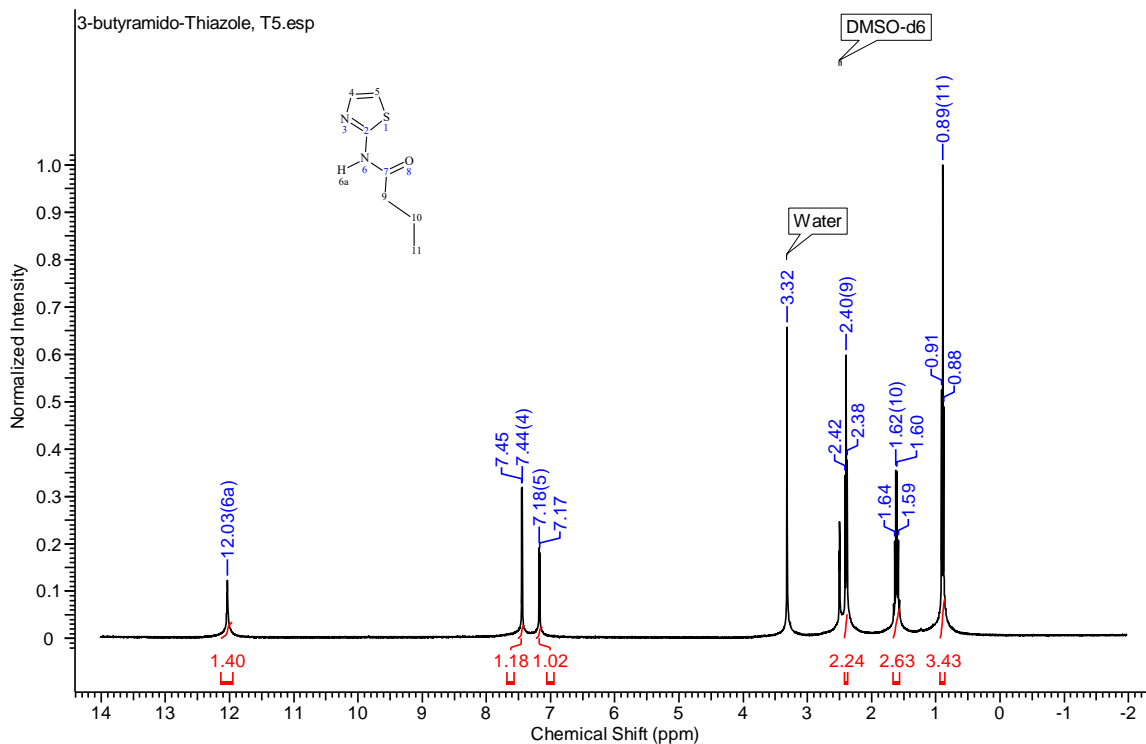
Figure B.2.5 NMR spectrum of 3-Butyramido-thiazole, **T5**
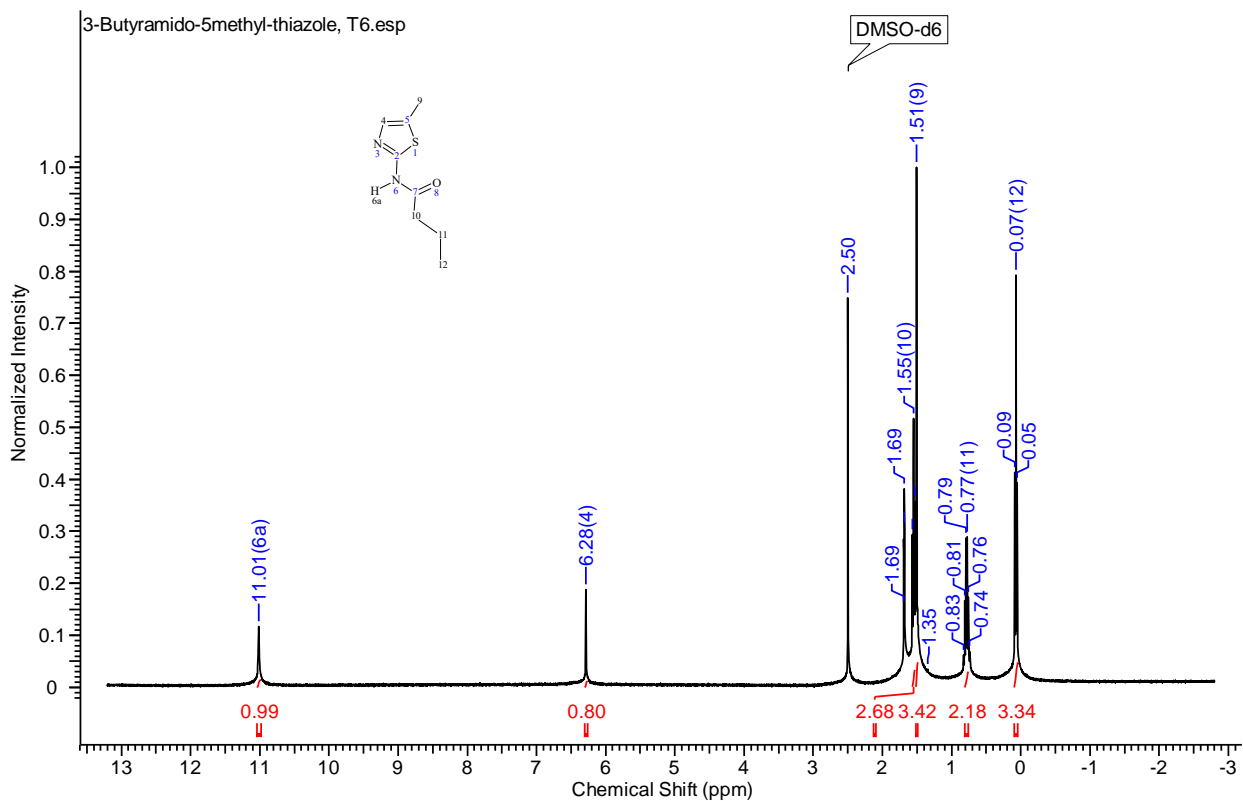
Figure B.2.6 NMR spectrum of 3-Butyramido-5methyl-thiazole, **T6**
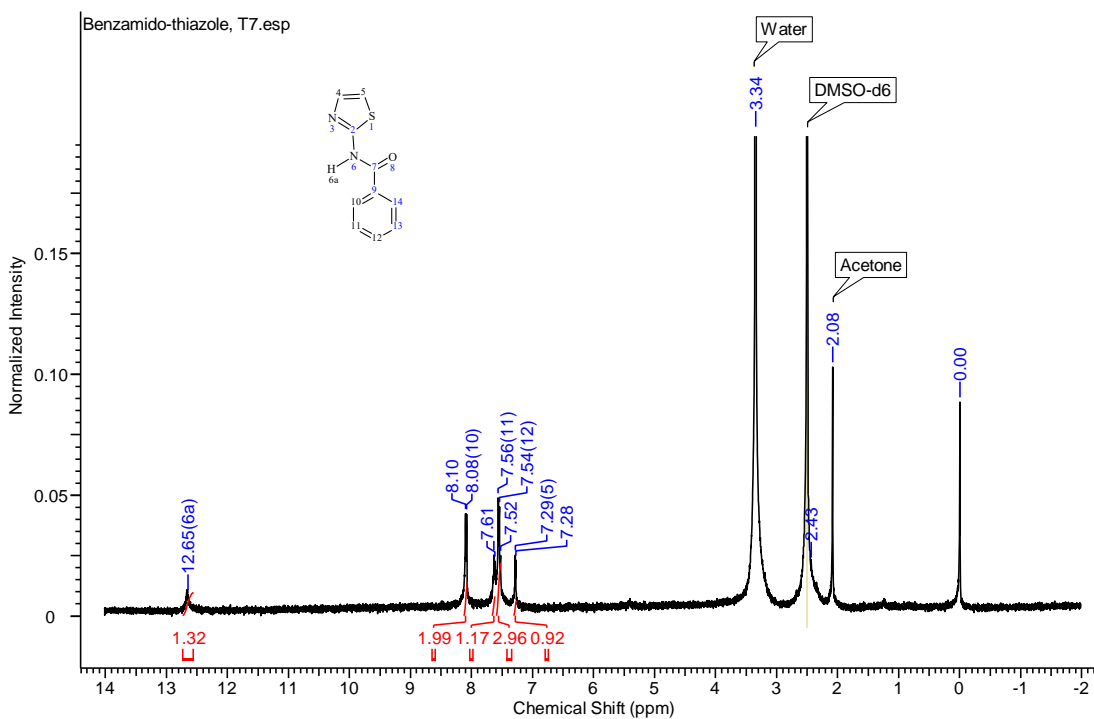
Figure B.2.7 NMR spectrum of Benzamido-thiazole, **T7**
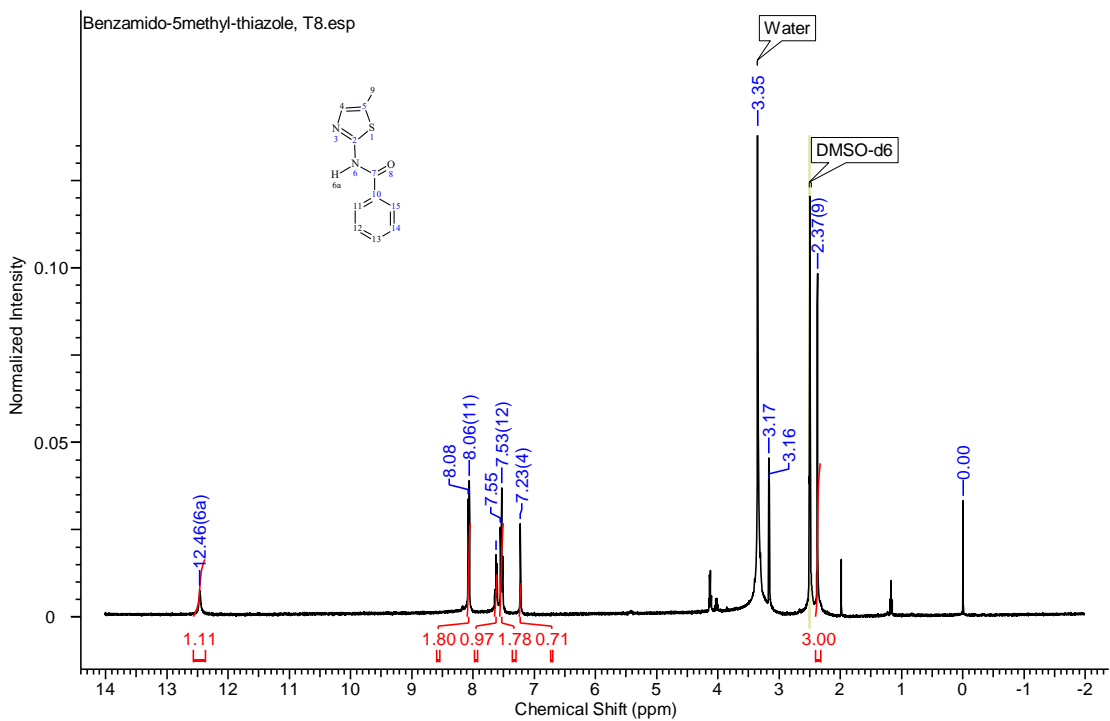
Figure B.2.8 NMR spectrum of Benzamido-5methyl-thiazole, **T8**
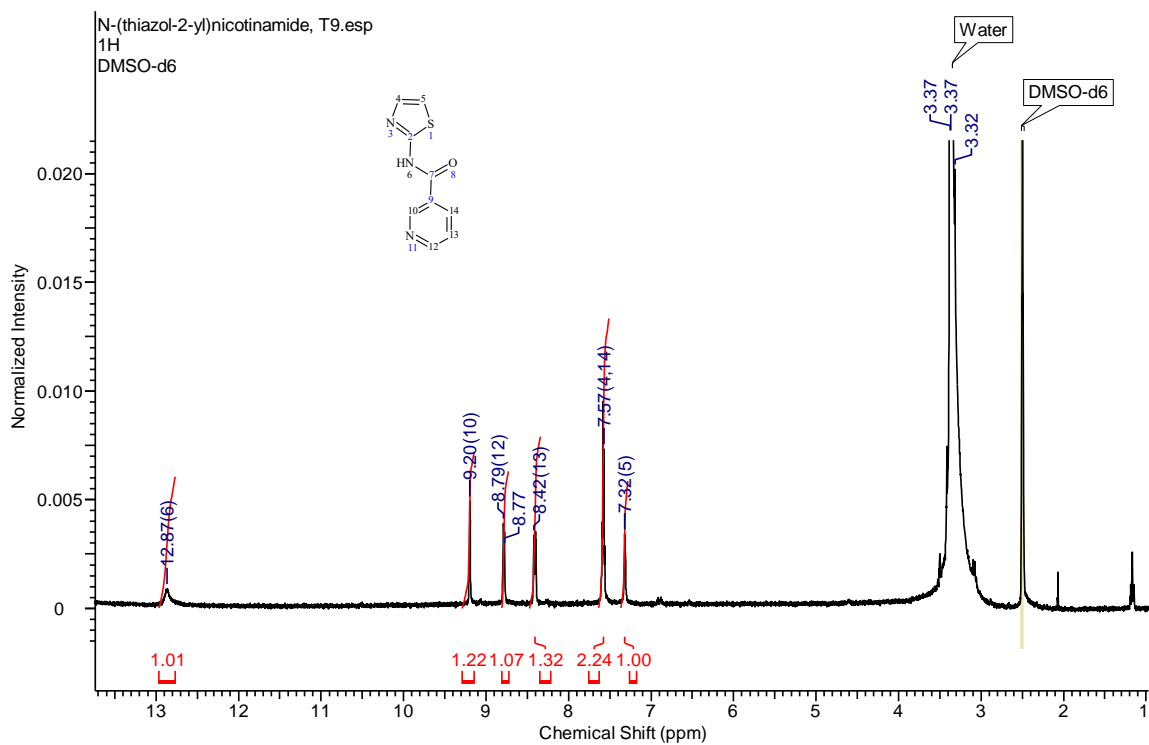
Figure B.2.9 NMR spectrum of N-(thiazol-2-yl)nicotinamide, **T9**
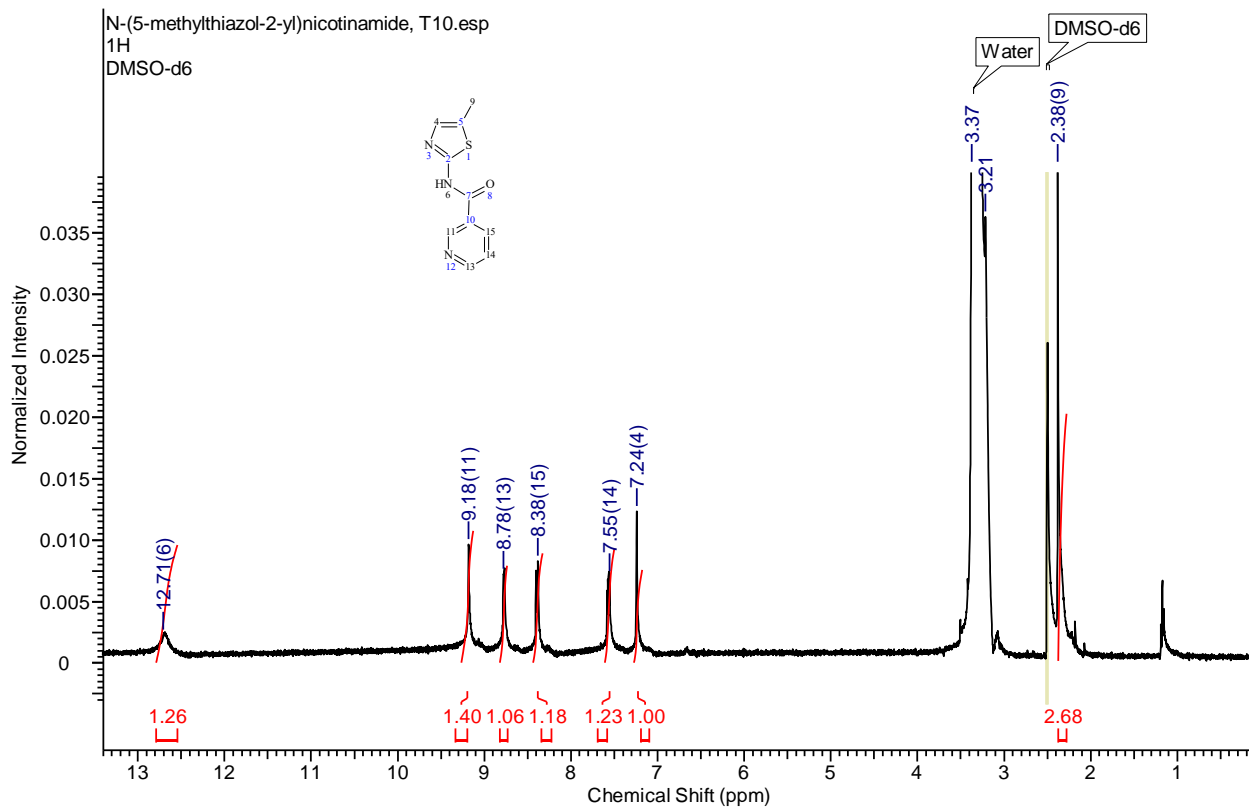
Figure B.2.10 NMR spectrum of N-(5-methylthiazol-2-yl)nicotinamide, **T10**
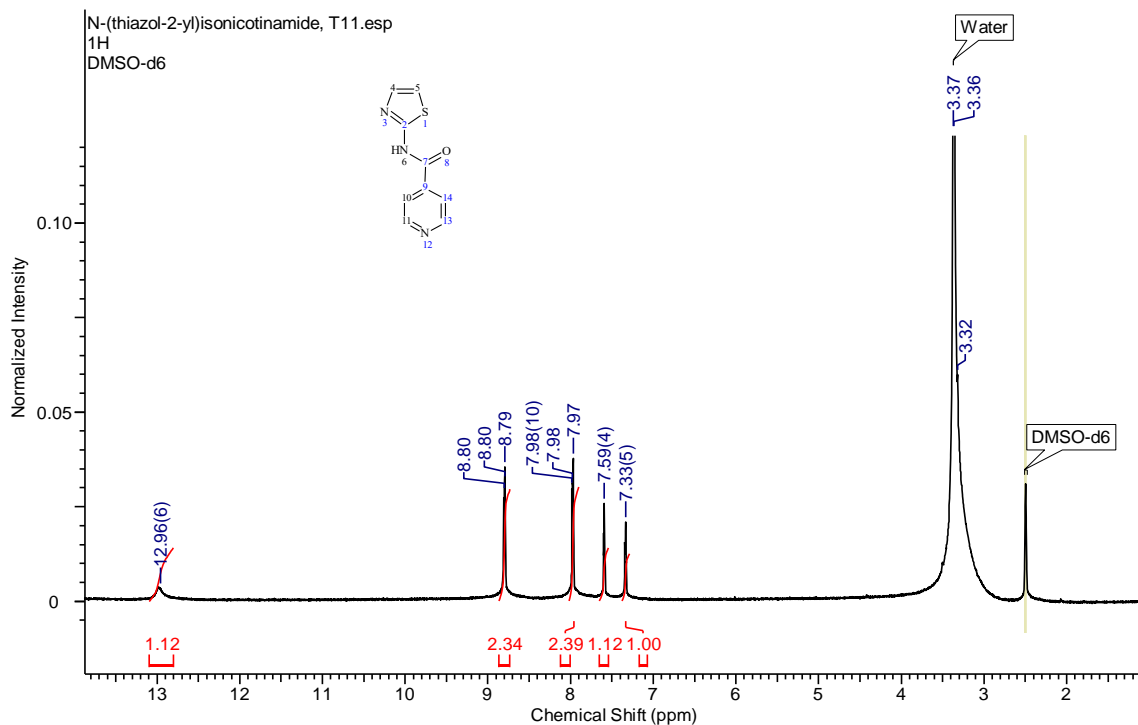
Figure B.2.11 NMR spectrum of N-(thiazol-2-yl)isonicotinamide, **T11**
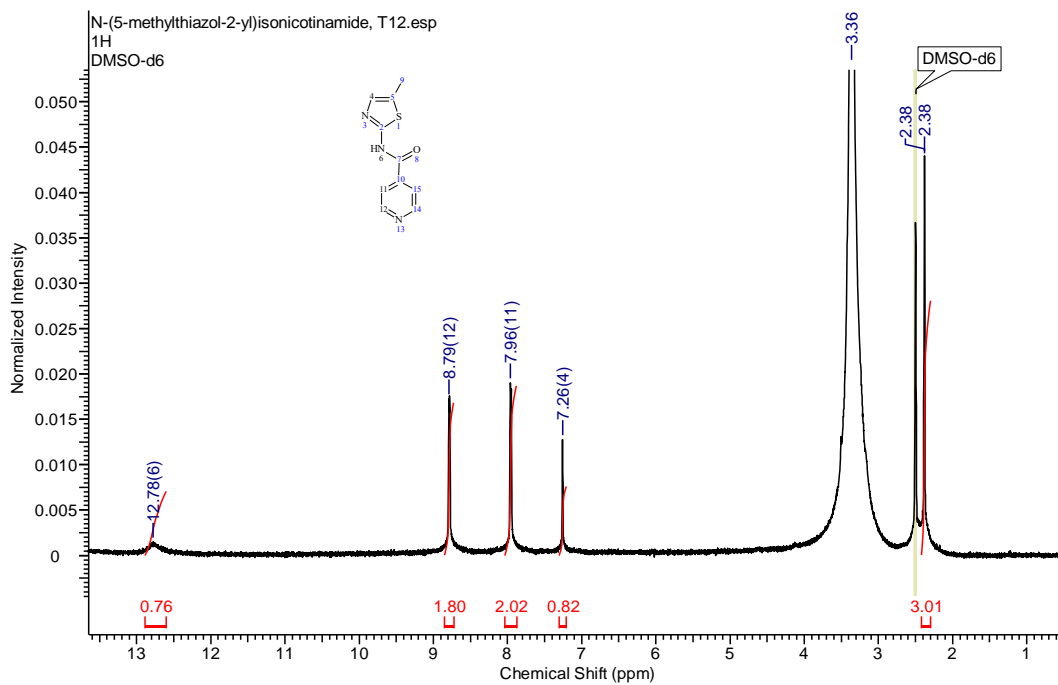
Figure B.2.12 NMR spectrum of N-(5-methylthiazol-2-yl)isonicotinamide, **T12**

