# Brønsted interactions in organic crystal structures: towards crystallographic structure refinement using core level spectroscopies

Paul Timothy Edwards

The University of Leeds
School of Chemical and Process Engineering

Diamond Light Source Ltd

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

August 2023

## Intellectual Property and Publication Statement

The candidate confirms that the work submitted is his own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

The published material forming chapter 2 has been published in the Crystal Growth and Design –

Edwards P.T., Saunders L.K., Pallipurath A.R., Britton A.J., Willneff E.A., Shotton E.J. and Schroeder S.L.M., Proton Transfer on the Edge of the Salt/Cocrystal Continuum: X-ray Photoelectron Spectroscopy (XPS) of Three Isonicotinamide Salts, Cryst. Growth Des. **21**, 6332 (2021)

The candidate carried out all of the XPS experiments, data calibration and analysis and drafted the manuscript. Dr Lucy Saunders kindly prepared the samples ahead of measurement. Dr Anuradha Pallipurath initiated the collaboration between the candidate and Dr Lucy Saunders. Dr Andrew Britton provided assistance during the experiment, and Dr Elizabeth Willneff provided support with XPS data analysis. Dr Elizabeth Shotton and Prof. Sven Schroeder carried out a review of the draft paper in preparation for submission.

The published material forming chapter 3 has been published in the Journal of Physical Chemistry A –

Edwards P.T., Saunders L.K., Grinter D.C., Ferrer P., Held G., Shotton E.J. and Schroeder S.L.M., Determination of H-atom Positions in Organic Crystal Structures by NEXAFS Combined with Density Functional Theory: A Study of Two-Component Systems Containing Isonicotinamide, J. Phys. Chem. A **126**, 2889 (2022)

The candidate planned and carried out the NEXAFS experiment and DFT calculations. The data analysis, including full comparison of NEXAFS and DFT data was also carried out, including a new procedure to normalise for the gas phase on the new beamline endstation. A draft manuscript was then produced. Dr Lucy Saunders prepared the samples in advance of the experiment. Dr David Grinter, Dr Pilar Ferrer and Prof. Georg Held provided support during the two week beamtime on the first commissioning experiment at B07-B. Dr Elizabeth Shotton and Prof. Sven Schroeder also supported during the experiments and carried out a review of the manuscript ahead of submission.

The material forming chapter 4 has been submitted for publication in Physical Review Letters –

Edwards P.T., Shotton E.J. and Schroeder S.L.M., Hydrogen Bonds in Molecular Crystals: Strong Correlation between Proton Positions and 1s Core Level Binding Energies

The candidate carried out the theoretical calculations, developed the model and carried out the data analysis including comparison to experimental data. A draft manuscript was also produced. Dr Elizabeth Shotton and Prof. Sven Schroeder carried out a review of the manuscript prior to submission.

The material forming chapter 5 has been drafted for future submission –

Edwards P.T., Shotton E.J. and Schroeder S.L.M., Core Level Spectroscopy to Determine H-atom Positions in Organic Crystal Structures Combined With Density Functional Theory: Short Strong Hydrogen Bonds and Proton Disorder

The candidate carried out the planning for the experiment, identification of suitable samples, sample preparation, theoretical calculations, experiments and carried out the data analysis incorporating all techniques. A draft manuscript was then produced. Dr Elizabeth Shotton assisted in sample preparation and the experiment, including collecting XRD data to confirm the structure of the samples and also reviewed the draft manuscript. Prof. Sven Schroeder helped during the experiment and carried out a review of the manuscript prior to submission.

# Acknowledgements

I would like to express my greatest appreciation to my supervisors, Sven and Elizabeth, for their inspiration, encouragement and support throughout my project, without which my project would not have been nearly as successful or enjoyable. Their belief in me and enthusiasm for my project has kept my motivation and determination. I do hope we will keep in touch as I enter the next chapter of life. I would also like to express my gratitude to the Engineering and Physical Sciences Research Council and Diamond Light Source for funding my research endeavours (EPSRC Grant EP/R513258/1). Many thanks to my colleagues at Diamond Light Source for the informal chats and advice over the four years. A special thanks to Lucy for initially providing me with samples and continued collaboration throughout my project.

I would also like to extend my thanks to the Versatile Soft X-ray Facility at the University of Leeds and beamline B07-B at Diamond Light Source for access to first class experimental facilities. I am particularly grateful to Beth and Andrew from Leeds, and Dave and Pilar from Diamond for their help with experiments and for teaching me about the techniques and data analysis process.

Thanks should also go to the rest of the University of Leeds Molecular Process and Materials Engineering group over the years for the many catch ups and discussions on areas related to our research objectives. I would also like to recognise those who participated in my beamtime for their excellent company and late-night problem solving: Elizabeth, Sven, Anna and Nathan. It really was an enjoyable couple of weeks.

