

# CrystEngComm

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the RSC Publishing peer review process and has been accepted for publication.

*Accepted Manuscripts* are published online shortly after acceptance, which is prior to technical editing, formatting and proof reading. This free service from RSC Publishing allows authors to make their results available to the community, in citable form, before publication of the edited article. This *Accepted Manuscript* will be replaced by the edited and formatted *Advance Article* as soon as this is available.

To cite this manuscript please use its permanent Digital Object Identifier (DOI®), which is identical for all formats of publication.

More information about *Accepted Manuscripts* can be found in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics contained in the manuscript submitted by the author(s) which may alter content, and that the standard [Terms & Conditions](#) and the [ethical guidelines](#) that apply to the journal are still applicable. In no event shall the RSC be held responsible for any errors or omissions in these *Accepted Manuscript* manuscripts or any consequences arising from the use of any information contained in them.

# Synthon transferability probed with IR spectroscopy: cytosine salts as models for salts of lamivudine

Shaunak Chakraborty, Somnath Ganguly and Gautam R. Desiraju\*

Solid State and Structural Chemistry Unit, Indian Institute of Science, Bangalore 560 012, India

E-mail: desiraju@sscu.iisc.ernet.in

Fax: +91 80 23602306

## Abstract

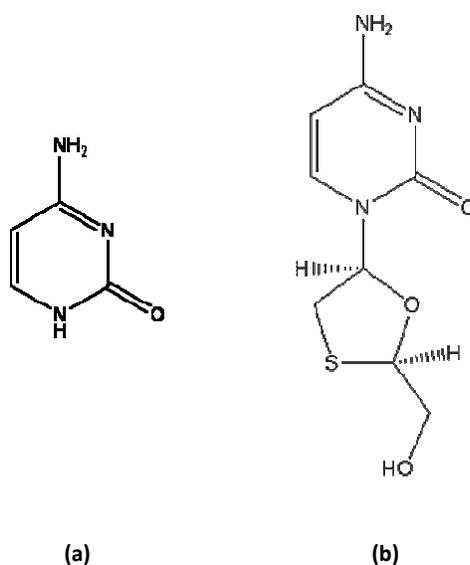
Co-crystal screening of the anti-HIV drug lamivudine was carried out with dicarboxylic acids as co-formers, and three of the resulting crystalline solids, two salts and a co-crystal, were studied with SCXRD, PXRD and FTIR spectroscopy. Salts of cytosine, a molecule that incorporates critical structural features of lamivudine, with the same co-formers, were taken as model systems for IR spectroscopic studies of the synthons in the salts of lamivudine. It is shown that different systems with the same synthon show very similar spectral signatures in the regions corresponding to the synthon absorptions. This reveals again the modular nature of the supramolecular synthon.

## Introduction

Co-crystals or multicomponent molecular crystals,<sup>1,2</sup> although known since 1844 when Wöhler described quinhydrone,<sup>3</sup> continue to offer fundamental and practical crystal engineering challenges.<sup>4,5</sup> Much has been said and done to explain why these multicomponent crystals form,<sup>6</sup> but their design and synthesis have considerable therapeutic and commercial benefits as they often lead to improvement of physicochemical properties like solubility, stability, dissolution rate and bioavailability. This is of particular interest in the scenario of the pharmaceutical industry, where co-crystallisation and salt formation of Active Pharmaceutical Ingredients (APIs) may lead to superior properties.<sup>7</sup> The idea of designing the structures of multi-component crystals using retrosynthesis<sup>8</sup> leads to the concept of the heterosynthon<sup>9</sup>— which is a kinetic unit that effectively embodies the primary, or most important, chemical and geometrical recognition between two different molecular components.<sup>10,11</sup> Design and synthesis of new co-crystals has progressed considerably over the years using this retrosynthetic approach.<sup>5,6</sup> However, there is a difference between such logical design of a co-crystal<sup>12</sup> and the actual *isolation* of such a substance in the laboratory. The reasons why a particular co-crystal is or cannot be obtained could lie in kinetics and energetics, with solvent choice and solubility more often than not playing a significant role in the co-crystallisation events. Since the supramolecular synthon is a central theme in a crystal structure, it is important to monitor its formation in multicomponent systems. While SCXRD has been traditionally

employed as the method for studying structures and synthons, obtaining high quality crystals for SCXRD studies can sometimes be difficult. We have previously described a simple, quick and robust IR method for synthon detection that does not require single crystals.<sup>13</sup> The present paper describes a further application of this method.

Lamivudine ( $\beta$ -L-2',3'-dideoxy-3'-thiacytidine) is the most common nucleoside reverse transcriptase inhibitor (NTRI) drug used in anti-HIV therapy. It is known to form co-crystals.<sup>14</sup> Cytosine, which is one of the four main nucleobases in DNA, has been used in this study as a model for lamivudine. In this paper, we use our IR method for synthon detection in related structures. The idea is that when two related systems, like lamivudine and cytosine co-crystals, contain the same synthon, there are unmistakable common signatures of its presence in both spectra, in regions of absorption pertaining to the synthon.



Scheme 1. (a) Cytosine and (b) Lamivudine

A cytosine derivative with the capability of inhibiting the growth of DNA, lamivudine is known to be effective against the Hepatitis-B virus as an NTRI.<sup>15</sup> The salt-forming ability of the drug has been extensively investigated. Of the 18 crystalline phases reported, 12 are salts and co-crystals formed with various co-formers, while the rest are hydrates.<sup>14</sup> Lamivudine exists in two forms in the solid state, one of which (form 1) is a hydrate and the other one is the pure API. The cytosine fragment on lamivudine is a potential site for the formation of hydrogen bonded supramolecular synthons with co-formers possessing sufficiently acidic protons, as evident from the structures of the recently reported lamivudine hydrogen phthalate and the salicylate monohydrate,<sup>14a</sup> lamivudine hydrogen phthalate hemihydrate and lamivudine hydrogen 4,5-dichlorophthalate<sup>14i</sup> and from the slightly older reports of 3, 5-dinitrosalicylate lamivudine monohydrate,<sup>14b</sup> the hydrogen maleate<sup>14c</sup> and the saccharinate salts.<sup>14g</sup> It may be pointed out that all these substances, with the exception of lamivudine saccharinate, form the same synthon as cytosine does, hereafter called synthon **I**, shown in Figure 1. The aim of this study is to assess the transferability of the IR bands

corresponding to synthon **I** from the lamivudine systems to the cytosine systems. If this is possible, structure analysis by IR could be simplified to an examination of the characteristic synthon bands.

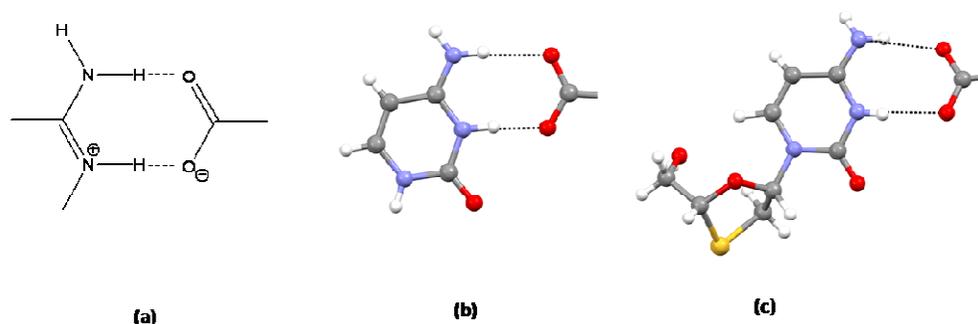
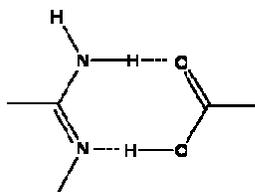


Figure 1. Synthon **I**. (a) Schematic depiction (b) In carboxylates of cytosine (c) In carboxylates of lamivudine. Note the similar synthon environment in (b) and (c).