B.3 ¹H NMR for activated halogen-bond donors
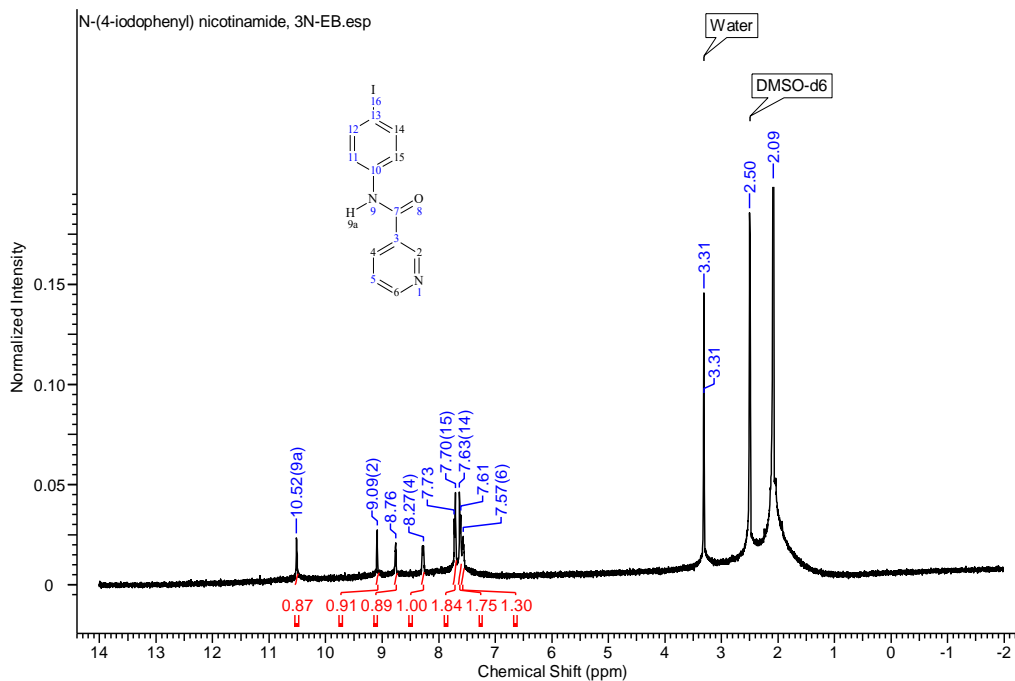
Figure B.3.1 NMR spectrum of N-(4-iodophenyl) nicotinamide, **3N-EB**
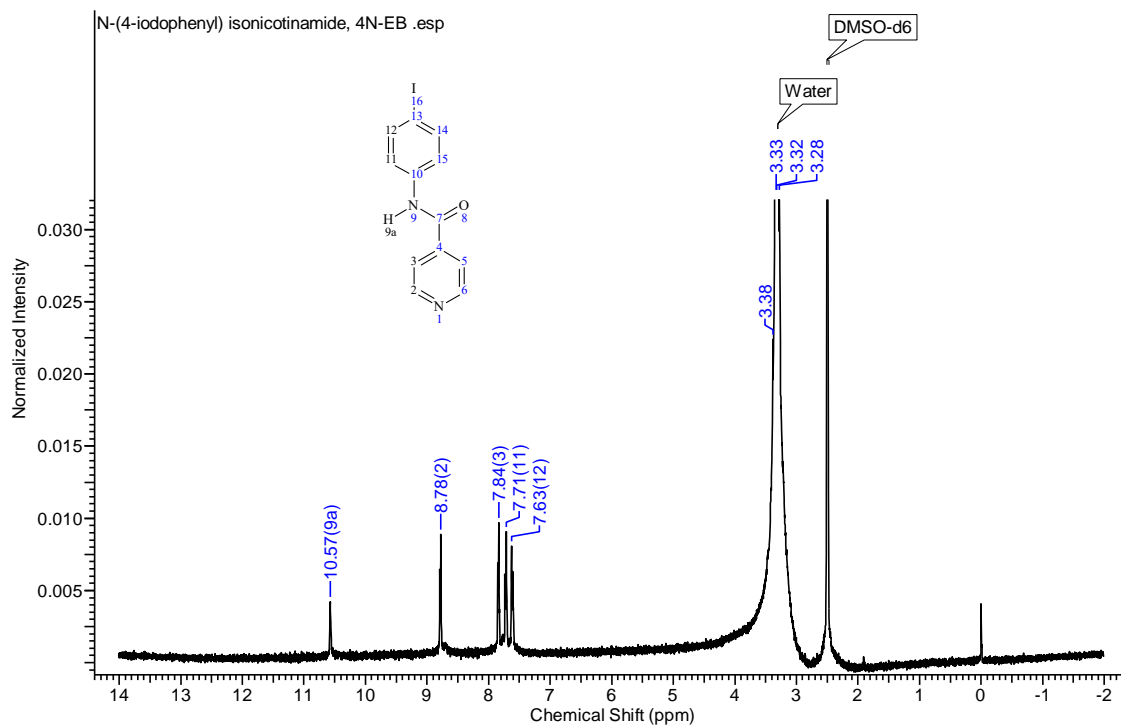
Figure B.3.2 ^{1}H NMR spectrum of N-(4-iodophenyl) isonicotinamide, **4N-EB**
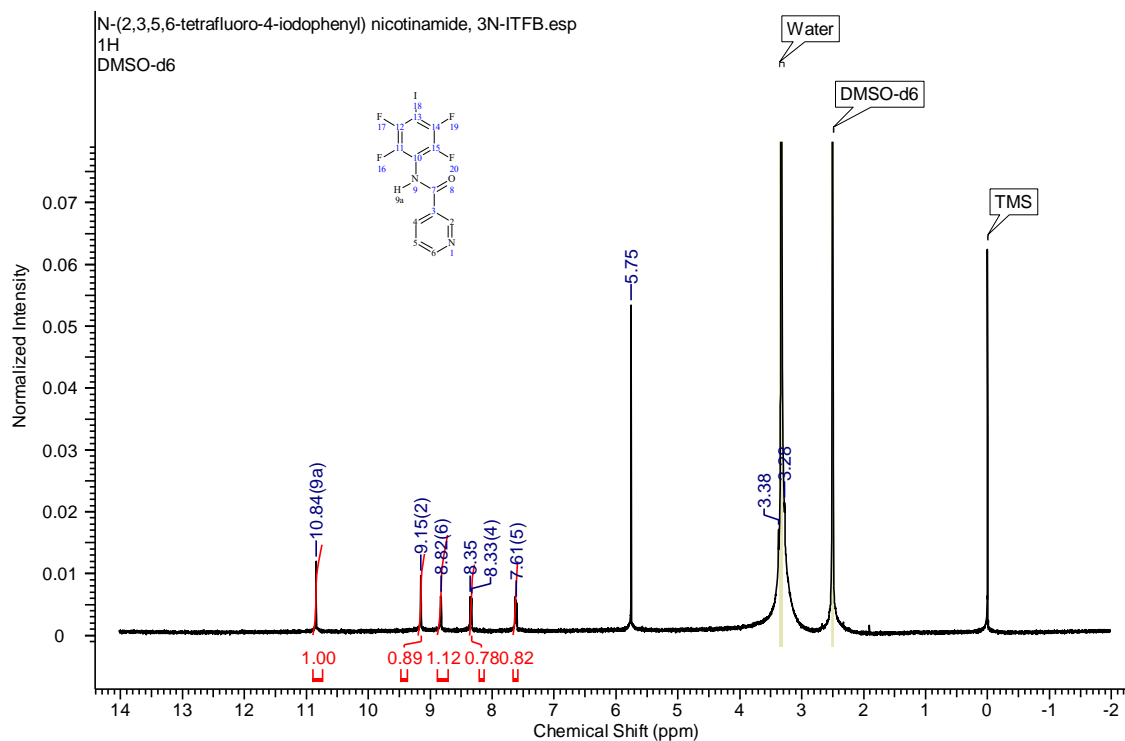
Figure B.3.3 ^{1}H NMR spectrum of N-(2,3,5,6-tetrafluoro-4-iodophenyl)nicotinamide, **3N-ITFB**
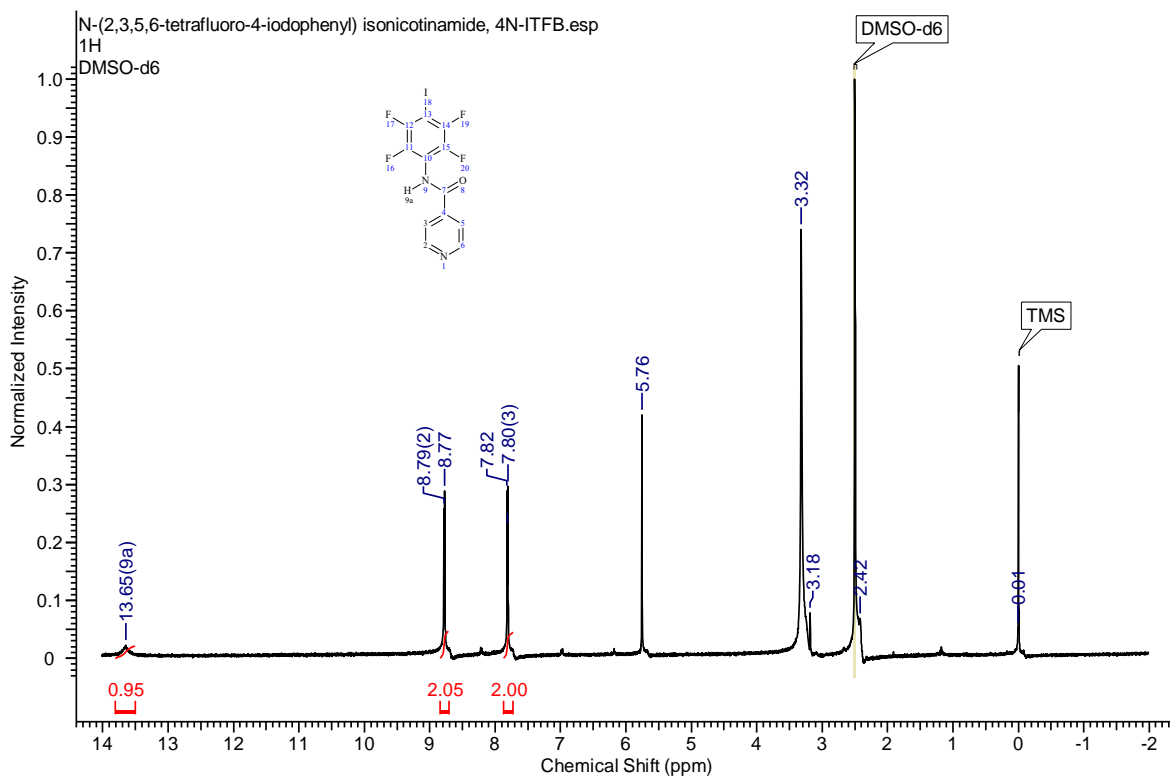
Figure B.3.4 NMR spectrum of N-(2,3,5,6-tetrafluoro-4-iodophenyl)isonicotinamide, **4N-ITFB**
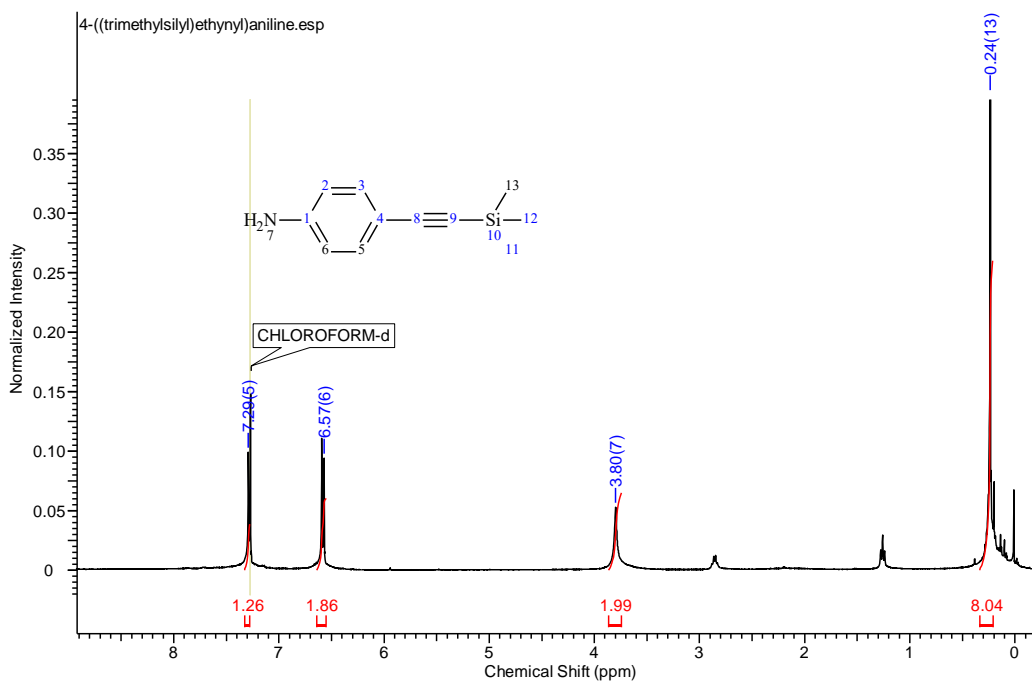
Figure B.3.5 ^{1}H NMR spectrum of 4-((trimethylsilyl)ethynyl)aniline
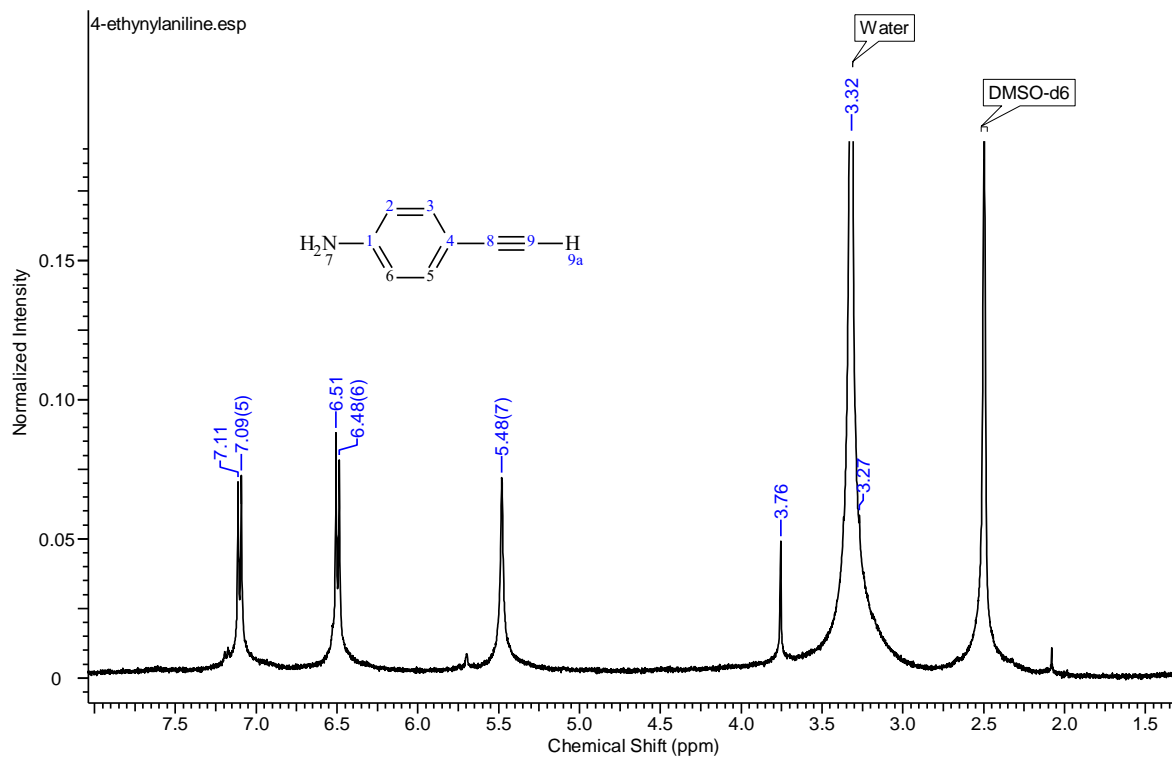
Figure B.3.6 ^{1}H NMR spectrum of 4-ethynylaniline.
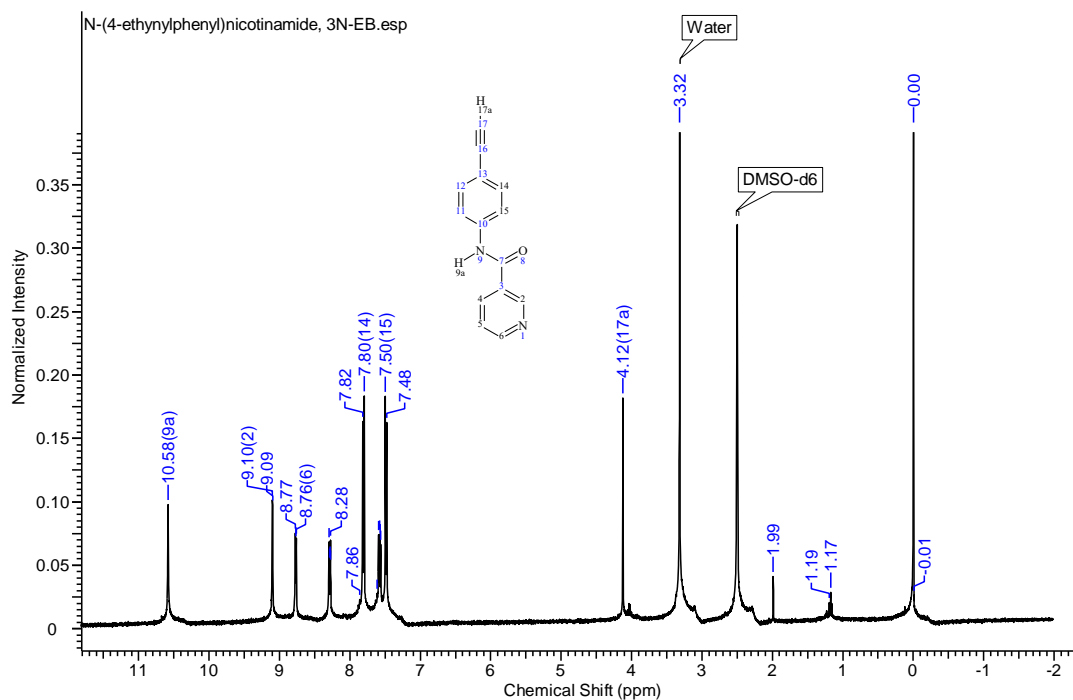
Figure B.3.7 ^{1}H NMR spectrum of N-(4-ethynylphenyl)nicotinamide, **3N-EB**
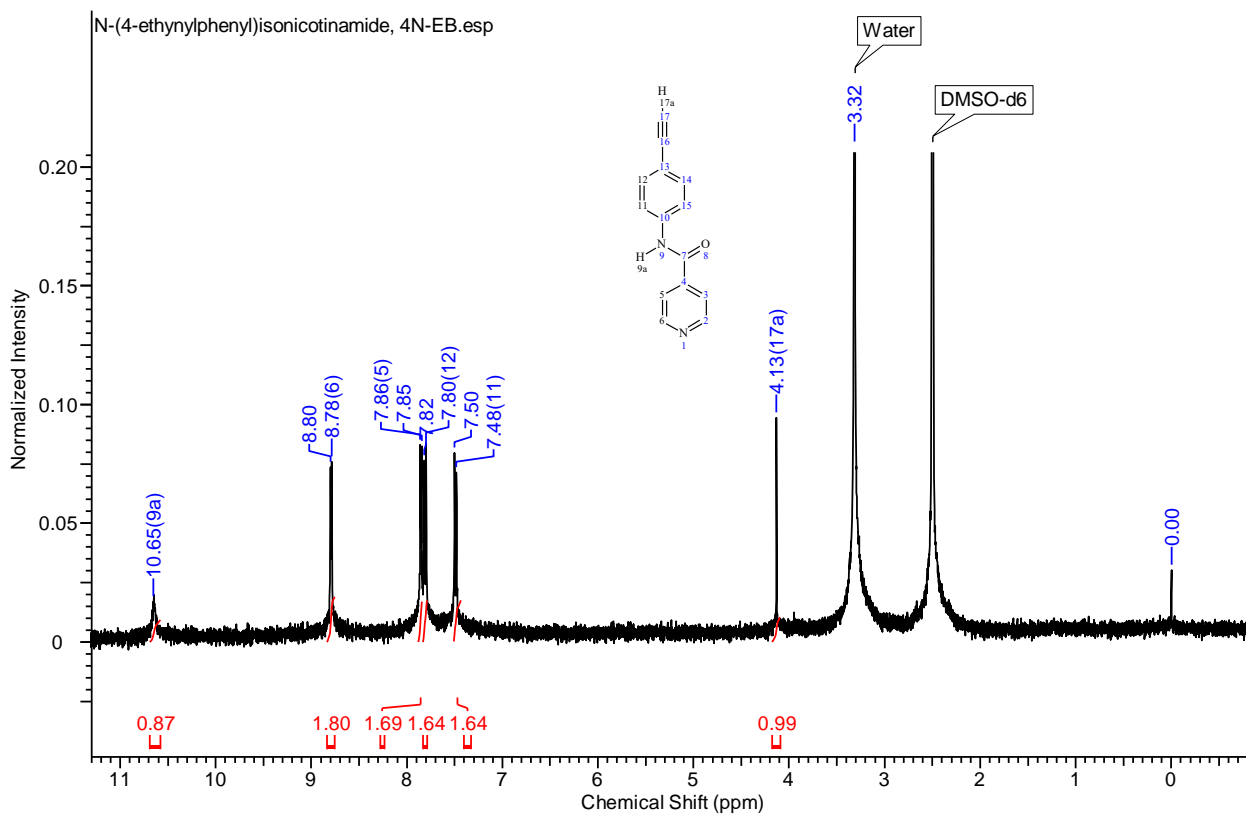
Figure B.3.8 ^{1}H NMR spectrum of **N**-(4-ethynylphenyl)isonicotinamide, **4N-EB**
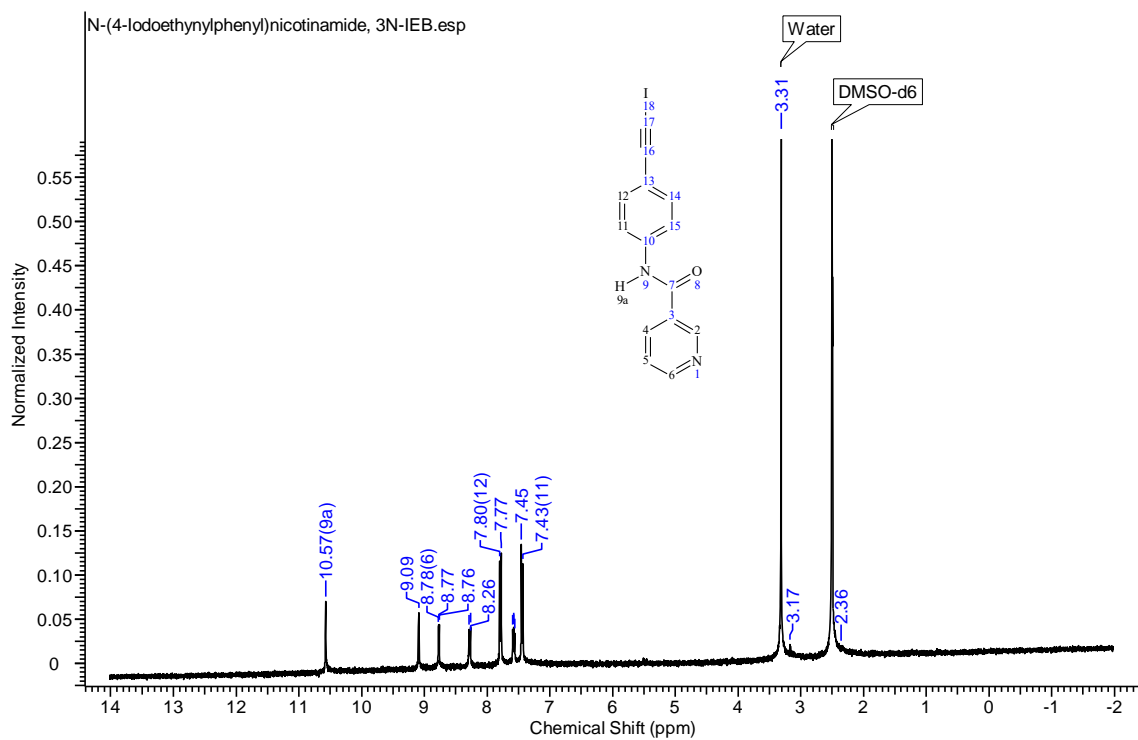
Figure B.3.9 ^{1}H NMR spectrum of N-(4-Iodoethynylphenyl)nicotinamide, **3N-IEB**
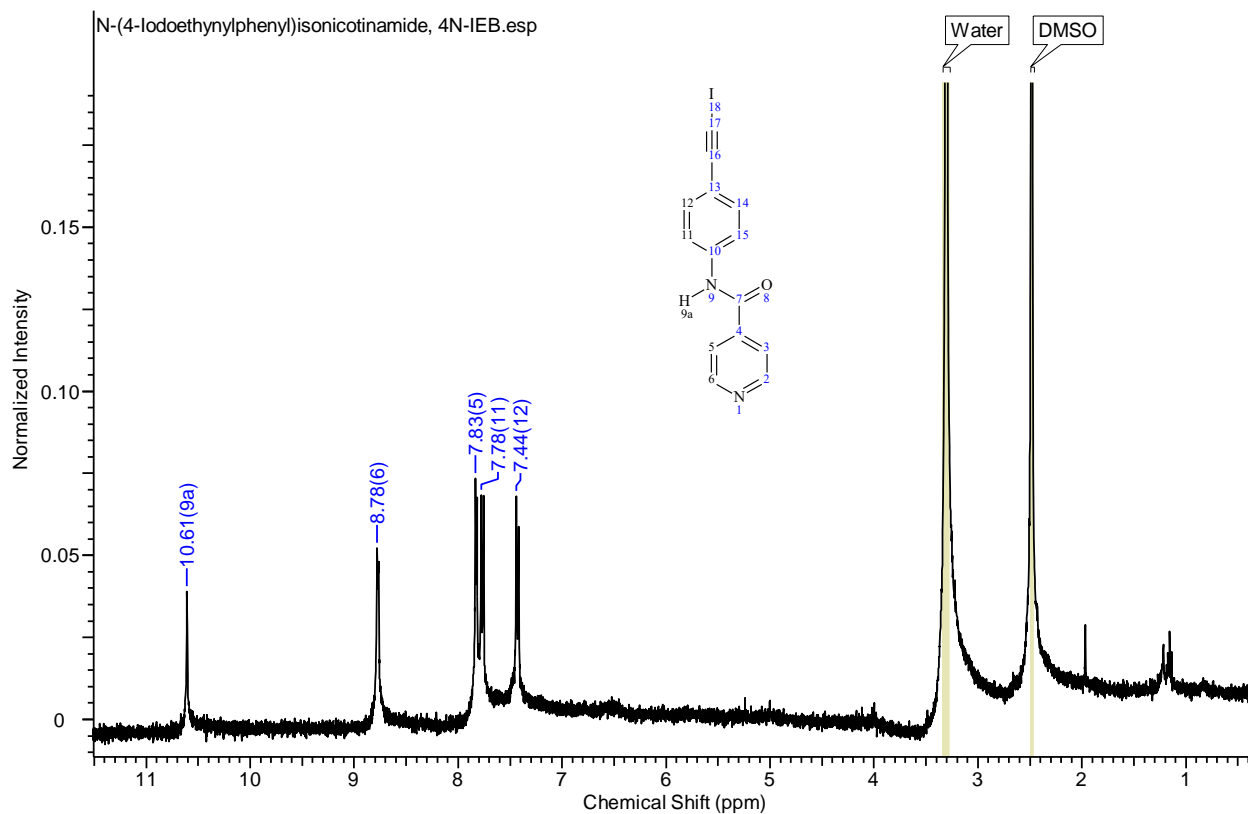
Figure B.3.10 ^{1}H NMR spectrum of N-(4-Iodoethynylphenyl)isonicotinamide, **4N-IEB**

Appendix-IR

Appendix C - IR

C.1 IR analysis of pyrazole based molecules

Table C.1.1 IR analysis of pyrazole based molecules with 20 carboxylic acids

Mixture	O-H...N Grinded mixture	Result	Mixture	O-H...N Grinded mixture	Result
P1-Suc	-	No co-crystal	P1-3-HydroxyBA	2588, 1885	Co-crystal
P1-Adi	-	No co-crystal	P1-4-HydroxyBA	2534, 1916	Co-crystal
P1-sub	-	No co-crystal	P1-3-AminoBA	2397, 1979	No co-crystal
P1-Seb	2459, 1730	Co-crystal	P1-4-AminoBA	2542, 1902	Co-crystal
P1-Dod	2340, 1866	Co-crystal	P1-3-NitroBA	2361,1883	Co-crystal
P1-Fum	-	No co-crystal	P1-4-NitroBA	-	No co-crystal
P1-Mal	-	No co-crystal	P1-BA	-	No co-crystal
P1-Glu	-	No co-crystal	P1-4-IodoBA	-	No co-crystal
P1-Pim	-	No co-crystal	P1-4-BromoBA	-	No co-crystal
P1-Aze	-	No co-crystal	P1-PentafluoroBA	2360,1851	co-crystal
P2-Suc	2342, 1877	co-crystal	P2-3-HydroxyBA	2574, 1959	co-crystal
P2-Adi	2479, 1830	co-crystal	P2-4-HydroxyBA	2502, 1803	co-crystal

307

P2-sub	2506, 1865	co-crystal	P2-3-AminoBA		No co-crystal
P2-Seb	2302, 1981	co-crystal	P2-4-AminoBA	2348, 1929	No co-crystal
P2-Dod	2434, 1864	co-crystal	P2-3-NitroBA	2342,1847	co-crystal
P2-Fum	2358, 1838	co-crystal	P2-4-NitroBA	-	No co-crystal
P2-Mal	2476, 1984	co-crystal	P2-BA	2328,1894	co-crystal
P2-Glu	2475, 1893	co-crystal	P2-4-IodoBA	-	No co-crystal
P2-Pim	2514, 1917	co-crystal	P2-4-BromoBA	-	No co-crystal
P2-Aze	2331, 1884	co-crystal	P2-PentafluoroBA	2346,1854	co-crystal
P3-Suc	2456, 1838	co-crystal	P3-3-HydroxyBA	-	No co-crystal
P3-Adi	2485, 1829	co-crystal	P3-4-HydroxyBA	-	No co-crystal
P3-sub	2502, 1860	co-crystal	P3-3-AminoBA		no co-crystal
P3-Seb	-	No co-crystal	P3-4-AminoBA	2445,1881	co-crystal
P3-Dod	-	No co-crystal	P3-3-NitroBA	2334,1873	co-crystal
P3-Fum	2440, 1804	co-crystal	P3-4-NitroBA		no co-crystal
P3-Mal	2438, 1895	co-crystal	P3-BA	2523,1854	co-crystal
P3-Glu	2574, 1909	co-crystal	P3-4-IodoBA	-	No co-crystal
P3-Pim	2532, 1893	co-crystal	P3-4-BromoBA	-	No co-crystal
P3-Aze	-	No co-crystal	P3-PentafluoroBA	2380,1865	co-crystal
P4-Suc	2467, 1917	Co-crystal	P4-3-HydroxyBA	-	No co-crystal
P4-Adi	2458, 1880	Co-crystal	P4-4-HydroxyBA	-	No co-crystal
P4-sub	2583, 1988	Co-crystal	P4-3-AminoBA		No co-crystal
P4-Seb	2366, 1885	No co-crystal	P4-4-AminoBA	-	No co-crystal
P4-Dod	-	No co-crystal	P4-3-NitroBA	2358,1854	Co-crystal
P4-Fum	2364, 1858	Co-crystal	P4-4-NitroBA	-	No co-crystal
P4-Mal	2427, 1953	Co-crystal	P4-BA	2492,1868	Co-crystal
P4-Glu	2452, 1872	Co-crystal	P4-4-IodoBA	-	No co-crystal
P4-Pim	2364, 1924	Co-crystal	P4-4-BromoBA	-	No co-crystal
P4-Aze	2565, 1984	Co-crystal	P4-PentafluoroBA	2349,1864	Co-crystal
P5-7-Suc	-	No co-crystal	P5-3-HydroxyBA	1879, 2355	Co-crystal
P5-8-Adi	1887, 2362	Co-crystal	P5-4-HydroxyBA	-	No co-crystal
P5-9-Sub	1883, 2512	Co-crystal	P5-3-AminoBA	2001, 2294	Co-crystal
P5-10-Seb	-	No co-crystal	P5-4-AminoBA	-	No co-crystal
P5-11-Dod	-	No co-crystal	P5-3-NitroBA	1853, 2475	Co-crystal
P5-12-Fum	1865, 2520	Co-crystal	P5-4-NitroBA	-	No co-crystal
P5-13-Mal	1860, 2357	Co-crystal	P5-BA	1865, 2450	Co-crystal
P5-14-Glu	1895, 2432	Co-crystal	P5-4-IodoBA	-	No co-crystal
P5-15-Pim	1884, 2364	Co-crystal	P5-4-BromoBA	-	No co-crystal
P5-16-Aze	1868, 2355	Co-crystal	P5-PentafluoroBA	1871, 2417	Co-crystal
P6-7-Suc	-	No co-crystal	P6-3-HydroxyBA	-	No co-crystal
P6-8-Adi	-	No co-crystal	P6-4-HydroxyBA	-	No co-crystal
P6-9-Sub	1899, 2326	Co-crystal	P6-3-AminoBA	1891, 2346	Co-crystal
P6-10-Seb	1882, 2360	Co-crystal	P6-4-AminoBA	-	No co-crystal
P6-11-Dod	-	No co-crystal	P6-3-NitroBA	1894, 2367	Co-crystal
P6-12-Fum	-	No co-crystal	P6-4-NitroBA	-	No co-crystal
P6-13-Mal	1892, 2359	Co-crystal	P6-BA	1917, 2366	Co-crystal

P6-14-Glu	1918, 2304	Co-crystal	P6-4-IodoBA	-	No co-crystal
P6-15-Pim	-	No co-crystal	P6-4-BromoBA	-	No co-crystal
P6-16-Aze	-	No co-crystal	P6-PentafluoroBA	1837, 2355	Co-crystal
P7-Suc	2435, 1909	Co-crystal	P7-3-HydroxyBA		No co-crystal
P7-Adi	2517, 1877	Co-crystal	P7-4-HydroxyBA		No co-crystal
P7-sub	2480, 1874	Co-crystal	P7-3-AminoBA	2358, 2004	No co-crystal
P7-Seb	2317, 1886	Co-crystal	P7-4-AminoBA	2361,1872	Co-crystal
P7-Dod	-	No co-crystal	P7-3-NitroBA	2367,1865	Co-crystal
P7-Fum	2327, 1886	Co-crystal	P7-4-NitroBA	2331,1879	Co-crystal
P7-Mal	2537, 2024	Co-crystal	P7-BA	2557, 1858	no co-crystal
P7-Glu	2340, 1834	Co-crystal	P7-4-IodoBA	2554, 1927	no co-crystal
P7-Pim	2441, 1892	Co-crystal	P7-4-BromoBA	2361, 1929	no co-crystal
P7-Aze	2350, 1866	Co-crystal	P7-PentafluoroBA	2348,1874	Co-crystal
P8-Suc	2525, 2033	Co-crystal	P8-3-HydroxyBA	-	No co-crystal
P8-Adi	-	No co-crystal	P8-4-HydroxyBA	-	No co-crystal
P8-sub	2392, 1869	Co-crystal	P8-3-AminoBA	2425,1969	Co-crystal
P8-Seb	2482, 1901	Co-crystal	P8-4-AminoBA	-	No co-crystal
P8-Dod	2415, 1975	Co-crystal	P8-3-NitroBA	2361,1886	Co-crystal
P8-Fum	2354, 1871	Co-crystal	P8-4-NitroBA	-	No co-crystal
P8-Mal	2516, 1969	Co-crystal	P8-BA	2239,1906	Co-crystal
P8-Glu	2562, 1956	Co-crystal	P8-4-IodoBA	-	No co-crystal
P8-Pim	-	No co-crystal	P8-4-BromoBA	-	No co-crystal
P8-Aze	2342, 1869	Co-crystal	P8-PentafluoroBA	2378,1771	Co-crystal
P9-Suc	2492, 1929	Co-crystal	P9-3-HydroxyBA	2516, 1877	No co-crystal
P9-Adi	2525, 1895	Co-crystal	P9-4-HydroxyBA	2314, 1873	No co-crystal
P9-sub	2366, 1914	Co-crystal	P9-3-AminoBA	2375, 2014	Co-crystal
P9-Seb	2357, 1923	Co-crystal	P9-4-AminoBA	-	No co-crystal
P9-Dod	2522, 1874	Co-crystal	P9-3-NitroBA	2329, 1910	Co-crystal
P9-Fum	2446, 1889	Co-crystal	P9-4-NitroBA	-	No co-crystal
P9-Mal		Co-crystal	P9-BA	2437, 1837	Co-crystal
P9-Glu	2525, 1898	Co-crystal	P9-4-IodoBA	-	No co-crystal
P9-Pim	-	No co-crystal	P9-4-BromoBA	-	No co-crystal
P9-Aze	2443, 1926	Co-crystal	P9-PentafluoroBA	2400, 1926	Co-crystal
P10-Suc	2443, 1932	Co-crystal	P10-3-HydroxyBA	-	No co-crystal
P10-Adi	2342, 1896	Co-crystal	P10-4-HydroxyBA	-	No co-crystal
P10-sub	2375, 1929	Co-crystal	P10-3-AminoBA	2411, 2070	Co-crystal
P10-Seb	2482, 1895	Co-crystal	P10-4-AminoBA	-	No co-crystal
P10-Dod	2236, 2366	Co-crystal	P10-3-NitroBA	2452, 1944	Co-crystal
P10-Fum	2345, 1892	Co-crystal	P10-4-NitroBA	-	No co-crystal
P10-Mal	See below	Co-crystal	P10-BA	See below	Co-crystal
P10-Glu	2354, 1883	Co-crystal	P10-4-IodoBA	-	No co-crystal
P10-Pim	-	No co-crystal	P10-4-BromoBA	-	No co-crystal
P10-Aze	2354, 1914	Co-crystal	P10-PentafluoroBA	2455, 1917	Co-crystal
P11-Suc	2366, 1842	Co-crystal	P11-3-HydroxyBA	-	No co-crystal
P11-Adi	2360, 1895	Co-crystal	P11-4-HydroxyBA	-	No co-crystal

P11-sub	2375, 1929	Co-crystal	P11-3-AminoBA	2161, 1914	Co-crystal
P11-Seb	2473, 1917	Co-crystal	P11-4-AminoBA	-	No co-crystal
P11-Dod	2366, 1962	Co-crystal	P11-3-NitroBA	2513, 1929	Co-crystal
P11-Fum	2323, 1941	Co-crystal	P11-4-NitroBA	-	No co-crystal
P11-Mal	2366, 1935	Co-crystal	P11-BA	2440, 1871	Co-crystal
P11-Glu	2363, 1871	Co-crystal	P11-4-IodoBA	-	No co-crystal
P11-Pim	-	No co-crystal	P11-4-BromoBA	-	No co-crystal
P11-Aze	2452, 1883	Co-crystal	P11-PentafluoroBA	2366, 1877	Co-crystal
P12-Suc	2470, 1862	Co-crystal	P12-3-HydroxyBA	-	No co-crystal
P12-Adi	2350, 1889	Co-crystal	P12-4-HydroxyBA	-	No co-crystal
P12-sub	2324, 1911	Co-crystal	P12-3-AminoBA	See below	Co-crystal
P12-Seb	2375, 1855	Co-crystal	P12-4-AminoBA	-	No co-crystal
P12-Dod	2461, 1910	Co-crystal	P12-3-NitroBA	2378, 1950	Co-crystal
P12-Fum	2356, 1865	Co-crystal	P12-4-NitroBA	2409, 1956	No co-crystal
P12-Mal	2370, 1889	Co-crystal	P12-BA	2427, 1874	Co-crystal
P12-Glu	2358, 1893	Co-crystal	P12-4-IodoBA	-	No co-crystal
P12-Pim	-	No co-crystal	P12-4-BromoBA	-	No co-crystal
P12-Aze	2372, 1920	Co-crystal	P12-PentafluoroBA	2446, 1944	Co-crystal

	Ligand	**Acid**	**Grinded mixture**	**Result**
P7-3-HydroxyBA	1724, 1651, 1583, 1485	1695, 1552, 1493, 1457	1679, 1557, 1461, 1284	Co-crystal
P7-4-HydroxyBA	1724, 1651, 1583, 1485	1668, 1592, 1286	1661, 1592, 1488, 1420, 1293	Co-crystal
P10-Mal	1662, 1587, 1480, 1419	1433, 1300, 1214, 1167	1712, 1673, 1574, 1484	Co-crystal
P10-BA	1662, 1587, 1480, 1419		1671, 1587, 1481, 1313	Co-crystal
P12-3-AminoBA	1699, 1658, 1589, 1481	1623, 1388	1631, 1560, 1383, 1032	Co-crystal

C.2 IR analysis of thiazole based molecules

Table C.2.1 IR analysis of thiazole based molecules with 20 carboxylic acids

Mixture	**IR bands in mixture**	**Result**	**Mixture**	**IR bands in mixture**	**Result**
T1-Suc	2253, 1935	Co-crystal	T1-3HydroxyBA	2358, 1865	Co-crystal
T1-Adi	-	No Co-crystal	T1-4HydroxyBA	2353, 1949	Co-crystal
T1-Sub	-	No co-crystal	T1-3AminoBA	2346, 1917	Co-crystal
T1-Seb	-	No co-crystal	T1-4AminoBA	2440, 1822	Co-crystal
T1-Dod	-	No co-crystal	T1-3NitroBA	2359, 1895	Co-crystal
T1-Fum	-	No co-crystal	T1-4NitrooBA	-	No co-crystal