I would also like to acknowledge the part played by my family, for their love and support throughout including attempting to understand what I do and what it means. Lastly, and most importantly, I would like to thank my partner Matthew for putting up with the ups and downs and providing a constant source of support and encouragement.

# Abstract

Hydrogen Bonds are vital to the crystal structure of organic crystals, and the nature of the hydrogen bonding interaction has an effect on the properties of these crystals. Being able to accurately define the position of a hydrogen atom in a molecular structure is difficult with X-ray diffraction based techniques, and more accurate techniques often require onerous sample preparation or long measurement times. As an alternative, core level X-ray spectroscopy is investigated as a technique for hydrogen atom position refinement, with a particular focus on hydrogen bonding in organic crystals.

Through the use of XPS and NEXAFS spectroscopy, an analysis procedure has been developed which consistently achieves the absolute core level binding energies and excitation energies. From these, a direct dependence on the hydrogen atom position in relation to a proton acceptor has been determined through a shift in the core level binding energy of +2 eV. These shifts are also reproduced in *ab initio* DFT calculations and allow hydrogen bonding interactions to be easily classified as salt or cocrystal.

An investigation to identify the cause of the +2 eV shift was successful in demonstrating this shift using electrostatics and can be entirely attributed to the electrostatic potential difference between the two hydrogen atom positions as measured at the core level of the proton acceptor.

Furthermore, by utilising the different dependencies of XPS and NEXAFS, the molecular structure of a number of crystals has been refined through comparison to theoretical spectra, demonstrating that highly accurate structures can be obtained due to this direct sensitivity to hydrogen atom position.

# List of Abbreviations

| | |
|---|---|
| BE | Binding Energy |
| CSD | Cambridge Structural Database |
| DFT | Density Functional Theory |
| eV | Electron Volt |
| FWHM | Full Width Half Maximum |
| NAP | Near Ambient Pressure |
| NEXAFS | Near Edge X-ray Absorption Fine Structure |
| NMR | Nuclear Magnetic Resonance |
| SSHB | Short Strong Hydrogen Bond |
| TDDFT | Time Dependent Density Functional Theory |
| UHV | Ultra High Vacuum |
| XPS | X-ray Photoelectron Spectroscopy |
| XRD | X-ray Diffraction |

# Table of Contents

**Intellectual Property and Publication Statement....ii**
**Acknowledgements....v**
**Abstract....vi**
**List of Abbreviations....vii**
**Table of Contents....viii**
**List of Tables....xiii**
**List of Figures....xv**
**1 Hydrogen Bonding, Core Level Spectroscopy and Density functional Theory....1**
1.1 Context of research....1
1.1.1 Hydrogen Bonding....1
1.1.1.1 Overview of Intermolecular Interactions....2
1.1.1.2 Short Strong Hydrogen Bonds....4
1.1.1.3 Salts, cocrystals and the salt-cocrystal continuum....4
1.1.2 Pharmaceuticals....7
1.1.3 Traditional Methods of Characterisation....7
1.1.3.1 $\Delta pK_a$ approach....7
1.1.3.2 X-ray Diffraction....8
1.1.3.3 Neutron Diffraction....10
1.1.3.4 Solid State Nuclear Magnetic Resonance (ss-NMR)12
1.2 Aims and Objectives....14
1.3 Experimental Techniques....15
1.3.1 XPS....15
1.3.1.1 Experimental set up....17
1.3.1.2 Ultra High Vacuum XPS....19
1.3.1.3 Near Ambient Pressure XPS....19
1.3.1.4 Spectrum processing and calibration....21
1.3.1.5 Peak Fitting and analysis....22
1.3.1.6 Literature using XPS to investigate hydrogen bonding....23
1.3.2 NEXAFS....27
1.3.2.1 Experimental set up....29

1.3.2.2 Spectrum processing and calibration .....29
1.3.2.3 Peak fitting and analysis .....31
1.3.2.4 Literature using NEXAFS to investigate hydrogen bonding .....31
1.4 Computational Techniques .....35
1.4.1 Density Functional Theory .....35
1.4.1.1 Experimental set up and details .....37
1.4.1.2 Calculation of XP spectra .....38
1.4.2 TDDFT .....39
1.4.2.1 Calculation of NEXAFS .....39
1.4.2.2 Literature using TDDFT to investigate hydrogen bonding .....40
1.5 Conclusions .....43
1.6 References .....44
**2 Proton Transfer on the Edge of the Salt/Cocrystal Continuum: X-ray Photoelectron Spectroscopy (XPS) of Three Isonicotinamide Salts .....53**
2.1 Introduction .....54
2.2 Experimental Methods .....57
2.2.1 Materials .....57
2.2.2 X-ray Photoelectron Spectroscopy (XPS) .....58
2.2.3 Quantitative Elemental Analysis .....58
2.2.4 Binding Energy Scale Calibration .....60
2.3 Results .....60
2.3.1 C 1s Emission and Binding Energy Scale Calibration .....60
2.3.2 N 1s Emission .....64
2.3.3 O 1s Emission .....66
2.4 Discussion .....67
2.5 Conclusions .....72
2.6 References .....73
**3 Determination of H-atom Positions in Organic Crystal Structures by NEXAFS Combined with Density Functional Theory: A Study of Two-Component Systems Containing Isonicotinamide .....77**
3.1 Introduction .....78
3.2 Modelling and Experimental Methods .....80
3.2.1 Density Functional Theory Calculations .....80