### Experimental section

Co-crystallisation screening of lamivudine was done with oxalic, malonic, succinic, glutaric, adipic and pimelic acids, and cytosine was screened with trifluoro, trichloro and tribromoacetic acids in addition to with these six. Form 2 of the API was exclusively used in this study, form 1 being a hydrate.<sup>14</sup> We have exclusively used form 2 of the API, obtained from Mylan, for our study. All the dicarboxylic acids were obtained from Spectrochem, while trifluoro, trichloro and tribromoacetic acids were obtained from SD Fine Chemicals. All the chemicals were used as received. We obtained and studied an oxalate (**A**), an oxalate monohydrate (**B**) and a pimelic acid co-crystal (**C**) of lamivudine, along with a tribromoacetate of cytosine (**D**) with single crystal X-ray diffraction (SCXRD), powder X-ray diffraction (PXRD) and IR spectroscopy. **C** is seen to form a different synthon, hereafter called synthon **II**. Three other solid phases, referred to hereafter as **E**, **F** and **G**, obtained from co-crystallisation of lamivudine with succinic, glutaric and adipic acids respectively, were also studied using PXRD and FTIR methods.



Scheme 2. Synthon **II**

Table 1 shows the names of the compounds along with their letter codes for easy reference.

**Table 1:** The compounds and their letter codes

Compound	Letter code
(Lamivudine) <sub>2</sub> oxalate	<b>A</b>
(Lamivudine) <sub>2</sub> (oxalate) <sub>2</sub> hydrate	<b>B</b>
(Lamivudine) <sub>2</sub> . pimelic acid co-crystal	<b>C</b>
(Cytosine) <sub>2</sub> (tribromoacetate)	<b>D</b>
Lamivudine-succinic acid co-crystal	<b>E</b>
Lamivudine-glutaric acid co-crystal	<b>F</b>
Lamivudine-adipic acid co-crystal	<b>G</b>
Cytosine hydrochloride	<b>P</b>
Lamivudine hydrochloride	<b>Q</b>
Cytosine oxalate hydrate	<b>R</b>

### Sample preparation and crystallisation

- a. **(Lamivudine)<sub>2</sub> oxalate (A):** 0.5 mmol each of lamivudine form 2 and oxalic acid dihydrate were ground together with a few drops of ethanol and dissolved in a variety of solvents. Colourless needle-like crystals were obtained from methanol after three days.
- b. **(Lamivudine)<sub>2</sub>(oxalate)<sub>2</sub>hydrate (B):** 0.5 mmol each of lamivudine form 2 and oxalic acid dihydrate were ground together with a few drops of ethanol and dissolved in 3 ml of distilled water. Colourless block shaped crystals were obtained after two weeks.
- c. **(Lamivudine)<sub>2</sub>. pimelic acid co-crystal (C):** 0.5 mmol each of lamivudine form 2 and pimelic acid were ground together with a few drops of ethanol and dissolved in a number of different solvents. Colourless block shaped crystals grew from nitromethane after around five days.
- d. **(Cytosine)<sub>2</sub>(tribromoacetate) (D):** 0.5 mmol each of cytosine and tribromoacetic acid were ground together with a few drops of ethanol and dissolved in a number of solvents. Colourless needle shaped crystals grew from methanol over a period of a week.

- e. **Lamivudine-succinic acid co-crystal (E):** 0.5 mmol each of lamivudine form 2 and succinic acid were ground together with a few drops of ethanol and dissolved in a range of solvents. All of them yielded white fibres of the new solid form after complete evaporation.
- f. **Lamivudine-glutaric acid co-crystal (F):** 0.5 mmol each of lamivudine form 2 and glutaric acid were ground together with a few drops of ethanol and dissolved in a number of different solvents for crystallisation. A white powder was obtained after complete evaporation of the solvent in each case.
- g. **Lamivudine-adipic acid co-crystal (G):** 0.5 mmol each of form 2 of the API and adipic acid were ground together, again with a few drops of ethanol, and dissolved for crystallisation in a variety of solvents. A white powder was obtained after the solvents evaporated from each vial.

The hydrochlorides of cytosine<sup>16</sup> and lamivudine<sup>17</sup> (**P** and **Q** respectively) and the oxalate of cytosine (**R**) were prepared according to literature procedures.<sup>18</sup>

**Powder X-ray Diffraction:** All PXRD data were collected on a Philips X'Pert powder diffractometer equipped with an X'cellerator detector within the  $2\theta$  scan range 5-45 degrees.

**FTIR spectroscopy:** All IR spectra were collected on a Bruker alpha FTIR spectrometer, from 4000 to 500  $\text{cm}^{-1}$ , at a resolution of  $\pm 2 \text{ cm}^{-1}$  using an ATR attachment.

**Differential Scanning Calorimetry:** All differential scanning calorimetric measurements were done on a Mettler Toledo DSC instrument, using aluminium pans at a heating rate of  $3^\circ\text{C}/\text{min}$ .

**Single Crystal X-ray Diffraction:** SCXRD data for **A**, **C** and **D** were collected on a Rigaku Mercury375R/M CCD (XtaLAB mini) diffractometer using graphite monochromated  $\text{Mo K}\alpha$  radiation, equipped with a low temperature gas spray cooler. The data were processed with Rigaku CrystalClear 2.0.<sup>19</sup> SCXRD data for **B** were collected at room temperature on a Bruker D8 QUEST diffractometer equipped with an Oxford  $\text{N}_2$  open flow cryostat. The crystal to detector distance was fixed at 50 mm with a scan width of  $0.5^\circ$  per frame. Cell refinements, data integration and data reduction were performed using SAINTPLUS<sup>20</sup> implemented in the Apex2 suite of programs. The structure solution and refinements were performed using SHELX97<sup>21</sup> using the WinGX<sup>22</sup> suite. The ORTEP<sup>23</sup> diagrams are given in the supplementary information. Some hydrogen atoms were assigned using a riding model, and the rest were located on the difference Fourier map. The structural data are available at CCDC numbers 965966-965969. The structural data and refinement parameters are given in table 1. Proton transfer and formation of synthon **I** in **A** were detected from the difference Fourier map. We also checked the C–O bond length on the oxalate in both **A** and **B** and found them to be equal confirming proton transfer. Formation of synthon **II** was confirmed likewise from the difference Fourier map and by checking that the C–O bond lengths on the pimelic acid molecule are unequal.