T1-Mal	-	No co-crystal	T1-BA	-	No co-crystal
T1-Glut	-	No co-crystal	T1-4BromoBA	-	No co-crystal
T1-Pim	-	No co-crystal	T1-4IodoBA	-	No co-crystal
T1-Aze	-	No co-crystal	T1-PentaBA	2360, 1865	Co-crystal
T2-Suc	-	No co-crystal	T2-3HydroxyBA	2513, 2033	Co-crystal
T2-Adi	-	No co-crystal	T2-4HydroxyBA	2461, 1841	Co-crystal
T2-Sub	-	No co-crystal	T2-3AminoBA	2351, 1926	Co-crystal
T2-Seb	-	No co-crystal	T2-4AminoBA	2354, 1910	Co-crystal
T2-Dod	-	No co-crystal	T2-3NitroBA	2349, 1927	Co-crystal
T2-Fum	2311, 1926	Co-crystal	T2-4NitrooBA	-	No co-crystal
T2-Mal	2381, 1877	Co-crystal	T2-BA	-	No co-crystal
T2-Glut	-	No co-crystal	T2-4BromoBA	-	No co-crystal
T2-Pim	-	No co-crystal	T2-4IodoBA	-	No co-crystal
T2-Aze	-	No co-crystal	T2-PentaBA	2567, 1870	Co-crystal
T3-Suc	-	No co-crystal	T3-3HydroxyBA	2468, 1870	Co-crystal
T3-Adi	-	No co-crystal	T3-4HydroxyBA	-	No co-crystal
T3-Sub	-	No co-crystal	T3-3AminoBA	2363, 2005	Co-crystal
T3-Seb	-	No co-crystal	T3-4AminoBA	-	No co-crystal
T3-Dod	-	No co-crystal	T3-3NitroBA	2366, 1759	Co-crystal
T3-Fum	-	No co-crystal	T3-4NitrooBA	-	No co-crystal
T3-Mal	-	No co-crystal	T3-BA	-	No co-crystal
T3-Glut	-	No co-crystal	T3-4BromoBA	-	No co-crystal
T3-Pim	-	No co-crystal	T3-4IodoBA	-	No co-crystal
T3-Aze	-	No co-crystal	T3-PentaBA	2504, 1917	Co-crystal
T4-Suc	-	No co-crystal	T4-3HydroxyBA	2363, 1899	Co-crystal
T4-Adi	-	No co-crystal	T4-4HydroxyBA	2301, 1845	Co-crystal
T4-Sub	-	No co-crystal	T4-3AminoBA	-	No Co-crystal
T4-Seb	-	No co-crystal	T4-4AminoBA	-	No co-crystal
T4-Dod	-	No co-crystal	T4-3NitroBA	2366, 1759	Co-crystal
T4-Fum	-	No co-crystal	T4-4NitrooBA	-	No co-crystal
T4-Mal	-	No co-crystal	T4-BA	2358, 1825	Co-crystal
T4-Glut	-	No co-crystal	T4-4BromoBA	-	No co-crystal
T4-Pim	-	No co-crystal	T4-4IodoBA	-	No co-crystal
T4-Aze	-	No co-crystal	T4-PentaBA	2504, 1917	Co-crystal
T5-Suc	-	No co-crystal	T5-3HydroxyBA	1865,2367	Co-crystal
T5-Adi	-	No co-crystal	T5-4HydroxyBA	-	No co-crystal
T5-Sub	-	No co-crystal	T5-3AminoBA	1977, 2365	Co-crystal
T5-Seb	-	No co-crystal	T5-4AminoBA	1906,2305	No co-crystal
T5-Dod	-	No co-crystal	T5-3NitroBA	1870,2357	Co-crystal
T5-Fum	-	No co-crystal	T5-4NitroBA	-	No co-crystal
T5-Mal	1878, 2314	Co-crystal	T5-BA	-	No co-crystal
T5-Glut	-	No co-crystal	T5-4BromoBA	-	No co-crystal
T5-Pim	-	No co-crystal	T5-4IodoBA	-	No co-crystal
T5-Aze	-	No co-crystal	T5-PentaBA	1902,2492	Co-crystal
T6-Suc	-	No co-crystal	T6-3HydroxyBA	1840,2375	Co-crystal

T6-Adi	1866, 2400	Co-crystal	T6-4HydroxyBA	1854,2543	Co-crystal
T6-Sub	1918, 2284	Co-crystal	T6-3AminoBA	1985,2305	Co-crystal
T6-Seb	-	No co-crystal	T6-4AminoBA	-	No co-crystal
T6-Dod	-	No co-crystal	T6-3NitroBA	1839,2306	Co-crystal
T6-Fum	1862, 2371	Co-crystal	T6-4NitroBA	-	No co-crystal
T6-Mal	1942, 2369	Co-crystal	T6-BA	1846,2357	Co-crystal
T6-Glut	1877, 2361	Co-crystal	T6-4BromoBA	-	No co-crystal
T6-Pim	-	No co-crystal	T6-4IodoBA	-	No co-crystal
T6-Aze	-	No co-crystal	T6-PentaBA	1897,2567	Co-crystal
T7-Suc	2455, 1898	Co-crystal	T7-3HydroxyBA	2468, 1870	Co-crystal
T7-Adi	2336, 1907	Co-crystal	T7-4HydroxyBA	2361, 1769	Co-crystal
T7-Sub	2339, 1892	Co-crystal	T7-3AminoBA	2406, 1983	Co-crystal
T7-Seb	2366, 1929	Co-crystal	T7-4AminoBA	-	No co-crystal
T7-Dod	2378, 1969	Co-crystal	T7-3NitroBA	2360, 1849	Co-crystal
T7-Fum	2348, 1871	Co-crystal	T7-4NitroBA	2364, 1897	Co-crystal
T7-Mal	2363, 1865	Co-crystal	T7-BA	2364, 1906	Co-crystal
T7-Glut	2360, 1904	Co-crystal	T7-4BromoBA	2337, 1909	Co-crystal
T7-Pim	2339, 1886	Co-crystal	T7-4IodoBA	2320, 1906	Co-crystal
T7-Aze	-		T7-PentaBA	2328, 1866	Co-crystal
T8-Suc	2345, 1868	Co-crystal	T8-3HydroxyBA	2489, 1852	Co-crystal
T8-Adi	2360, 1880	Co-crystal	T8-4HydroxyBA	2539, 1900	Co-crystal
T8-Sub	2354, 1912	Co-crystal	T8-3AminoBA	2367, 1977	Co-crystal
T8-Seb	2357, 1862	Co-crystal	T8-4AminoBA	2364, 1891	Co-crystal
T8-Dod	2370, 1854	Co-crystal	T8-3NitroBA	2334, 1885	Co-crystal
T8-Fum	2345, 1914	Co-crystal	T8-4NitroBA	2547, 1815	Co-crystal
T8-Mal	2363, 1892	Co-crystal	T8-BA	2408, 1875	Co-crystal
T8-Glut	2362, 1885	Co-crystal	T8-4BromoBA	-	No co-crystal
T8-Pim	2326, 1898	Co-crystal	T8-4IodoBA	-	No co-crystal
T8-Aze	2345, 1900	Co-crystal	T8-PentaBA	2359, 1892	Co-crystal
T9-Suc	1871, 2433	Co-crystal	T9-3HydroxyBA	1950, 2556	Co-crystal
T9-Adi	1901, 2363	Co-crystal	T9-4HydroxyBA	1849, 2492	Co-crystal
T9-Sub	1840, 2339	Co-crystal	T9-3AminoBA	1788, 2323	Co-crystal
T9-Seb	1944, 2541	Co-crystal	T9-4AminoBA	1898, 2495	Co-crystal
T9-Dod	1880, 2345	Co-crystal	T9-3NitroBA	1892, 2369	Co-crystal
T9-Fum	1868, 2357	Co-crystal	T9-4NitroBA	1956, 2363	Co-crystal
T9-Mal	1929, 2259	Co-crystal	T9-BA	1898, 2348	Co-crystal
T9-Glut	1914, 2357	Co-crystal	T9-4BromoBA	1935, 2357	Co-crystal
T9-Pim	1929, 2259	Co-crystal	T9-4IodoBA	-	No Co-crystal
T9-Aze	-	No Co-crystal	T9-PentaBA	1944, 2412	Co-crystal
T10-Suc	1842, 2366	Co-crystal	T10-3HydroxyBA	1935, 2397	Co-crystal
T10-Adi	1886, 2354	Co-crystal	T10-4HydroxyBA	1898, 2336	Co-crystal
T10-Sub	1865, 2354	Co-crystal	T10-3AminoBA	1914, 2161	Co-crystal
T10-Seb	1941, 2213	Co-crystal	T10-4AminoBA	1904, 2336	Co-crystal
T10-Dod	1920, 2339	Co-crystal	T10-3NitroBA	1929, 2513	Co-crystal
T10-Fum	1871, 2363	Co-crystal	T10-4NitroBA	1953, 2430	Co-crystal
T10-Mal	1825, 2329	Co-crystal	T10-BA	1871, 2440	Co-crystal

T10-Glut	1825, 2329	Co-crystal	T10-4BromoBA	-	No Co-crystal
T10-Pim	1751, 2360	Co-crystal	T10-4IodoBA	-	No Co-crystal
T10-Aze	1883, 2452	Co-crystal	T10-PentaBA	1877, 2366	Co-crystal
T11-Suc	1865, 2427	Co-crystal	T11-3HydroxyBA	1889, 2345	Co-crystal
T11-Adi	1889, 2516	Co-crystal	T11-4HydroxyBA	1895, 2342	Co-crystal
T11-Sub	1889, 2369	Co-crystal	T11-3AminoBA	1947, 2351	Co-crystal
T11-Seb	1881, 2475	Co-crystal	T11-4AminoBA	2014, 2336	Co-crystal
T11-Dod	1889, 2420	Co-crystal	T11-3NitroBA	2516, 1929	Co-crystal
T11-Fum	1825, 2339	Co-crystal	T11-4NitroBA	2427, 1950	Co-crystal
T11-Mal	2372, 1828	Co-crystal	T11-BA	2351, 1910	Co-crystal
T11-Glut	2357, 2005	Co-crystal	T11-4BromoBA	-	No Co-crystal
T11-Pim	1843, 2354	Co-crystal	T11-4IodoBA	-	No Co-crystal
T11-Aze	1874, 2342	Co-crystal	T11-PentaBA	2437, 1883	Co-crystal
T12-Suc	1863, 2466	Co-crystal	T12-3HydroxyBA	1849, 2357	Co-crystal
T12-Adi	1929, 2506	Co-crystal	T12-4HydroxyBA	1892, 2336	Co-crystal
T12-Sub	See below	Co-crystal	T12-3AminoBA	See below	Co-crystal
T12-Seb	2384, 1856	Co-crystal	T12-4AminoBA	1846, 2357	Co-crystal
T12-Dod	See below	Co-crystal	T12-3NitroBA	1950, 2378	Co-crystal
T12-Fum	2356, 1865	Co-crystal	T12-4NitroBA	2409, 1956	Co-crystal
T12-Mal	1889, 2370	Co-crystal	T12-BA	2427, 1874	Co-crystal
T12-Glut	1893, 2358	Co-crystal	T12-4BromoBA	-	No co-crystal
T12-Pim	1907, 2358	Co-crystal	T12-4IodoBA	-	No co-crystal
T12-Aze	1923, 2372	Co-crystal	T12-PentaBA	2446, 1944	Co-crystal

	Ligand	Acid	Grinded mixture	Co-crystal?
T9-3AminoBA	1549, 1593	1623, 1388	1633, 1543, 1392, 1590	Co-crystal
T9-4AminoBA	1549, 1668	1420	1423, 1543, 1670	Co-crystal
T10-4HydroxyBA	1679	1286	1294, 1673	Co-crystal
T10-3AminoBA	1543	1388, 1623	1385, 1546, 1635	Co-crystal
T10-4AminoBA	1153, 1301, 1543, 1679	1285	1159, 1294, 1558, 1689	Co-crystal
T12-sub	1305, 1538, 1556	1187, 1248	1183, 1244, 1309, 1544, 1562	Co-crystal
T12-dod	1538, 1556, 1672	1222, 1686	1220, 1548, 1561, 1676	Co-crystal
T12-3AminoBA	1538	1623	1627, 1541	Co-crystal
T12-4AminoBA	1412, 1597	1285	1603, 1408	Co-crystal

C.3 IR of co-crystals of ternary systems

Table C.3.1 IR analysis of thiazole based molecules with 15 carboxylic acids and one halogen bond donor

Ternary co-crystals	Hydrogen-bond donor	Grinded mixture	Ternary Co-crystal?
T9-H1-D1	1678, 1410, 1306	1883, 2519, 1453, 1415, 937, 756	Yes
T9-H2-D1	1683, 1406, 1272	1904, 2455, 1461, 1418, 939	Yes
T9-H3-D1	1684, 1404, 1248	1886, 2287, 1464, 1409, 940	Yes
T9-H4-D1	1686, 1405, 1297, 1232	2366, 1950, 1680, 1543, 1455	Yes
T9-H5-D1	1686, 1427, 1408, 1279	1941, 2492, 1463, 1408, 937	Yes
T9-H6-D1	1657, 1421, 1270	2339, 1956, 1673, 1595, 1522	Yes
T9-H7-D1	1693, 1433	2329, 1880, 1670, 1559, 1460	Yes
T9-H8-D1	1682, 1406, 1301	2314, 1819, 1556, 1486, 1303	Yes
T9-H9-D1	1683, 1407, 1267	1681, 1544, 1455, 1407	No (only binary XB)
T9-H10-D1	1682, 1408, 1250	1681, 1542, 1456, 1407, 1292	No (only binary XB)
T9-H11-D1	1695, 1552, 1493, 1457	1456, 1429, 1862, 1460, 1427 (Only XB)	No (only binary XB)
T9-H12-D1	1668, 1592, 1286	1865, 2357, 1460, 1427, 938	Yes
T9-H13-D1	1686, 1528, 1351	1461, 1422, 1870, 2424 (Only XB)	No (only binary XB)
T9-H14-D1	1685, 1599, 1536, 1275	2332, 1956, 1463, 1420, 939	Yes
T9-H15-D1	1708, 1650	2427, 1883, 1678, 1531	Yes
T10-H1-D1	1678, 1410, 1306	1883, 2342, 1458, 1410, 939	Yes
T10-H2-D1	1683, 1406, 1272	1883, 2489, 1458, 1424, 939	Yes
T10-H3-D1	1684, 1404, 1248	1889, 2467, 1457, 1423, 934	Yes
T10-H4-D1	1686, 1405, 1297, 1232	1458, 1434, 939 (Only XB)	No (only binary XB)
T10-H5-D1	1686, 1427, 1408, 1279	2342, 1914, 1457, 1426, 936	Yes
T10-H6-D1	1657, 1421, 1270	1455, 1434, 939 (Only XB)	No (only binary XB)
T10-H7-D1	1693, 1433	1455, 1434, 939 (Only XB)	No (only binary XB)
T10-H8-D1	1682, 1406, 1301	1455, 1434, 939 (Only XB)	No (only binary XB)
T10-H9-D1	1683, 1407, 1267	1655, 1553, 1451, 1553, 1411, 1134 (Only XB)	No (only binary XB)
T10-H10-D1	1682, 1408, 1250	2366, 1962, 1683, 1520, 1465	Yes
T10-H11-D1	1695, 1552, 1493, 1457	1867, 2268, 1458, 1409, 940, 1215	Yes
T10-H12-D1	1668, 1592, 1286	2402, 1872, 1459, 1429, 938	Yes

T10-H13-D1	1686, 1528, 1351	1940, 2321, 1461, 1438, 940	Yes
T10-H14-D1	1685, 1599, 1536, 1275	2357, 1858, 1664, 1598, 1572	Yes
T10-H15-D1	1708, 1650	1957, 2360, 1465, 1421, 940	Yes
T11-H1-D1	1678, 1410, 1306	1910, 2476, 1455, 1413, 935	Yes
T11-H2-D1	1683, 1406, 1272	1456, 142, 937, 1907, 2489	Yes
T11-H3-D1	1684, 1404, 1248	1892, 2363, 1458, 1424, 933	Yes
T11-H4-D1	1686, 1405, 1297, 1232	1950, 2319, 1458, 1421, 936	Yes
T11-H5-D1	1686, 1427, 1408, 1279	1910, 2371, 1682, 1544, 1457	Yes
T11-H6-D1	1657, 1421, 1270	1681, 1544, 1455, 1289, 939 (only XB)	No (only binary XB)
T11-H7-D1	1693, 1433	1681, 1543, 1455, 1291, 939 (Only XB)	No (only binary XB)
T11-H8-D1	1682, 1406, 1301	2143, 1846, 1681, 1540, 1455	Yes
T11-H9-D1	1683, 1407, 1267	2354, 1846, 1665, 1572	Yes
T11-H10-D1	1682, 1408, 1250	2339, 1895, 1680, 1552, 1460	Yes
T11-H11-D1	1695, 1552, 1493, 1457	1456, 976, 1996, 2447(Only XB)	No (only binary XB)
T11-H12-D1	1668, 1592, 1286	1875, 2472, 939, 1455, 1429	Yes
T11-H13-D1	1686, 1528, 1351	1894, 2340, 1456, 936	Yes
T11-H14-D1	1685, 1599, 1536, 1275	1953, 2363, 1465, 1421, 940	Yes
T11-H15-D1	1708, 1650	946, 1466, 1684(Only XB)	No (only binary XB)
T12-H1-D1	1678, 1410, 1306	1898, 2482, 1458, 1409, 940	Yes
T12-H2-D1	1683, 1406, 1272	1887, 2487, 1460, 1435, 941	Yes
T12-H3-D1	1684, 1404, 1248	1946, 2533, 1461, 1426, 940	Yes
T12-H4-D1	1686, 1405, 1297, 1232	1897, 2267, 1230, 1435, 941	Yes
T12-H5-D1	1686, 1427, 1408, 1279	1461, 1430, 938, 1890, 2325 (Only XB)	No (only binary XB)
T12-H6-D1	1657, 1421, 1270	2366, 1966, 1522, 1467, 1424	Yes
T12-H7-D1	1693, 1433	2492, 1920, 1672, 1568, 1450	Yes
T12-H8-D1	1682, 1406, 1301	2357, 1709, 1579, 1409	Yes
T12-H9-D1	1683, 1407, 1267	2305, 1941, 1681, 1540, 1455	Yes
T12-H10-D1	1682, 1408, 1250	2396, 1946, 1677, 1492, 1326	Yes
T12-H11-D1	1695, 1552, 1493, 1457	1463, 1419 (Only XB)	No (only binary XB)
T12-H12-D1	1668, 1592, 1286	1833, 2490 1450, 1412, 1429	Yes

T12-H13-D1	1686, 1528, 1351	1458, 938, 1938, 2360 (Only XB)	No (only binary XB)		
T12-H14-D1	1685, 1599, 1536, 1275	1960, 2363, 1455, 1424, 940	Yes		
T12-H15-D1	1708, 1650	1946, 2396, 1469, 1417, 941	Yes		

C.4 IR of co-crystals of halogen bond donors

Table C.4.1 IR analysis of 10 halogen bond donors with 20 carboxylic acids

Mixture	IR bands in mixture	Result	Mixture	IR bands in mixture	Result
3N-EB-Suc	2525, 1901	Co-crystal	3N-IB-3OHBA	2363, 1901	Co-crystal
3N-EB-Adi	2468, 1907	Co-crystal	3N-IB-4OHBA	2485, 1901	Co-crystal
3N-EB-Sub	2482, 1907	Co-crystal	3N-IB-3AminoBA	2461, 2005	Co-crystal
3N-EB-Seb	2498, 1889	Co-crystal	3N- IB -4AminoBA	2556, 1898	Co-crystal
3N-EB-Dod	2351, 1901	Co-crystal	3N- IB -3NitroBA	2476, 1938	Co-crystal
3N-EB-Fum	2534, 1901	Co-crystal	3N- IB -4NitroBA	2387, 1896	Co-crystal
3N-EB-Mal	2580, 1938	Co-crystal	3N- IB -BA	2351, 1895	Co-crystal
3N-EB-Glut	2553, 1920	Co-crystal	3N- IB -4BromoBA	2498, 1895	Co-crystal
3N-EB-Pim	2537, 1901	Co-crystal	3N- IB -4IodoBA	-	No co-crystal
3N-EB-Aze	2510, 1904	Co-crystal	3N- IB -PentaFBA	2333, 1907	Co-crystal
4N-EB-Suc	2504, 1907	Co-crystal	4N-IB-3OHBA	2332, 1904	Co-crystal
4N-EB-Adi	-	No co-crystal	4N-IB-4OHBA	See below	Co-crystal
4N-EB-Sub	-	No co-crystal	4N- IB -3AminoBA	See below	Co-crystal
4N-EB-Seb	-	No co-crystal	4N- IB -4AminoBA	2340, 1903	Co-crystal
4N-EB-Dod	-	No co-crystal	4N- IB -3NitroBA	2316, 1901	Co-crystal
4N-EB-Fum	2360, 1895	Co-crystal	4N- IB -4NitrooBA	2364, 1886	Co-crystal
4N-EB-Mal	2556, 1898	Co-crystal	4N- IB -BA		No co-crystal
4N-EB-Glut	-	No co-crystal	4N- IB -4BromoBA		No co-crystal
4N-EB-Pim	-	No co-crystal	4N- IB-4IodoBA		No co-crystal
4N-EB-Aze	-	No co-crystal	4N- IB -PentaFBA	2345, 1886	Co-crystal
3N-ITFB-Suc	-	No co-crystal	3N-ITFB:3OHBA	-	No co-crystal
3N-ITFB-Adi	-	No co-crystal	3N-ITFB:4OHBA	-	No co-crystal
3N-ITFB-Sub	-	No co-crystal	3N-ITFB-3AminoBA	-	No co-crystal

3N-ITFB-Seb	-	No co-crystal	3N-ITFB-4AminoBA	-	No co-crystal
3N-ITFB-Dod	-	No co-crystal	3N-ITFB-3NitroBA	-	No co-crystal
3N-ITFB-Fum	-	No co-crystal	3N-ITFB-4NitrooBA	-	No co-crystal
3N-ITFB-Mal	-	No co-crystal	3N-ITFB-BA	-	No co-crystal
3N-ITFB-Glut	-	No co-crystal	3N-ITFB-4BromoBA	-	No co-crystal
3N-ITFB-Pim	-	No co-crystal	3N-ITFB-4IodoBA	-	No co-crystal
3N-ITFB-Aze	-	No co-crystal	3N-ITFB-PentaBA	-	No co-crystal
4N-ITFB-Suc	-	No co-crystal	4N-ITFB-3HydroxyBA	-	No co-crystal
4N-ITFB-Adi	-	No co-crystal	4N-ITFB-4HydroxyBA	-	No co-crystal
4N-ITFB-Sub	-	No co-crystal	4N-ITFB-3AminoBA	-	No co-crystal
4N-ITFB-Seb	-	No co-crystal	4N-ITFB-4AminoBA	-	No co-crystal
4N-ITFB-Dod	-	No co-crystal	4N-ITFB-3NitroBA	-	No co-crystal
4N-ITFB-Fum	-	No co-crystal	4N-ITFB-4NitrooBA	-	No co-crystal
4N-ITFB-Mal	-	No co-crystal	4N-ITFB-BA	-	No co-crystal
4N-ITFB-Glut	-	No co-crystal	4N-ITFB-4BromoBA	-	No co-crystal
4N-ITFB-Pim	-	No co-crystal	4N-ITFB-4IodoBA	-	No co-crystal
4N-ITFB-Aze	-	No co-crystal	4N-ITFB-PentaBA	-	No co-crystal
				-	No co-crystal
3N-EB-Suc	2522, 1907	Co-crystal	3N-EB-3HydroxyBA	2364, 1886	Co-crystal
3N-EB-Adi	2482, 1907	Co-crystal	3N-EB-4HydroxyBA	2498, 1882	Co-crystal
3N-EB-Sub	2525, 1914	Co-crystal	3N-EB-3AminoBA	-	No co-crystal
3N-EB-Seb	-	No co-crystal	3N-EB-4AminoBA	2369, 1923	Co-crystal
3N-EB-Dod	-	No co-crystal	3N-EB-3NitroBA	2354, 1950	Co-crystal
3N-EB-Fum	2357, 1917	Co-crystal	3N-EB-4NitroBA	-	No co-crystal
3N-EB-Mal	2354, 1959	Co-crystal	3N-EB-BA	2559, 1920	Co-crystal
3N-EB-Glut	2485, 1914	Co-crystal	3N-EB-4BromoBA	2458, 1907	Co-crystal
3N-EB-Pim	2516, 1920	Co-crystal	3N-EB-4IodoBA	2544, 1920	Co-crystal
3N-EB-Aze	2366, 1898	Co-crystal	3N-EB-PentaBA	2342, 1932	Co-crystal
4N-EB-Suc	2501, 1969	Co-crystal	4N-EB-3HydroxyBA	2362, 1906	Co-crystal
4N-EB-Adi	2369, 1929	Co-crystal	4N-EB-4HydroxyBA	2350, 1985	Co-crystal
4N-EB-Sub	2366, 1926	Co-crystal	4N-EB-3AminoBA	2446, 1956	Co-crystal
4N-EB-Seb	See below	Co-crystal	4N-EB-4AminoBA	2544, 2097	Co-crystal

317

4N-EB-Dod	2320, 1932	Co-crystal	4N-EB-3NitroBA	2492, 1947	Co-crystal
4N-EB-Fum	2498, 1920	Co-crystal	4N-EB-4NitroBA	2562, 1956	Co-crystal
4N-EB-Mal		Co-crystal	4N-EB-BA	2437, 1959	Co-crystal
4N-EB-Glut	2372, 1941	Co-crystal	4N-EB-4BromoBA	2550, 1932	Co-crystal
4N-EB-Pim	2534, 1926	Co-crystal	4N-EB-4IodoBA	2544, 1914	Co-crystal
4N-EB-Aze	2550, 1920	Co-crystal	4N-EB-PentaBA	2421, 1923	Co-crystal
3N-IEB-Suc			3N-IEB-3HydroxyBA	-	No co-crystal
3N-IEB-Adi	2348, 1926	Co-crystal	3N-IEB-4HydroxyBA	-	No co-crystal
3N-IEB-Sub	-	No co-crystal	3N-IEB-3AminoBA	-	No co-crystal
3N-IEB-Seb	-	No co-crystal	3N-IEB-4AminoBA	-	No co-crystal
3N-IEB-Dod	-	No co-crystal	3N-IEB-3NitroBA	2291, 1951	Co-crystal
3N-IEB-Fum	-	No co-crystal	3N-IEB-4NitroBA	-	No co-crystal
3N-IEB-Mal	2362, 1931	Co-crystal	3N-IEB-BA	-	No co-crystal
3N-IEB-Glut	-	No co-crystal	3N-IEB-4BromoBA	-	No co-crystal
3N-IEB-Pim	-	No co-crystal	3N-IEB-4IodoBA	-	No co-crystal
3N-IEB-Aze	-	No co-crystal	3N-IEB-PentaBA	2505, 1892	Co-crystal
	-	No co-crystal		-	No co-crystal
4N-IEB-Suc	-	No co-crystal	4N-IEB-3HydroxyBA	-	No co-crystal
4N-IEB-Adi	-	No co-crystal	4N-IEB-4HydroxyBA	-	No co-crystal
4N-IEB-Sub	-	No co-crystal	4N-IEB-3AminoBA	-	No co-crystal
4N-IEB-Seb	-	No co-crystal	4N-IEB-4AminoBA	-	No co-crystal
4N-IEB-Dod	-	No co-crystal	4N-IEB-3NitroBA	2511, 1912	Co-crystal
4N-IEB-Fum	-	No co-crystal	4N-IEB-4NitroBA	-	No co-crystal
4N-IEB-Mal	2513, 1886	Co-crystal	4N-IEB-BA	-	No co-crystal
4N-IEB-Glut	-	No co-crystal	4N-IEB-4BromoBA	-	No co-crystal
4N-IEB-Pim	-	No co-crystal	4N-IEB-4IodoBA	-	No co-crystal
4N-IEB-Aze	-	No co-crystal	4N-IEB-PentaBA	2483, 1884	Co-crystal

	Ligand	Acid	Grinded mixtures	Co-crystal?
4N-IB:4OH	1654, 1508, 1391, 1310	1668, 1606, 1592, 1446	1653, 1589, 1509, 1389, 1311	Co-crystal
4N-IB:3AminoBA	1654, 1508, 1391, 1310	1622, 1553, 1386, 1223	1654, 1507, 1388,	Co-crystal
4N-EB-Seb	1657, 1586, 1514, 1405, 1317	1686, 1186	1692, 1659, 1518	Co-crystal

C.5 IR of co-crystals of urea

Table C.5.1 IR results of grinding experiments for co-crystallizations with aliphatic acids.

Urea	Co-former	Ground mixture	Results	Urea	Co-former	Ground mixture	Results
Urea	CA	U:CA		Urea	CmA	U:CmA	
1708, 1672	1694	1697		1708, 1672	1732	1701	
1587, 1555	1571	1569, 1527		1587, 1555	-	1611, 1554	
1455	1408	1427		1455	1471, 1426	1456	
-	1262	1258	Co-crystal	-	1369, 1325	1338, 1304	Co-crystal
1145, 1038	1113, 1094	1142, 1079		1145, 976	1189, 1115	1175, 1127	
976	-	951		1038	-	1024	
-	882	876			961	966	
Urea	MA	U:MoA		Urea	SA	U:ScA	
3429, 3333	2989, 2906	3456, 3198		3429, 3333	2956	3464, 3323	
2354		2487, 1857		2354		2373, 1908	
1708, 1672	1693	1658		1708, 1672	1678	1615	
1587	-	1580		1587	-	-	
1455	1433	1463, 1404	Co-crystal	1455	1410,	1459	Co-crystal
-	1300	-		-	1306	1326	
1145	1214, 1167	1291, 1162		1145	1197	1197	
976	916	935, 876		976	635	1005	
788, 699	767, 652	747		788, 699	-	906	
Urea	FA	U:FA		Urea	GA	U:GA	
3429, 3333		3472, 3218		3429, 3333		3325	
2354	2910, 2843	2362, 1869		2354	2911, 2823	2362, 1975	
1708, 1672	1686	1617		1708, 1672	1686	1701, 1623	
1587		1521		1587	1427	1575	
1455	1427, 1405	1460	Co-crystal	1455	1332	1463	Co-crystal
-	1350	1300		-	1279, 1222		
	1297, 1232	-		1145	1185	1138	
1145	1186	1146, 1013		976	922	970	
976	923	936		788, 699	680	782	
788, 699	676	773, 621					
Urea	AA	U:AdA		Urea	PA	U:PA	
3429, 3333	2950	3472, 3225		3429, 3333		3468, 3210	
2354		2389, 1898		2354	2911, 2823	2420, 1889	
1708, 1672	1683	1687, 1636		1708, 1672	1686	1693, 1658	
1587	1461	1563		1587	1427	1572	
1455	1406	1495	Co-crystal	1455	1332	1434, 1407	Co-crystal
-	1314	1338		-	1279, 1222	1349, 1264	
1145	1272, 1189	1258, 1182		1145	1185	1189, 1091	
976		971		976	922	976	
788,699	733	731, 680		788, 699	680	778, 615	
Urea	SbA	U:SbA		Urea	AzA	U:AzA	
3429, 3333		3338, 3203		3429, 3333		3464, 3222	
2354	2934, 2868	2353, 1818		2354	2911, 2823	2483, 1904	
1708, 1672	1684	1702, 1649		1708, 1672	1686	1786, 1657	
1587	1422, 1407	1592		1587	1427	1591, 1435	
1455	1330	1487	Co-crystal	1455	1332	1314	Co-crystal
-	1248	1331		-	1279, 1222	1242	
		1229		1145	1185	1162	
1145, 976	1187	1186		976	922	983, 904	
788, 699	922	953		788, 699	680	771, 613	
Urea	TA	U:TA		Urea	MeA	U:MeA	
1708	1736, 1716	1712		1708, 1672	1686	1709	
1587	-	1622		1587		1611	
1301,1252	1219	1300, 1233		1455	1427, 1405	1463	
1038	1085	1098	Co-crystal	-	1350	1396	Co-crystal
979	936	976		1145	1186	1194, 1110	
699	685	702		976	923	982	
Urea	SeA	U:SeA					
3429, 3333		3343, 3206					
2354	2910, 2843	1830	Co-crystal				
1708, 1672	1686	1693, 1650					
1587		1592					

1145	1186	1189					
976	923	919					

Table C.5.2 IR results of grinding experiments for co-crystallizations with aromatic acids.

Urea	Co-former	Ground mixture	Results	Urea	Co-former	Ground mixture	Results
Urea	3HBA	U:3HBA		Urea	4HBA	U:3HBA	
3429, 3333	3350	3432, 3320		3429, 3333	3368	3421, 3347	
2354		2589		2354		2502, 1893	
1708, 1672	1678	1702		1708, 1672	1668, 1606	1663, 1631	
1587	1594	1597	Co-crystal	1587	1592	1584	Co-crystal
1455	1459	1452		1455	1446	1467	
-	1226	1269		-	1286	1374, 1254	
1145, 976		1157, 1106		1145, 976		1154	
788, 699	753	751		788, 699	767	841, 743	
Urea	3ABA	U:3ABA		Urea	4-ABA	U:4ABA	
3429	-	3429		3429, 3333	3460, 3355	3480, 3362	
1672	1622	1619		2354		2549, 1924	
-	1553	1533, 1514		1708, 1672	1658,1623	1618	
-	1386	1378	No Co-crystal	1587	1597	1588	Co-crystal
	1223	1217		1455	1420	1408	
1145	-	1138		-	1309	1313	
788	786, 757	786, 755		1145, 976	1285	1271, 1166	
699	666	673		788, 699	768	912, 837, 771	
Urea	PtA	U:PtA		Urea	Py2_CA	U:Py2_CA	
3429, 3333				1672		1666	
2354	1667	1636		1587, 1455		1582, 1450	
1587	1583	1541			1386	1395	
1455	1399	1370	Co-crystal		1309, 1270	1310, 1269	Co-crystal
1145	1268	1286		1145	1152	1152	
976				1038	1052	1049	
788, 699				715	715	715	
				699	698	690	
Urea	Py2,3_DCA	U:Py2,3_DCA		Urea	Py2,5_DCA	U:Py2,5_DCA	
1672	2437, 1865	2418, 1865		3429, 3333	1696		
1587, 1455	1711	1716		2354			
	1687	1644		1708, 1672	1605	1707	No co-crystal
	1576	1298	Co-crystal	1587			
1145	1258	109		1455	1451		
1038	1094			-	1241, 1173	1234	
Urea	Py2,6_DCA	U:SbA		Urea	Py3,5_DCA	U:Py3,5_DCA	
3429, 3333		2361, 1865		3429, 3333		1721	
2354				2354			
1708, 1672	1697	1687		1708, 1672	1697, 1637	1669	No co-crystal
1587	1637, 1583	1646	Co-crystal	1587	1583	1602	
1455		1455		-	1244	1152	
-	1244	1254					
Urea	HipA	U:HipA		Urea	2,5DHBA	U:2,5DHBA	
	1736	1744		1670	1669	1669	
1672		1678		1587	1593	1520	
	1596	1596	No co-crystal	1455	1453	1453	Co-crystal
1555	1552	1553			1219	1220	
1455	1489	1463					
1145	1175	1173					
Urea	3,4DHBA	U:3,4DHBA		Urea	3,5DHBA	U:3,5DHBA	

1587	1667	1678		1672	1680, 1605	1680	Co-crystal
1455	1596	1594	Co-crystal	1587, 1455	1479	1587	
	1289	190		1145	1158	1149	
	760	740		-			

Table C.5.3 IR results of grinding experiments for co-crystallizations with nitro susbtituents.