3.2.2 Materials ....81
3.2.3 Near Edge X-ray Absorption Fine Structure (NEXAFS) ....82
3.3 Results ....83
3.3.1 Isonicotinamide 2,4-dinitrobenzoic acid (IN24DNBA) ....84
3.3.2 Isonicotinamide 3,5-dinitrobenzoic acid (IN35DNBA) ....86
3.3.3 Isonicotinamide phthalic acid (INPA) ....87
3.4 Discussion ....89
3.4.1 Isonicotinamide phthalic acid peak intensities ....89
3.4.2 DFT of cluster structures – isonicotinamide phthalic acid ....90
3.4.3 The effect of H-atom position on peak position. ....91
3.4.4 Difference Spectra – complexes compared to individual components ....92
3.5 Conclusions ....94
3.6 References ....95
**4 Hydrogen Bonds in Molecular Crystals: Strong Correlation between Proton Positions and 1s Core Level Binding Energies ....99**
4.1 Introduction ....99
4.2 Electrostatic Model ....100
4.3 Results and Discussion ....101
4.4 Conclusions ....103
4.5 References ....103
**5 Crystallographic Refinement of Proton Positions and Site Occupancies by Combining Core Level Spectroscopies with Density Functional Theory ....105**
5.1 Introduction ....106
5.2 Materials and Methods ....109
5.3 Results ....112
5.3.1 XPS ....112
5.3.2 NEXAFS ....115
5.3.3 DFT ....116
5.4 Discussion ....118
5.4.1 4,4’-Bipyridine Succinic Acid ....118
5.4.2 INHFA ....119
5.4.3 INOA ....121
5.4.4 4EP35DNBA ....123

5.4.5 35L35DNBA ..... 124
5.5 Conclusions ..... 127
5.6 References ..... 128
**6 Discussion, Additional Results and Analysis ..... 132**
6.1 Overview of the papers ..... 132
6.1.1 Proton Transfer on the Edge of the Salt/Cocrystal Continuum: X-Ray Photoelectron Spectroscopy of Three Isonicotinamide Salts ..... 133
6.1.2 Determination of H-Atom Positions in Organic Crystal Structures by NEXAFS Combined with Density Functional Theory: a Study of Two-Component Systems Containing Isonicotinamide. ..... 136
6.1.3 Classical Electrostatics Predicts the Core Level Binding Energy Shifts at Brønsted Acceptor Sites ..... 138
6.1.4 Core Level Spectroscopy to Determine H-atom Positions in Organic Crystal Structures Combined With Density Functional Theory: Short Strong Hydrogen Bonds and Proton Disorder ..... 142
6.1.4.1 Applying the Electrostatics model to these additional systems investigated. ..... 144
6.1.4.2 Checking for consistency between XPS and NEXAFS for the isonicotinamide samples investigated in Chapters 2 and 3 ..... 146
6.1.5 Accuracy of XPS and NEXAFS ..... 147
6.1.5.1 Experimental issues with UHV-XPS ..... 148
6.1.5.2 Beam Damage and Vacuum conditions ..... 149
6.1.5.3 Commissioning of the new B07-B high throughput NEXAFS endstation ..... 150
6.1.6 Accuracy and Reliability of DFT calculations ..... 151
6.1.6.1 Choice of Exchange Correlation Functional and Basis set ..... 152
6.2 Additional complementary investigations ..... 154
6.2.1 Deconvolution of broadened XPS peaks ..... 155
6.2.2 Histidine angle dependent NEXAFS ..... 157
6.3 Conclusions and Future Work ..... 158
6.3.1 Relevance of findings to industry ..... 159
6.3.2 Future work ..... 160
6.3.3 Conclusions ..... 162

6.4 References.....163

**Appendix A: Proton Transfer on the Edge of the Salt/Co-Crystal Continuum: X-ray Photoelectron Spectroscopy (XPS) of Three Isonicotinamide Salts - Supporting Information.....166**

**Appendix B: Determination of H-atom Positions in Organic Crystal Structures by NEXAFS Combined with Density Functional Theory: A study of Two-Component Systems Containing Isonicotinamide – Supporting information.....170**