Table 2: Crystallographic data and structural refinement parameters

Compound	A	B	C	D
Structural formula	$(C_8H_{12}N_3O_3S)_2(C_2O_4)$	$(C_8N_3O_3S)_2(C_2O_4)_2O$	$(C_8H_{11}N_3O_3S)_2(C_7H_{12}O_4)$	$(C_4N_3O)_2(C_2Br_3O_2)$
Formula weight (g/mol)	548.57	628.37	618.70	507.86
Crystal system	orthorhombic	monoclinic	monoclinic	triclinic
Space group	$P2_12_12_1$	$P2$	$P2_1$	$P1$
T(K)	150	298	150	150
a (Å)	6.7021(15)	12.8936(11)	6.9440(7)	6.0904(15)
b (Å)	9.134(2)	6.6821(5)	28.891(3)	10.3943(15)
c (Å)	37.734(10)	17.1763(13)	6.9559(8)	13.5423(19)
$\alpha$ (°)	90	90	90	73.19(5)
$\beta$ (°)	90	108.031(2)	95.357(4)	79.70(5)
$\gamma$ (°)	90	90	90	78.31(6)
V(Å <sup>3</sup> )	2309.9(10)	1407.17(19)	1389.4(3)	797.0(3)
Z	4	2	2	2
$\rho_{calc}$ (g/cm <sup>-3</sup> )	1.577	1.483	1.479	2.116
$\theta$ range for data collection	1.08, 27.49	3.17, 27.63	1.41, 27.50	1.58, 27.43
F(000)	1140	684	652	570
$\mu$ (mm <sup>-1</sup> )	0.299	0.270	0.258	7.619
$R_1$ [ $I > 2\sigma(I)$ ]	0.0588	0.1174	0.0412	0.0703
wR <sub>2</sub>	0.1465	0.3468	0.1176	0.1747
Goodness of fit	1.092	1.687	1.105	1.189
Reflections collected	5279	6484	6245	3593
Unique reflections	4720	5629	3233	2796
Observed reflections	20786	42195	13530	8346
$h_{max}$	8	16	9	7
$h_{min}$	-8	-16	-8	-7
$k_{max}$	11	8	37	13
$k_{min}$	-11	-8	-37	-13
$l_{max}$	49	22	9	17
$l_{min}$	-49	-22	-9	-17
Flack x parameter	0.01(11)	0.2(2)	0.01(2)	-
CCDC no.	965969	965967	965968	965966

## Results and discussion

The cytosine fragment of lamivudine is involved in the formation of synthon **I** in salts of lamivudine with carboxylic acids.<sup>15</sup> Accordingly it may be possible to pick out common features in the IR spectra of salts of lamivudine and salts of cytosine itself. The cytosine salt is a model for the lamivudine salt, and is expected to contain the same synthon **I**. Cytosine is indeed reported to form synthon **I** in its salt with trichloroacetic acid.<sup>24</sup>

The initial part of the FTIR spectroscopic investigation is a study of compounds **A**, **B**, **P** and **Q**. Changes in the N–H/O–H and C=O stretching modes of the salts (protonated lamivudine and cytosine) were studied against the parent cytosine and lamivudine spectra. These observations, along with changes in ring vibrations due to protonation of the cytosine fragments in both systems, were used to interpret synthon structures. Next, **C**, **E**, **F** and **G** were monitored using IR and the changes occurring in the N–H/ O–H stretching regions and the C=O stretching absorption were examined. The spectra of **E**, **F** and **G** were compared with the results from co-crystal **C**.

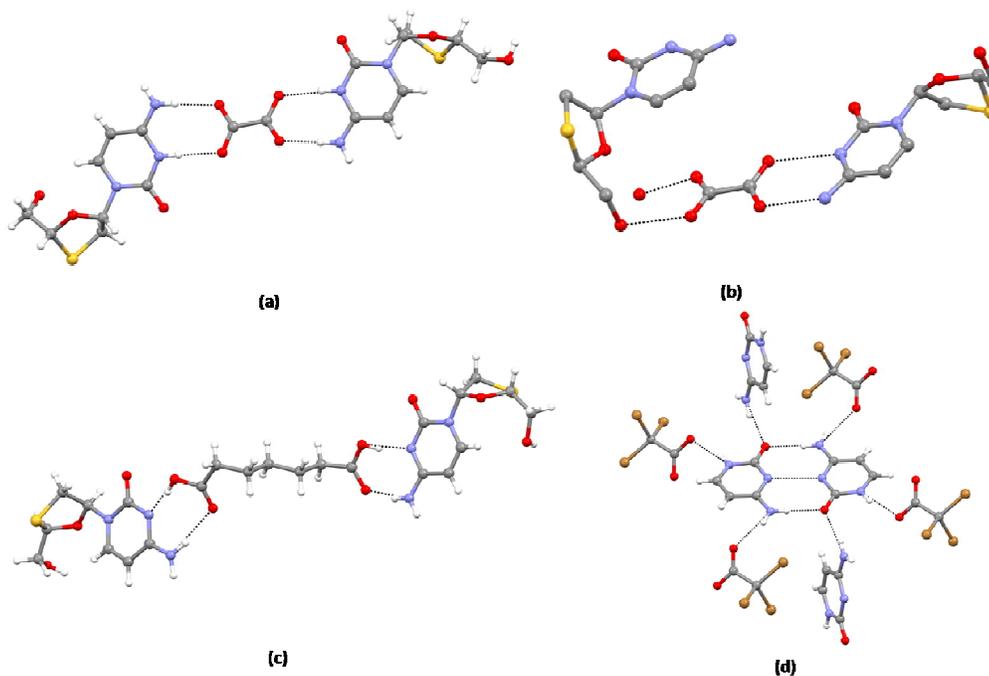


Figure 2. (a) Lamivudine oxalate (**A**), (b) lamivudine oxalate monohydrate (**B**), (c) lamivudine-pimelic acid co-crystal (**C**) and (d) cytosine-tribromoacetate (**D**). The hydrogen atoms could not be located in **B** because of solvent disorder.

Salt **A** crystallises in the space group  $P2_12_12_1$  and has two molecules of the protonated API in the asymmetric unit, along with a fully deprotonated oxalate anion. It is perhaps not very surprising that lamivudine should form a salt with oxalic acid, since it has a rather low  $pK_a$  of 1.25. Synthons **I** is seen on both sides of the oxalate dianion, and the structure extends in the third dimension by means of strong N–H $\cdots$ O and O–H $\cdots$ O interactions. There is also a short S $\cdots$ S contact (3.543 Å).

Salt hydrate **B** crystallises in the space group  $P2$  and has two molecules of protonated lamivudine, two oxalate dianions and a molecule of water in the asymmetric unit. Synthons **I** is again the main synthon, but this time on only one side of the oxalate dianion. The other side is involved in O–H $\cdots$ O hydrogen bonds with the water molecules and the O–H group on the oxathiolane ring of another lamivudine molecule.

It is rather surprising that tribromoacetic acid, with its high acidity, does not form the expected synthon **I** with cytosine. **D** crystallises in the space group  $P1$ , with one tribromoacetate anion and two molecules of cytosine in the asymmetric unit. Two molecules of cytosine form the centrosymmetric dimer commonly seen in many structures involving the nucleobase, and seemingly robust N–H $\cdots$ O interactions with tribromoacetate anions and other cytosine dimer units extend the structure in three dimensions.

Co-crystal **C** crystallises in the space group  $P2_1$  and contains two molecules of lamivudine and a molecule of pimelic acid in the asymmetric unit, with two molecules of lamivudine attached to both sides of the pimelic acid with synthon **II** formed of robust N–H $\cdots$ O hydrogen bonds. This unit repeats by a  $2_1$  screw operation along the  $b$  direction. The structure grows in the third dimension with the help of N–H $\cdots$ O, O–H $\cdots$ O hydrogen bonds and notably an O $\cdots$ S contact (3.302 Å). Pimelic acid is not deprotonated possibly because of its relatively high  $pK_a$  ( $pK_a=4.48$ ). Diagrams of the four structures discussed above are shown in Figure 2.

### IR spectroscopy of lamivudine salts

FTIR spectroscopy has been used extensively in the study of pharmaceutical solids<sup>25</sup> and polymorphism (known to have an effect on biological activity).<sup>26-28</sup> Our strategy in the FTIR study lies in the identification of changes that arise during protonation of cytosine, and thereafter identification and mapping of the same bands in protonated lamivudine. We have accordingly studied the IR spectra of pure cytosine, **A**, **B**, **P**, **Q** and **R**. The latter part of the FTIR study of synthon **I** identifies bands corresponding to the carboxylate ion in the presence of protonated cytosine and lamivudine.