Urea	Co-former	Ground mixture	Results	Urea	Co-former	Ground mixture	Results
Urea	2NP	U:2NP		Urea	3NP	U:3NP	
2354					1672	1676	
1708, 1672		1671		1587	1594	1585	
1587	1583	1584	No co-crystal	1455	1463	1477	Co-crystal
1455	1473	1473		1145, 976	1144	1148	
1145, 976	1131	1131		788, 699	811, 791	808, 772	
788, 699	868	866		788, 699	811, 791	808, 772	
Urea	4NP	U:4NP		Urea	1,2DNB	U:1,2DNB	
3429		3120		1672	1526	1674, 1523	
1672	1457	1445		1587	1353	1585, 1350	
-	1328	1334	Co-crystal	1455	790	1461, 787	Co-crystal
-	1110	1109					
	850, 816	844, 825					
Urea	1,3DNB	U:1,3DNB		Urea	1,4DNB	U:1,4DNB	
1672	1639, 1611	1596		1672	1548	1602	
1587	1521	1504	Co-crystal	1587	1526	1545	Co-crystal
1455	1424			1455	1341, 1105	1338	
	1326	1336					
Urea	2NBA	U:2NBA		Urea	3NBA	U:3NBA	
1587, 1670	1672	1685, 1608			1504	1509	
1455	1587	1528	Co-crystal	1455	1469	1459	Co-crystal
	1455	1484			1244	1242	
1133, 1290	1145	1257		788	792	789	
Urea	4NBA	U:4NBA					
1587	1684, 1599	1680, 1598					
1455	1423	1419	Co-crystal				
	1274	1275					

Table C.5.4 IR results of grinding experiments for co-crystallization of amino acids.

323

Urea	Co-former	Grounded mixture	Results	Urea	Co-former	Grounded mixture	Results
Urea	**His**	**U: His**		**Urea**	**Arg**	**U: Arg**	
1672	1628	1676, 1620		1587	1602	1607	
1587	1580	1597		1455	1471	1455	
1455	1453	1463	Co-crystal		1344	1327	Co-crystal
	1340	1342		1133	1142	1145	
	1246	1244					
Urea	**Ala**	**U: Ala**		Urea	**Asp**	**U: Asp**	
1587	1586	1581			1504	1509	
1455	1455	1458		1455	1469	1459	
	1235	1226	No Co-crystal		1244	1242	No Co-crystal
	788	781		788	792	789	
	678, 659	680, 659					
Urea	**Glu**	**U: Glu**		Urea	Lys	**U: Lys**	
1708	-	1716		1587		1681, 1623	Co-crystal
1455	-	1462	Co-crystal		1574	1583	
-	1255	1256		1455		1450	
-	864	864		1133			

Table C.5.5 IR results of grinding experiments for co-crystallization of N-oxides.

Urea	Co-former	Grounded mixture	Results	Urea	Co-former	Grounded mixture	Results
Urea	**Py_NO**	**U: Py_NO**		**Urea**	**TmPy**	**U: TmPy_NO**	
1672	1649	1615	Co-crystal	1587	1581	1618	Co-crystal
1587	1592	1597		1455	1472	1451	
1455	1467, 1430	1467			1382,1318	1382, 1310	
	1308	1385		1133	1136	1130	
1145	1211, 1002	1091			1010	1029	
	884	858		803		811	
Urea	**Py_NNO**	**U: Py_NNO**		**Urea**	**TmPy_NNO**	**U: TmPy_NNO**	
1587	1588	1662, 1598	Co-crystal		1523	1509	Co-crystal
1455	1481, 1440	1459		1455	1444, 1427	1459	
	1255	1298			1387, 1333	1242	
1145	1027	1108, 1090			1303		
	859	803		788	1113	789	
788	796	669			864		
Urea	**BiP_NO**	**U: BiP_NO**					
1708	1728	1857, 1666	Co-crystal				
1455	1585, 1530	1576, 1474					
-	1405						
-	1217	1237					
	1074, 988	1176, 1033					
	802, 732	844, 728					

Table C.5.6 IR results of grinding experiments for co-crystallization of nicotinamides**.**

Urea	IN	U:IN	Results	Urea	Me_N	U:Me_N	Results
	1660, 1619	1656, 1621		1587	1585	1595	
	1550	1550		1555	1548	1558	
1455		1457	Co-crystal	1455		1466	Co-crystal
	1407, 1389	1408, 1391		1038	1026	1021	
788		783					
Urea	2Cl_N	U:2Cl_N	Results	Urea	6Cl_N	U:6Cl_N	Results
	1669, 1622	1662, 1624			1650, 1611	1649, 1615	
1587	1576	1585			1397	1402	
1455	-	1450	No co-crystal	1145	1142	1142	No-cocrystal
	1385	1383			1017	1018	
	1156	1158		803, 788	807, 783	805, 779	
	1070	1070					

Table C.5.7 IR results of grinding experiments for co-crystallization of various**.**

Urea	Co-former	Grounded mixture	Results	Urea	Co-former	Grounded mixture	Results
Urea	Biu	U: Biu		Urea	Ur	U:Ur	
1672	1569	1695	No co-crystal	1672	1650	1652	No Co-crystal
1587		1587		1555	1574	1557	
1455	-	1457		-	1345, 1297	1326	
	1406, 1356	1412, --1356		-	1117	1115	
	1221	1218			987	987	
1145		1135					
	761, 708	762, 709					
Urea	Ma	U: Ma		Urea	Glu	U: Glu	
	1826	1828	Co-crystal	1672		1682, 1630	Co-crystal
1672		1678		1587		1604	
1587, 1555		1616, 1595		1455	1442, 1424	1459	
1455		1459			1332	1339	
	1373	1378				1021	
	1134	1140		1145		994	
Urea	Im	U: Im		Urea	Re	U: Re	
1672	1649	1647	Co-crystal	1672	1604	1598	Co-crystal
1587				1587		1578	
1455	1505	1507, 1448		1455	1486	1480	
	1268	1270			1166, 1143	1168, 1147	
Urea	Lut	U: Lut		Urea	Phen	U: Phen	
1672	1691	1683	Co-crystal	1672	1644		Co-crystal
1587	1596	1627, 1605		1587	1502, 1420	1500	
1455	1447			1455		1418	
	1383				1343	1342	
	1194				1090	1088	

Appendix D - Single Crystal X-ray diffraction data

D.1 Chapter 2-Pyrazole target molecules

D.1.1Experimental details

All datasets except those previously reported were collected on a Bruker Kappa APEX II system using MoKα radiation. Data were collected using APEX2[5] software. Initial cell constants were found by small widely separated "matrix" runs. Data collection strategies were determined using COSMO[6]. Scan speed and scan widths were chosen based on scattering power and peak rocking curves.

The unit cell constants and orientation matrix were improved by least-squares refinement of reflections thresholded from the entire dataset. Integration was performed with SAINT[7], using this improved unit cell as a starting point. Precise unit cell constants were calculated in SAINT from the final merged dataset. Lorenz and polarization corrections were applied. Multi-scan absorption corrections were performed with SADABS[8].

Data were reduced with SHELXTL[9]. The structures were solved in all cases by direct methods without incident. Except as noted, hydrogen atoms were located in idealized positions and were treated with a riding model. All non-hydrogen atoms were assigned anisotropic thermal parameters. Refinements continued to convergence, using the recommended weighting schemes.

D.1.2 Hydrogen bond geometries

Co-crystal	D-H...A/A°	D-H/ A°	H...A// A°	D...A// A°	D-H...O/ °
P1	N1-H1...O17	0.97(3)	1.90(4)	2.862(3)	171.(3)
	N6-H6...N11	0.88(3)	2.15(3)	3.033(3)	173.(3)
	N10-H10...O8	0.91(3)	2.03(3)	2.930(3)	169.(3)
	N15-H15...N2	0.83(3)	2.26(3)	3.087(4)	175.(3)
P2	N1-H1...O18	0.96(4)	1.82(4)	2.771(3)	170.(3)
	N6 -H6 ...N12	0.92(4)	1.99 (4)	2.903(4)	173.(3)
	N11- H11... O8	0.91(3)	1.91(3)	2.787(3)	161.(3)
	N16 -H16 ...N2	0.90(3)	2.12(3)	3.014(4)	170.(3)
P3	N1 -H1... O22	0.95(3)	1.91(3)	2.831(3)	161.(2)
	N6 -H6 ...N16	0.93(3)	2.12(3)	3.041(3)	171.(2)
	N15- H15 ...O8	0.88(3)	2.20(3)	2.866(3)	133.(2)
	N20 -H20 ...N2	0.93(3)	2.07(3)	2.981(3)	168.(3)
P4	N11 -H11 ...O21	0.913(18)	1.937(18)	2.8103(14)	159.7(16)
	N13 -H13 ...N12	0.886(18)	2.045(18)	2.9225(14)	170.2(15)

P7	N1 -H1 …O20	0.91(5)	1.89(7)	2.755(16)	159.(12)
	N7 -H7 …N13	0.91(5)	2.21(8)	2.946(15)	138.(8)
	N12 -H12 …O9	0.91(5)	1.86(6)	2.755(15)	168.(12)
	N18 -H18… N2	0.88(10)	2.06(11)	2.939(15)	174.(9)
P10	N1 -H1 …N2	0.89(6)	2.19(7)	3.047(9)	163.(13)
	N7 -H7 …N12	0.91(9)	2.23(10)	3.112(9)	164.(8)
P11	N1 -H1… N12	0.92(2)	2.13(2)	2.910(2)	142.3(16)
	N6 -H6 …N2	0.869(19)	2.15(2)	3.0002(19)	166.6(16)

D.1.3 Crystallographic data

Code	P1	P2	P3	P4	P7	P10	P11
Formula moiety	$C_5H_7N_3O$	C6H9N3O	$C_{10}H9N3O$	C14 H22 N6 O2	C14H22N6O2	C10H10N4O	C9H8N4O
Empirical formula	C5H7N3O	C6H9N3O	C_{10}H9N3O	C14 H22 N6 O2	C14H22N6O2	C10H10N4O	C9H8N4O
Molecular weight	125.14	139.16	187.20	306.37	306.37	3 202.22 37.35	188.19
Color, Habit	blocks, colorless	plates, colorless	Blocks, colorless	Blocks, colorless	Bronze, Rod	Colorless, Plate	Blocks, Colorless
Crystal system	Monoclinic	orthorhombic	orthorhombic	monoclinic	Monoclinic	orthorhombic	monoclinic
Space group, Z	P2(1)/c, 8	pbca, 16	pbca, 16	C 1 c 1, 4	C-2yc, 4	Pna21, 4	C-2yc, 8
a, Å	15.212(4)	9.962(4)	10.586(3)	7.570(2)	7.570(2)	21.043(11)	11.261(3)
b, Å	11.424(3)	10.024(4)	14.558(4)	21.799(6)	21.799(6)	8.832(5)	11.476(3)
c, Å	7.214(2)	28.792(12)	25.523(7)	10.659(3)	10.659(3)	5.500(3)	13.654(3)
α, º	90.00	90.00	90.00	90	90.00	90.00	90
β, º	102.42(2)	90.00	90.00	108.247(17)	108.247(17)	90.00	103.67(2)
γ, º	90.00	90.00	90.00	90	90.00	90.00	90
Volume, Å³	1224.3(6)	2875.(2)	3933.4(19)	1670.5(8)	1670.5(8)	1022.2(10)	1714.5(7)
Density, g/cm³	1.358	1.286	1.265	1.218	1.218	1.426	1.458
T, ºK	296(2)	296 (2)	200 (2)	130(2)	120(2)	180.(2)	130(2)
Crystal size, min x mid x max	0.128X0.208 X0.278	0.064X0.168X 0.304	0.108X0.252X 0.394	0.136X0.202 X0.252	0.136X0.202 X0.252	0.075x0.352 x0.502	0.112X0.218 X0.432
X-ray wavelength, Å	0.71073	0.71073	0.71073	0.71073	0.71073	0.71073	0.71073
μ, mm^{-1}	0.100	0.092	0.086	0.086	0.086	0.091	0.102
Trans min / max	0.97/0.99	0.97, 0.99	0.97/0.99	0.98, 0.99	0.98/ 0.99	0.96/ 0.99	0.96/0.99
ϑ_{min}, º	2.25	2.83	2.50	1.87	1.87	1.94	2.57
ϑ_{max}, º	21.98	21.20	22.05	25.39	25.39	25.30	24.88
Reflections							
collected	30410	30040	69812	1523	1523	23993	21370
independent	2263	2543	3486	998	1523	1800	1565
observed	1384	1534	2093	219	998	1184	1260
R_{int}	0.1016	0.1286	0.1294	0.1197	0.0000	0.4966	0.0487
Threshold expression	> 2σ(l)	> 2σ(l)	> 2σ(l)	> 2σ(l)	> 2σ(l)	> 2σ(l)	> 2σ(l)
No. parameters	181	200	270	219	219	145	135
No. restraints	0	0	0	5	0	0	0
R_1 (observed)	0.0650	0.0580	0.0518	0.1197	0.1197	0.0883	0.0381
wR_2 (all)	0.1730	0.1555	0.1518	0.2232	0.2557	0.2194	0.1037
Goodness of fit (all)	1.051	1.034	1.037	1.075	1.075	0.976	1.073
ρ_{max}, ρ_{min}, e Å^{-3}	0.215,-0.241	0.641, -0.491	0.195,-0.179	0.552, -0.600	0.552, -0.600	0.265,-0.341	0.226, -0.206
Completeness to 2ϑ limit	0.972	0.993	0.988	0.989	0.989	0.976	0.987

D.2 Chapter 3-Pyrazole co-crystals

Code	P1-PentafluoroBA	P2-Fum	P2-Suc	P2-Adi	P2-Pim	P3-3NitroBA	P3-4-amino BA
Formula moiety	C7 H F5 O2, C5 H7 N3 O	$(C_{12}H_{10}N_4O)$ $(C_3H_4O_4)$	$(C_6H_9N_3O)_2$ $(C_4 H_6 O_4)$	(C6 H9 N3 O)2 (C6 H10 O4)	(C6 H9 N3 O)2 (C7 H12 O4)	C7 H5 N O4, C6 H9 N3 O	C6 H9 N3 O, C7 H7 N O2
Empirical formula	C12 H8 F5 N3 O3	$C_{16} H_{22} N_6 O_6$	$C_{16} H_{24} N_6 O_6$	C18 H28 N6 O6	C19 H30 N6 O6	C13 H14 N4 O5	C13 H16 N4 O3
Molecular weight	337.21	394.40	396.41	424.46	438.49	306.28	276.30
Color, Habit	Colorless, blocks	Colorless, Plate	Colorless, Plate	Colorless, Plate	Colorless, Prism	Colorless, Rhombohedral	Orange, blocks
Crystal system	monoclinic	Triclinic	Triclinic	Monoclinic	monoclinic	Monoclinic	orthorhombic
Space group, Z	P2(1)/c, 4	$P\bar{1}$, 4	$P\bar{1}$, 1	P2(1)/n, 2	'C 2/c', 4	$P2_1/n$, 4	$P b c a$, 8
a, Å	7.729(2)	10.649(2)	14.103(3)	5.2338(5)	14.1481(9)	8.484(4)	13.764(6)
b, Å	14.281(4)	14.078(3)	16.797(4)	13.6144(12)	13.1406(9)	11.985(6)	8.082(3)
c, Å	11.865(3)	13.409(3)	20.357(5)	14.8734(14)	12.4456(8)	14.344(7)	24.375(10)
α, º	90	90.00	94.460(3)	90.00	90	90.00	90.00
β, º	93.293(14)	112.837(19)	90.011(3)	100.113(3)	99.494(2)	106.86(3)	90.00
γ, º	90	90.00	93.296(4)	90.00	90	90.00	90.00
Volume, Å³	1307.5(6)	1852.6(7)	4800(2)	1043.34 (17)	2282.1(3)	1395.8(11)	2711.5(19)
Density, g/cm³	130.(2)	1.414	1.371	1.351	1.276	1.457	1.354
T, ºK	130.(2)	120(2)	120(2)	120(2)	120(2)	130.(2)	130.(2)
Crystal size, min x mid x max	0.14 x 0.37 x 0.54	0.18 x 0.24 x 0.28	0.12 x 0.30 x 0.42	0.12 x 0.26 x 0.46	0.20 x 0.24 x 0.32	0.08 x 0.26 x 0.29	0.10 x 0.16 x 0.27
X-ray wavelength, Å	0.71073	0.71073	0.71073	0.71073	0.71073	0.7107	0.7107
μ, mm^{-1}	0.170	0.110	0.107	0.103	0.096	0.114	
Trans min / max	0.91/0.98	0.9698/ 0.9804	0.9566/ 0.9873	0.9542/ 0.9878	0.681/ 0.981	0.97 / 0.99	0.97 / 0.99
ϑ_{min}, º	2.64	2.08	6.13	3.16	2.129	2.26	2.23
ϑ_{max}, º	25.88	32.07	30.15	31.39	31.081	26.02	22.76
Reflections							
collected	35456	23564	122238	12032	7114	13782	55537
independent	2568	5910	27543	3334	7114	2167	2543
observed	2183	3947	14260	2570	5633	1507	1725
R_{int}	0.0209	0.0307	0.0991	0.0354	0.0360	0.0725	0.0409
Threshold expression	> $2\sigma(I)$	> $2\sigma(I)$	> $2\sigma(I)$	> $2\sigma(I)$	> $2\sigma(I)$	> $2\sigma(I)$	> $2\sigma(I)$
No. parameters	221	275	1261	147	153	212	202
No. restraints	0	0	0	0	0	0	1
R_1 (observed)	0.0395	0.0532	0.1017	0.0463	0.1495	0.0816	0.0866
wR_2 (all)	0.0789	0.1763	0.2694	0.1304	0.1747	0.1360	0.1190
Goodness of fit (all)	1.031	1.185	1.303	1.096	1.138	1.002	1.037
ρ_{max}, ρ_{min}, e Å^{-3}	0.242, -0.244	0.435, -0.383	1.297, -0.548	0.343, -0.288	0.280, -0.290	0.218, -0.216	0.181, -0.229
2ϑ limit, º	26.08	30.00	30.00	31.47	67.50	30.00	
Completeness to 2ϑ limit	0.989	0.998	0.980	0.960	0.96	0.789	

Code	P3-pentafluoroB	P4-Fum	P4-adi	P7-sub	P8-Aze	P8-sub
Formula moiety	C7 H F5 O2, C6 H9 N3 O	C7 H11 N3 O, 0.5(C4 H4 O4)	$(C_{12}H_{10}N_4O)$ $(C_3H_4O_4)$	$(C_6H_9N_3O)_2$ $(C_4 H_6 O_4)$	(C11 H11 N3 O),2 C9 H16 O4	(C11 H11 N3 O), C8 H14 O4
Empirical formula	C13 H10 F5 N3 O3	C9 H13 N3 O3	$C_{16} H_{22} N_6 O_6$	$C_{16} H_{24} N_6 O_6$	C31 H38 N6 O6	C19 H25 N3 O5
Molecular weight	351.24	211.22	394.40	396.41	590.67	375.42
Color, Habit		Colorless, Rhombohedral	Colorless, Plate	Colorless, Plate	Colorless, needles	
Crystal system	triclinic	Monoclinic	Triclinic	Triclinic	orthorhombic	Monoclinic
Space group, Z	$P-1$, 4	$P2_1/c$, 4	$P\bar{1}$, 4	$P\bar{1}$, 1	$P c a 21$, 4	$P2_1/c$, 4
a, Å	11.563(3)	8.391(3)	10.649(2)	14.103(3)	27.095(5)	9.247(4)
b, Å	12.171(3)	14.273(5)	14.078(3)	16.797(4)	5.0230(9)	22.132(8)
c, Å	12.261(3)	9.443(4)	13.409(3)	20.357(5)	22.862(4)	10.159(4)
α, º	70.680(11)	90	90.00	94.460(3)	90.00	90.00
β, º	77.235(12)	111.223(19)	112.837(19)	90.011(3)	90.00	90.00
γ, º	62.474(11)	95.615(3)	90.00	93.296(4)	90.00	90.00

Volume, Å³	1439.2(5)	1054.2(6)	1852.6(7)	4800(2)	3111.5(10)	1867.3(12)
Density, g/cm³	1.621	1.331	1.414	1.371	1.261	1.335
T, ºK	130.(2)	130.(2)	120(2)	120(2)	130.(2)	120(2)
Crystal size, min x mid x max	0.14 x 0.28 x 0.436	0.176 x 0.303 x 0.458	0.18 x 0.24 x 0.28	0.12 x 0.30 x 0.42	0.12 x 0.15 x 0.41	0.12 x 0.32 x 0.42
X-ray wavelength, Å	0.71073	0.7107	0.71073	0.71073	0.71073	0.71073
μ, mm⁻¹	0.158	0.102	0.110	0.107	0.089	0.098
Trans min / max	0.94/0.98	0.95 / 0.98	0.9698/ 0.9804	0.9566/ 0.9873	0.96/0.99	
ϑ_{min}, º	2.26	2.60	2.08	6.13	2.33	6.29
ϑ_{max}, º	25.68	26.57	32.07	30.15	25.26	30.92
Reflections						
collected	43427	37645	23564	122238	26807	8600
independent	5566	11890	5910	27543	6268	4910
observed	4295	7481	3947	14260	4578	3653
R_{int}	0.0326	0.0410	0.0307	0.0991	0.0739	
Threshold expression	> 2σ(I)	> 2σ(I)	> 2σ(I)	> 2σ(I)	> 2σ(I)	> 2σ(I)
No. parameters	459	548	275	1261	414	257
No. restraints	2	0	0	0	6	0
R_1 (observed)	0.0367	0.0580	0.0532	0.1017	0.0812	0.0563
wR_2 (all)	0.0961	0.1839	0.1763	0.2694	0.1025	
Goodness of fit (all)	1.029	1.023	1.185	1.303	1.066	1.081
ρ_{max}, ρ_{min}, e Å⁻³		0.243, -0.228	0.435, -0.383	1.297, -0.548	0.197, -0.235	
2ϑ limit, º		30.00	30.00	30.00	26.26	
Completeness to 2ϑ limit		0.98	0.998	0.980	0.999	

Code	P10-Fum	P10-Mal	P11-Aze	P11-Dod
Formula moiety	C10 H10 N4 O, C5 H8 O4, H2 O	C10 H10 N4 O, 0.5(C4 H4 O4	2(C9 H8 N4 O), C9 H16 O4	(C9H8N4O) (C12 H22 O4)
Empirical formula	C15 H20 N4 O6	C12 H12 N4 O3	C27 H32 N8 O6	C21 H30 N4 O5
Molecular weight	352.35	260.26	564.61	418.49
Color, Habit	Yellow, chunk	Colorless, rectangular	Colorless, rectangular	Colorless, needle
Crystal system	monoclinic	triclinic	monoclinic	orthorhombic
Space group, Z	P21/c,4	P -1,4	P21/c,4	P n a 21, 8
a, Å	11.9858(2)	5.44270(10)	16.106(5)	19.5258(4)
b, Å	13.3594(2)	6.8606(2)	17.267(5)	5.14640(10)
c, Å	10.7736(2)	17.4166(4)	10.260(3)	42.2729(10)
α, º	90	89.5020(10)	90	90
β, º	90.7100(10)	89.0240(10)	100.303(17)	90
γ, º	90	70.6290(10)	90	90
Volume, Å³	1724.97(5)	613.43(3)	2807.3(15)	4247.90(16)
Density, g/cm³	1.357	1.409	1.336	1.309
T, ºK	296.(2)	296.(2)	130.(2)	200(2)
Crystal size, min x mid x max	0.04x 0.050x 0.050	0.020x0.05x0.06	0.12X0.24X0.43	0.02 x 0.04 x 0.06
X-ray wavelength, Å	0.71073	0.71073	0.71073	0.71073
μ, mm⁻¹	0.899	0.877	0.097	0.776
Trans min / max	0.96, 0.96	0.95, 0.98	0.96, 0.99	0.65/ 0.98
ϑ_{min}, º	3.69	5.08	1.28	2.09
ϑ_{max}, º	69.86	69.70	26.63	70.42
Reflections				
collected	20722	8530	17199	28297
independent	3229	2170	5122	7413
observed	2752	1959	3035	6268
R_{int}	0.0367			0.0689
Threshold expression	> 2σ(I)	> 2σ(I)	> 2σ(I)	> 2σ(I)
No. parameters	251	185	394	573

No. restraints	0	1	5	4
R_1 (observed)	0.0376	0.0388	0.0531	0.0522
wR_2 (all)	0.1042	0.1000	0.1236	0.1279
Goodness of fit (all)	1.051	1.073	0.970	1.057
ρ_{max}, ρ_{min}, e Å^{-3}	0.357, -0.0194	0.219, -0.231	0.290, -0.380	0.223, -0.261
2ϑ limit, º	30.00	30.00	30.00	30.00
Completeness to 2ϑ limit	0.983	0.935	0.870	0.988

Co-crystal	D-H...A/A°	D-H/ A°	H...A// A°	D...A// A°	D-H...O/ °
P1-PentafluoroBA	N1 H1 O8	0.89(2)	1.85(2)	2.7201(18)	166.0(18)
	N6 H6 O17	0.845(18)	2.005(19)	2.8407(17)	169.7(16)
	O18 H18 N2	0.93(2)	1.72(2)	2.6522(16)	177.(2)
P2-Fum	N11 H11 O21	0.864(17)	1.896(17)	2.6958(16)	153.3(14)
	N13 H13 O52	0.902(16)	1.986(16)	2.8829(15)	173.1(14)
	N31 H31 O41	0.842(17)	1.956(17)	2.7674(16)	161.6(15)
	N33 H33 O55	0.902(16)	1.955(16)	2.8498(14)	171.7(14)
	O51 H51 N12	0.989(18)	1.624(19)	2.6106(14)	174.5(15)
	O54 H54 N32	0.930(19)	1.69(2)	2.6153(14)	173.5(15)
P2-Suc	N11_1 H11_1O26_1	0.88	1.95	2.765(3)	152.4
	N11_2 H11_2 O26_2	0.88	1.94	2.745(3)	152.3
	N11_3 H11_3 O26_3	0.88	1.95	2.768(3)	153.7
	N11_4 H11_4 O26_4	0.88	1.96	2.785(3)	154.8
	N11_5 H11_5 O26_5	0.88	1.96	2.771(3)	153.6
	N21_1 H21_1 O16_1	0.88	1.94	2.753(3)	152.3
	N21_2 H21_2 O16_2	0.88	1.94	2.746(3)	152.1
	N21_3 H21_3 O16_3	0.88	1.96	2.773(3)	152.6
	N21_4 H21_4 O16_4	0.88	1.96	2.779(3)	154.3
	N21_5 H21_5 O16_5	0.88	1.96	2.778(3)	155.0
	N13_1 H13_1 O32_1	0.88	2.00	2.865(3)	168.7
	N13_2 H13_2 O32_2	0.88	2.01	2.875(3)	169.5
	N13_3 H13_3 O32_3	0.88	2.00	2.873(3)	168.2
	N13_4 H13_4 O32_4	0.88	2.00	2.862(3)	167.4
	N13_5 H13_5 O32_5	0.88	2.01	2.867(3)	165.5
	N23_1 H23_1 O35_1	0.88	1.99	2.864(3)	169.4
	N23_2 H23_2 O35_2	0.88	2.02	2.885(3	169.2
	N23_3 H23_3 O35_3	0.88	2.01	2.874(3)	168.3
	N23_4 H23_4 O35_4	0.88	2.00	2.860(3)	165.8
	N23_5 H23_5 O35_5	0.88	1.99	2.860(3)	169.2
	O31_1 H31_1 N12_1	0.84	1.80	2.639(3)	176.6
	O31_2 H31_2 N12_2	0.84	1.80	2.638(3)	176.2
	O31_3 H31_3 N12_3	0.84	1.80	2.643(3)	176.9
	O31_4 H31_4 N12_4	0.84	1.80	2.640(3)	176.0
	O31_5 H31_5 N12_5	0.84	1.82	2.657(3)	177.3
	O34_1 H34_1 N22_1	0.84	1.80	2.640(3)	177.1
	O34_2 H34_2 N22_2	0.84	1.79	2.629(3)	175.7
	O34_3 H34_3 N22_3	0.84	1.82	2.656(3)	177.1
	O34_4 H34_4 N22_4	0.84	1.80	2.644(3)	176.7
	O34_5 H34_5 N22_5	0.84	1.80	2.637(3)	175.9
P2-Adi	N11 H11 O21	0.890(18)	1.893(18)	2.7655(14)	166.3(14)
	N13 H13 O32	0.888(16)	2.039(17)	2.9142(13)	168.3(15)
	O31 H31 N12	0.930(18)	1.785(18)	2.7134(13)	175.7(16)
P2-Pim	O31 H31 N12	0.96(2)	1.80(2)	2.7590(14)	175.1(18)
	N11 H11 O21	0.883(19)	1.87(2)	2.7419(16)	168.1(18)
	N13 H13 O32	0.936(19)	1.874(19)	2.8051(14)	172.9(16)

P3-3NitroBA	O19 H19 N2	0.90(3)	1.74(3)	2.641(3)	173.(3)
	N6 H6 O18	0.90(3)	1.90(3)	2.789(3)	171.(3)
	N1 H1 O8	0.96(3)	1.82(3)	2.770(3)	172.(3)
P3-4-amino BA	N1 H1 O8	0.94(3)	1.77(3)	2.714(3)	176.(2)
	N6 H6 O18	0.95(3)	1.95(3)	2.893(3)	170.(2)
	O19 H19 N2	0.98(3)	1.68(3)	2.655(2)	174.(3)
	N20 H20A O18	0.92(3)	2.25(3)	3.127(3)	160.(3)
P3-pentafluoroB	N1 H1 O18 1_565	0.89(2)	1.87(2)	2.755(2)	170.(2)
	N6 H6 O42 1_465	0.82(2)	2.04(2)	2.849(2)	167.5(19)
	N11 H11 O8 1_554	0.87(2)	1.87(2)	2.739(2)	173.(2)
	N16 H16 O28 1_554	0.87(2)	2.03(2)	2.889(2)	170.9(19)
	O29 H29 N12 1_556	0.96(2)	1.67(2)	2.6272(19)	174.(2)
	O43 H43 N2 1_645	0.93(2)	1.71(2)	2.6478(19)	178.(2)
P4-Fum	O14 H14 N2	0.92(2)	1.77(2)	2.6882(16)	179.(2)
	N6 H6 O15	0.916(16)	2.121(16)	3.0229(16)	167.8(15)
	N1 H1 O10	0.902(19)	1.865(19)	2.7624(17)	173.3(16)
P4-adi	O21 H21 N2	0.90(3)	1.79(3)	2.683(4)	173.(4)
	O13 H13 O9	0.87(3)	1.77(4)	2.636(4)	175.(6)
	N7 H7 O20	0.91(4)	2.01(4)	2.909(4)	167.(3)
	N1 H1 O12	0.88(5)	1.94(5)	2.811(4)	167.(5)
P5-sub	N11H11 O27	0.896(17)	1.908(18)	2.7212(16)	149.9(15)
	N27H27 O52	0.910(15)	2.024(16)	2.9154(15)	166.1(14)
	N31H31 O47	0.881(17)	1.924(17)	2.7421(15)	153.8(15)
	N47H47 O59	0.904(15)	2.086(15)	2.9744(15)	167.1(13)
	O51H51 N12	0.965(19)	1.756(19)	2.7205(16)	177.7(17)
	O58H58 N32	0.966(17)	1.756(18)	2.7150(15)	171.6(15)
P6-Aze	N1 H1 O24	0.87(4)	1.89(4)	2.747(5)	165.(5)
	N7 H7 O43	0.94(4)	2.00(4)	2.940(5)	175.(5)
	N16 H16 O9	0.89(5)	1.99(5)	2.826(5)	157.(4)
	N22 H22 O32	0.88(4)	2.11(4)	2.962(5)	165.(5)
	31 H31 N17	0.81(4)	1.82(4)	2.622(5)	170.(8)
	O42 H42 N2	0.86(4)	1.79(4)	2.654(5)	176.(6)
P6-sub	N13 H13 O32	0.90(2)	2.00(2)	2.8819(19)	166.0(17)
	O31 H31 N12	0.92(2)	1.76(2)	2.6757(18)	175(2)
	O38 H38 O27	0.91(2)	1.77(2)	2.6738(17)	176(2)
P10-Fum	N1 H1 N12	0.92(2)	1.96(2)	2.8701(17)	169.9(17)
	N7 H7 O16	0.855(18)	2.040(19)	2.8839(16)	168.9(16)
	O17 H17 N2	0.93(2)	1.72(2)	2.6501(16)	173.(2)
	O24 H24 O25	0.95(3)	1.63(3)	2.5714(18)	171.(3)
	O25 H25A O23	0.88(3)	1.91(3)	2.772(2)	168.(3)
	O25 H25B O9	0.86(3)	1.90(3)	2.7396(17)	165.(3)
P10-Mal	N1 H1 N26	0.90(2)	1.95(2)	2.850(3)	175.(2
	N6 H6 O29	0.89(2)	2.00(3)	2.882(3)	169.(2)
	N15 H15 O8	0.92(2)	1.89(2)	2.790(3)	165.(3)
	N20 H20 O41	0.903(19)	2.03(2)	2.911(3)	166.(2)
	O30 H30 N2	0.92(2)	1.76(2)	2.673(3)	172.(3)
	O40 H40 N16	0.92(2)	1.77(2)	2.685(3)	175.(3)
P11-Aze	N1 H1 N26	0.90(2)	1.95(2)	2.850(3)	175.(2)
	N6 H6 O29	0.89(2)	2.00(3)	2.882(3)	169.(2)
	N15 H15 O8	0.92(2)	1.89(2)	2.790(3)	165.(3)
	N20 H20 O41	0.903(19)	2.03(2)	2.911(3)	166.(2)
	O30 H30 N2	0.92(2)	1.76(2)	2.673(3)	172.(3)
	O40 H40 N16	0.92(2)	1.77(2)	2.685(3)	175.(3)
P11-Dod	O60 H60 N16	0.81(6)	1.88(6)	2.682(5)	170.(6)
	O45 H45 N12	0.95(5)	1.77(5)	2.701(5)	165.(5)
	O43 H43 N26	0.89(7)	1.81(7)	2.695(5)	169.(7)
	O30 H30 N2	0.89(5)	1.79(6)	2.675(5)	169.(8)
	N20 H20 O59	0.81(6)	2.35(6)	3.131(5)	164.(6)
	N17 H17 O44	0.97(5)	1.96(5)	2.801(5)	144.(5)
	N6 H6 O29	0.81(6)	2.28(7)	3.070(5)	164.(6)
	N3 H3 O46	0.81(6)	2.02(5)	2.797(5)	161.(5)