**Appendix C: Hydrogen Bonds in Molecular Crystals: Strong Correlation between Proton Positions and 1s Core Level Binding Energies – Supporting Information.....180**

# List of Tables

**Table 2.1 Element and adventitious carbon ($C_{adv}$) percentages compared to sample stoichiometries, including corrected values for N and O after removal of the $C_{adv}$ contributions.....59**

**Table 2.2 Best fit binding energies (eV), intensities (a.u.) and FWHMs (eV) for the fitted components to the C 1s emission shown in Figure 2.2.....61**

**Table 2.3 Best fit binding energies (eV), intensities (a.u.) and FWHMs (eV) for the fitted components to the N 1s emission. Single component fit describes both pyridinic and amide nitrogen indicated by *.....65**

**Table 2.4 Best fit binding energies (eV), intensities (a.u.) and FWHMs (eV) for the fitted components to the O 1s emission.....67**

**Table 2.5 ∆pKa values compared with N-H, N-O bond distances (from XRD) and N 1s Binding Energy.....68**

**Table 3.1 Best fit photon energies (eV) for the N 1s → π* transitions. DFT energies after uniform shift applied taken from the calculated excitation energies. *Additional changes to the molecular structure to reproduce peak shifts and intensities from the experimental spectrum. **Molecular Cluster calculation to account for longer range interactions. .....88**

**Table 5.1 Samples and their properties from the Cambridge Structural Database at room temperature.23 Molecular structures are shown in figure 5.1.....108**

**Table 5.2 Experimental Binding Energies. C 1s. Adventitious carbon proportion determined from the survey spectra. The C-X component includes all aromatic carbon covalently bonded to any other electron withdrawing group including nitro, amide, carboxylic acid. Expected % are calculated from the anticipated molecular structure, accounting for the known quantity of adventitious carbon. Very good agreement across all five samples. .....112**

**Table 5.3 Experimental N 1s Binding Energies. BE Calibration with respect to the adventitious carbon peak. Expected % calculated from the anticipated molecular structure. Good agreement for 44BPYSA and INHFA samples, but unexpected splitting of hydrogen bonded nitrogen peak observed in INOA, 4EP35DNBA and 35L35DNBA. .....113**

**Table 5.4 INHFA summary of results showing the change in proton position and the effect on the calculated binding energies and electron transition energies to better agree with experiment. .....120**

**Table 5.5 INOA (2:1) summary of results showing the change in half of the hydrogen bonded proton positions and the effect on the calculated binding energies and electron transition energies to better agree with experiment. .....122**

**Table 5.6 4EP35DNBA summary of results showing the change in a quarter of the hydrogen bonded proton positions and the effect on the calculated binding energies and electron transition energies to better agree with experiment. .....125**

**Table 5.7 35L35DNBA summary of results showing the change in 1/7 of the hydrogen bonded proton positions and the effect on the calculated binding energies and electron transition energies to better agree with experiment. The final structure was required to ensure both the transitions and binding energies agreed with experiment.....126**

**Table 6.1 Comparison of errors from electrostatic model grouped by closest neighbour group. This shows that the error observed in the electrostatic model is not random and there is a clear trend between the neighbouring group and the error observed. .....140**

**Table 6.2 The electrostatic model applied to the five systems in chapter 5. Errors are in line with those found in the original work with a mean of 0.15 eV.....145**

**Table 6.3 Comparison of DFT calculated binding energies with experimental XPS for the three isonicotinamide complexes INPA, IN35DNBA and IN24DNBA. Agreement of binding energies is good, except for an underestimate of the nitro group binding energy by 2 eV.....146**

# List of Figures

**Figure 1.1 A typical hydrogen bond interaction between an oxygen donor and nitrogen acceptor. The slight positive charge on the hydrogen is attracted to the lone pair of electrons on the electronegative nitrogen. .....3**

**Figure 1.2 Typical proton potential energy surfaces for hydrogen bonds. The most common situations are a typical double well and proton transfer, while the interesting characteristics of single well and dynamic disorder are observed in short strong hydrogen bonds. .....5**

**Figure 1.3 Change in proton position investigated using neutron diffraction. Blue line is 3,5-dicarboxylic acid, red line is deuterated 3,5-dicarboxylic acid. Error bars on deuterate sample are smaller due to greater certainty of the position, but the data suggests that deuterating the sample has an effect om the structure so not directly comparable to the non-deuterated sample. Figure reproduced from ref. 29. .....11**

**Figure 1.4 Processes occurring in XPS. (a) Photoemission on absorption of a photon, (b) Auger emission to refill the core hole and (c) fluorescence to refill the core hole. .....17**

**Figure 1.5 Schematic diagram of a NAP-XPS experimental set up showing the key components. The choice of gas phase and the gas phase pressure are decided to avoid peak clash with the sample. .....20**