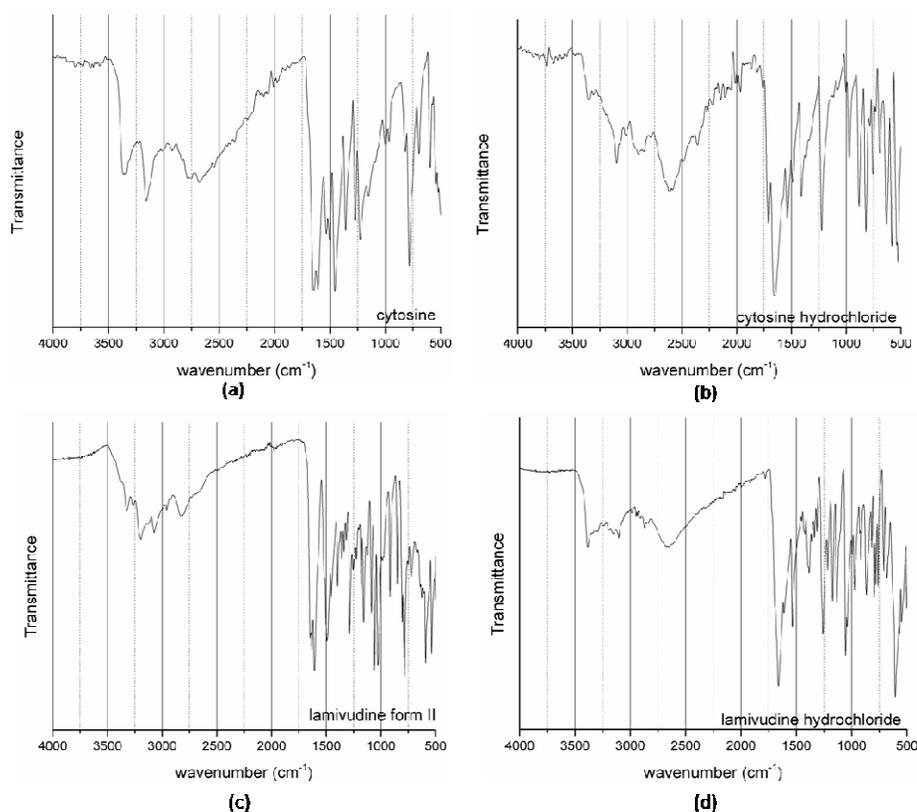


Figure 3. FTIR spectra of (a) cytosine, (b) cytosine hydrochloride (**P**), (c) lamivudine and (d) lamivudine hydrochloride (**Q**)

The spectrum of cytosine shows the following absorptions: The asymmetric N–H stretch appears at  $3390\text{ cm}^{-1}$  and the symmetric stretch occurs at  $3180\text{ cm}^{-1}$ . The N–H stretching vibration is seen at about  $3010\text{ cm}^{-1}$ . The C=O stretching mode group is seen at  $1667\text{ cm}^{-1}$  and a ring stretch is seen at  $1540\text{ cm}^{-1}$ . Ring modes appear at  $1469$  and  $1280\text{ cm}^{-1}$  respectively. The band at  $788\text{ cm}^{-1}$  is attributed to the ring breathing mode.

Lamivudine is a polymorphic solid which crystallises in two solid forms, one orthorhombic and the other tetragonal.<sup>15</sup> Strictly speaking, form 1 is a hydrate. The molecular structure of the API belongs to the  $C_1$  point group symmetry with all the vibrational modes active in IR and Raman. The C–H stretching bands are observed between  $3062$  and  $2936\text{ cm}^{-1}$ , while the weak bands between  $3400$  and  $3200\text{ cm}^{-1}$  correspond to the N–H/O–H stretch. The carbonyl stretching mode appears at  $1652\text{ cm}^{-1}$ , the C=N stretch at  $1636\text{ cm}^{-1}$  and the C–O at  $1060\text{ cm}^{-1}$ .<sup>29</sup>

The IR spectrum of protonated cytosine has been discussed by Floria et al.<sup>30</sup> and by Slosarek and Zamboni.<sup>31</sup> The spectrum of **P**<sup>32</sup> is dominated by bands at 3350 and 3180  $\text{cm}^{-1}$  due to the asymmetric and symmetric  $\text{NH}_2$  stretching vibrations, the former moving down on protonation. The N–H stretch appears at 3020  $\text{cm}^{-1}$ . The  $\text{C}(\text{sp}^3)\text{--H}$  stretching vibration bands are observed in the range 2800–3000  $\text{cm}^{-1}$ . Strong vibrations seen at 1720, 1710 and 1670  $\text{cm}^{-1}$  in the protonated cytosine are due to C=O, C=C and C=N stretches respectively.<sup>33</sup> Ring vibrations are seen at 1414 and 1260  $\text{cm}^{-1}$  in the spectrum, as well as at 1000  $\text{cm}^{-1}$  and at around 785  $\text{cm}^{-1}$ . All the values are lower than the corresponding vibration in the parent molecule.

The spectrum of **Q**, like that of **P**, (Figure 3) is dominated by bands at 3360 and 3160  $\text{cm}^{-1}$  due to the asymmetric and symmetric  $\text{NH}_2$  stretching vibrations. The N–H stretch appears at 3020  $\text{cm}^{-1}$ . Strong vibrations are seen at 1720, 1660 and 1620  $\text{cm}^{-1}$  due to the C=O, C=C and C=N stretching modes. The C=O stretch of protonated cytosine is known to appear at 1720  $\text{cm}^{-1}$ . Ring vibrations are seen at about 1425, 1220 and 1250  $\text{cm}^{-1}$  in the spectrum of **Q**, much like the equivalent cytosine salt and lower than the pure cytosine vibrations. Ring vibrations are also seen at 1050  $\text{cm}^{-1}$  and 760  $\text{cm}^{-1}$ .

Comparison of the spectra of **P** and **Q** shows remarkable overlap of bands due to the N–H/O–H groups of the molecules in the salts. The C=O region also exhibits striking band structure similarity and a clear overlap in the band positions. The modes of the protonated rings also show considerable overlap. These overlapping features point towards structural similarities between the API salt and cytosine salt systems.

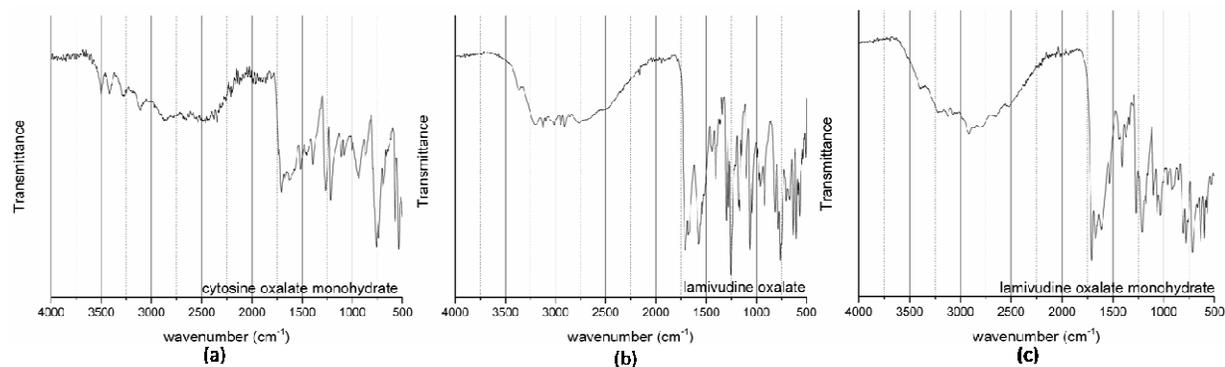


Figure 4. FTIR spectra of (a) cytosine oxalate monohydrate (**R**), (b) lamivudine oxalate (**A**) and (c) lamivudine oxalate monohydrate (**B**)

Having identified the effects of protonation of the cytosine ring fragment, we proceeded to analyse the effects of the presence of a carboxylate in the oxalates of cytosine and lamivudine. Because of the ionic character of **R**, its hydrogen bonds are stronger than those present in cytosine monohydrate and the anhydrous cytosine. The spectrum is expected to show bands due to a protonated ring and ionised carboxylate bands. Figure 4 shows the spectra of **A** and **B** along with that of **R**. In **R**, the N–H stretch region is dominated by the asymmetric  $\text{NH}_2$  stretch at 3415  $\text{cm}^{-1}$  while the symmetric  $\text{NH}_2$  stretch appears at 3210  $\text{cm}^{-1}$ .