D.3 Chapter 4-Thiazole target molecules

Code	T2	T4	T6	T10	T11
Formula moiety	$C_6H_8N_2OS$	$C_7H_{10}N_2OS$	$C_{11}H_{10}N_2OS$	C10 H9 N3 O S'	C9 H7 N3 O S
Empirical formula	$C_6H_8N_2OS$	$C_7H_{10}N_2OS$	$C_{11}H_{10}N_2OS$	C10 H9 N3 O S'	C9 H7 N3 O S
Molecular weight	156.20	170.23	218.27	219.26	205.24
Habit, color	Plate, colorless	Blocks, colorless	Blocks, colorless	Plate, yellow	Blocks, colorless
Crystal system	Monoclinic	orthorhombic	monoclinic	monoclinic	monoclinic
Space group, Z	P2(1)/c, 4	Ccca, 16	I 1 2/a 1	P21	C 2/c
a, Å	5.2743(9)	14.260(4)	21.635(6)	3.8482(5)	21.258(12)
b, Å	10.6364(19)	16.783(5)	3.9315(13)	9.8509(14)	3.7982(19)
c, Å	12.652(2)	13.999(6)	25.555(10)	12.6883(17)	23.045(12)
α, º	90.00	90.00	90.00	90	90
β, º	101.689(4)	90.00	109.975(14)	95.494(7)	115.72(3)
γ, º	90.00	90.00	90.00	90	90
Volume, Å³	695.1(2)	3350.(2)	2042.9(12)	478.781	1676.4(1
Density, g/cm³	1.493	1.350	1.419	1.521	1.626
T, ºK	120(2)	296 (2)	180 (2)	120(2)	180.(2)
Crystal size, min x mid x max	0.08X0.20X0.44	0.08X0.18X0.26	0.18X0.33X0.34	0.06x0.18x0.44	0.158x0.378x0.392
X-ray wavelength, Å	0.71073	0.71073	0.71073	0.71073	0.71073
μ, mm⁻¹	0.390	0.330	0.288	0.311	0.349
Trans min / max	0.8471/0.9695	0.92, 0.97	0.91/0.95	0.875/0.982	0.88, 0.95
ϑ_{min}, º	3.29	2.37	1.70	2.62	3.46
ϑ_{max}, º	31.95	25.10	25.84	27.28	25.03
Reflections					
collected	8738	27503	26731	4952	7943
independent	2191	1498	1951	2543	1473
observed	1937	910	1679	1930	1173
R_{int}	0.0241	0.0973	0.0501	0.0650	0.0754
Threshold expression	$> 2\sigma(I)$	$> 2\sigma(I)$	$> 2\sigma(I)$	$> 2\sigma(I)$	$> 2\sigma(I)$
No. parameters	96	106	141	136	131
No. restraints	0	0	0	1	1
R_1 (observed)	0.0317	0.0568	0.0336	0.0642	0.0934
wR_2 (all)	0.0859	0.1381	0.1018	0.1688	0.2333
Goodness of fit (all)	1.048	1.071	1.149	0.995	1.221
ρ_{max}, ρ_{min}, e Å⁻³	0.363, -0.241	0.176, -0.196	0.313, -0.346	0.676, -0.492	0.748, -0.415
Completeness to 2ϑ limit	0.997	1.000	0.990	0.979	0.969

Co-crystal	D-H...A/A°	D-H/ A°	H...A// A°	D...A// A°	D-H...O/ °
T2	N12–H12···N13	0.833(16)	2.110(16)	2.9415(14)	175.7(16)
T4	N7–H7···N3	0.90(3)	2.05(4)	2.947(4)	176.(3)
T6	N7–H7···N3	0.85(2)	2.13(2)	2.953(2)	163.(2)
T10	N12 H12 N21	0.88	2.10	2.946(6)	162.1
T11	N6 H6 N3	0.84(4)	2.16(4)	2.986(9)	167.(6)

Chapter 5-Thiazole co-crystals

Code	T1-3HydroxyBA	T2-4HydroxyBA	T4-3NitroBA	T8-Suc	T8-Sub	T8-Seb

Formula moiety	(C₅H₆N₂OS) (C₇H₆O₃)	(C₆H₈N₂OS) (C₇H₆O₃)	(C₇H₅NO₄) C₇H₁₀N₂OS)	(C₁₁H₁₀N₂OS)₂ (C₄H₆O₄	(C₁₁H₁₀N₂OS) 0.5(C₈H₁₄O₄)	(C₁₁H₁₀N₂OS)₂ (C₁₀H₁₈O₄)
Empirical formula	C₁₂H₁₂N₂O₄S	C₁₃H₁₄N₂O₄S	C₁₄H₁₅N₃O₅S	C₂₆H₂₆N₄O₆S₂	C₁₅H₁₇N₃O₃S₂	C₃₂H₃₈N₄O₆S₂
Molecular weight	280.30	294.32	337.35	554.63	305.37	638.78
Color, Habit	Bronze, Rod	Colorless, Plate	Colorless, Plate	Colorless, Plate	Colorless, Plate	Colorless, Plate
Crystal system	Monoclinic	monoclinic	Triclinic	Monoclinic	Monoclinic	Triclinic
Space group, Z	P2(1)/n, 4	P 21/n, 4	P$\bar{1}$, 2	P 21/c,2	P2(1)/n, 4	P-1, 2
a, Å	12.482(6)	8.4638(16)	7.845(3)	11.0595(16)	15.206(2)	8.529(2)
b, Å	5.082(3)	4.9284(9)	9.389(4)	6.7428(10)	5.7250(8)	8.887(2)
c, Å	19.848(9)	32.056(6)	12.093(5)	17.931(3)	17.016(3)	11.214(3)
α, º	90.00	90	98.623(13)	90	90.00	77.662(6)
β, º	101.265(9)	94.963(7)	106.647(13)	107.426(4)	90.591(7)	76.190(6)
γ, º	90.00	90	107.590(13)	90	90.00	89.982(6)
Volume, Å³	1234.6(10)	1332.1(4)	785.7(6)	1275.8(3)	1481.2(4)	805.2(3)
Density, g/cm³	1.508	1.468	1.426	1.444	1.369	1.373
T, ºK	120(2)	120(2)	130.(2)	120(2)	120(2)	120 (2)
Crystal size, min x mid x max	0.08X0.14X0.38	0.100X0.24X0.48	0.12x0.22x0.22	0.08X0.38X0.42	0.12X0.38X0.44	0.14X0.38X0.46
X-ray wavelength, Å	0.71073	0.71073	0.71073	0.71073	0.71073	0.71073
μ, mm⁻¹	0.274	0.258	0.235	0.259	0.230	0.215
Trans min / max	0.9029/ 0.9784	0.746/0.975	0.95/ 0.97	0.906/0.980	0.91/0.97	0.911/0.983
ϑ_{min}, º	3.33	2.444	2.36	1.930	2.39	2.35
ϑ_{max}, º	30.98	30.848	25.35	32.618	31.44	31.11
Reflections						
collected	18533	14370	11385	30608	16744	13784
independent	3784	4023	2907	4371	4618	4861
observed	3066	3313	2454	3940	3646	4161
R$_{int}$	0.0479	0.0387	0.0384	0.0249	0.0328	0.0398
Threshold expression	> 2σ(I)	> 2σ(I)	> 2σ(I)	> 2σ(I)	> 2σ(I)	> 2σ(I)
No. parameters	182	192	218	179	197	206
No. restraints	0	0	0	0	1	0
R₁ (observed)	0.0415	0.0445	0.0344	0.0331	0.0431	0.0554
wR₂ (all)	0.1107	0.1247	0.0904	0.0934	0.1255	0.1639
Goodness of fit (all)	1.124	1.023	1.051	1.044	1.077	1.270
ρ_{max}, ρ_{min}, e Å⁻³	0.472, -0.368	0.595, -0.370	0.268,-0.231	0.463, -0.217	0.480, -0.321	0.738, -0.445
Completeness to 2ϑ limit	0.986	0.984	0.970	0.988	0.929	0.987

Code	T8-Aze	T8-Dod	T11-Sub	T12-adi	T9-3HydroxyBA	T12-Suc	T12-Dod
Formula moiety	(C₁₁H₁₀N₂OS)₂ (C₉H₁₆O₄)	(C₁₁H₁₀N₂OS) 0.5(C₁₂H₂₂O₄)	C9H7N3OS, 0.5(C8H4O4	C₁₀H₉N₃OS, C₆H₁₀O₄	C₉H7NS, C₇H₆O₃	(C10 H 9 N3 O S) (C4 H6 O4)	C22 H31 N3 O5 S
Empirical formula	C₃₁H₃₆N₄O₆S₂	C₁₇H₂₁N₂O₃S	C₁₃H₁₄N₃O₃S	C₁₆H₁₉N₃O₅S	C₁₆H₁₃N₃O₄S	C14 H15 N3 O5 S	C22 H31 N3 O5 S
Molecular weight	624.76	333.42	292.33	365.40	343.35	337.35	449.56
Color, Habit	Colorless, Prism	Colorless, Prism	Colorless, rectangular	Yellow, prism	Yellow, plates	Bronze, prism	Colorless, needles
Crystal system	Monoclinic	Triclinic	Monoclinic	Monoclinic	Monoclinic	monoclinic	monoclinic
Space group, Z	Cc, 4	P-1, 2	P21/n	P21/c	P21/n	P 21/n, 4	P21/c,
a, Å	36.965(3)	8.8618(10)	14.1486(3)	12.429(5)	8.362(3)	4.9540(8)	19.249(11)
b, Å	5.0868(4)	10.0557(11)	5.35660(10)	19.565(7)	7.421(3)	15.979(3)	5.092(3)
c, Å	16.1184(12)	11.7950(13)	18.7308(4)	7.212(2)	24.785(8)	19.468(3)	23.614(14)
α, º	90.00	113.637(4)	90	90	90	90	90
β, º	94.167(3)	94.054(4)	109.0480(10)	104.506(16)	95.097(17)	93.103(5)	107.475(12)
γ, º	90.00	112.682(4)	90	90	90	90	90
Volume, Å³	3022.8(4)	856.28(17)	1341.85(5)	1697.9(10)	1531.9(9)	1538.9(4)	2208(2)
Density, g/cm³	1.508	1.293	1.447	1.429	1.489	1.456	1.353
T, ºK	120(2)	120(2)	200.(2)	130.(2)	130.(2)	120(2)	120(2)

Crystal size, min x mid x max	0.08X0.14X0.38	0.14X0.32X0.42	0.060 X 0.150 X 0.270	0.108 X0.254 X0.352	0.328 x 0.402 x 0.498	0.120x0.260x 0.420	0.080x0.12x0.480
X-ray wavelength, Å	0.71073	0.71073	0.71073	0.71073	0.71073	0.71073	0.71073
μ, mm^{-1}	0.227	0.205	0.71073	0.224	0.238	0.240	0.186
Trans min / max	0.904/ 0.969	0.92/0.97	0.58, 0.88	0.93, 0.98	0.89, 0.93	0.906, 0.972	0.916, 0.985
ϑ_{min}, º	2.210	2.34	3.43	2.68	2.51	2.45	
ϑ_{max}, º	32.030	31.01	70.01	25.70	25.82	31.61	
Reflections							
collected	17769	11911	9188	25849	32415	19679	9292
independent	8258	5137	2466	3214	2972	4879	5795
observed	7484	4198	2410	2668	2562	4091	4000
R_{int}	0.0294	0.0236	0.0349	0.0462	0.0404	0.0329	0.0364
Threshold expression	> 2σ(I)	> 2σ(I)	> 2σ(I)	> 2σ(I)	> 2σ(I)	> 2σ(I)	> 2σ(I)
No. parameters	402	217	190	239	229	218	290
No. restraints	0	0	0	0	2	0	0
R_1 (observed)	0.0417	0.0479	0.0355	0.035	0.0545	0.0375	0.0524
wR_2 (all)	0.0959	0.1448	0.0888	0.0847	0.1049	0.1057	0.1623
Goodness of fit (all)	1.008	1.056	1.098	1.044	1.190	1.079	0.972
ρ_{max}, ρ_{min}, e Å^{-3}	0.423, -0.229	0.840, -0.319	0.353, -0.344	0.333, -0.304	0.409, -0.492	0.382, -0.286	0.659, -0.529
Completeness to 2ϑ limit	0.976	0.939	0.965	0.989	0.889	0.901	

Co-crystal	D-H...A/A°	D-H/ A°	H...A// A°	D...A// A°	D-H...O/ °
T1-3HydroxyBA	N12–H12···O28	0.843(19)	2.130(19)	2.9589(18)	167.9(18)
	O23–H23···O28	0.93(2)	1.77(2)	2.6921(17)	169.7(17)
	O27–H27··· N13	0.86(2)	1.70(2)	2.5556(18)	171.9(19)
T2-4HydroxyBA	N12–H12···O28	0.88(2	1.89(2)	2.7618(16)	171.7(18)
	O27–H27···N13	0.82(2	1.88(2)	2.6972(17	170(2)
	O24–H24···O16	0.74(2)	1.93(2)	2.6710(16	175(2)
T4-3NitroBA	N7–H7···O19	0.864(19)	2.01(2)	2.865(2)	172.2(18)
	O20–H20···N3	0.93(3)	1.64(3)	2.577(2)	178.(3)
T8-Suc	N12–H12···O32	0.868(14)	1.960(14)	2.8153(11)	168.3(13)
	O31–H31···N13	0.862(16)	1.806(16)	2.6670(11)	176.0(15)
T8-Sub	N12–H12···O32	0.79(3)	2.17(3)	2.952(4)	169(2)
	O31–H31···N13	0.96(3)	1.66(3)	2.613(4)	175(2)
T8-Seb	N12–H12···O32	0.960(17)	1.892(17)	2.8435(16)	170.5(14)
	O31–H31···N13	0.90(2)	1.77(2)	2.6720(16)	175.3(17
T8-Aze	N12–H12···O52	0.85(3)	1.96(3)	2.802(3)	173(3)
	N32–H32···O60	0.86(3)	2.05(3)	2.891(3)	166(3)
	O51–H51···N13	0.79(4)	1.90(4)	2.693(3)	176(4)
	O59–H59···N33	0.80(4)	1.89(4)	2.652(3)	159(4)
T8-Dod	N12–H12···O32	0.870(14)	1.990(15)	2.8530(15)	171.0(14)
	O31–H31···N13	0.893(16)	1.747(17)	2.6377(15)	176.0(17)
T12-Suc	N12 H12 O32	0.851(15)	2.020(15)	2.8620(13)	169.8(14)
	O31 H31 N13	0.900(18)	1.761(18)	2.6589(14)	176.3(17)
	O34 H34 N21	0.918(17)	1.762(17)	2.6795(13)	178.4(15)
T12-Dod	N12 H12 O32	0.85(2)	2.10(2)	2.931(2)	168(2)

	O31 H31 N13	0.89(3)	1.75(3)	2.633(2)	173(2)
	O42 H42 N21	0.86(3)	1.87(3)	2.697(2)	163(2)

D.4 Chapter 7 (ternary)

Code	T10-D1	T12-D1	T11-H10-D1	T11-D1
Formula moiety	C10H9N3OS, 0.5(C6F4I2)	C10 H9 N3 O S, 0.5(C6 F4 I2)'	C27 H27 F3 I1.50 N4.50O4.50 S1.50	C9 H7 N3 O S, 0.5(C6 F4 I2
Empirical formula	C13H9F2IN3OS	C13 H9 F2 I N3 O S'	C27 H27 F3 I1.50 N4.50O4.50 S1.50	C12 H7 F2 I N3 O S
Molecular weight	420.19	420.19	781.97	406.17
Color, Habit	Colorless, rectangular	Colorless, Parallelepiped	Colorless, blocks	Colorless, plates
Crystal system	Triclinic	Triclinic	triclinic	monoclinic
Space group, Z	P-1, 2	P-1, 2	P -1, 4	P21/n, 4
a, Å	4.8363(5)	6.4078(7)	14.350(3)	9.334(3)
b, Å	9.6588(10)	8.7487(9)	14.558(4)	7.392(3)
c, Å	15.5868(16)	13.1516(14)	15.377(3)	19.794(8)
α, º	92.7220(10)	93.4920(10)	71.972(5)	90
β, º	96.3790(10)	92.6420(10)	74.463(6)	102.19(2)
γ, º	91.0120(10)	103.0880(10)	80.417(6)	90
Volume, Å³	722.59(13)	715.43(13)	2925.9(10)	1334.9(8)
Density, g/cm³	1.931	1.951	1.775	2.021
T, ºK	228 (2)	228.(2)	130.(2)	130.0(2)
Crystal size, min x mid x max	0.128X0.208X0.278	0.120x0.220x0.450	0.07x0.228x0.272	0.128x0.298x0.332
X-ray wavelength, Å	0.71073	0.71073	0.71073	0.71073
μ, mm^{-1}	0.100	2.407	1.792	2.577
Trans min / max	0.97/0.99	0.41, 0.76	0.64, 0.89	0.48, 0.73
ϑ_{min}, º	2.25	2.40		2.26
ϑ_{max}, º	21.98	32.93		25.60
Reflections				
collected	30410	9970	57433	2535
independent	2263	5074	11979	2535
observed	1384	4333	3846	2217
R_{int}	0.1016	0.0287	0.0527	0.00
Threshold expression	> 2σ(I)	> 2σ(I)	> 2σ(I)	> 2σ(I)
No. parameters	181	195	781	185
No. restraints	0	1	6	1
R_1 (observed)	0.0650	0.0641	0.0938	0.2210
wR_2 (all)	0.1730	0.0689	0.1604	0.2369
Goodness of fit (all)	1.051	1.028	0.803	1.198
ρ_{max}, ρ_{min}, e Å^{-3}	0.215,-0.241	0.765, -0.987	0.690, -0.886	3.143, -2.549
Completeness to 2ϑ limit	0.972	0.913	0.985	0.975

Co-crystal	D-H...A/A°	D-H/ A°	H...A// A°	D...A// A°	D-H...O/ °
37-D7	N7 H7 N5	0.82(2)	2.12(2)	2.942(3)	179.(3)
38-D4	N7 H7 N5	0.82(2)	2.12(2)	2.936(2)	169.(3)
BS1603	N6 H6 O43	0.81(4)	2.10(4)	2.891(7)	165.(7)
	N20 H20 O60	0.80(3)	2.14(4)	2.892(8)	156.(6)
	N34 H34 O57	0.90(4)	1.98(4)	2.866(8)	168.(7)
	O44 H44 N5	0.84(4)	1.84(4)	2.673(8)	175.(6)
	O58 H58 N33	0.85(5)	1.80(5)	2.646(8)	171.(10)
	O59 H59 N19	0.84(5)	2.16(14)	2.662(7)	118.(14)
BS1615	N6 H6 N3	0.88(5)	2.13(5)	2.99(10)	165.(8)
BS1713	N7 H7 N3	0.82(3)	2.16(3)	2.967(3)	168.(3)

D.5 Chapter 8 (Activated halogens)

Code	3N-ITFB	3N-EB:Fum	3N-EB	4N-EB
Formula moiety	$C_{12}H_5F_4IN_2O$	C16 H12 N2 O3.50	C14 H10 N2 O	C14 H10 N2 O
Empirical formula	C12 H5 F4 I N2 O	C16 H12 N2 O3.50	C14 H10 N2 O'	C14 H10 N2 O
Molecular weight	396.08	288.28	222.24	222.24
Color, Habit	Colorless, blocks		Yellow, needles	Yellow, blocks
Crystal system	orthorhombic	monoclinic	monoclinic	monoclinic
Space group, Z	P 21 21 21, 4	P 1 21/c, 4	C 1 2/c, 8	P 1 21/n, 4
a, Å	4.8009(17)	8.444(3)	25.764(5)	15.117(4)
b, Å	9.695(3)	9.231(3)	5.3338(9)	5.2846(15)
c, Å	26.315(9)	17.275(7)	16.701(3)	16.102(4)
α, º	90	90	90	90
β, º	90	97.74(2)	103.011(14)	114.406(9)
γ, º	90	90	90	90
Volume, Å³	1224.8(7)	1334.3(8)	2236.1(8)	1171.4(6)
Density, g/cm³	2.148	1.435	1.320	1.260
T, ºK	293(2)	130.(2)	296.(2)	296.(2)
Crystal size, min x mid x max		0.114x0.228x0.402	0.135x0.262x0.324	0.126x0.194x0.261
X-ray wavelength, Å	0.71073	0.71073	0.71073	0.71073
μ, mm^{-1}	2.661	0.103	0.085	0.082
Trans min / max		0.96, 0.99	0.97, 0.99	0.98, 0.99
ϑ_{min}, º	2.239	2.38	2.5033	2.41
ϑ_{max}, º	25.996	26.13	24.7894	25.04
Reflections				
collected	21062	32167	14333	21569
independent	2380	2641	2010	2067
observed	2225	1995	1488	1145
R_{int}		0.0609	0.0508	0.0972
Threshold expression	$> 2\sigma(I)$	$> 2\sigma(I)$	$> 2\sigma(I)$	$> 2\sigma(I)$
No. parameters	182	204	158	158
No. restraints		1	0	0
R_1 (observed)		0.0546	0.0450	0.0605
wR_2 (all)		0.1704	0.1298	0.1686
Goodness of fit (all)		1.045	1.042	1.020
ρ_{max}, ρ_{min}, e Å^{-3}		1.367, -0.241	2.5033, 24.7894	0.144, -0.186
Completeness to 2ϑ limit		0.989	0.982	1.020

Code	3N-IB	4N-ITFB	3N-IEB	3N-IB: Sub
Formula moiety	C12 H9 I N2 O	$C_{12}H_5F_4IN_2O$	C14H9IN2O	C10 H9 N3 O S, C7 H6 O3
Empirical formula	C12 H9 I N2 O	C12H5F4IN2O	C14H9IN2O	C17 H15 N3 O4 S'
Molecular weight	324.11	396.08	348.13	357.38
Color, Habit	Colorless, blocks	Colorless, blocks		Colorless, plate
Crystal system	monoclinic	triclinic	monoclinic	monoclinic
Space group, Z	C 1 2/c, 8	'P -1, 6	Pc, 2	P 1 21/n, 4
a, Å	19.262(4)	10.377(4)	11.5898(6)	20.9506(4)
b, Å	5.3201(10)	13.547(5)	6.1640(3)	3.78608(5)
c, Å	22.246(4)	14.302(6)	9.1027(6)	22.5283(4)
α, º	90	93.78(2)	90	90
β, º	95.545(10)	92.66(2)	99.225(6)	112.879(2)
γ, º	90	103.28(2)	90	90
Volume, Å³	2269.0(8)	1948.7(13)	641.88(6)	1646.36(5)
Density, g/cm³	1.898	2.025	1.801	1.442

T, ºK	296.(2)	130(2)	296(2)	100.00(10)
Crystal size, min x mid x max	0.077x0.232x0.303	0.056x0.228x0.494		0.06x0.176x0.182
X-ray wavelength, Å	0.71073	0.71073	0.71073	0.71073
μ, mm⁻¹	2.801	2.509	2.483	2.003
Trans min / max	0.48, 0.81	0.370, 0.872		1.000, 0.570
ϑ_{min}, º	2.67	1.430	4.237	4.2670
ϑ_{max}, º	24.24	25.500	32.635	73.4990
Reflections				
collected	18437	40986	7187	7722
independent	2185	6952	3841	2846
observed	1602	1904	3367	2994
R_{int}	0.0468	0.1431	0.0249	0.0198
Threshold expression	> 2σ(I)	> 2σ(I)	> 2σ(I)	> 2σ(I)
No. parameters	149	231	164	230
No. restraints	1	0	2	0
R_1 (observed)	0.0337	0.1201	0.0301	0.0833
wR_2 (all)	0.0693	0.3896	0.0649	0.0843
Goodness of fit (all)	1.053	0.964	1.034	1.090
ρ_{max}, ρ_{min}, e Å⁻³	0.425, -0.709	3.142, -3.405	0.681, -0.428	0.250, -0.245
Completeness to 2ϑ limit	0.995	0.959	0.918	0.997

Co-crystal	D-H...A/A°	D-H/ A°	H...A// A°	D...A// A°	D-H...O/ °
3N-EB:Fum	O19 H19 N4	0.94(3)	1.68(3)	2.599(2)	163.(3)
	N9 H9 O18	0.91(3)	2.43(3)	3.326(3)	170.(2)
3N-IB: Sub	N9 H9 O11	0.85(2)	2.33(2)	3.135(2)	157.6(16)
4N-EB	N9 H9 O8	0.88(3)	2.26(3)	3.096(3)	157.(2)
	C17 H17 N4	0.93	2.44	3.352(5)	166.7
3N-IB	N9 H9 O8	0.80(3)	2.37(3)	3.145(4)	162.(4)
3N-IEB	N1 H1N O1	0.84	2.08	2.921(5)	175.6
	C7 H7 O1	0.93	2.42	2.900(6)	112.3
	C13 H12 I1	0.93	3.17	3.889(5)	135.9
	C14 H14 O1	0.93	2.50	3.255(6)	138.9

Appendix E- Hydrogen-bond energies of co-crystal synthons

E.1 Molecular electrostatic potentials for 20 acids

Table E.1.1 Electrostatics (in kJ/mol) and alpha ranked as (I) and (II) and beta ranked as (I) and (II) values for **aliphatic acids** calculated using equations 1 and 2 for donor and acceptor group. Each aliphatic acid has two acceptors; C=O ranked as (I) and (II) and two donors; O-H ranked as (I) and (II) based on molecular electrostatic potentials.

Hydrogen bond donors	C=O (1)	OH (1)	C=O(2)	OH(2)	Alpha (1)	beta (1)	Alpha(2)	beta(2)	Eacid1	Eacid2
Succinic	-177.00	306.00	-131.00	283.00	4.46	6.22	4.02	3.72	27.75	14.97
Adipic	-157.00	268.00	-157.00	263.00	3.74	5.06	3.65	5.06	18.93	18.47
Suberic	-164.00	264.00	-164.00	262.00	3.67	5.45	3.63	5.45	20.00	19.80
Sebacic	-186.00	294.00	-161.00	261.00	4.23	6.78	3.61	5.28	28.67	19.09

Fumaric	-162.00	282.00	-159.00	270.00	4.00	5.34	3.78	5.17	21.36	19.53
Malonic	-154.00	286.00	-154.00	284.00	4.08	4.89	4.04	4.89	19.95	19.77
Glutaric	-162.00	272.00	-153.00	265.00	3.82	5.34	3.69	4.84	20.37	17.85
Pimelic	-167.00	261.00	-164.00	260.00	3.61	5.62	3.60	5.45	20.33	19.61
Azelaic	-169.00	262.00	-160.00	258.00	3.63	5.74	3.56	5.23	20.86	18.60
Dodeca	-197.00	296.00	-164.00	260.00	4.27	7.50	3.60	5.45	31.99	19.61

Table E.1.2 Electrostatics (in kJ/mol) and alpha ranked as (I) and (II) and beta ranked as (I) and (II) values for **aromatic acids** calculated using equations 1 and 2 for donor and acceptor group. Each aliphatic acid has two acceptors; C=O ranked as (I) and O(OH)/N(NH$_2$) ranked as (II) and two donors; O-H ranked as (I) and OH/NH$_2$ ranked as (II) based on molecular electrostatic potentials.