**Figure 1.6 N 1s Binding energy plotted against N-H bond distance for a range of crystal structures with salt and co-crystal interactions[47] . A clear distinction between salt and co-crystal is evident with a shift in binding energy. Reproduced from ref. 51...24**

**Figure 1.7 N 1s XP spectra of isonicotinamide and theophylline complexes with different hydrogen bonding characteristics[35] Nitrogen of interest is labelled. No intermolecular hydrogen bond in bottom structure, but nitrogen is protonated leading to peak shift. Reproduced from ref. 35. .....25**

**Figure 1.8 N 1s XP spectra for the three PDCA structures displaying significant differences in both peak position and peak shape.[19] Hydrogen bonds are between the pyridine nitrogen and carboxylic acid group in each case. Reproduced from ref. 19. .....26**

**Figure 1.9 Schematic diagram of the NAP-NEXAFS experimental set up at B07-B showing the key components and processes. Everything to the left of the SiN membrane is within the sample chamber. Both measurement of electrons and fluorescence are possible through total fluorescence yield (TFY) and total electron yield (TEY) measurement. .....30**

**Figure 1.10 Simulated NEXAFS spectrum showing a theoretical change in binding energy with bond length.[55] This demonstrates the theoretical sensitivity of the NEXAFS technique to individual bond lengths. Reproduced from ref. 59 .....32**

**Figure 1.11 N K-edge NEXAFS for both polymorphs of para-amino benzoic acid.[44] This shows how sensitive NEXAFS is to very minor changes in molecular structure, in particular where hydrogen bonding is present. Reproduced from ref. 44. .....33**

**Figure 1.12 N K-edge NEXAFS for the three pyridine dicarboxylic acid complexes.[19] Demonstrating the sensitivity of NEXAFS to proton position with broadened peak observed identifying the intermediate state. Reproduced from ref. 19. .....34**

**Figure 1.13 Experimental spectra (upper line) and computed spectra (lower line) for four molecules using the SRC-1 range separated hybrid functional[64] It is clear that peak positions are in agreement with experiment. Reproduced from ref. 68. .....41**

**Figure 1.14 Calculated molecular orbital visualisation determined using DFT (left) and experimental and calculated N K-edge NEXAFS (right) for para-amino benzoic acid[48] Demonstrates the dependence of NEXAFS spectra on the overlap between molecular orbitals. Reproduced from ref. 52. .....42**

**Figure 2.1 Chemical structures investigated comprising three two component isonicotinamide-acid (base-acid) systems with (a) 2,4-dinitrobenzoic acid (1:1), (b) 3,5-dinitrobenzoic acid (1:1) and (c) phthalic acid (2:1). N-H-O distances can be found in Appendix A Table A1. .....55**

**Figure 2.2 C 1s emission lines with best-fit components for the carboxylic acid (O−C=O) group, the amide (N−C=O) group from IN, aromatic carbons with electronegative substituents in DNBAs and PA (−(CX)=, with X = COOH, NO2), unsubstituted aromatic carbons (−CH=), and adventitious carbon (CAdv). (a) IN24DNBA, (b) IN35DNBA, (c) INPA and (d) pure isonicotinamide. The binding energy scale was calibrated at 285.0 eV with respect to the adventitious carbon contamination (Adv) emission, which is superimposed over the emission from the aromatic carbons (−C=). .....62**

**Figure 2.3 Nitrogen 1s XPS with fitted components for (a) IN24DNBA, (b) IN35DNBA, (c) INPA and (d) pure isonicotinamide. Binding energy scale calibration is taken with respect to the fitted adventitious carbon component peak. .....63**

**Figure 2.4 Oxygen 1s XPS with fitted components for (a) IN24DNBA, (b) IN35DNBA, (c) INPA and (d) pure isonicotinamide. Binding energy scale calibration is taken with respect to the fitted adventitious carbon peak.....66**

**Figure 2.5 Relationship between the N 1s core level binding energies (XPS) with the ∆pKa for a range of salt, co-crystal and short strong hydrogen bonded crystal structures, including the three complexes investigated here. .....69**

**Figure 2.6 Relationship between the N 1s core level binding energies (XPS) with the distance between the Brønsted proton and the acceptor nitrogen in a range of crystal structures including the three complexes investigated here. All distances determined using X-ray diffraction and sourced from the Cambridge Structural Database. .....71**

**Figure 3.1 Schematic of the principal transitions for the pyridinium nitrogen showing the energy levels of the 1s, π* and σ* orbitals, and the ionization potential. The effect of hydrogen bonding and proton transfer increasing the photon energy of the N 1s → π* transition is shown due to stabilizing of the N 1s orbital in the hydrogen bonded complex. .....79**