The N–H stretch appears at  $3125\text{ cm}^{-1}$ . Strong bands are seen at  $1740$  and  $1710\text{ cm}^{-1}$  due to the free C=O stretch. Carboxylate ion bands<sup>33</sup> are seen at about  $1630$  and at  $1440\text{ cm}^{-1}$ . The band at  $1260\text{ cm}^{-1}$  can be assigned to the protonated ring. A strong band at  $780\text{ cm}^{-1}$  is attributed to the  $\delta_{CCO}$ /ring breathing mode of the protonated cytosine fragment.

The spectrum is expected to show bands due to a protonated ring and ionised carboxylate bands.<sup>33</sup> The N–H stretch region is dominated by the asymmetric  $\text{NH}_2$  stretch at  $3370\text{ cm}^{-1}$  while the symmetric  $\text{NH}_2$  appears at  $3210\text{ cm}^{-1}$ , as seen in the two hydrochlorides discussed above. The N–H stretch appears at  $3120\text{ cm}^{-1}$ . A strong band is seen at  $1710\text{ cm}^{-1}$  due to the free C=O stretch. The absorptions at  $1630\text{ cm}^{-1}$  and  $1440\text{ cm}^{-1}$  are due to the ionised carboxylate. The bands at  $1262\text{ cm}^{-1}$  and  $1050\text{ cm}^{-1}$  and  $760\text{ cm}^{-1}$  can likewise be assigned to the protonated ring. All the bands observed above indicate proton transfer to the nitrogen.

To summarise, **A** and **B** show clear overlap of bands with **R** due to the stretching vibrations of N–H/O–H groups of the molecules in the salts, pointing towards synthon similarity in the API and cytosine salt systems. Also showing overlap are bands due to the C=O stretch and the modes of the rings along with signatures of the presence of the carboxylate anion. Here, again, are features that point towards synthon similarity associated with the API salt and backbone salt systems.

#### Synthon differences monitored with IR spectroscopy

We had hoped that given the very high acidity of tribromoacetic acid ( $\text{pK}_a=0.72$ ), it would form synthon **I** with cytosine. But analysis of the IR spectra led us to suspect that **I** may not be present, only to be proved correct afterwards when the structure was elucidated with SCXRD.

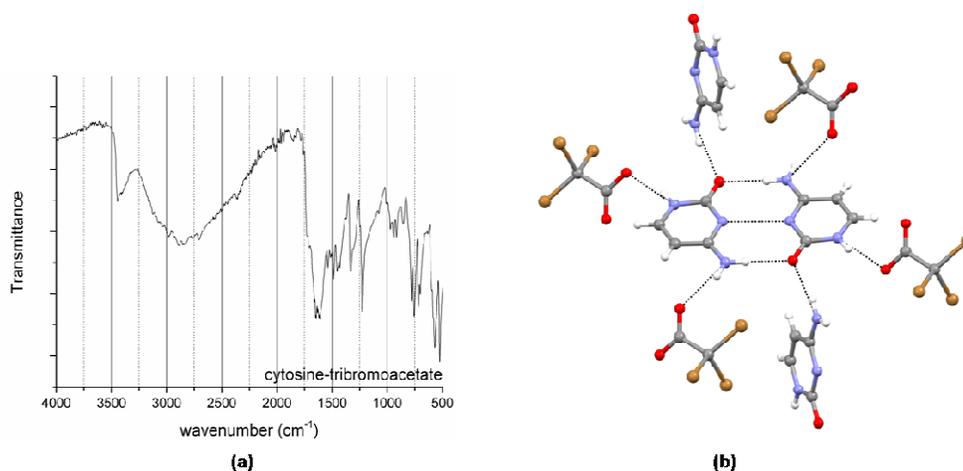


Figure 5. (a) FTIR spectrum of cytosine tribromoacetate (**D**) and (b) diagram of cytosine tribromoacetate (**D**)

A comparison of cytosine salts with differing synthons can be done by comparing **Q** and **D**. The IR spectrum of **D** in the 4000-500  $\text{cm}^{-1}$  region is shown in Figure 5. The spectrum of **Q** is dominated by bands at 3440 and 2900  $\text{cm}^{-1}$  due to the asymmetric and symmetric  $\text{NH}_2$  stretching vibrations. However, in **D**, a medium intensity band at 3440  $\text{cm}^{-1}$  is seen lying on an intense broad band centered at 2900  $\text{cm}^{-1}$ . Also, the C=O stretch appears at lower values, indicating that the carbonyl is strongly H-bonded in this system. The IR bands in this region of cytosine and its salts and co-crystals are known to be due to the O–H/N–H and C=O stretches and the differences in the spectra are a clear illustration of the difference in the synthons in **Q** and **D**.

The above discussion shows that the bonding patterns in **P** and **Q** are equivalent and mapped through a comparison of the IR spectra of the two salts. Comparison of the spectral assignments of the carboxylic acid salts of lamivudine shows clear resemblances, especially in the bands due to the asymmetric and symmetric N–H/O–H stretch, those due to the  $\text{COO}^-$  group as well as the bands due to the protonated lamivudine ring. This draws an interesting parallel between lamivudine and cytosine salts, an indication that they have the same interactions in the crystal structure, evidenced from the SCXRD studies as well. The simple IR method can be advantageously used to elucidate synthon information in systems where single crystals are difficult to obtain. Conversely, IR spectroscopy (confirmed by SCXRD) showed that non-matching synthons give different spectral signatures.