Hydrogen bond donors	C=O (I)	β (I)	OH (I)	α (I)	O(OH)	β (II)	OH/NH$_2$ (II)	α (II)
3HydroxyBA	-154	4.89	267	3.72	-116	3.04	276	3.89
4HydroxyBA	-169	5.74	255	3.51	-106	2.62	297	4.28
3AminoBA	-168	5.68	251	3.44	-117	3.08	203	2.62
4AminoBA	-184	6.65	236	3.17	-98	2.31	222	2.93
3NitroBA	-138	4.06	299	4.32				
4NitroBA	-126	3.49	304	4.42				
BA	-162	5.34	245	3.33				
4BromoBA	-152	4.79	275	3.87				
4IodoBA	-152	4.79	274	3.85				
PentafluroBA	-136	3.97	305	4.44				

E.2 Chapter-3: Pyrazole heteromeric interactions

Table E.2.1 Hydrogen bond energies (kJ/mol) of **synthon I** in **P1-P12** aliphatic acid co-crystals

synthon I										
	SUC	ADI	SUB	SEB	FUM	MAL	GLU	PIM	AZE	DOD
P1	34.02	36.79	37.21	36.74	38.57	36.61	36.98	37.50	37.63	35.84
P2	34.86	37.40	37.73	37.21	39.27	37.50	37.65	38.05	38.10	36.35
P3	34.14	36.85	37.25	36.77	38.65	36.74	37.06	37.55	37.66	35.88
P4	35.52	38.02	38.33	37.78	39.94	38.21	38.29	38.66	38.68	36.93
P5	34.08	36.74	37.12	36.63	38.54	36.67	36.95	37.42	37.52	35.75
P6	35.15	37.61	37.91	37.37	39.50	37.81	37.87	38.24	38.25	36.52
P7	33.82	36.59	37.02	36.55	38.36	36.40	36.78	37.30	37.44	35.65
P8	34.86	37.40	37.73	37.21	39.27	37.50	37.65	38.05	38.10	36.35

P9	31.42	34.13	34.58	34.16	35.75	33.81	34.28	34.83	35.01	33.29
P10	30.67	33.40	33.86	33.46	34.97	33.01	33.53	34.10	34.30	32.60
P11	32.85	35.35	35.70	35.22	37.09	35.34	35.56	35.99	36.07	34.39
P12	31.63	34.23	34.63	34.19	35.88	34.04	34.40	34.90	35.03	33.35

Table E.2.2 Hydrogen bond energies (kJ/mol) of **synthon I** in **P1-P12** aromatic acid co-crystals

	synthon I											
	3-hydroxy	4-hydroxy	3-amino	4-amino	3-nitro	4-nitro	ba	4-bromo	4iodo	pentaF	arom-w/o	arom-with
P1	31.32	29.34	29.13	26.79	32.23	31.68	29.09	31.93	28.79	34.14	31.31	29.15
P2	32.26	30.21	29.96	27.53	33.65	33.22	29.82	32.96	30.00	35.54	32.53	29.99
P3	31.47	29.48	29.27	26.91	32.50	31.98	29.20	32.10	29.01	34.40	31.53	29.28
P4	32.93	30.84	30.57	28.09	34.50	34.11	30.40	33.67	30.74	36.41	33.31	30.61
P5	31.45	29.46	29.23	26.88	32.54	32.05	29.15	32.09	29.05	34.43	31.55	29.25
P6	32.59	30.52	30.25	27.80	34.16	33.77	30.08	33.32	30.44	36.05	32.97	30.29
P7	31.12	29.16	28.96	26.63	32.00	31.44	28.91	31.72	28.59	33.90	31.09	28.97
P8	32.26	29.96	29.96	27.53	33.65	33.22	29.82	32.96	30.00	35.54	32.53	29.93
		29.68±0.60					32.10±0.83					
P9	28.83	26.84	26.84	24.69	29.42	28.84	26.85	29.35	26.31	31.22	28.66	26.80
P10	28.10	26.18	26.18	24.08	28.55	27.95	26.20	28.58	25.55	30.33	27.86	26.13
P11	30.34	28.19	28.19	25.92	31.48	31.02	28.09	30.97	28.09	33.29	30.49	28.16
P12	29.10	27.08	27.08	24.90	29.91	29.39	27.04	29.66	26.72	31.69	29.07	27.04
		27.03±0.84					29.02±1.10					

Table E.2.3 Hydrogen bond energies (kJ/mol) of **synthon II** in **P1-P12** aliphatic acid co-crystals

	synthon II									
	SUC	ADI	SUB	SEB	FUM	MAL	GLU	PIM	AZE	DOD
P1	30.05	32.07	32.30	31.82	33.70	32.32	32.31	32.58	32.57	31.12
P2	31.49	33.40	33.56	33.04	35.14	33.86	33.68	33.88	33.80	32.35
P3	30.12	32.08	32.28	31.79	33.72	32.40	32.33	32.57	32.53	31.10
P4	31.93	33.75	33.88	33.33	35.53	34.33	34.06	34.21	34.09	32.66
p5	30.55	32.55	32.75	32.26	34.21	32.86	32.80	33.05	33.01	31.56
P6	32.03	33.90	34.04	33.50	35.68	34.44	34.20	34.37	34.27	32.81
P7	28.71	30.52	30.69	30.22	32.10	30.88	30.77	30.98	30.93	29.58
P8	30.21	31.88	31.98	31.46	33.57	32.48	32.18	32.30	32.17	30.83
P9	27.83	29.87	30.13	29.72	31.35	29.94	30.06	30.39	30.43	29.03
P10	27.08	29.13	29.42	29.02	30.57	29.14	29.31	29.66	29.72	28.34
P11	28.64	30.35	30.49	30.01	31.94	30.79	30.61	30.78	30.70	29.39
P12	28.18	30.13	30.36	29.92	31.65	30.31	30.34	30.62	30.63	29.25

Table E.2.4 Hydrogen bond energies (kJ/mol) of **synthon II** in **P1-P12** aromatic acid co-crystals

	Synthon II									
	3-hydroxy	4-hydroxy	3-amino	4-amino	3-nitro	4-nitro	ba	4-bromo	4iodo	pentaF
P1	27.91	26.14	25.89	23.79	29.39	29.10	25.72	28.56	26.17	30.98
P2	29.37	27.50	27.21	24.99	31.23	31.03	26.97	30.10	27.78	32.86
P3	28.03	26.24	25.98	23.87	29.62	29.36	25.79	28.70	26.36	31.20
P4	29.85	27.95	27.64	25.38	31.93	31.77	27.36	30.63	28.37	33.55
P5	28.42	26.61	26.35	24.21	30.01	29.75	26.16	29.10	26.72	31.62
P6	29.91	28.01	27.70	25.44	31.92	31.74	27.44	30.68	28.37	33.56
P7	26.74	25.04	24.79	22.77	28.34	28.12	24.59	27.39	25.21	29.84
P8	28.27	26.47	26.17	24.02	30.32	30.19	25.89	29.02	26.93	31.84
	26.52±1.00				**24.37±0.96**					
P9	25.75	24.12	23.91	21.98	26.85	26.50	23.81	26.31	23.94	28.36
P10	25.02	23.44	23.25	21.37	25.98	25.62	23.16	25.54	23.18	27.47
P11	26.73	25.02	24.76	25.31	28.46	28.29	24.53	27.40	25.31	29.94
P12	26.14	24.48	24.26	24.44	27.44	27.14	24.12	26.74	24.44	28.95
	29.15±1.18				**26.23±0.91**					

Table E.2.5 Hydrogen bond energies (kJ/mol) of **synthon III** in **P1-P12** aliphatic acid co-crystals

	Synthon III									
	SUC	ADI	SUB	SEB	FUM	MAL	GLU	PIM	AZE	DOD
P1	29.41	28.93	28.30	27.51	30.88	31.54	29.57	28.77	27.98	27.35
P2	30.44	29.29	29.29	28.47	31.96	32.64	30.61	29.78	28.96	28.31
P3	28.65	26.80	27.56	26.80	30.08	30.72	28.81	28.03	27.25	26.64
P4	29.67	31.14	28.54	27.75	31.14	31.81	29.83	29.02	28.22	27.59
p5	29.18	30.63	28.07	27.29	30.63	31.29	29.34	28.54	27.76	27.14
P6	29.97	31.46	28.83	28.03	31.46	32.13	30.13	29.31	28.51	27.87
P7	27.65	29.02	26.60	25.86	29.02	29.64	27.80	27.05	26.30	25.71
P8	28.65	30.08	27.56	26.80	30.08	30.72	28.81	28.03	27.25	26.64
P9	24.75	25.98	23.81	23.15	25.98	26.53	24.88	24.21	23.54	23.01
P10	22.00	23.10	21.17	20.58	23.10	23.59	22.12	21.52	20.93	20.46
P11	23.35	24.52	22.47	21.84	24.52	25.04	23.48	22.85	22.22	21.72
P12	22.90	24.04	22.03	21.42	24.04	24.55	23.02	22.40	21.78	21.30

Table E.2.6 Hydrogen bond energies (kJ/mol) of **synthon III** in **P1-P12** aromatic acid co-crystals

	synthon III											
	3-hydroxy	4-hydroxy	3-amino	4-amino	3-nitro	4-nitro	ba	4-bromo	4iodo	pentaF	arom-w/o	arom-with
P1	28.77	26.89	26.27	24.02	34.07	34.93	25.36	30.06	29.90	35.11	31.57	26.49
P2	29.78	27.83	27.19	24.86	35.26	36.16	26.25	31.11	30.95	36.34	32.68	27.42
P3	28.03	26.19	25.59	23.39	33.19	34.03	24.70	29.28	29.12	34.20	30.75	25.80
P4	29.02	27.12	26.50	24.23	34.37	35.24	25.58	30.32	30.16	35.41	31.85	26.72
p5	28.54	26.67	26.06	23.83	33.80	34.66	25.16	29.82	29.66	34.83	31.32	26.28
P6	29.31	27.39	26.77	24.47	34.71	35.59	25.84	30.62	30.46	35.77	32.16	26.98
P7	27.05	25.27	24.70	22.58	32.02	32.84	23.84	28.26	28.10	33.00	29.68	24.90
P8	28.03	26.19	25.59	23.39	33.19	34.03	24.70	29.28	29.12	34.20	30.75	25.80

	26.61±0.67				**31.34±0.94**						
P9	24.21	22.62	22.10	20.21	28.66	29.39	21.34	25.29	25.15	29.54	26.56	22.29
P10	21.52	20.11	19.65	17.97	25.48	26.13	18.97	22.49	22.36	26.26	23.62	19.81
P11	22.85	21.35	20.86	19.07	27.05	27.74	20.14	23.87	23.74	27.88	25.07	21.03
P12	22.40	20.93	20.45	18.70	26.52	27.20	19.74	23.40	23.28	27.33	24.58	20.62
	20.94±1.03				**24.96±1.23**						

Table E.2.7 Hydrogen bond energies (kJ/mol) of **synthon IV** in **P1-P12** aliphatic acid co-crystals

	Synthon IV									
	SUC	**ADI**	**SUB**	**SEB**	**FUM**	**MAL**	**GLU**	**PIM**	**AZE**	**DOD**
P9	27.49	27.04	26.45	25.71	28.86	29.48	27.64	26.89	26.15	25.57
P10	25.47	25.05	24.50	23.82	26.74	27.31	25.61	24.92	24.23	23.69
P11	29.59	29.11	28.47	27.68	31.06	31.73	29.75	28.95	28.15	27.52
P12	26.47	26.04	25.47	24.76	27.79	28.38	26.62	25.90	25.18	24.62

Table E.2.8 Hydrogen bond energies (kJ/mol) of **synthon IV** in **P9-P12** aromatic acid co-crystals

	synthon IV											
	3-hydroxy	**4-hydroxy**	**3-amino**	**4-amino**	**3-nitro**	**4-nitro**	**ba**	**4-bromo**	**4iodo**	**pentaF**	**arom-w/o**	**arom-with**
P9	26.89	25.13	24.56	22.45	31.85	32.65	23.71	28.10	27.95	32.82	29.51	24.76
P10	24.92	23.28	22.75	20.80	29.50	30.25	21.96	26.03	25.89	30.40	27.34	22.94
P11	28.95	27.05	26.43	24.16	34.28	35.15	25.51	30.24	30.08	35.32	31.76	26.65
P12	25.90	24.20	23.65	21.62	30.66	31.44	22.82	27.05	26.91	31.60	28.42	23.84
	24.55±1.59			**29.26±1.89**								

Table E.2.9 Hydrogen bond energies (kJ/mol) of **synthon V/VI/VIII/IX** in **P9-P12** aromatic acid co-crystals

	Synthon V/VIII						Synthon VI/IX			
	3-hydroxy	**4-hydroxy**	**3-amino**	**4-amino**			**3-hydroxy**	**4-hydroxy**	**3-amino**	**4-amino**
P1	18.26	20.38	11.70	13.29	15.91	**P1**	30.22	33.73	19.36	21.99
P2	20.16	22.50	12.92	14.67	17.56	**P2**	31.28	34.91	20.04	22.76
P3	18.68	20.84	11.97	13.59	16.27	**P3**	29.44	32.85	18.86	21.42
P4	21.04	23.47	13.48	15.31	18.32	**P4**	30.48	34.02	19.53	22.18
p5	18.89	21.08	12.10	13.74	16.45	**p5**	29.98	33.46	19.21	21.82
P6	20.86	23.28	13.37	15.18	18.17	**P6**	30.79	34.36	19.73	22.41
P7	18.06	20.15	11.57	13.14	15.73	**P7**	28.41	31.70	18.20	20.67
P8	20.16	22.50	12.92	14.67	17.56	**P8**	29.44	32.85	18.86	21.42
	17.00±1.03						**26.14±0.79**			
P9	16.06	17.93	10.29	11.69	13.99	**P9**	25.43	28.38	16.29	18.50
P10	15.30	17.07	9.80	11.13	13.32	**P10**	22.61	25.23	14.49	16.45

341

P11	18.47	20.61	11.83	13.44	16.09	**P11**	24.00	26.78	15.38	17.46
P12	16.85	18.80	10.79	12.26	14.68	**P12**	23.53	26.26	15.08	17.12
	14.52±1.18						20.81±1.02			

E.3 Chapter 5: Thiazoles heteromeric interactions

Table E.3.1 Hydrogen bond energies (kJ/mol) of **synthon F** in **T1-T12** aliphatic acid co-crystals

Ligands--- Acid	Synthon F									
	suc	adi	sub	seb	fum	mal	glut	pim	aze	dod
T1	-40.76	-33.66	-34.65	-41.60	-35.75	-34.60	-34.92	-34.98	-35.45	-44.17
T2	-41.43	-34.25	-35.15	-42.10	-36.40	-35.35	-35.50	-35.44	-35.90	-44.57
T3	-40.70	-33.63	-34.56	-41.43	-35.73	-34.65	-34.87	-34.86	-35.32	-43.92
T4	-41.52	-34.34	-35.19	-42.09	-36.50	-35.51	-35.57	-35.45	-35.90	-44.50
T5	-41.43	-34.25	-35.15	-42.10	-36.39	-35.34	-35.49	-35.44	-35.90	-44.57
T6	-40.68	-33.61	-34.54	-41.41	-35.71	-34.63	-34.85	-34.85	-35.31	-43.90
T7	-39.31	-32.50	-33.34	-39.92	-34.54	-33.55	-33.68	-33.61	-34.04	-42.25
T8	-40.47	-33.48	-34.26	-40.94	-35.59	-34.68	-34.66	-34.50	-34.93	-43.23
T9	-37.73	-31.12	-32.14	-38.67	-33.05	-31.88	-32.33	-32.49	-32.95	-41.18
T10	-39.06	-32.24	-33.22	-39.91	-34.25	-33.12	-33.46	-33.55	-34.01	-42.41
T11	-38.51	-31.77	-32.81	-39.50	-33.73	-32.52	-33.00	-33.18	-33.65	-42.07
T12	-40.21	-33.20	-34.19	-41.06	-35.26	-34.11	-34.45	-34.53	-34.99	-43.62
				average	-36.47	1.08				

Table E.3.2 Hydrogen bond energies (kJ/mol) of **synthon H** in **T1-T12** aliphatic acid co-crystals

Ligands---Acid	synthon H										
	suc	adi	sub	seb	fum	mal	glut	pim	aze	dod	
T1	-28.02	-23.50	-23.04	-26.56	-25.13	-25.61	-23.96	-22.70	-22.82	-26.80	-24.81
T2	-29.68	-24.90	-24.90	-28.14	-26.62	-27.13	-25.39	-24.05	-24.17	-28.39	-26.34
T3	-26.40	-22.15	-21.71	-25.03	-23.68	-24.13	-22.58	-21.39	-21.50	-25.26	-23.38
T4	-28.02	-23.50	-23.04	-26.56	-25.13	-25.61	-23.96	-22.70	-22.82	-26.80	-24.81
T5	-29.66	-24.88	-24.40	-28.12	-26.61	-27.11	-25.37	-24.04	-24.16	-28.37	-26.27
T6	-26.41	-22.15	-21.72	-25.03	-23.69	-24.13	-22.59	-21.40	-21.50	-25.26	-23.39
T7	-26.40	-22.15	-21.71	-25.03	-23.68	-24.13	-22.58	-21.39	-21.50	-25.26	-23.38
T8	-27.47	-23.05	-22.60	-26.04	-24.64	-25.11	-23.50	-22.26	-22.37	-26.28	-24.33
T9	-20.65	-17.32	-16.99	-19.58	-18.53	-18.87	-17.66	-16.73	-16.82	-19.76	-18.29
T10	-22.57	-18.93	-18.56	-21.40	-20.25	-20.63	-19.30	-18.29	-18.38	-21.59	-19.99
T11	-23.06	-19.34	-18.97	-21.86	-20.69	-21.08	-19.72	-18.69	-18.78	-22.06	-20.42
T12	-24.04	-20.17	-19.77	-22.79	-21.57	-21.97	-20.56	-19.48	-19.58	-23.00	-21.29
				AVERAGE	-23.06	2.547277					

Table E.3.3 Hydrogen bond energies (kJ/mol) of **synthon I** in **T9-T12** aliphatic acid co-crystals

Ligands--- Acid	Synthon I									
	suc	adi	sub	seb	fum	mal	glut	pim	aze	dod
T9	-32.53	-26.46	-28.51	-35.46	-27.92	-25.60	-27.92	-29.42	-30.03	-39.22	-30.31

T10	-35.33	-28.73	-30.97	-38.52	-30.32	-27.80	-30.32	-31.95	-32.61	-42.59	-32.91
T11	-35.70	-29.04	-31.29	-38.92	-30.64	-28.10	-30.64	-32.29	-32.96	-43.04	-33.26
T12	-35.70	-29.04	-31.29	-38.92	-30.64	-28.10	-30.64	-32.29	-32.96	-43.04	-33.26
										average	-32.44

Table E.3.4 Hydrogen bond energies (kJ/mol) of **synthon F** in **T1-T12** aromatic acid co-crystals

Ligands---Acid	Synthon F									
	3HydroxyBA	4HydroxyBA	3AminoBA	4AminoBA	3NitroBA	4NitroBA	BA	4BromoBA	4IodoBA	PentafluoroBA
T1	-32.99	-34.90	-34.39	-36.43	-32.90	-31.44	-32.76	-33.33	-33.24	-33.13
T2	-33.61	-35.30	-34.77	-36.56	-33.86	-32.53	-33.15	-34.01	-33.92	-34.16
T3	-32.98	-34.75	-34.23	-36.11	-33.08	-31.71	-32.63	-33.35	-33.26	-33.35
T4	-33.71	-35.28	-34.75	-36.39	-34.15	-32.89	-33.15	-34.15	-34.05	-34.47
T5	-33.60	-35.30	-34.77	-36.56	-33.86	-32.53	-33.15	-34.01	-33.91	-34.15
T6	-32.97	-34.73	-34.22	-36.09	-33.07	-31.70	-32.61	-33.34	-33.25	-33.34
T7	-31.89	-33.47	-32.97	-34.63	-32.18	-30.93	-31.44	-32.28	-32.19	-32.46
T8	-32.88	-34.31	-33.78	-35.26	-33.46	-32.30	-32.24	-33.34	-33.24	-33.80
T9	-30.48	-32.47	-32.00	-34.15	-30.08	-28.59	-30.46	-30.74	-30.66	-30.24
T10	-31.60	-33.49	-33.00	-35.02	-31.42	-29.99	-31.43	-31.91	-31.82	-31.63
T11	-31.11	-33.16	-32.68	-34.90	-30.66	-29.12	-31.10	-31.36	-31.28	-30.82
T12	-32.54	-34.45	-33.95	-36.00	-32.40	-30.94	-32.33	-32.86	-32.78	-32.62
				average	-33.09	1.07				

Table E.3.5 Hydrogen bond energies (kJ/mol) of **synthon H** in **T1-T12** aromatic acid co-crystals

Ligands---Acid	synthon H									
	3HydroxyBA	4HydroxyBA	3AminoBA	4AminoBA	3NitroBA	4NitroBA	BA	4BromoBA	4IodoBA	PentafluoroBA
T1	-23.37	-22.05	-21.61	-19.91	-19.91	-27.14	-27.76	-20.92	-24.18	-27.89
T2	-24.75	-23.36	-22.89	-21.09	-21.09	-28.75	-29.41	-22.16	-25.62	-29.54
T3	-22.02	-20.78	-20.36	-18.76	-18.76	-25.57	-26.16	-19.71	-22.79	-26.28
T4	-23.37	-22.05	-21.61	-19.91	-19.91	-27.14	-27.76	-20.92	-24.18	-27.89
T5	-24.74	-23.34	-22.88	-21.08	-21.08	-28.73	-29.39	-22.14	-25.60	-29.53
T6	-22.02	-20.78	-20.36	-18.77	-18.77	-25.57	-26.17	-19.71	-22.79	-26.28
T7	-22.02	-20.78	-20.36	-18.76	-18.76	-25.57	-26.16	-19.71	-22.79	-26.28
T8	-22.91	-21.62	-21.19	-19.52	-19.52	-26.61	-27.22	-20.51	-23.71	-27.35
T9	-17.22	-16.25	-15.93	-14.68	-14.68	-20.00	-20.46	-15.42	-17.83	-20.56

T10	-18.82	-17.76	-17.41	-16.04	-16.04	-21.86	-22.37	-16.85	-19.48	-22.47
T11	-19.23	-18.15	-17.78	-16.39	-16.39	-22.33	-22.85	-17.22	-19.90	-22.95
T12	-20.05	-18.92	-18.54	-17.09	-17.09	-23.28	-23.82	-17.95	-20.75	-23.93
				average	-21.81	2.40				

Table E.3.6 Hydrogen bond energies (kJ/mol) of **synthon I** in **T9-T12** aromatic acid co-crystals

Ligands---Acid	Synthon I									
	3HydroxyBA	4HydroxyBA	3AminoBA	4AminoBA	3NitroBA	4NitroBA	BA	4BromoBA	4IodoBA	PentafluoroBA
T9	-19.46	-18.36	-17.99	-16.58	-22.59	-23.12	-17.42	-20.24	-20.14	-23.22
T10	-21.13	-19.94	-19.54	-18.01	-24.54	-25.11	-18.91	-21.98	-21.87	-25.22
T11	-21.35	-20.15	-19.75	-18.20	-24.80	-25.37	-19.11	-22.21	-22.10	-25.49
T12	-21.35	-20.15	-19.75	-18.20	-24.80	-25.37	-19.11	-22.21	-22.10	-25.49
				average	-21.31	0.94				

Table E.3.7 Hydrogen bond energies (kJ/mol) of **synthon J, K and L** in **T1-T12** aromatic acid co-crystals

	Synthon J		Synthon K		Synthon L	
E	3HydroxyBA	4HydroxyBA	3HydroxyBA	4HydroxyBA	3HydroxyBA	4HydroxyBA
T1	24.43	26.88	17.39	19.13	N/A	N/A
T2	25.88	28.48	18.83	20.72	N/A	N/A
T3	23.02	25.33	18.00	19.80	N/A	N/A
T4	24.43	26.88	19.46	21.42	N/A	N/A
T5	25.87	28.46	18.83	20.72	N/A	N/A
T6	23.03	25.34	18.01	19.82	N/A	N/A
T7	23.02	25.33	18.00	19.80	N/A	N/A
T8	23.96	26.36	19.46	21.42	Synthon L	
T9	18.01	19.82	15.05	16.56	20.34	22.38
T10	19.68	21.66	16.38	18.02	22.10	24.31
T11	20.11	22.13	15.25	16.78	22.33	24.57
T12	20.97	23.07	17.00	18.70	22.33	24.57

Table E.3.8 Hydrogen bond energies (kJ/mol) of **synthon M, N and L** in **T1-T12** aromatic acid co-crystals

	Synthon M		Synthon N		Synthon O	
E	4AminoBA	3AminoBA	4AminoBA	3AminoBA	3HydroxyBA	4HydroxyBA
T1	16.46	18.40	11.71	13.10	N/A	N/A
T2	17.43	19.50	12.68	14.18	N/A	N/A
T3	15.51	17.34	12.12	13.56	N/A	N/A

T4	16.46	18.40	13.11	14.66	N/A	N/A
T5	17.42	19.48	12.68	14.18	N/A	N/A
T6	15.51	17.35	12.13	13.57	N/A	N/A
T7	15.51	17.34	12.12	13.56	N/A	N/A
T8	16.14	18.05	13.11	14.66	Synthon O	
T9	12.13	13.57	10.14	11.34	13.70	15.32
T10	13.26	14.83	11.03	12.34	14.88	16.64
T11	13.55	15.15	10.27	11.49	15.04	16.82
T12	14.12	15.79	11.45	12.80	15.04	16.82

Appendix F: Descriptors

Table F.1 Information for molecular descriptors.

ID	MW	HAC	HC	CHA	hete roat om	carbon s	LogP	Vol ume	SA	RB	Polariz ability	Densit y	crystall izabilit y
4-Biphenylcarboxyli c acid	198	15	10	6.50	2	13	3.27	167	37	2	24	1.28	Yes
4-Biphenylmethanol	184	14	12	13.00	1	13	3.13	169	20	2	23	1.25	Yes
4-Phenylphenol	170	13	10	12.00	1	12	3.32	153	20	1	21	1.24	Yes
Anthranilic acid	147	10	7	2.33	3	7	0.79	104	63	1	15	1.42	Yes
Benzamide	121	9	7	3.50	2	7	0.65	108	43	1	14	1.18	Yes
Benzocaine	165	12	11	3.00	3	9	1.39	146	52	3	19	1.21	Yes
Carbamazepine	236	18	12	5.00	3	15	2.93	186	46	0	28	1.34	Yes
Chlorpropamide	276	17	13	1.43	7	10	1.78	207	83	4	26	1.42	Yes
Chlorzoxazone	169	11	4	1.75	4	7	1.72	114	38	0	15.5	1.65	Yes
Felbinac	212	16	12	7.00	2	14	3.21	182	37	3	25	1.23	Yes
Flufenamic acid	281	20	10	2.33	6	14	3.98	201	49	3	27	1.48	Yes
Haloperidol	376	26	23	4.20	5	21	3.49	303	41	6	40	1.31	Yes
Indoprofen	281	21	15	4.25	4	17	2.84	215	58	3	31	1.32	Yes
Phenacetin	180	13	13	3.33	3	10	1.56	163	38	3	20	1.23	Yes
Theophylline	180	13	8	1.17	6	7	-0.06	123	69	0	17	1.49	Yes
Tolbutamide	270	18	18	2.00	6	12	2.13	228	84	5	28	1.25	Yes
Lidocaine	234	17	22	4.67	3	14	2.41	228	32	5	29	1.09	Yes
Tolfenamic acid	262	18	12	3.50	4	14	4.1	196	49	3	29	1.33	Yes
Celecoxib	381	26	14	1.89	9	17	4.34	266	86	3	36	1.52	No
Clofoctol	365	24	26	7.00	3	21	8.19	324	20	5	41	1.22	No
Dibucaine	343	25	29	4.00	5	20	3.77	320	55	10	41	1.15	No
Droperidol	379	28	22	3.67	6	22	3.77	299	53	6	41	1.31	No
Flurbiprofen	244	18	13	5.00	3	15	3.94	204	37	3	26	1.29	No
Nifedipine	346	25	18	2.13	8	17	2.31	272	110	5	35	1.38	No
Salicin	286	20	18	1.86	7	13	-0.77	190	120	4	28	1.41	No
Tolazamide	311	21	21	2.00	7	14	1.41	241	87	3	32	1.34	No
Aceclofenac	354	23	13	2.29	7	16	3.65	243	76	7	35	1.51	No
Felodipine	384	25	19	2.57	7	18	2.24	301	65	6	38	2.24	No
Ibuprofen	206	15	18	6.50	2	13	3.75	200	37	4	24	1.18	No
Indomethacin	358	25	16	3.17	6	19	3.58	270	69	4	37	1.41	No
Nimesulide	308	21	12	1.63	8	13	3.08	212	110	4	30	1.48	No
Ketoprofen	254	19	14	5.33	3	16	3.31	212	54	4	28	1.28	No
Pimozide	461	34	29	4.67	6	28	5.34	377	36	7	51	1.23	No
Probucol	517	35	48	7.75	4	31	11.62	481	91	8	63	1.05	No
Procaine	236	17	20	3.25	4	13	1.32	219	56	7	27	1.17	No
Ritonavir	720	50	46	2.85	13	37	5.98	582	202	18	79	1.25	No

Table F.2 Information for conformational descriptors.

ID	conformers	crystallizability
4-Biphenylcarboxylic acid	3	Yes

4-Biphenylmethanol	2	Yes
4-Phenylphenol	2	Yes
Anthranilic acid	2	Yes
Benzamide	1	Yes
Benzocaine	19	Yes
Carbamazepine	4	Yes
Chlorpropamide	1872	Yes
Chlorzoxazone	0	Yes
Felbinac	24	Yes
Flufenamic acid	95	Yes
Haloperidol	2000	Yes
Indoprofen	48	Yes
Phenacetin	40	Yes
Theophylline	0	Yes
Tolbutamide	2000	Yes
Lidocaine	2000	Yes
Tolfenamic acid	46	Yes
Tolazamide	2000	No
Flurbiprofen	38	No
Clofoctol	384	No
Celecoxib	72	No
Dibucaine	2000	No
Droperidol	2000	No
Nifedipine	223	No
Salicin	258	No
Aceclofenac	2000	No
Felodipine	546	No
Ibuprofen	46	No
Indomethacin	384	No
Nimesulide	71	No
Ketoprofen	153	No
Pimozide	2000	No
Probucol	174	No
Procaine	2000	No
Ritonavir	2000	No

Table F.3 Information for thermodynamic descriptors.

ID	Tm	Tg	TmTg	Enthalpy	entropy	crystallizability
4-Biphenylcarboxylic acid	227	69	158	32.26	0.06	Yes
4-Biphenylmethanol	99	-22	121	25.06	0.07	Yes
4-Phenylphenol	166	21	145	31.5	0.07	Yes
Anthranilic acid	147	5	142	22.86	0.05	Yes
Benzamide	127	-10	137	21.69	0.05	Yes
Benzocaine	89	-31	120	22.55	0.06	Yes
Carbamazepine	192	61	131	25.54	0.06	Yes
Chlorpropamide	124	16	108	27.44	0.07	Yes
Chlorzoxazone	191	38	153	25.62	0.06	Yes
Felbinac	164	24	140	29.76	0.07	Yes
Flufenamic acid	135	17	118	27.13	0.07	Yes
Haloperidol	152	33	119	54.26	0.13	Yes
Indoprofen	212	50	162	36.04	0.07	Yes
Phenacetin	136	2	134	31.49	0.08	Yes

Theophylline	272	94	178	29.61	0.05	Yes
Tolbutamide	129	4	125	26.24	0.07	Yes
Lidocaine	68	-39	107	16.7	0.05	Yes
Tolfenamic acid	213	63	150	38.83	0.08	Yes
Celecoxib	163	58	105	37.42	0.09	No
Clofoctol	88	-4	92	35.15	0.10	No
Dibucaine	68	-39	107	29.23	0.09	No
Droperidol	143	29	114	29.23	0.09	No
Flurbiprofen	115	-6	121	27.41	0.07	No
Nifedipine	173	45	128	38.19	0.09	No
Salicin	201	58	143	52.76	0.11	No
Tolazamide	172	18	154	43.44	0.10	No
Aceclofenac	153	10	143	42.25	0.10	No
Felodipine	147	45	102	30.98	0.07	No
Ibuprofen	77	-45	122	26.48	0.08	No
Indomethacin	162	45	117	37.56	0.09	No
Nimesulide	150	21	129	33.38	0.08	No
Ketoprofen	95	-3	98	28.31	0.08	No
Pimozide	219	54	165	42.74	0.09	No
Probucol	127	27	100	34.22	0.09	No
Procaine	62	-39	101	26.2	0.08	No
Ritonavir	126	49	77	65.34	0.16	No

Table F.4 Information for intermolecular descriptors.

ID	Donor	Acceptor	Ratio	Propensity	crystallizability
4-Biphenylcarboxylic acid	1	1	1	0.38	Yes
4-Biphenylmethanol	1	1	1.00	0.16	Yes
4-Phenylphenol	1	1	1.00	0.16	Yes
Anthranilic acid	2	3	0.67	0.57	Yes
Benzamide	1	1	1.00	0.7	Yes
Benzocaine	1	3	0.33	0.65	Yes
Carbamazepine	1	1	1.00	0.788	Yes
Chlorpropamide	2	3	0.67	0.677	Yes
Chlorzoxazone	1	2	0.50	0.722	Yes
Felbinac	1	2	0.50	0.375	Yes
Flufenamic acid	2	6	0.33	0.399	Yes
Haloperidol	1	4	0.25	0.39	Yes
Indoprofen	1	3	0.33	0.3	Yes
Phenacetin	1	2	0.50	0.65	Yes
Theophylline	1	3	0.33	0.38	Yes
Tolbutamide	2	3	0.67	0.71	Yes
Lidocaine	1	2	0.50	0.36	Yes
Tolfenamic acid	2	3	0.67	0.91	Yes
Celecoxib	1	7	0.14	0.7	No
Clofoctol	1	1	1.00	0.09	No
Dibucaine	1	4	0.25	0.3	No
Droperidol	1	4	0.25	0.26	No
Flurbiprofen	1	3	0.33	0.32	No
Nifedipine	1	7	0.14	0.26	No
Salicin	5	7	0.71	0.59	No
Tolazamide	2	4	0.50	0.7	No
Aceclofenac	2	5	0.40	0.57	No

Felodipine	1	5	0.20	0.32	No
Ibuprofen	1	2	0.50	0.37	No
Indomethacin	1	4	0.25	0.36	No
Nimesulide	1	6	0.17	0.68	No
Ketoprofen	1	3	0.33	0.19	No
Pimozide	1	4	0.25	0.27	No
Probucol	2	4	0.50	0.02	No
Procaine	1	4	0.25	0.59	No
Ritonavir	4	9	0.44	0.4	No

Table F.5 Information for intermolecular descriptors.