**Figure 3.2 Modelled and experimental N K-edge NEXAFS spectrum of IN24DNBA showing excellent agreement of π* resonance peak positions. Experimental NEXAFS spectrum is fitted with Gaussian peaks identifying transitions and error functions at the core electron binding energies determined using XPS. N 1s → π* transitions are identified based on the origin of the core level electron and the π* orbital related to the relevant ring structure. Gaussian peaks at photon energies above 405 eV are to allow data fitting only, and we extract no physical meaning from these. .....83**

**Figure 3.3 Modelled and experimental N K-edge NEXAFS spectrum of IN35DNBA showing excellent agreement of π* resonance peak positions. Overall peak shapes also fully consistent with salt structure with low energy amide peak (398.8 eV) and higher energy protonated nitrogen peak (399.7 eV). Additional feature compared with IN24DNBA acid above nitro peak (404.4 eV) present in both spectra visible as a slight broadening. .....85**

**Figure 3.4 Modelled and experimental N K-edge NEXAFS spectrum of INPA showing excellent agreement of π* resonance peak positions. However, the DFT calculated spectrum has altered peak intensities for the peaks at 398.8 eV and 399.7 eV, suggesting a difference in structure. .....87**

**Figure 3.5 Modelled and experimental N K-edge NEXAFS spectrum of INPA showing excellent agreement of π* resonance peak positions and peak intensities. N-H distances at the amide groups have been reduced by 0.15 Å compared to the original database geometry to correctly match the experimental spectrum.....89**

**Figure 3.6 Modelled and experimental N K-edge NEXAFS spectrum of INPA showing excellent agreement of π* resonance peak positions and peak intensities. Calculation based on a cluster of molecules to simulate longer range interactions involved in the crystal structure.....90**

**Figure 3.7 Series of DFT calculated NEXAFS spectra moving the H-atom across the hydrogen bond from the structure in the crystal database to the "optimized hydrogen" structure. The peak positions for the amide nitrogen is different for the "optimized hydrogen" structure calculation since all hydrogen atom positions were optimized. .....92**

**Figure 3.8 Difference spectra between experimental spectra and sum of the individual component's experimental spectra (See Figure S3). All three complexes exhibit a clear shift of both the amide and pyridine ring nitrogen species from the pure components to the crystal complex. Negative peaks (solid lines) correspond to the peak position in the pure compounds and arrows indicate the peak shift from pure compound to crystal complex. .....93**

**Figure 4.1 Schematic representation of the primary interac- tions accounted for by the simple electrostatic model. These consist of interactions between the core level electron and i) the atomic nucleus ii) the valence electrons iii) the hydrogen atom and iv) the donor oxygen.....100**

**Figure 4.2 Core level binding energies predicted by electrostat- ics modelling (full line) for a range of N-H distances, compris- ing salts (N-H distance ~ 1.0 ˚A) and cocrystals (N-H distance ~ 1.8 ˚A). The datapoints are experimental N 1s core level binding energies for salts and cocrystals ([6] and references therein). ...101**

**Figure 4.3 DFT-calculated core level electron binding energy shifts (line) for a range of N-H distances compared to ex- perimental results ([6] and references therein) for salts (N-H distance ~ 1.0 ˚A) and cocrystals (N-H distance ~ 1.8 ˚A).....102**

**Figure 4.4 Comparison of N 1s core level binding energies predicted by the electrostatics model with experimentally ob- served values. The model calculations were performed with crystallographically determined r(N−H) and r(N−O) distances (see text) .....103**

**Figure 5.1. Molecular structures of the 5 samples investigated as found in the crystallography database. (a) 44BPYSA, (b) INHFA, (c) INOA, (d) 4EP35DNBA and (e) 35L35DNBA. .....107**

**Figure 5.2 Fitted Experimental XPS spectra of the 5 samples, left C 1s and right N 1s. (a) 44BPYSA, (b) INHFA, (c) INOA, (d) 4EP35DNBA and (e) 35L35DNBA. Peak positions and relative intensities are reported in table 5.2 (C 1s) and table 5.3 (N 1s). Peak fits consistent with fixed FWHM and peak intensities to match expected stoichiometry.. .....114**

**Figure 5.3 Experimental NEXAFS spectra of the 5 samples. (a) 44BPYSA, (b) INHFA, (c) INOA, (d) 4EP35DNBA and (e) 35L35DNBA. Important peak positions labelled showing equivalent peaks over the five structures. Non proton transferred peaks (C-N and $NO_2$) very consistent in energy while $NH^+$ shows some variation between samples. .....116**

**Figure 5.4 Simulated NEXAFS spectra of the 5 samples based on the literature geometry. (a) 44BPYSA, (b) INHFA, (c) INOA, (d) 4EP35DNBA and (e) 35L35DNBA. Some significant differences in peak shapes and positions compared to experimental data suggesting incorrect input structures for the calculations.....117**