#### IR spectroscopy of lamivudine co-crystals

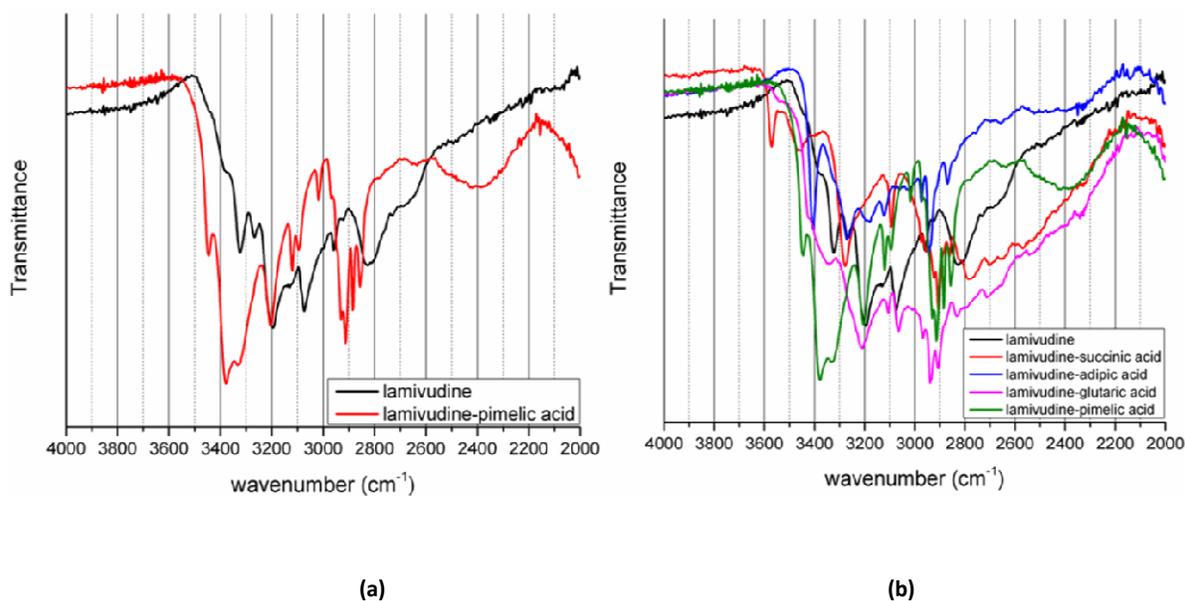


Figure 6. (a) The FTIR spectra of lamivudine and the pimelic acid co-crystal (C), superposed and (b) the FTIR spectra of lamivudine and all the co-crystalline solid phases(C, E, F, G) superposed.

The well known concepts of H-bonding and IR spectroscopy are used to interpret changes in the spectra of lamivudine with dicarboxylic acids.<sup>34-36</sup> Figure 6 shows the spectra in the N–H/O–H stretch region of **C**. The N–H and O–H stretching bands show variations with respect to lamivudine in all of these cases, which may not be possible to explain in a straightforward manner in the absence of crystal structures.

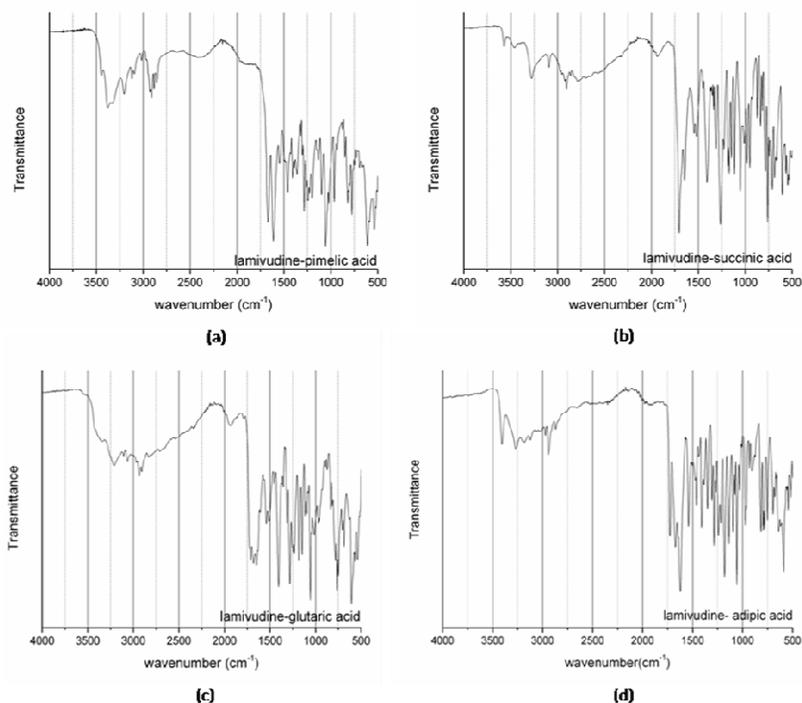


Figure 7. IR spectra of (a) lamivudine–pimelic acid co-crystal (**C**), (b) lamivudine–succinic acid co-crystal (**E**), (c) lamivudine–glutaric acid co-crystal (**F**) and (d) lamivudine–adipic acid co-crystal (**G**).

Figure 7 shows the IR spectra of **C**, **E**, **F** and **G**. In **C**, the N–H stretch bands show remarkable intensity increase at 3400 and 3220  $\text{cm}^{-1}$  indicating hydrogen bonding interactions between the API and co-former. The C–H stretch in lamivudine is a split band at 3130 and 3100  $\text{cm}^{-1}$  which changes to a single broad band at 3080  $\text{cm}^{-1}$  in the co-crystal and a C–H stretching band at 2920 and 2900  $\text{cm}^{-1}$  in lamivudine weakens to a single band at 2910  $\text{cm}^{-1}$  in the co-crystal, all these possibly indicating C–H $\cdots$ O interactions. In **E**, new intense bands are seen at 3460 and 3280  $\text{cm}^{-1}$ , both of which are due to the N–H stretch. No significant changes were noticed for **F** except for some broadening of the N–H stretching bands at 3330 and 3200  $\text{cm}^{-1}$  and an intense and split C–H stretch at 2920–2900  $\text{cm}^{-1}$ . The spectra of **C**, **E**, **F** and **G** show the appearance of a band at 2400  $\text{cm}^{-1}$  due to the acid-pyridine synthon. The spectrum in Figure 5 of **C** in the 2000–500  $\text{cm}^{-1}$  region shows bands at 1950  $\text{cm}^{-1}$  due to the acid-pyridine hydrogen bond, again indicating co-crystal formation. The carbonyl region in **C** is characterised by two bands of medium intensity at around 1670 and 1640  $\text{cm}^{-1}$ , the former due to the carbonyl of the dicarboxylic acid and the latter to the

carbonyl of lamivudine. The carbonyl of the acid shows a significant low frequency shift indicating strong H-bonding interactions and formation of a co-crystal.

In the 2000-500  $\text{cm}^{-1}$  region of the FTIR spectrum of the **E** (Figure 5), we see the appearance of the carbonyl stretching band of the dicarboxylic acid at 1700  $\text{cm}^{-1}$ , and the carbonyl of lamivudine at around 1650  $\text{cm}^{-1}$ . The carbonyl shows smaller shift in **E** compared to the shift of that in **C**, which indicates a weaker H-bonded system compared to **C**. The carbonyls in **F** and **G** show little shift and a triplet structure not seen either in **E** or **B**. Here again the system may have weaker H-bonding compared to **C**. Although the absolute structures of these solid forms could not be determined for lack of single crystals suitable for structure elucidation, the FTIR spectra of **E**, **F** and **G** in the 2000-500  $\text{cm}^{-1}$  region show remarkable similarities, indicating possible formation of co-crystalline phases. The PXRD patterns and DSC curves are also indicative of this (supporting information). Since in all these three solids there are no bands at 1570  $\text{cm}^{-1}$  and 1420  $\text{cm}^{-1}$  due to deprotonated  $\text{COOH}$  and due to the asymmetric and symmetric  $\text{COO}^-$ , we may conclude that they are all co-crystals and not salts. The  $\text{pK}_a$  values of the acids (supporting information) also suggest this.

The spectra of all the co-crystal and the uncharacterised solid systems show very similar spectral features indicating similar, but slightly varying types of bonding and structures, allowing us to conclude that **E**, **F** and **G** may yield the same synthon as seen in **C**. The observed spectral changes in the co-crystal and the solids are in line with the synthon observed by X-ray crystallography in **C**. Thus, the SCXRD and FTIR studies of **C** show the formation of a co-crystal and the results are not different for the solids formed between the API and the other acids.