ID	Tm (Online Software)	TPSA (Tot)	S axis (Å)	S/L axis ratio	crystallizability
4-Biphenylcarboxylic acid	122.8	37.3	4.608666667	0.337666667	Yes
4-Biphenylmethanol	86.97	20.23	4.9148	0.3732	Yes
4-Phenylphenol	86.56	20.23	4.227333333	0.342666667	Yes
Anthranilic acid	94	75.63	4.74425	0.54275	Yes
Benzamide	81.22	43.09	3.4	0.359	Yes
Benzocaine	66.17	63.32	4.818666667	0.386	Yes
Carbamazepine	162	52.32	7.252	0.631	Yes
Chlorpropamide	179.9	48.02	6.1579	0.4188	Yes
Chlorzoxazone	93	86.36	5.4946	0.39255	Yes
Felbinac	130	83.65	5.4945	0.39255	Yes
Flufenamic acid	139	46	6.67225	0.521	Yes
Haloperidol	194	20.23	6.49175	0.38165	Yes
Indoprofen	198	54.46	6.56155	0.42705	Yes
Phenacetin	115	58.1	5.02925	0.38155	Yes
Theophylline	193	37.3	4.161	0.424	Yes
Tolbutamide	182	64.63	6.62595	0.4787	Yes
Lidocaine	144	49.33	6.4858	0.4933	Yes
Tolfenamic acid	150	37.3	6.88925	0.5488	Yes
Celecoxib	212	40.54	6.532222222	0.405111111	No
Clofoctol	171	37.3	7.95265	0.524	No
Dibucaine	217	68.53	7.6447	0.45505	No
Droperidol	229	57.61	6.74135	0.3273	No
Flurbiprofen	133	54.37	6.33635	0.46205	No
Nifedipine	176	32.34	9.05515	0.7494	No
Salicin	208	110.45	6.8581	0.5697	No
Tolazamide	203	109.6	6.89295	0.4738	No
Aceclofenac	200	38.33	7.9244	0.58995	No
Felodipine	166	41.03	9.2632	0.7568	No
Ibuprofen	94	55.56	6.48265	0.52055	No
Indomethacin	219	119.61	6.77875	0.4188	No

Nimesulide	191	72.68	6.6523	0.5263	No
Ketoprofen	149	86.89	6.7587	0.5162	No
Pimozide	252	83.65	8.6274	0.44765	No
Probucol	245	49.33	4.193	0.231	No
Procaine	113	91.06	6.6091	0.4563	No
Ritonavir	349	202.26	4.381	0.246	No

Table F.6 Information for intermolecular descriptors for bergstrom study

ID	MW	Tm	Tm using online software	TPSA(Tot)	Tg	SA	RB	heteroatom	S axis (Å)	S/L axis ratio	crystallizability
Bucetin	223	158	149	58.56	-3	58	5	4	5.6122	0.3522	Yes
Bufexamac	223	167	161	58.56	2	59	6	4	6.13385	0.38085	Yes
Diflunisal	250	213	169	57.53	43	57	2	5	5.57575	0.4335	Yes
Fenbufen	254	186	162	54.37	28	54	5	3	5.6306	0.353	Yes
Flumequine	261	258	167	59.3	78	58	1	5	5.9036667	0.534	Yes
Mefenamic acid	241	231	152	49.33	51	49	3	3	7.04445	0.5566	Yes
Naproxen	230	153	137	46.53	-3	46	3	3	6.3718	0.4626	Yes
Pindolol	248	169	145	57.28	14	57	6	4	6.4751	0.45105	Yes
Primidone	218	283	195	58.2	72	58	2	4	5.849	0.513	Yes
Saccharin	183	226	159	71.62	20	72	0	5	3.751	0.389	Yes
Spiperone	395	207	238	52.65	106	53	6	6	7.24255	0.40845	Yes
Sulfamethoxazole	253	169	172	106.6	16	107	3	7	5.222	0.359	Yes
Trimethoprim	290	197	189	105.51	51	106	5	7	6.6568	0.4368	Yes
Tyramine	137	161	60	46.25	-37	46	2	2	5.057	0.4815	Yes
Zoxazolamine	168	183	93.8	52.05	11	52	0	4	3.5	0.319	Yes
Acetohexamide	324	190	211	100.72	26	101	4	7	6.5139	0.432	No
Bezafibrate	361	184	230	75.63	73	76	7	6	6.7565	0.32955	No
Testosterone	288	153	144	37.3	42	37	0	2	4.381	0.295	No
Glafenine	372	164	229	91.68	63	92	7	7	6.2761	0.4086	No
Hydrochlorothiazide	297	263	202	135.12	110	135	1	10	4.955	0.451	No
Hydrocortisone	362	211	214.5	94.83	86	95	2	5	7.60525	0.51875	No
Warfarin	308	162	197	67.51	68	64	4	4	7.7584	0.58205	No

Table F.7 Information for intermolecular descriptors for Kohrenon study

ID	MW	Tm	Tm from source	TPSA(Tot)	S axis (Å)	S/L axis ratio	Tg	SA	RB	heteroatom	crystallizability
Nabumetone	228	81	107.84	26.3	5.8124	0.3766	51	26	4	2	Yes
Salcylamide	137	140	118	63.32	4.742	0.50075	123	63	1	3	Yes
Salicyclic acid	138	159	93	57.53	4.74925	0.54325	107	58	1	3	Yes
Thiosalycyclic acid	154	164	85.72	76.1	4.979333	0.567333	146	76	1	3	Yes
Sulfanilamide	172	164	124	94.56	3.6	0.315	128	95	1	5	Yes
Pyrazinecarboxamide	123	189	106.77	68.87	3.464	0.373	178	69	1	4	Yes
Sulfamerazine	264	237	184.38	106.35	5.712667	0.381667	174	106	3	7	Yes

								-				
Acetylsalicylic acid	180	136	93	63.6	5.6951	0.56335	26.92	64	3	4	Yes
Sulfadimidine	278	197	189	106.35	6.61255	0.4355	78.12	106	3	7	No
Perphenazine	404	94	223	59.88	7.3087	0.45585	15.31	55	6	6	No
Cimetidine	252	139	194	114.19	6.69895	0.52875	46.52	114	8	7	No
Pyridoxine	169	157	124	73.58	5.220167	0.54	66.26	74	2	4	No

Table F.8 Information for intermolecular descriptors for Taylor remaining dataset study

ID	MW	Tm	Tm using software	Tg	S axis (Å)	S/L axis ratio	TPSA (Tot)	heteroatom	SA	RB	crystallizability
Antipyrin	188	111	119	-22	5.83975	0.489	26.93	3	24	1	yes
Caffeine	194	237	180	-31	4.164	0.429	61.82	6	58	0	yes
4-Biphenyl carboxaldehyde	182	59	73	-50	4.382667	0.333667	17.07	1	17	2	Yes
Griseofulvin	353	218	189	89	9.553	0.688333	71.06	7	71	3	yes
Nilutamide	317	155	202	33	6.924875	0.522	95.23	10	95	3	No
Bifonazole	310	151	197	17	7.45185	0.49655	17.82	2	18	4	No
Felodipine	384	147	166	45	8.4902	0.6496	64.63	7	65	6	No
Cinnarizine	368	121	198	7	7.7209	0.4385	6.48	2	6	6	No
Loratadine	382	136	201	37	7.46815	0.45445	42.43	5	42	2	No
Clotrimazole	345	145	206.55	30	3.5	0.283	17.82	3	18	4	No
Itraconazole	705	168	349.84	58	7.9829	0.27455	104.7	14	101	11	No
Ketoconazole	531	150	291.82	45	7.81685	0.33825	69.06	10	69	7	No
Miconazole	416	86	215.55	1	7.5195	0.5026	27.05	7	27	6	No

Appendix-G: HBP and HBE results of co-crystal screens

G.1 Chapter 3-Pyrazole with 20 carboxylic acids

Table G.3.1 The multicomponent energy (MCE) score of **P1-P12** with 20 co-formers.

Co-crystal	MCE Score	Co-crystal	MCE Score	Co-crystal	MCE Score	Co-crystal	MCE Score
P1-Suc	2.17	P2-Suc	3.82	P3-Suc	2.56	P4-Suc	3.94
P1-Adipic	1.75	P2-Adipic	3.38	P3-Adipic	2.15	P4-Suc	3.52
P1-Seb	1.20	P2-Seb	2.81	P3-Seb	1.61	P4-Adipic	2.97
P1-Sub	0.52	P2-Sub	2.10	P3-Sub	0.95	P4-Seb	2.28
P1-Dod	0.38	P2-Dod	1.96	P3-Dod	0.81	P4-Sub	2.14
P1-Fum	3.43	P2-Fum	5.12	P3-Fum	3.79	P4-Dod	5.22
P1-Mal	4.00	P2-Mal	5.71	P3-Mal	4.34	P4-Fum	5.79
P1-Glut	2.31	P2-Glut	3.96	P3-Glut	2.69	P4-Mal	4.08
P1-Pim	1.62	P4-Pim	3.24	P3-Pim	2.02	P4-Glut	3.38
P1-Aze	0.93	P2-Aze	2.53	P3-Aze	1.34	P4-Pim	2.69
P1-3OHBA	2.88	P2-3OHBA	0.00	P3-3OHBA	3.24	P4-Aze	4.65
P1-4OHBA	5.83	P2-4OHBA	7.60	P3-4OHBA	6.12	P4-3OHBA	7.63
P1-3AminoBA	-0.53	P2-3AminoBA	1.02	P3-3AminoBA	-0.07	P4-4OHBA	1.22
P1-4AminoBA	-2.57	P2-4AminoBA	-1.10	P3-4AminoBA	-2.06	P4-3AminoBA	-0.84
P1-3NitroBA	6.13	P2-3NitroBA	7.92	P3-3NitroBA	6.41	P4-4AminoBA	6.41
P1-4NitroBA	6.89	P2-4NitroBA	8.70	P3-4NitroBA	7.15	P4-3NitroBA	8.70
P1-BA	-1.36	P2-BA	0.16	P3-BA	-0.88	P4-4NitroBA	0.38
P1-4-IodoBA	2.57	P2-4-IodoBA	4.23	P3-4-IodoBA	2.95	P4-BA	4.35
P1-4-bromoBA	2.73	P2-4-bromoBA	4.39	P3-4-bromoBA	3.10	P4-4-IodoBA	4.50
P1-PentaBA	7.04	P2-PentaBA	8.86	P3-PentaBA	7.30	P4-4-bromoBA	8.85
P5-Suc	2.43	P6-Suc	3.94	P7-Suc	3.94	P8-Suc	2.04
P5-Adipic	2.02	P6-Adipic	3.52	P7-Adipic	3.52	P8-Adipic	1.65
P5-Seb	1.48	P6-Seb	2.97	P7-Seb	2.97	P8-Seb	1.13
P5-Sub	0.80	P6-Sub	2.28	P7-Sub	2.28	P8-Sub	0.48
P5-Dod	0.67	P6-Dod	2.14	P7-Dod	2.14	P8-Dod	0.36
P5-Fum	3.67	P6-Fum	5.22	P7-Fum	5.22	P8-Fum	3.22
P5-Mal	4.23	P6-Mal	5.79	P7-Mal	5.79	P8-Mal	3.75
P5-Glut	2.56	P6-Glut	4.08	P7-Glut	4.08	P8-Glut	2.17
P5-Pim	1.88	P6-Pim	3.38	P7-Pim	3.38	P8-Pim	1.52
P5-Aze	1.21	P6-Aze	2.69	P7-Aze	2.69	P8-Aze	0.87
P5-3OHBA	3.12	P6-3OHBA	4.65	P7-3OHBA	2.70	P8-3OHBA	4.27
P5-4OHBA	6.02	P6-4OHBA	7.63	P7-4OHBA	5.47	P8-4OHBA	7.15
P5-3AminoBA	-0.22	P6-3AminoBA	1.22	P7-3AminoBA	-0.50	P8-3AminoBA	0.96
P5-4AminoBA	-2.23	P6-4AminoBA	-0.84	P7-4AminoBA	-2.41	P8-4AminoBA	-1.03
P5-3NitroBA	6.32	P6-3NitroBA	7.94	P7-3NitroBA	5.75	P8-3NitroBA	7.44
P5-4NitroBA	7.06	P6-4NitroBA	8.70	P7-4NitroBA	6.46	P8-4NitroBA	8.18
P5-BA	-1.04	P6-BA	0.38	P7-BA	-1.28	P8-BA	0.15
P5-4-IodoBA	2.82	P6-4-IodoBA	4.35	P7-4-IodoBA	2.41	P8-4-IodoBA	3.98
P5-4-bromoBA	2.97	P6-4-bromoBA	4.50	P7-4-bromoBA	2.56	P8-4-bromoBA	4.13
P5-PentaBA	7.21	P6-PentaBA	8.85	P7-PentaBA	6.60	P8-PentaBA	8.33

Co-crystal		Co-crystal		Co-crystal		Co-crystal	
P9-Suc	3.59	P10-Suc	2.46	P11-Suc	3.92	P12-Suc	2.28
P9-Adipic	3.18	P10-Adipic	2.10	P11-Adipic	3.53	P12-Adipic	1.94
P9-Seb	2.64	P10-Seb	1.62	P11-Seb	3.02	P12-Seb	1.50
P9-Sub	1.98	P10-Sub	1.03	P11-Sub	2.38	P12-Sub	0.96
P9-Dod	1.84	P10-Dod	0.92	P11-Dod	2.25	P12-Dod	0.85
P9-Fum	4.82	P10-Fum	3.55	P11-Fum	5.09	P12-Fum	3.28
P9-Mal	5.37	P10-Mal	4.04	P11-Mal	5.62	P12-Mal	3.74
P9-Glut	3.72	P10-Glut	2.58	P11-Glut	4.05	P12-Glut	2.39
P9-Pim	3.05	P10-Pim	1.98	P11-Pim	3.40	P12-Pim	1.83
P9-Aze	2.38	P10-Aze	1.39	P11-Aze	2.76	P12-Aze	1.28
P9-3OHBA	3.07	P10-3OHBA	4.58	P11-3OHBA	2.84	P12-3OHBA	3.64
P9-4OHBA	5.62	P10-4OHBA	7.32	P11-4OHBA	5.19	P12-4OHBA	6.09
P9-3AminoBA	0.13	P10-3AminoBA	1.41	P11-3AminoBA	0.12	P12-3AminoBA	0.82
P9-4AminoBA	-1.63	P10-4AminoBA	-0.49	P11-4AminoBA	-1.51	P12-4AminoBA	-0.88
P9-3NitroBA	5.88	P10-3NitroBA	7.60	P11-3NitroBA	5.44	P12-3NitroBA	6.34
P9-4NitroBA	6.53	P10-4NitroBA	8.31	P11-4NitroBA	6.04	P12-4NitroBA	6.97
P9-BA	-0.59	P10-BA	0.63	P11-BA	-0.54	P12-BA	0.13
P9-4-IodoBA	2.81	P10-4-IodoBA	4.29	P11-4-IodoBA	2.60	P12-4-IodoBA	3.39
P9-4-bromoBA	2.94	P10-4-bromoBA	4.44	P11-4-bromoBA	2.72	P12-4-bromoBA	3.52
P9-PentaBA	6.66	P10-PentaBA	8.45	P11-PentaBA	6.16	P12-PentaBA	7.10

Table G.3.2 The multicomponent propensity (MCP) score of **P1-P12** with 20 co-formers.

Co-crystal	MCP Score	Co-crystal	MCP Score	Co-crystal	MCP Score	Co-crystal	MCP Score
P1-Suc	-0.09	P2-Suc	-0.10	P3-Suc	-0.15	P4-Suc	-0.15
P1-Adipic	-0.09	P2-Adipic	-0.10	P3-Adipic	-0.18	P4-Suc	-0.18
P1-Seb	-0.09	P2-Seb	-0.10	P3-Seb	-0.17	P4-Adipic	-0.17
P1-Sub	-0.09	P2-Sub	-0.09	P3-Sub	-0.19	P4-Seb	-0.18
P1-Dod	-0.08	P2-Dod	-0.09	P3-Dod	-0.18	P4-Sub	-0.18
P1-Fum	-0.11	P2-Fum	-0.09	P3-Fum	-0.05	P4-Dod	-0.05
P1-Mal	-0.08	P2-Mal	-0.10	P3-Mal	-0.12	P4-Fum	-0.12
P1-Glut	-0.09	P2-Glut	-0.10	P3-Glut	-0.16	P4-Mal	-0.17
P1-Pim	-0.09	P4-Pim	-0.09	P3-Pim	-0.17	P4-Glut	-0.17
P1-Aze	-0.09	P2-Aze	2.97	P3-Aze	-0.18	P4-Pim	-0.18
P1-3OHBA	-0.19	P2-3OHBA	-0.17	P3-3OHBA	-0.19	P4-Aze	-0.16
P1-4OHBA	-0.20	P2-4OHBA	-0.17	P3-4OHBA	-0.17	P4-3OHBA	-0.17
P1-3AminoBA	0.02	P2-3AminoBA	0.04	P3-3AminoBA	-0.10	P4-4OHBA	0.03
P1-4AminoBA	0.02	P2-4AminoBA	0.05	P3-4AminoBA	-0.02	P4-3AminoBA	0.04
P1-3NitroBA	-0.16	P2-3NitroBA	-0.17	P3-3NitroBA	-0.09	P4-4AminoBA	-0.10
P1-4NitroBA	-0.11	P2-4NitroBA	-0.13	P3-4NitroBA	0.05	P4-3NitroBA	-0.03
P1-BA	-0.14	P2-BA	-0.13	P3-BA	-0.07	P4-4NitroBA	-0.10
P1-4-IodoBA	-0.14	P2-4-IodoBA	-0.14	P3-4-IodoBA	-0.10	P4-BA	-0.15
P1-4-bromoBA	-0.11	P2-4-bromoBA	-0.11	P3-4-bromoBA	-0.16	P4-4-IodoBA	-0.11
P1-PentaBA	-0.11	P2-PentaBA	-0.10	P3-PentaBA	-0.04	P4-4-bromoBA	-0.05
P5-Suc	-0.11	P6-Suc	-0.12	P7-Suc	0.05	P8-Suc	0.05
P5-Adipic	-0.13	P6-Adipic	-0.14	P7-Adipic	0.04	P8-Adipic	0.06
P5-Seb	-0.14	P6-Seb	-0.14	P7-Seb	0.04	P8-Seb	0.07
P5-Sub	-0.14	P6-Sub	-0.14	P7-Sub	0.05	P8-Sub	0.06
P5-Dod	-0.15	P6-Dod	-0.15	P7-Dod	0.07	P8-Dod	0.06

353

P5-Fum	-0.04	**P6-Fum**	-0.03	**P7-Fum**	0.00	**P8-Fum**	0.06
P5-Mal	-0.09	**P6-Mal**	-0.08	**P7-Mal**	0.04	**P8-Mal**	0.06
P5-Glut	-0.13	**P6-Glut**	-0.13	**P7-Glut**	0.05	**P8-Glut**	0.07
P5-Pim	-0.14	**P6-Pim**	-0.14	**P7-Pim**	0.04	**P8-Pim**	0.07
P5-Aze	-0.14	**P6-Aze**	-0.15	**P7-Aze**	0.06	**P8-Aze**	0.07
P5-3OHBA	-0.09	**P6-3OHBA**	-0.09	**P7-3OHBA**	-0.20	**P8-3OHBA**	-0.16
P5-4OHBA	-0.08	**P6-4OHBA**	-0.07	**P7-4OHBA**	-0.19	**P8-4OHBA**	-0.15
P5-3AminoBA	0.02	**P6-3AminoBA**	0.02	**P7-3AminoBA**	0.05	**P8-3AminoBA**	0.09
P5-4AminoBA	0.03	**P6-4AminoBA**	-0.01	**P7-4AminoBA**	0.07	**P8-4AminoBA**	0.10
P5-3NitroBA	0	**P6-3NitroBA**	0	**P7-3NitroBA**	-0.09	**P8-3NitroBA**	-0.06
P5-4NitroBA	0.01	**P6-4NitroBA**	0.01	**P7-4NitroBA**	0.05	**P8-4NitroBA**	0.08
P5-BA	-0.11	**P6-BA**	-0.12	**P7-BA**	-0.07	**P8-BA**	-0.04
P5-4-IodoBA	-0.09	**P6-4-IodoBA**	-0.1	**P7-4-IodoBA**	-0.08	**P8-4-IodoBA**	-0.02
P5-4-bromoBA	-0.07	**P6-4-bromoBA**	-0.06	**P7-4-bromoBA**	-0.03	**P8-4-bromoBA**	0.01
P5-PentaBA	-0.05	**P6-PentaBA**	-0.06	**P7-PentaBA**	-0.05	**P8-PentaBA**	-0.01
P9-Suc	-0.10	**P10-Suc**	-0.08	**P11-Suc**	-0.09	**P12-Suc**	-0.08
P9-Adipic	-0.10	**P10-Adipic**	-0.08	**P11-Adipic**	-0.10	**P12-Adipic**	-0.08
P9-Seb	-0.10	**P10-Seb**	-0.08	**P11-Seb**	-0.09	**P12-Seb**	-0.08
P9-Sub	-0.09	**P10-Sub**	-0.08	**P11-Sub**	-0.09	**P12-Sub**	-0.07
P9-Dod	-0.10	**P10-Dod**	-0.07	**P11-Dod**	-0.09	**P12-Dod**	-0.07
P9-Fum	-0.16	**P10-Fum**	-0.14	**P11-Fum**	-0.16	**P12-Fum**	-0.13
P9-Mal	-0.09	**P10-Mal**	-0.07	**P11-Mal**	-0.09	**P12-Mal**	-0.07
P9-Glut	-0.10	**P10-Glut**	-0.09	**P11-Glut**	-0.10	**P12-Glut**	-0.08
P9-Pim	-0.10	**P10-Pim**	-0.08	**P11-Pim**	-0.10	**P12-Pim**	-0.08
P9-Aze	-0.10	**P10-Aze**	-0.08	**P11-Aze**	-0.09	**P12-Aze**	-0.07
P9-3OHBA	-0.17	**P10-3OHBA**	-0.15	**P11-3OHBA**	-0.16	**P12-3OHBA**	-0.14
P9-4OHBA	-0.17	**P10-4OHBA**	-0.15	**P11-4OHBA**	-0.16	**P12-4OHBA**	-0.14
P9-3AminoBA	0.03	**P10-3AminoBA**	0.05	**P11-3AminoBA**	0.03	**P12-3AminoBA**	0.05
P9-4AminoBA	0.04	**P10-4AminoBA**	0.06	**P11-4AminoBA**	0.03	**P12-4AminoBA**	0.06
P9-3NitroBA	-0.19	**P10-3NitroBA**	-0.17	**P11-3NitroBA**	-0.16	**P12-3NitroBA**	-0.16
P9-4NitroBA	-0.19	**P10-4NitroBA**	-0.16	**P11-4NitroBA**	-0.16	**P12-4NitroBA**	-0.14
P9-BA	-0.22	**P10-BA**	-0.19	**P11-BA**	-0.20	**P12-BA**	-0.19
P9-4-IodoBA	-0.19	**P10-4-IodoBA**	-0.17	**P11-4-IodoBA**	-0.18	**P12-4-IodoBA**	-0.16
P9-4-bromoBA	-0.19	**P10-4-bromoBA**	-0.08	**P11-4-bromoBA**	-0.09	**P12-4-bromoBA**	-0.08
P9-PentaBA	-0.18	**P10-PentaBA**	-0.08	**P11-PentaBA**	-0.10	**P12-PentaBA**	-0.08

Table G.3.3 Comparison of MCE score for **P1-P12** with 20 co-formers with the actual outcome of the experimental co-crystal screening (check mark means co-crystal formation and a cross means no co-crystal formation). The green box shows true positive, yellow box shows true negatives, the red box shows false positives and blue box shows true negatives.

	MCE	Exp.	Match?		MCE	Exp.	Match?		MCE	Exp.	Match?
P1-Suc	2.17	✗	No	**P2-Suc**	3.82	✓	yes	**P3-Suc**	2.56	✓	yes
P1-Adipic	1.75	✗	No	**P2-Adipic**	3.38	✓	yes	**P3-Adipic**	2.15	✓	yes
P1-Seb	1.20	✗	No	**P2-Seb**	2.81	✓	yes	**P3-Seb**	1.61	✓	yes
P1-Sub	0.52	✗	No	**P2-Sub**	2.10	✓	yes	**P3-Sub**	0.95	✓	yes
P1-Dod	0.38	✗	No	**P2-Dod**	1.96	✓	yes	**P3-Dod**	0.81	✓	yes
P1-Fum	3.43	✗	No	**P2-Fum**	5.12	✓	yes	**P3-Fum**	3.79	✓	yes
P1-Mal	4.00	✗	No	**P2-Mal**	5.71	✓	yes	**P3-Mal**	4.34	✓	yes
P1-Glut	2.31	✗	No	**P2-Glut**	3.96	✓	yes	**P3-Glut**	2.69	✓	yes
P1-Pim	1.62	✗	No	**P4-Pim**	3.24	✓	yes	**P3-Pim**	2.02	✓	yes
P1-Aze	0.93	✗	No	**P2-Aze**	2.53	✓	yes	**P3-Aze**	1.34	✓	yes

Name	MCE	Exp.	Match?	Name	MCE	Exp.	Match?	Name	MCE	Exp.	Match?
P1-3OHBA	2.88	✓	yes	P2-3OHBA	0.00	✗	yes	P3-3OHBA	3.24	✗	no
P1-4OHBA	5.83	✓	yes	P2-4OHBA	7.60	✓	yes	P3-4OHBA	6.12	✗	no
P1-3AminoBA	-0.53	✓	no	P2-3AminoBA	1.02	✓	yes	P3-3AminoBA	-0.07	✓	no
P1-4AminoBA	-2.57	✓	no	P2-4AminoBA	-1.10	✓	no	P3-4AminoBA	-2.06	✓	no
P1-3NitroBA	6.13	✓	yes	P2-3NitroBA	7.92	✓	yes	P3-3NitroBA	6.41	✓	yes
P1-4NitroBA	6.89	✗	no	P2-4NitroBA	8.70	✓	yes	P3-4NitroBA	7.15	✓	yes
P1-BA	-1.36	✗	yes	P2-BA	0.16	✓	yes	P3-BA	-0.88	✓	no
P1-4-IodoBA	2.57	✗	no	P2-4-IodoBA	4.23	✗	no	P3-4-IodoBA	2.95	✗	no
P1-4-bromoBA	2.73	✗	no	P2-4-bromoBA	4.39	✗	no	P3-4-bromoBA	3.10	✗	no
P1-PentaBA	7.04	✓	yes	P2-PentaBA	8.86	✓	yes	P3-PentaBA	7.30	✓	yes
	MCE	Exp.	Match?		MCE	Exp.	Match?		MCE	Exp.	Match?
P4-Suc	3.94	✓	yes	P5-Suc	2.43	✗	No	P6-Suc	3.94	✗	no
P4-Adipic	3.52	✓	yes	P5-Adipic	2.02	✓	No	P6-Adipic	3.52	✗	no
P4-Seb	2.97	✓	yes	P5-Seb	1.48	✓	yes	P6-Seb	2.97	✓	yes
P4-Sub	2.28	✓	yes	P5-Sub	0.80	✗	No	P6-Sub	2.28	✓	yes
P4-Dod	2.14	✓	yes	P5-Dod	0.67	✗	no	P6-Dod	2.14	✗	no
P4-Fum	5.22	✓	yes	P5-Fum	3.67	✓	yes	P6-Fum	5.22	✗	no
P4-Mal	5.79	✓	yes	P5-Mal	4.23	✓	yes	P6-Mal	5.79	✓	yes
P4-Glut	4.08	✓	yes	P5-Glut	2.56	✓	yes	P6-Glut	4.08	✓	yes
P4-Pim	3.38	✓	yes	P5-Pim	1.88	✓	yes	P6-Pim	3.38	✗	no
P4-Aze	2.69	✓	yes	P5-Aze	1.21	✓	yes	P6-Aze	2.69	✗	no
P4-3OHBA	4.65	✓	yes	P5-3OHBA	3.12	✓	yes	P6-3OHBA	4.65	✗	no
P4-4OHBA	7.63	✓	yes	P5-4OHBA	6.02	✗	no	P6-4OHBA	7.63	✗	no
P4-3AminoBA	1.22	✓	yes	P5-3AminoBA	-0.22	✓	no	P6-3AminoBA	1.22	✓	yes
P4-4AminoBA	-0.84	✓	no	P5-4AminoBA	-2.23	✗	yes	P6-4AminoBA	-0.84	✗	yes
P4-3NitroBA	6.41	✓	yes	P5-3NitroBA	6.32	✓	yes	P6-3NitroBA	7.94	✓	yes
P4-4NitroBA	8.70	✗	no	P5-4NitroBA	7.06	✗	no	P6-4NitroBA	8.70	✗	no
P4-BA	0.38	✓	yes	P5-BA	-1.04	✓	no	P6-BA	0.38	✓	yes
P4-4-IodoBA	4.35	✗	no	P5-4-IodoBA	2.82	✗	no	P6-4-IodoBA	4.35	✗	no
P4-4-bromoBA	4.50	✗	no	P5-4-bromoBA	2.97	✗	no	P6-4-bromoBA	4.50	✗	no
P4-PentaBA	8.85	✓	yes	P5-PentaBA	7.21	✓	yes	P6-PentaBA	8.85	✓	yes
	MCE	Exp.	Match?		MCE	Exp.	Match?		MCE	Exp.	Match?
P7-Suc	3.94	✓	yes	P8-Suc	2.04	✓	yes	P9-Suc	3.59	✓	yes
P7-Adipic	3.52	✓	yes	P8-Adipic	1.65	✗	no	P9-Adipic	3.18	✓	yes
P7-Seb	2.97	✓	yes	P8-Seb	1.13	✓	yes	P9-Seb	2.64	✓	yes
P7-Sub	2.28	✓	yes	P8-Sub	0.48	✓	yes	P9-Sub	1.98	✓	yes
P7-Dod	2.14	✗	no	P8-Dod	0.36	✓	yes	P9-Dod	1.84	✓	yes
P7-Fum	5.22	✓	yes	P8-Fum	3.22	✓	yes	P9-Fum	4.82	✓	yes
P7-Mal	5.79	✓	yes	P8-Mal	3.75	✓	yes	P9-Mal	5.37	✓	yes
P7-Glut	4.08	✓	yes	P8-Glut	2.17	✓	yes	P9-Glut	3.72	✓	yes
P7-Pim	3.38	✓	yes	P8-Pim	1.52	✗	no	P9-Pim	3.05	✓	yes
P7-Aze	2.69	✓	yes	P8-Aze	0.87	✓	yes	P9-Aze	2.38	✓	yes
P7-3OHBA	2.70	✗	no	P8-3OHBA	4.27	✗	no	P9-3OHBA	3.07	✓	yes
P7-4OHBA	5.47	✗	no	P8-4OHBA	7.15	✓	yes	P9-4OHBA	5.62	✓	Yes
P7-3AminoBA	-0.50	✓	no	P8-3AminoBA	0.96	✓	yes	P9-3AminoBA	0.13	✓	yes
P7-4AminoBA	-2.41	✗	yes	P8-4AminoBA	-1.03	✗	yes	P9-4AminoBA	-1.63	✗	yes
P7-3NitroBA	5.75	✓	yes	P8-3NitroBA	7.44	✓	yes	P9-3NitroBA	5.88	✓	yes
P7-4NitroBA	6.46	✗	no	P8-4NitroBA	8.18	✓	yes	P9-4NitroBA	6.53	✗	no
P7-BA	-1.28	✓	no	P8-BA	0.15	✓	yes	P9-BA	-0.59	✓	no
P7-4-IodoBA	2.41	✗	no	P8-4-IodoBA	3.98	✗	no	P9-4-IodoBA	2.81	✗	no
P7-4-bromoBA	2.56	✗	no	P8-4-bromoBA	4.13	✗	no	P9-4-bromoBA	2.94	✓	yes
P7-PentaBA	6.60	✓	yes	P8-PentaBA	8.33	✓	yes	P9-PentaBA	6.66	✓	yes
	MCE	Exp.	Match?		MCE	Exp.	Match?		MCE	Exp.	Match?
P10-Suc	2.46	✓	yes	P11-Suc	3.92	✓		P12-Suc	2.28	✓	yes
P10-Adipic	2.10	✓	yes	P11-Adipic	3.53	✓		P12-Adipic	1.94	✓	yes
P10-Seb	1.62	✓	yes	P11-Seb	3.02	✓		P12-Seb	1.50	✓	yes
P10-Sub	1.03	✓	yes	P11-Sub	2.38	✓		P12-Sub	0.96	✓	yes
P10-Dod	0.92	✓	yes	P11-Dod	2.25	✓		P12-Dod	0.85	✓	yes
P10-Fum	3.55	✓	yes	P11-Fum	5.09	✓		P12-Fum	3.28	✓	yes
P10-Mal	4.04	✓	yes	P11-Mal	5.62	✓		P12-Mal	3.74	✓	yes

P10-Glut	2.58	✓	yes	P11-Glut	4.05	✓	(green)	P12-Glut	2.39	✓	yes
P10-Pim	1.98	✓	yes	P11-Pim	3.40	✓	(green)	P12-Pim	1.83	✓	yes
P10-Aze	1.39	✓	yes	P11-Aze	2.76	✓	(green)	P12-Aze	1.28	✓	yes
P10-3OHBA	4.58	✓	yes	P11-3OHBA	2.84	✗	(red)	P12-3OHBA	3.64	✓	yes
P10-4OHBA	7.32	✓	yes	P11-4OHBA	5.19	✗	(red)	P12-4OHBA	6.09	✗	no
P10-3AminoBA	1.41	✓	yes	P11-3AminoBA	0.12	✗	(red)	P12-3AminoBA	0.82	✗	no
P10-4AminoBA	-0.49	✓	no	P11-4AminoBA	-1.51	✗	(blue)	P12-4AminoBA	-0.88	✗	yes
P10-3NitroBA	7.60	✓	yes	P11-3NitroBA	5.44	✓	(green)	P12-3NitroBA	6.34	✓	yes
P10-4NitroBA	8.31	✓	yes	P11-4NitroBA	6.04	✓	(green)	P12-4NitroBA	6.97	✗	no
P10-BA	0.63	✓	yes	P11-BA	-0.54	✗	(blue)	P12-BA	0.13	✓	yes
P10-4-IodoBA	4.29	✗	no	P11-4-IodoBA	2.60	✓	(green)	P12-4-IodoBA	3.39	✓	yes
P10-4-bromoBA	4.44	✗	no	P11-4-bromoBA	2.72	✓	(green)	P12-4-bromoBA	3.52	✓	yes
P10-PentaBA	8.45	✓	yes	P11-PentaBA	6.16	✓	(green)	P12-PentaBA	7.10	✓	yes

Table G.3.4 Comparison of MCP score for **P1-P12** with 20 co-formers with the actual outcome of the experimental co-crystal screening (check mark means co-crystal formation and a cross means no co-crystal formation). The green box shows true positive, yellow box shows true negatives, the red box shows false positives and blue box shows true negatives.