**Figure 5.5 Simulated NEXAFS spectrum of INHFA with the hydrogen bond proton moved 0.4 Å towards the proton donor showing improved agreement to experimental data. .....120**

**Figure 5.6 Simulated NEXAFS spectrum of the INOA with hydrogen bond proton moved 0.4 Å towards the proton acceptor, demonstrating improved agreement with experiment.....122**

**Figure 5.7 S Simulated NEXAFS spectrum of the 4EP35DNBA with hydrogen bond proton moved 0.5 Å towards the proton donor in ¼ of molecules. This proportion was determined from the relative intensities of XPS peaks. Agreement with experimental spectrum is improved.....124**

**Figure 5.8 Simulated NEXAFS spectra of the 35L35DNBA combining the two structures found in the database to model the peak shape observed experimentally.....126**

**Figure 6.1 N k-edge NEXAFS spectra of half of the amino acid samples. Our experimental data (left) compared to a literature set of data[15] (right) exhibiting excellent agreement in peak position and shape. .....150**

**Figure 6.2 Comparison of calculated NEXAFS spectra from three different exchange correlation functionals compared with experimental data. Photon energy scale calibrated to the $NO_2$ peak. 0.2 eV shift for SRC-1, 12.3 eV for CAM-B3LYP and 14.2 eV for B3LYP. The B3LYP functional clearly best represents the experimental data. .....153**

**Figure 6.3 Broadening function for broadened N 1s emission line observed in 3,5-pyridine dicarboxylic acid calculated using Fourier Deconvolution technique. This appears to relate to the population density of states. .....156**

**Figure 6.4 Experimental C k-edge NEXAFS of histidine showing dependence on orientation (left) and x, y and z components of DFT calculated NEXAFS spectrum of Histidine (right).....158**

# 1 Hydrogen Bonding, Core Level Spectroscopy and Density Functional Theory

This chapter will give an overview of hydrogen bonding and its importance in the structure of organic crystals and cover a range of analytical techniques used to identify crystal structures. It will then cover the rationale of the research carried out and provide background to the techniques used. Finally it will identify the aims of the research and expected outcomes. Throughout, attention will be given to the background underlying science in addition to the unique approach taken in crystal structure refinement.

## 1.1 Context of research

The primary aim of the project was to make use of core level spectroscopy as an alternative method of crystal structure refinement with several advantages over traditional methods used in the field of organic hydrogen bonded crystals. In industry, the importance of easy, cheap and accurate crystal structure characterisation is very important to minimise costs whilst ensuring compliance with patents. Traditional methods of crystal structure refinement have been thoroughly tested in literature and we have good knowledge of the accuracy with which we can determine the crystal structure. For organic pharmaceutical crystals, traditional methods are excellent for determining the structure of atoms with atomic number greater than 1. However, the most commonly used are relatively insensitive to hydrogen atoms resulting in a low accuracy. In structures containing hydrogen bonds, this lack of sensitivity becomes important in understanding the interactions relevant to the crystalline structure.

### 1.1.1 Hydrogen Bonding

Hydrogen bonds are a vitally important form of intermolecular interaction responsible for the chemical and physical properties of many substances. In organic crystals, hydrogen bonds between crystal components play an

important role in chemical characteristics which affect the pharmaceutical properties making our understanding and characterisation of hydrogen bonds an important area of research and highlights the importance of accurate crystal structure characterisation.

#### 1.1.1.1 Overview of Intermolecular Interactions

Intermolecular interactions are primarily electrostatic interactions which occur between molecules. Van der Waals interactions, or dispersion forces, are present in every system.[1] Van der Waals forces are the result of random fluctuations in the electron density around atoms or molecules. These random fluctuations result in a temporary dipole being formed[1]; electrons on nearby molecules then react to this temporary dipole by moving towards a positive charge, resulting in an equivalent temporary dipole on neighbouring molecules. These temporary dipoles are therefore synchronised, and the resulting attraction between positive and negative charges on adjacent molecules result in an attractive force responsible for the majority of intermolecular interaction.[1] A second form of intermolecular interactions are dipole-dipole interactions. These follow the same principles as van der Walls interactions, except the origin is from permanent dipoles, such as those in an ionic solid.[2] In this case, the interactions occur in addition to van der Waals, and while individual dipole-dipole interactions are stronger than a van der Waals interaction, there are only as many dipole-dipole interactions as there are permanent dipoles, meaning any effect on the structural or thermal properties of a molecule is generally less than the sum of the van der Waals interactions.[1,2]