## Conclusions

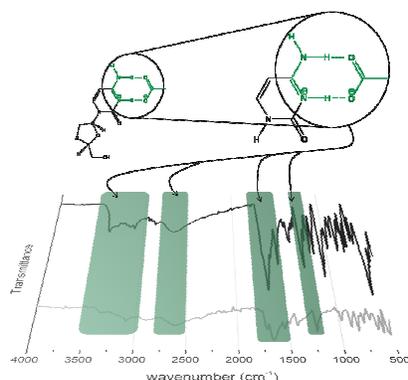


Figure 8. Pictorial summary of the IR approach for synthon identification

Our results emphasise the validity of the idea of transferability of synthons and the synthon theory as a whole, based on the assumption that these are modules within crystal structures. The essence of our argument is summarised in Figure 8 above. In the

present study, these modules can be specifically probed using IR spectroscopy. Our studies show that in related structures possessing the same synthon, the signatures pertaining to a particular synthon clearly show up in the IR spectra, and the simpler of the systems at hand can be taken as a model for the more complex ones. This approach can be used in high throughput crystal screening wherein the interactions of an API with various co-formers can be monitored with IR spectroscopy by identifying and comparing the spectral signatures of the interactions of its basic building unit with the same co-formers. We have also shown using IR spectroscopy, with SCXRD as a backup, that non-matching synthons give different spectral signatures. The FTIR method can well be used in online crystallisation monitoring<sup>37</sup> in multi-well crystallisation systems.

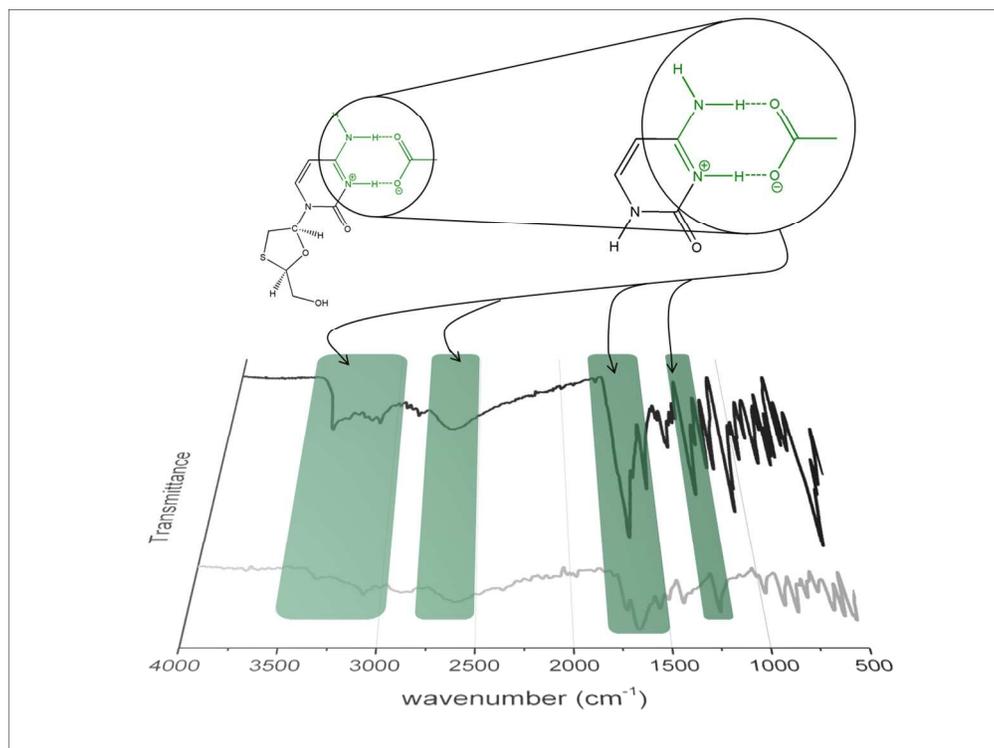
**Acknowledgements:** S.C. thanks the I.I.Sc for a scholarship; S.G. thanks the I.I.Sc for a fellowship. G.R.D. thanks the D.S.T. for the award of a J. C. Bose fellowship. We thank Mr. Srinu Tothadi for helpful discussions pertaining to the crystallographic aspects of this work.

#### References:

1. A. D. Bond, *CrystEngComm* **2007**, *9*, 833
2. (a) G.R. Desiraju, *CrystEngComm* **2003**, *5*, 466; (b) J.D. Dunitz, *CrystEngComm* **2003**, *5*, 506; (c) J. Zuckerman-Schpector, E.R.T.Z. Tiekink, *Kristallogr.* **2008**, *223*, 233
3. F. Wöhler, *Annalen Chem. Pharm* **1844**, *51*, 145
4. N. Rodriguez-Hornedo, *Mol. Pharmaceutics* **2007**, *4*, 299
5. Ö. Almarsson, M.J. Zaworotko, *Chem. Commun.* **2004**, 1889
6. (a) N. Shan, F. Toda, W. Jones, *Chem. Commun.* **2002**, 2372; (b) S. G. Fleishman, S.S. Kuduva, J.A. MacMahon, B. Moulton, R.B. Walsh, N. Rodriguez-Hornedo, M. J. Zaworotko, *Cryst. Growth Des.* **2003**, *3*, 909; (c) J.F. Remanar, S.L. Morisette, M.L. Peterson, B. Moulton, J.M. MacPhee, H.R. Guzman, Ö. Almarsson, *J. Am. Chem. Soc.* **2003**, *125*, 8456; (d) B. Olenik, T. Smolka, R. Boese, R. Sustmann, *Cryst. Growth Des.* **2003**, *3*, 183; (e) C. B. Aakeröy, A. M. Beatty, B.A. Helfrich, M. Nieuwenhuyzen, *Cryst. Growth Des.* **2003**, *3*, 159; (f) R.D. Walsh, M.W. Bradner, S. Fleischman, L.A. Morales, B. Moulton, N. Rodriguez-Hornedo, M.J. Zaworotko, *Chem Commun.* **2003**, 186; (g) B.R. Bhogala, A. Nangia, *Cryst. Growth Des.* **2003**, *3*, 547 (h) S.L. Childs, L.J. Chyall, J.T. Dunlap, V.N. Smolenskaya, B.C. Stahly, G.P. Stahly, *J. Am. Chem. Soc.* **2004**, *126*, 13335, (i) X. Gao, T. Fričić, L. R. McGillivray, *Angew. Chem. Int. Ed.* **2004**, *43*, 232 (j) B.Q. Ma, P. Coppens, *Cryst. Growth Des.* **2004**, *4*, 211 (k) L.S. Reddy, A. Nangia, V. M. Lynch, *Cryst. Growth Des.* **2004**, *4*, 89 (l) A. V. Trask, N. Shan, W.D.S. Motherwell, W. Jones, V. Feng, R.B.H. Tan, K.J. Carpenter, *Chem. Commun.* **2005**, 880 (m) A.V. Trask, W. Jones, *Top. Curr. Chem.* **2005**, *254*, 41 (n) P. Vishweshwar, J.A. MacMahon, M.L. Peterson, M.B. Hickey, T.R. Shattock, M.J. Zaworotko, *Chem. Commun.* **2005**, 4601; (o) T. Fričić, L. Fábíán, J. C. Burley, W. Jones, W.D.S. Motherwell, *Chem. Commun.* **2006**, 5009; (p) T. Fričić, A.V. Trask, W. Jones, W.D.S.