	MCE	Exp.	Match?		MCE	Exp.	Match?		MCE	Exp.	Match?
P1-Suc	-0.09	✗	(blue)	P2-Suc	-0.10	✓	(yellow)	P3-Suc	-0.15	✓	(yellow)
P1-Adipic	-0.09	✗	(blue)	P2-Adipic	-0.10	✓	(yellow)	P3-Adipic	-0.18	✓	(yellow)
P1-Seb	-0.09	✗	(blue)	P2-Seb	-0.10	✓	(yellow)	P3-Seb	-0.17	✓	(yellow)
P1-Sub	-0.09	✗	(blue)	P2-Sub	-0.09	✓	(yellow)	P3-Sub	-0.19	✓	(yellow)
P1-Dod	-0.08	✗	(blue)	P2-Dod	-0.09	✓	(yellow)	P3-Dod	-0.18	✓	(yellow)
P1-Fum	-0.11	✗	(blue)	P2-Fum	-0.09	✓	(yellow)	P3-Fum	-0.05	✓	(yellow)
P1-Mal	-0.08	✗	(blue)	P2-Mal	-0.10	✓	(yellow)	P3-Mal	-0.12	✓	(yellow)
P1-Glut	-0.09	✗	(blue)	P2-Glut	-0.10	✓	(yellow)	P3-Glut	-0.16	✓	(yellow)
P1-Pim	-0.09	✗	(blue)	P4-Pim	-0.09	✓	(yellow)	P3-Pim	-0.17	✓	(yellow)
P1-Aze	-0.09	✗	(blue)	P2-Aze	2.97	✓	(yellow)	P3-Aze	-0.18	✓	(yellow)
P1-3OHBA	-0.19	✓	(yellow)	P2-3OHBA	-0.17	✗	(blue)	P3-3OHBA	-0.19	✗	(blue)
P1-4OHBA	-0.20	✓	(yellow)	P2-4OHBA	-0.17	✓	(yellow)	P3-4OHBA	-0.17	✗	(blue)
P1-3AminoBA	0.02	✓	(green)	P2-3AminoBA	0.04	✓	(yellow)	P3-3AminoBA	-0.10	✓	(yellow)
P1-4AminoBA	0.02	✓	(green)	P2-4AminoBA	0.05	✓	(yellow)	P3-4AminoBA	-0.02	✓	(yellow)
P1-3NitroBA	-0.16	✓	(yellow)	P2-3NitroBA	-0.17	✓	(yellow)	P3-3NitroBA	-0.09	✓	(yellow)
P1-4NitroBA	-0.11	✗	(blue)	P2-4NitroBA	-0.13	✓	(yellow)	P3-4NitroBA	0.05	✓	(yellow)
P1-BA	-0.14	✗	(blue)	P2-BA	-0.13	✓	(yellow)	P3-BA	-0.07	✓	(yellow)
P1-4-IodoBA	-0.14	✗	(blue)	P2-4-IodoBA	-0.14	✗	(blue)	P3-4-IodoBA	-0.10	✗	(blue)
P1-4-bromoBA	-0.11	✗	(blue)	P2-4-bromoBA	-0.11	✗	(blue)	P3-4-bromoBA	-0.16	✗	(blue)
P1-PentaBA	-0.11	✓	(yellow)	P2-PentaBA	-0.10	✓	(yellow)	P3-PentaBA	-0.04	✓	(yellow)
	MCE	Exp.	Match?		MCE	Exp.	Match?		MCE	Exp.	Match?
P4-Suc	-0.15	✓	(yellow)	P5-Suc	-0.11	✗	(blue)	P6-Suc	-0.12	✗	(blue)
P4-Adipic	-0.18	✓	(yellow)	P5-Adipic	-0.13	✓	(yellow)	P6-Adipic	-0.14	✗	(blue)
P4-Seb	-0.17	✓	(yellow)	P5-Seb	-0.14	✓	(yellow)	P6-Seb	-0.14	✓	(yellow)
P4-Sub	-0.18	✓	(yellow)	P5-Sub	-0.14	✗	(blue)	P6-Sub	-0.14	✓	(yellow)
P4-Dod	-0.18	✓	(yellow)	P5-Dod	-0.15	✗	(blue)	P6-Dod	-0.15	✗	(blue)
P4-Fum	-0.05	✓	(yellow)	P5-Fum	-0.04	✓	(yellow)	P6-Fum	-0.03	✗	(blue)
P4-Mal	-0.12	✓	(yellow)	P5-Mal	-0.09	✓	(yellow)	P6-Mal	-0.08	✓	(yellow)
P4-Glut	-0.17	✓	(yellow)	P5-Glut	-0.13	✓	(yellow)	P6-Glut	-0.13	✓	(yellow)
P4-Pim	-0.17	✓	(yellow)	P5-Pim	-0.14	✓	(yellow)	P6-Pim	-0.14	✗	(blue)
P4-Aze	-0.18	✓	(yellow)	P5-Aze	-0.14	✓	(yellow)	P6-Aze	-0.15	✗	(blue)
P4-3OHBA	-0.16	✓	(yellow)	P5-3OHBA	-0.09	✓	(yellow)	P6-3OHBA	-0.09	✗	(blue)

	MCE	Exp.	Match?		MCE	Exp.	Match?		MCE	Exp.	Match?
P4-4OHBA	-0.17	✓	yellow	P5-4OHBA	-0.08	✗	blue	P6-4OHBA	-0.07	✗	blue
P4-3AminoBA	0.03	✓	yellow	P5-3AminoBA	0.02	✓	green	P6-3AminoBA	0.02	✓	green
P4-4AminoBA	0.04	✓	yellow	P5-4AminoBA	0.03	✗	red	P6-4AminoBA	-0.01	✗	blue
P4-3NitroBA	-0.10	✓	yellow	P5-3NitroBA	0	✓		P6-3NitroBA	0	✓	yellow
P4-4NitroBA	-0.03	✗	blue	P5-4NitroBA	0.01	✗	red	P6-4NitroBA	0.01	✗	red
P4-BA	-0.10	✓	yellow	P5-BA	-0.11	✓	yellow	P6-BA	-0.12	✓	yellow
P4-4-IodoBA	-0.15	✗	blue	P5-4-IodoBA	-0.09	✗	blue	P6-4-IodoBA	-0.1	✗	blue
P4-4-bromoBA	-0.11	✗	blue	P5-4-bromoBA	-0.07	✗	blue	P6-4-bromoBA	-0.06	✗	blue
P4-PentaBA	-0.05	✓	yellow	P5-PentaBA	-0.05	✓	yellow	P6-PentaBA	-0.06	✓	yellow

	MCE	Exp.	Match?		MCE	Exp.	Match?		MCE	Exp.	Match?
P7-Suc	0.05	✓	green	P8-Suc	0.05	✓	green	P9-Suc	-0.10	✓	yellow
P7-Adipic	0.04	✓	green	P8-Adipic	0.06	✗	red	P9-Adipic	-0.10	✓	yellow
P7-Seb	0.04	✓	green	P8-Seb	0.07	✓	green	P9-Seb	-0.10	✓	yellow
P7-Sub	0.05	✓	green	P8-Sub	0.06	✓	green	P9-Sub	-0.09	✓	yellow
P7-Dod	0.07	✗	red	P8-Dod	0.06	✓	green	P9-Dod	-0.10	✓	yellow
P7-Fum	0.00	✓	green	P8-Fum	0.06	✓	green	P9-Fum	-0.16	✓	yellow
P7-Mal	0.04	✓	green	P8-Mal	0.06	✓	green	P9-Mal	-0.09	✓	yellow
P7-Glut	0.05	✓	green	P8-Glut	0.07	✓	green	P9-Glut	-0.10	✓	yellow
P7-Pim	0.04	✓	green	P8-Pim	0.07	✗	red	P9-Pim	-0.10	✓	yellow
P7-Aze	0.06	✓	green	P8-Aze	0.07	✓	green	P9-Aze	-0.10	✓	yellow
P7-3OHBA	-0.20	✗	blue	P8-3OHBA	-0.16	✗	blue	P9-3OHBA	-0.17	✓	yellow
P7-4OHBA	-0.19	✗	blue	P8-4OHBA	-0.15	✓	yellow	P9-4OHBA	-0.17	✓	yellow
P7-3AminoBA	0.05	✓	green	P8-3AminoBA	0.09	✓	green	P9-3AminoBA	0.03	✓	green
P7-4AminoBA	0.07	✗	red	P8-4AminoBA	0.10	✗	red	P9-4AminoBA	0.04	✗	red
P7-3NitroBA	-0.09	✓	yellow	P8-3NitroBA	-0.06	✓	yellow	P9-3NitroBA	-0.19	✓	yellow
P7-4NitroBA	0.05	✗	red	P8-4NitroBA	0.08	✓	green	P9-4NitroBA	-0.19	✗	blue
P7-BA	-0.07	✓	yellow	P8-BA	-0.04	✓	yellow	P9-BA	-0.22	✓	yellow
P7-4-IodoBA	-0.08	✗	blue	P8-4-IodoBA	-0.02	✗	blue	P9-4-IodoBA	-0.19	✗	blue
P7-4-bromoBA	-0.03	✗	blue	P8-4-bromoBA	0.01	✗	red	P9-4-bromoBA	-0.19	✓	yellow
P7-PentaBA	-0.05	✓	yellow	P8-PentaBA	-0.01	✓	yellow	P9-PentaBA	-0.18	✓	yellow

	MCE	Exp.	Match?		MCE	Exp.	Match?		MCE	Exp.	Match?
P10-Suc	-0.08	✓	yellow	P11-Suc	-0.09	✓	yellow	P12-Suc	-0.08	✓	yellow
P10-Adipic	-0.08	✓	yellow	P11-Adipic	-0.10	✓	yellow	P12-Adipic	-0.08	✓	yellow
P10-Seb	-0.08	✓	yellow	P11-Seb	-0.09	✓	yellow	P12-Seb	-0.08	✓	yellow
P10-Sub	-0.08	✓	yellow	P11-Sub	-0.09	✓	yellow	P12-Sub	-0.07	✓	yellow
P10-Dod	-0.07	✓	yellow	P11-Dod	-0.09	✓	yellow	P12-Dod	-0.07	✓	yellow
P10-Fum	-0.14	✓	yellow	P11-Fum	-0.16	✓	yellow	P12-Fum	-0.13	✓	yellow
P10-Mal	-0.07	✓	yellow	P11-Mal	-0.09	✓	yellow	P12-Mal	-0.07	✓	yellow
P10-Glut	-0.09	✓	yellow	P11-Glut	-0.10	✓	yellow	P12-Glut	-0.08	✓	yellow
P10-Pim	-0.08	✓	yellow	P11-Pim	-0.10	✓	yellow	P12-Pim	-0.08	✓	yellow
P10-Aze	-0.08	✓	yellow	P11-Aze	-0.09	✓	yellow	P12-Aze	-0.07	✓	yellow
P10-3OHBA	-0.15	✓	yellow	P11-3OHBA	-0.16	✗	blue	P12-3OHBA	-0.14	✓	yellow
P10-4OHBA	-0.15	✓	yellow	P11-4OHBA	-0.16	✗	blue	P12-4OHBA	-0.14	✗	blue
P10-3AminoBA	0.05	✓	green	P11-3AminoBA	0.03	✗	red	P12-3AminoBA	0.05	✗	red
P10-4AminoBA	0.06	✓	green	P11-4AminoBA	0.03	✗	red	P12-4AminoBA	0.06	✗	red
P10-3NitroBA	-0.17	✓	yellow	P11-3NitroBA	-0.16	✓	yellow	P12-3NitroBA	-0.16	✓	yellow
P10-4NitroBA	-0.16	✓	yellow	P11-4NitroBA	-0.16	✓	yellow	P12-4NitroBA	-0.14	✗	blue
P10-BA	-0.19	✓	yellow	P11-BA	-0.20	✗	blue	P12-BA	-0.19	✓	yellow
P10-4-IodoBA	-0.17	✗	blue	P11-4-IodoBA	-0.18	✓	yellow	P12-4-IodoBA	-0.16	✓	yellow
P10-4-bromoBA	-0.17	✗	blue	P11-4-bromoBA	-0.18	✓	yellow	P12-4-bromoBA	-0.16	✓	yellow
P10-PentaBA	-0.17	✓	yellow	P11-PentaBA	-0.18	✓	yellow	P12-PentaBA	-0.16	✓	yellow

G.2 Chapter 5-Thiazole with 20 carboxylic acids

Table G.5.1 MC score of **T1-T12** against 20 carboxylic acids (HBP for **T1-T8** were determined manually whereas HBP for **T9-T12** were determined using new automated HBP tool, HBP tool is very sensitive to the method and number of data used and therefore, MC score could change as the CCDC updates in the future)

Reactants	MC Score	Reactants	MC Score	Reactants	MC Score	Reactants	MC Score
T1-Suc	-0.06	T4-Suc	0	T3-Suc	-0.03	T10-Suc	0.01
T1-Adipic	-0.05	T4-Adipic	0	T3-Adipic	-0.03	T4-Suc	0
T1-Seb	-0.04	T4-Seb	0	T3-Seb	-0.02	T4-Adipic	0
T1-Sub	-0.05	T4-Sub	-0.01	T3-Sub	-0.03	T4-Seb	0
T1-Dod	-0.04	T4-Dod	0	T3-Dod	-0.03	T4-Sub	-0.01
T1-Fum	-0.05	T4-Fum	-0.01	T3-Fum	-0.03	T4-Dod	0
T1-Mal	-0.03	T4-Mal	0.01	T3-Mal	0	T4-Fum	-0.01
T1-Glut	-0.05	T4-Glut	-0.01	T3-Glut	-0.03	T4-Mal	0.01
T1-Pim	-0.05	T4-Pim	-0.01	T3-Pim	-0.02	T4-Glut	-0.01
T1-Aze	-0.05	T4-Aze	0	T3-Aze	-0.03	T4-Pim	-0.01
T1-3OHBA	-0.12	T4-3OHBA	-0.04	T3-3OHBA	-0.05	T4-Aze	0
T1-4OHBA	-0.11	T4-4OHBA	-0.04	T3-4OHBA	-0.07	T4-3OHBA	-0.04
T1-3AminoBA	0.08	T4-3AminoBA	0.15	T3-3AminoBA	0.11	T4-4OHBA	-0.04
T1-4AminoBA	0.07	T4-4AminoBA	0.15	T3-4AminoBA	0.13	T4-3AminoBA	0.15
T1-3NitroBA	-0.14	T4-3NitroBA	-0.04	T3-3NitroBA	-0.05	T4-4AminoBA	0.15
T1-4NitroBA	-0.11	T4-4NitroBA	-0.02	T3-4NitroBA	-0.05	T4-3NitroBA	-0.04
T1-BA	-0.14	T4-BA	-0.05	T3-BA	-0.07	T4-4NitroBA	-0.02
T1-4-IodoBA	-0.13	T4-4-IodoBA	-0.05	T3-4-IodoBA	-0.06	T4-BA	-0.05
T1-4-bromoBA	-0.12	T4-4-bromoBA	-0.03	T3-4-bromoBA	-0.03	T4-4-IodoBA	-0.05
T1-PentaBA	-0.1	T4-PentaBA	-0.02	T3-PentaBA	-0.04	T4-4-bromoBA	-0.03
T5-Suc	0.01	T6-Suc	0.01	T7-Suc	0.06	T8-Suc	0.02
T5-Adipic	-0.01	T6-Adipic	0.01	T7-Adipic	0.06	T8-Adipic	0.01
T5-Seb	0	T6-Seb	0.01	T7-Seb	0.04	T8-Seb	-0.01
T5-Sub	0	T6-Sub	0.02	T7-Sub	0.05	T8-Sub	0.01
T5-Dod	0.01	T6-Dod	0.02	T7-Dod	0.02	T8-Dod	-0.01
T5-Fum	-0.02	T6-Fum	0	T7-Fum	0.05	T8-Fum	0.07
T5-Mal	0.01	T6-Mal	0.01	T7-Mal	0.05	T8-Mal	0.03
T5-Glut	-0.01	T6-Glut	0	T7-Glut	0.05	T8-Glut	0.03
T5-Pim	0	T6-Pim	0.01	T7-Pim	0.06	T8-Pim	0.02
T5-Aze	0	T6-Aze	0.01	T7-Aze	0.05	T8-Aze	0.02
T5-3OHBA	-0.06	T6-3OHBA	-0.03	T7-3OHBA	0.01	T8-3OHBA	0.05
T5-4OHBA	-0.06	T6-4OHBA	-0.03	T7-4OHBA	0.02	T8-4OHBA	0.06
T5-3AminoBA	0.14	T6-3AminoBA	0.26	T7-3AminoBA	0.22	T8-3AminoBA	0.24
T5-4AminoBA	0.15	T6-4AminoBA	0.17	T7-4AminoBA	0.24	T8-4AminoBA	0.25
T5-3NitroBA	-0.02	T6-3NitroBA	0	T7-3NitroBA	0.01	T8-3NitroBA	0.05
T5-4NitroBA	-0.02	T6-4NitroBA	0	T7-4NitroBA	0.01	T8-4NitroBA	0.05
T5-BA	-0.06	T6-BA	-0.03	T7-BA	0.03	T8-BA	0.01
T5-4-IodoBA	-0.06	T6-4-IodoBA	-0.04	T7-4-IodoBA	0.07	T8-4-IodoBA	0.09
T5-4-bromoBA	-0.02	T6-4-bromoBA	-0.02	T7-4-bromoBA	0.09	T8-4-bromoBA	0.1
T5-PentaBA	-0.02	T6-PentaBA	0.01	T7-PentaBA	0.07	T8-PentaBA	0.1
T9-Suc	0.04	T10-Suc	0.01	T11-Suc	0.01	T12-Suc	0.04
T9-Adipic	0	T10-Adipic	0.05	T11-Adipic	0.04	T12-Adipic	0.05
T9-Seb	0.01	T10-Seb	0.05	T11-Seb	0.05	T12-Seb	0.02

T9-Sub	0.05	**T10-Sub**	0.05	**T11-Sub**	0.05	**T12-Sub**	0.05
T9-Dod	0.05	**T10-Dod**	0.06	**T11-Dod**	0.05	**T12-Dod**	0.02
T9-Fum	0.04	**T10-Fum**	-0.01	**T11-Fum**	-0.04	**T12-Fum**	-0.01
T9-Mal	-0.04	**T10-Mal**	0.05	**T11-Mal**	0.01	**T12-Mal**	0.05
T9-Glut	0.04	**T10-Glut**	0.04	**T11-Glut**	0	**T12-Glut**	0.01
T9-Pim	0	**T10-Pim**	0.05	**T11-Pim**	0.04	**T12-Pim**	0.02
T9-Aze	0.05	**T10-Aze**	0.01	**T11-Aze**	0.05	**T12-Aze**	0.05
T9-3OHBA	-0.06	**T10-3OHBA**	-0.05	**T11-3OHBA**	-0.05	**T12-3OHBA**	-0.04
T9-4OHBA	-0.03	**T10-4OHBA**	-0.02	**T11-4OHBA**	-0.05	**T12-4OHBA**	-0.04
T9-3AminoBA	0.13	**T10-3AminoBA**	0.15	**T11-3AminoBA**	0.12	**T12-3AminoBA**	0.14
T9-4AminoBA	0.14	**T10-4AminoBA**	0.15	**T11-4AminoBA**	0.13	**T12-4AminoBA**	0.15
T9-3NitroBA	-0.07	**T10-3NitroBA**	-0.04	**T11-3NitroBA**	-0.03	**T12-3NitroBA**	-0.03
T9-4NitroBA	-0.03	**T10-4NitroBA**	-0.07	**T11-4NitroBA**	-0.03	**T12-4NitroBA**	-0.07
T9-BA	-0.06	**T10-BA**	-0.05	**T11-BA**	-0.07	**T12-BA**	-0.04
T9-4-IodoBA	-0.05	**T10-4-IodoBA**	-0.01	**T11-4-IodoBA**	-0.02	**T12-4-IodoBA**	-0.01
T9-4-bromoBA	-0.03	**T10-4-bromoBA**	-0.02	**T11-4-bromoBA**	-0.02	**T12-4-bromoBA**	-0.02
T9-PentaBA	0.01	**T10-PentaBA**	0.02	**T11-PentaBA**	0.01	**T12-PentaBA**	0.02

Table G.5.2 Comparison of hydrogen-bond propensity MC score for **T1-T8** with aliphatic and aromatic acid co-crystals with the actual outcome of the experimental co-crystal screening (check mark means co-crystal formation and a cross means no co-crystal formation). The green box shows true positive, the blue box shows true negatives, the red box shows false positives and yellow box shows false negatives.

Group 1							
MC score	**MC Score**	**experimental**	**match?**	**MC score**	**MC Score**	**experimental**	**match ?**
T1-Suc	-0.06	✓	🟧	**T4-Suc**	0	✗	🟦
T1-Adipic	-0.05	✗	🟦	**T4-Adipic**	0	✗	🟦
T1-Seb	-0.04	✗	🟦	**T4-Seb**	0	✗	🟦
T1-Sub	-0.05	✗	🟦	**T4-Sub**	-0.01	✗	🟦
T1-Dod	-0.04	✗	🟦	**T4-Dod**	0	✗	🟦
T1-Fum	-0.05	✗	🟦	**T4-Fum**	-0.01	✗	🟦
T1-Mal	-0.03	✗	🟦	**T4-Mal**	0.01	✗	🟥
T1-Glut	-0.05	✗	🟦	**T4-Glut**	-0.01	✗	🟦
T1-Pim	-0.05	✗	🟦	**T4-Pim**	-0.01	✗	🟦
T1-Aze	-0.05	✗	🟦	**T4-Aze**	0	✗	🟦
T1-3OHBA	-0.12	✓	🟧	**T4-3OHBA**	-0.04	✓	🟧
T1-4OHBA	-0.11	✓	🟧	**T4-4OHBA**	-0.04	✗	🟦
T1-3AminoBA	0.08	✓	🟩	**T4-3AminoBA**	0.15	✓	🟩
T1-4AminoBA	0.07	✓	🟩	**T4-4AminoBA**	0.15	✗	🟥
T1-3NitroBA	-0.14	✓	🟧	**T4-3NitroBA**	-0.04	✓	🟧
T1-4NitroBA	-0.11	✓	🟧	**T4-4NitroBA**	-0.02	✗	🟦
T1-BA	-0.14	✗	🟦	**T4-BA**	-0.05	✓	🟧
T1-4-IodoBA	-0.13	✗	🟦	**T4-4-IodoBA**	-0.05	✗	🟦
T1-4-bromoBA	-0.12	✗	🟦	**T4-4-bromoBA**	-0.03	✗	🟦
T1-PentaBA	-0.1	✓	🟧	**T4-PentaBA**	-0.02	✓	🟧
T2-Suc	-0.04	✗	🟦	**T7-Suc**	0.06	✓	🟩
T2-Adipic	-0.03	✗	🟦	**T7-Adipic**	0.06	✓	🟩

T2-Seb	-0.03	✘	🟦	**T7-Seb**	0.04	✓	🟩
T2-Sub	-0.03	✘	🟦	**T7-Sub**	0.05	✓	🟩
T2-Dod	-0.03	✘	🟦	**T7-Dod**	0.02	✓	🟩
T2-Fum	-0.06	✓	🟧	**T7-Fum**	0.05	✓	🟩
T2-Mal	-0.01	✓	🟧	**T7-Mal**	0.05	✓	🟩
T2-Glut	-0.03	✘	🟦	**T7-Glut**	0.05	✓	🟩
T2-Pim	-0.03	✘	🟦	**T7-Pim**	0.06	✓	🟩
T2-Aze	-0.04	✘	🟦	**T7-Aze**	0.05	✘	🟥
T2-3OHBA	-0.1	✓	🟧	**T7-3OHBA**	0.01	✓	🟩
T2-4OHBA	-0.08	✓	🟧	**T7-4OHBA**	0.02	✓	🟩
T2-3AminoBA	0.08	✓	🟩	**T7-3AminoBA**	0.22	✓	🟩
T2-4AminoBA	0.12	✓	🟩	**T7-4AminoBA**	0.24	✓	🟩
T2-3NitroBA	-0.12	✓	🟧	**T7-3NitroBA**	0.01	✓	🟩
T2-4NitroBA	-0.02	✘	🟦	**T7-4NitroBA**	0.01	✓	🟩
T2-BA	-0.18	✘	🟦	**T7-BA**	0.03	✓	🟩
T2-4-IodoBA	-0.11	✘	🟦	**T7-4-IodoBA**	0.07	✓	🟩
T2-4-bromoBA	-0.11	✘	🟦	**T7-4-bromoBA**	0.09	✓	🟩
T2-PentaBA	-0.09	✓	🟧	**T7-PentaBA**	0.07	✓	🟩
T3-Suc	-0.03	✘	🟦	**T8-Suc**	0.02	✓	🟩
T3-Adipic	-0.03	✘	🟦	**T8-Adipic**	0.01	✓	🟩
T3-Seb	-0.02	✘	🟦	**T8-Seb**	-0.01	✓	🟧
T3-Sub	-0.03	✘	🟦	**T8-Sub**	0.01	✓	🟩
T3-Dod	-0.03	✘	🟦	**T8-Dod**	-0.01	✓	🟧
T3-Fum	-0.03	✘	🟦	**T8-Fum**	0.07	✓	🟩
T3-Mal	0	✘	🟦	**T8-Mal**	0.03	✓	🟩
T3-Glut	-0.03	✘	🟦	**T8-Glut**	0.03	✓	🟩
T3-Pim	-0.02	✘	🟦	**T8-Pim**	0.02	✓	🟩
T3-Aze	-0.03	✘	🟦	**T8-Aze**	0.02	✓	🟩
T3-3OHBA	-0.05	✓	🟧	**T8-3OHBA**	0.05	✓	🟩
T3-4OHBA	-0.07	✘	🟦	**T8-4OHBA**	0.06	✓	🟩
T3-3AminoBA	0.11	✓	🟩	**T8-3AminoBA**	0.24	✓	🟩
T3-4AminoBA	0.13	✘	🟥	**T8-4AminoBA**	0.25	✓	🟩
T3-3NitroBA	-0.05	✓	🟧	**T8-3NitroBA**	0.05	✓	🟩
T3-4NitroBA	-0.05	✘	🟦	**T8-4NitroBA**	0.05	✓	🟩
T3-BA	-0.07	✘	🟦	**T8-BA**	0.01	✓	🟩
T3-4-IodoBA	-0.06	✘	🟦	**T8-4-IodoBA**	0.09	✘	🟥
T3-4-bromoBA	-0.03	✘	🟦	**T8-4-bromoBA**	0.1	✘	🟥
T3-PentaBA	-0.04	✓	🟧	**T8-PentaBA**	0.1	✓	🟩
T5-Suc	0.01	✘	🟥	**T6-Suc**	0.01	✘	🟥
T5-Adipic	-0.01	✘	🟦	**T6-Adipic**	0.01	✓	🟩
T5-Seb	0	✘	🟦	**T6-Seb**	0.01	✓	🟩
T5-Sub	0	✘	🟦	**T6-Sub**	0.02	✘	🟥
T5-Dod	0.01	✘	🟥	**T6-Dod**	0.02	✘	🟥
T5-Fum	-0.02	✘	🟦	**T6-Fum**	0	✓	🟧
T5-Mal	0.01	✓	🟩	**T6-Mal**	0.01	✓	🟩
T5-Glut	-0.01	✘	🟦	**T6-Glut**	0	✓	🟧
T5-Pim	0	✘	🟦	**T6-Pim**	0.01	✘	🟥
T5-Aze	0	✘	🟦	**T6-Aze**	0.01	✘	🟥
T5-3OHBA	-0.06	✓	🟧	**T6-3OHBA**	-0.03	✓	🟧
T5-4OHBA	-0.06	✘	🟩	**T6-4OHBA**	-0.03	✓	🟧
T5-3AminoBA	0.14	✓	🟩	**T6-3AminoBA**	0.26	✓	🟩

	MC Score	experimental	match?		MC Score	experimental	match?
T5-4AminoBA	0.15	✓	yellow	T6-4AminoBA	0.17	✗	red
T5-3NitroBA	-0.02	✓	yellow	T6-3NitroBA	0	✓	yellow
T5-4NitroBA	-0.02	✗	red	T6-4NitroBA	0	✗	blue
T5-BA	-0.06	✗	blue	T6-BA	-0.03	✓	yellow
T5-4-IodoBA	-0.06	✗	red	T6-4-IodoBA	-0.04	✗	blue
T5-4-bromoBA	-0.02	✗	red	T6-4-bromoBA	-0.02	✗	blue
T5-PentaBA	-0.02	✓	yellow	T6-PentaBA	0.01	✓	green

Table G.5.3 Comparison of hydrogen-bond propensity MC score for **T9-T12** with aliphatic and aromatic acid co-crystals with the actual outcome of the experimental co-crystal screening (check mark means co-crystal formation and a cross means no co-crystal formation). The green box shows true positive, the blue box shows true negatives, the red box shows false positives and yellow box shows false negatives.

Group 2							
MC score	MC Score	experimental	match?	MC score	MC Score	experimental	match ?
T9-Suc	0.04	✓	green	T10-Suc	0.01	✓	green
T9-Adipic	0	✓	green	T10-Adipic	0.05	✓	green
T9-Seb	0.01	✓	green	T10-Seb	0.05	✓	green
T9-Sub	0.05	✓	green	T10-Sub	0.05	✓	green
T9-Dod	0.05	✓	green	T10-Dod	0.06	✓	green
T9-Fum	0.04	✓	green	T10-Fum	-0.01	✓	yellow
T9-Mal	-0.04	✓	yellow	T10-Mal	0.05	✓	green
T9-Glut	0.04	✓	green	T10-Glut	0.04	✓	green
T9-Pim	0	✓	green	T10-Pim	0.05	✓	green
T9-Aze	0.05	✓	green	T10-Aze	0.01	✓	green
T9-3OHBA	-0.06	✓	yellow	T10-3OHBA	-0.05	✓	yellow
T9-4OHBA	-0.03	✓	yellow	T10-4OHBA	-0.02	✓	yellow
T9-3AminoBA	0.13	✓	green	T10-3AminoBA	0.15	✓	green
T9-4AminoBA	0.14	✓	green	T10-4AminoBA	0.15	✓	green
T9-3NitroBA	-0.07	✓	yellow	T10-3NitroBA	-0.04	✓	yellow
T9-4NitroBA	-0.03	✓	yellow	T10-4NitroBA	-0.07	✓	yellow
T9-BA	-0.06	✓	yellow	T10-BA	-0.05	✓	yellow
T9-4-IodoBA	-0.05	✗	blue	T10-4-IodoBA	-0.01	✗	blue
T9-4-bromoBA	-0.03	✓	yellow	T10-4-bromoBA	-0.02	✗	blue
T9-PentaBA	0.01	✓	green	T10-PentaBA	0.02	✓	green
T1-Suc	0.01	✓	green	T12-Suc	0.04	✓	green
T12-Adipic	0.04	✓	green	T12-Adipic	0.05	✓	green
T12-Seb	0.05	✓	green	T12-Seb	0.02	✓	green
T12-Sub	0.05	✓	green	T12-Sub	0.05	✓	green
T12-Dod	0.05	✓	green	T12-Dod	0.02	✓	green
T12-Fum	-0.04	✓	yellow	T12-Fum	-0.01	✓	yellow
T12-Mal	0.01	✓	green	T12-Mal	0.05	✓	green
T12-Glut	0	✓	green	T12-Glut	0.01	✓	green
T12-Pim	0.04	✓	green	T12-Pim	0.02	✓	green
T12-Aze	0.05	✓	green	T12-Aze	0.05	✓	green
T12-3OHBA	-0.05	✓	yellow	T12-3OHBA	-0.04	✓	yellow
T12-4OHBA	-0.05	✓	yellow	T12-4OHBA	-0.04	✓	yellow

T12-3AminoBA	0.12	✓	🟩	**T12-3AminoBA**	0.14	✓	🟩
T12-4AminoBA	0.13	✓	🟩	**T12-4AminoBA**	0.15	✓	🟩
T12-3NitroBA	-0.03	✓	🟧	**T12-3NitroBA**	-0.03	✓	🟧
T12-4NitroBA	-0.03	✓	🟧	**T12-4NitroBA**	-0.07	✓	🟧
T12-BA	-0.07	✓	🟧	**T12-BA**	-0.04	✓	🟧
T12-4-IodoBA	-0.02	✗	🟦	**T12-4-IodoBA**	-0.01	✗	🟦
T12-4-bromoBA	-0.02	✗	🟦	**T12-4-bromoBA**	-0.02	✗	🟦
T12-PentaBA	0.01	✓	🟩	**T12-PentaBA**	0.02	✓	🟩