- Motherwell, *Angew. Chem. Int. Ed.* **2006**, 45, 7546 (q) J.A. Bis, O.L. McLaughlin, P. Vishweshwar, M.J. Zaworotko, *Cryst. Growth Des.* **2006**, 6, 2648; (r) S.L. Childs, K.I. Hardcastle, *Cryst. Growth Des.* **2007**, 7, 1007; (s) G.P. Stahly, *Cryst. Growth Des.* **2007**, 7, 616; (u) M. Dabros, P.R. Emery, V.R. Thalladi, *Angew. Chem. Int. Ed.* **2007**, 46, 4132(v) W.W. Porter III, S.C. Ellie, A.J. Matzger, *Cryst. Growth Des.* **2008**, 8, 14
7. (a) A.V. Trask, *Mol. Pharmaceutics* **2007**, 4, 301; (b) W. Jones, W.D.S. Motherwell, A.V. Trask, *MRS Bull.* **2006**, 31, 875; (c) The first pharmaceutical co-crystal has been taken to Phase II clinical trials. See, <http://www.newswire.ca/en/story/1242581/estev-announces-positive-results-from-its-phase-ii-clinical-study-of-e-58425-in-acute-pain> and [http://www.esteve.es/EsteveFront/CargarPagina.do?pagina=idi\\_pa\\_portfolio\\_E-58425.jsp&div=idi&lng=en](http://www.esteve.es/EsteveFront/CargarPagina.do?pagina=idi_pa_portfolio_E-58425.jsp&div=idi&lng=en)
  8. E.J. Corey, *Pure Appl. Chem.* **1967**, 14, 19
  9. (a) G.R. Desiraju, *Angew. Chem. Int. Ed.* **1995**, 34, 2311 (b) Ö. Almarsson and M. J. Zaworotko, *Chem. Commun.*, **2004**, 17, 1889
  10. G.R. Desiraju, *Angew. Chem. Int. Ed.* **2007**, 8342
  11. G.R. Desiraju, *J. Am. Chem. Soc.* **2013**, 135, 9952
  12. (a) D.S. Reddy, D.C. Craig, G.R. Desiraju, *J. Chem. Soc. Chem. Commun.*, **1994**, 1457, (b) R. Foster, *Organic Charge-Transfer Complexes*, Academic: London 1969, (c) P. Metrangolo, H. Neukirch, T. Pilati, G. Resnati, *Acc. Chem. Res.* **2005**, 38, 386, (d) A. Priimagi, G. Cavallo, P. Metrangolo, G. Resnati, *Acc. Chem. Res.* **2013** (e) A.I. Kitaigorodskii, *Mixed Crystals*, Springer-Verlag: Berlin, 1984
  13. A. Mukherjee, S. Tothadi, S. Chakraborty, S. Ganguly, G.R. Desiraju, *CrystEngComm*, **2013**, 15, 4640
  14. (a) C.C. da Silva, R.R. Coelho, M. de Lima Cirqueira, A.C.C. de Melo, I.M.L. Rosa, J. Ellena, F.T. Martins, *CrystEngComm*, **2012**, 14, 4562 (b) P.M. Bhatt, Y. Azim, T.S. Thakur, G.R. Desiraju, *Cryst. Growth Des.* **2009**, 9, 951, (c) F.T. Martins, N. Paparidis, A.C. Doriguetto, J. Ellena, *Cryst. Growth Des.* **2009**, 9, 5283, (d) J. Ellena, N. Paparidis, F.T. Martins, *CrystEngComm* **2012**, 14, 2373, (e) A. Bhattacharya, B.N. Roy, G.P. Singh, D. Srivastava, A.K. Mukherjee, *Acta Crystallogr., Sect. C: Cryst. Struct. Commun.* **2010**, 66, 0329 (f) R.K. Harris, R.R. Yeung, R.B. Lamont, R.W. Lancaster, S.M. Lynn, S.E. Staniforth, *J. Chem. Soc., Perkin Trans. 2*, **1997**, 2653 (g) R. Banjeree, P.M. Bhatt, N.V. Ravindra, G.R. Desiraju, *Cryst. Growth Des.* **2005**, 5, 2299, (h) P.M. Bhatt, G.R. Desiraju, *CrystEngComm*, **2008**, 10, 1747, (i) C. C. da Silva, M. de Lima Cirqueira, F.T. Martins, *CrystEngComm*, **2013**, 15, 6311
  15. M.J. Koziel, M.G. Peters, *N. Engl. J. Med.* **2007**, 356, 1445
  16. N.S. Mandel, *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem.* **1977**, 33, 1079
  17. J. Ellena, N. Paparidis, F.T. Martins, *CrystEngComm*, **2012**, 14, 2373

18. K. Bouchouit, N. Benali-Cherif, S. Dahaoui, E. E. Bendeif, C. Lecomte, *Acta Crystallogr., Sect. E: Struct. Rep. Online* **2005**, 61, o2755
19. (a) *CrystalClear 2.0*; Rigaku Corporation: Tokyo, Japan, (b) J.W. Pflugarth, *Acta Crystallogr. Sect. D*, **1999**, 55, 1718
20. *Bruker APEX2 (Version 1.0.22)*, *BIS (Version 1.2.08)*, *COSMO (Version 1.48)* and *SAINT (Version 7.06)*, Bruker AXS Inc. : Madison, Wisconsin, USA, 2006
21. G. M. Sheldrick, *Acta Crystallogr., Sect A: Found Crystallogr.* **2008**, 64, 112
22. L. J. Farrugia, *J. Appl. Crystallogr.*, **1999**, 32, 837
23. L.J. Farrugia, *J. Appl. Crystallogr.*, **1997**, 30, 565
24. M. Gdaniec, B. Brycki, M. Szafran, *J. Chem. Soc., Perkin Trans.* **1988**, 1775
25. H. G. Brittain, *Physical Characterization of Pharmaceutical Solids*, Marcel Dekker: New York, **1995**
26. J. Bernstein, *Conformational polymorphism in Organic Solid State Chemistry*, ed. G. R. Desiraju, Elsevier: Amsterdam, **1987**, 485
27. R. Bouche, *J. Pharm. Belg.*, **1968**, 23, 449
28. I. P. S. Kang, C. E. Kendall, R. W. Lee, *J. Pharm. Pharmacol.*, **1974**, 26, 201
29. (a) R.K. Harris, R.R. Yeung, R.B. Lamont, R.W. Lancaster, S.M. Lynn, S.E. Stanifort, *J. Chem. Soc. Perkin Trans. 2.* **1997** (b) T.R. Ramkumar, S. Srinivasan, T.J. Bhoopathy, S. Gunasekharan, *Spectrochimica Acta Part A.* **2012**,
30. J. Floria', V. Baumruk, and J. Leszczynski, *J. Phys. Chem.* **1996**, 100, 5578
31. G. Slosarek, R. Zamboni, *Spectrochim Acta*, **1991**, 47A, 863
32. V. Mathew, S. Joseph, S. Jacob, K.E. Abraham, *Bull. Mater. Sci.*, **2010**, 33, 433.
33. (a) S.E. Cabaniss, I.F. McVey, *Spectrochim. Acta A*, **1995**, 51, 2385. (b) M. Ibrahim, A. Nada, D. E. Kamal, *Ind. J. Pure & Appl. Physics*, **2005**, 911
34. L. J. Bellamy, *Advances in Infrared Group Frequencies*, Methuen & Co.: London, **1968**
35. D. Hadži, N. Sheppard, *Proc. R. Soc. London, Ser. A*, **1953**, A216, 247
36. S. Bratož, D. Hadži, N. Sheppard, *Spectrochim. Acta*, **1956**, 8, 249
37. T. Kojima, S. Tsutsumi, K. Yamamoto, Y. Ikeda, T. Moriwaki, *Int. J. Pharm.*, **2010**, 399, 52.



254x190mm (150 x 150 DPI)