Hydrogen bonds are formed between two neighbouring electronegative atoms when one has a terminal hydrogen covalently bonded to it, as intermolecular or intramolecular interactions of strength greater than those of van der Waals or dipole-dipole interactions, but weaker than a covalent or ionic bond.[3] A schematic of a hydrogen bond is shown in Figure 1.1. The interaction is strongly directional, strongest when the D-H…A angle is around 180°, where D is the proton donor atom, and A is the proton acceptor.[4] Hydrogen bonding can occur between any two atoms with sufficient electronegativity. Electronegativity is

defined as the tendency of an atom to attract a bonding pair of electrons.[5] To have a high electronegativity, an atom typically has an almost full octet and attracts a bonding pair of electrons to minimise its energy, while also being relatively small with little electron shielding.[5] Therefore, the most electronegative atoms are found in the top right of the periodic table consisting of nitrogen, oxygen and fluorine.[3,5] In organic pharmaceutical crystal structures, the most common hydrogen bonds are found between nitrogen and oxygen atoms, where oxygen is the donor and the acceptor can be nitrogen or oxygen. Fluorine is less common in organic structures, but forms the strongest hydrogen bonds when present. Hydrogen bonds between a donor and acceptor nitrogen are less common and significantly weaker due to the relatively lower electronegativity of nitrogen.[5]

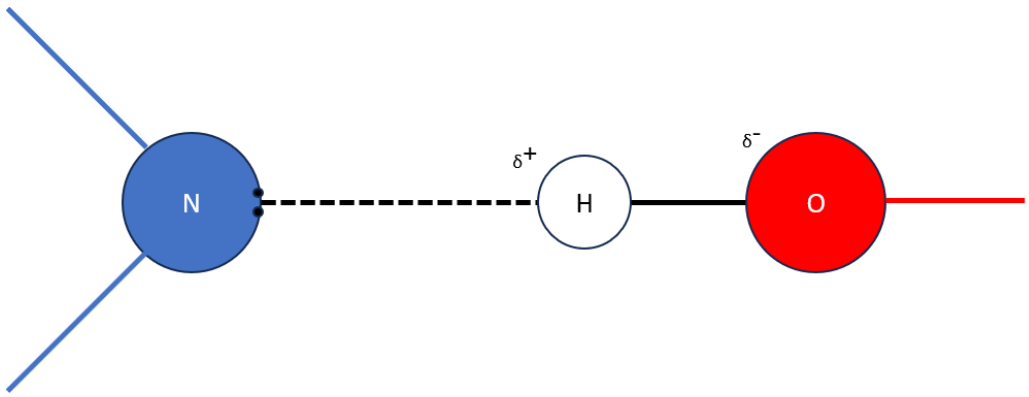



**Figure 1.1** A typical hydrogen bond interaction between an oxygen donor and nitrogen acceptor. The slight positive charge on the hydrogen is attracted to the lone pair of electrons on the electronegative nitrogen.

The interactions occurring where hydrogen bonding is present are the same as for the weaker interactions described above, an electrostatic interaction between a positive and negative charge.[4] However, in the case of hydrogen bonding, this electrostatic interaction originates from a covalent bond between an electronegative atom and a hydrogen atom, whereby the electron density is significantly shifted towards the electronegative atom, resulting in a slight (but more than for van der Waals) positive charge on the hydrogen.[4] This slight positive charge is then attracted to an electronegative atom in a neighbouring molecule (or in the case of intramolecular hydrogen bonds, the same molecule),

specifically the lone pairs of electrons.[6] This interaction, shown in Figure 1.1, is significantly stronger than a van der Waals or dipole-dipole interaction and can begin to look much like a covalent bond in certain situations.[1–5]

#### 1.1.1.2 Short Strong Hydrogen Bonds

A short strong hydrogen bond (SSHB) is a subset of hydrogen bonding often displaying interesting characteristics such as temperature dependence where the interaction is stronger than a standard hydrogen bond.[4] An SSHB is defined as a hydrogen bond with a very short donor-acceptor distance, typically less than 2.60 Å where the acceptor is nitrogen, or 2.50 Å where the proton acceptor is oxygen, resulting in the formation of a hydrogen bond with partial covalent character.[7,8] These bonds are known to have an increased bond energy over standard hydrogen bonds of up to 40 kJ $mol^{-1}$.[4] Figure 2.2 shows how the proton potential energy surface within a hydrogen bond changes as the donor-acceptor distance decreases, changing from a typical double well potential favouring a single site, to a broader potential energy surface with more possibility for environmental or chemically induced changes to the proton position.

#### 1.1.1.3 Salts, cocrystals and the salt-cocrystal continuum

A very important category of organic crystals are pharmaceutical drugs and their delivery mechanisms in the body.[9] Knowing the precise structure of drugs and their components is essential for the characterisation and predictive modelling of their chemical properties, which is a key stage in their development cycle.[9] The position of the hydrogen atom in the form of a cocrystal or salt has a large effect on the chemical properties of the drug, and can affect properties such as the drug release profile. [9–11]

Active pharmaceutical ingredients are often prepared through crystallisation with a second non-active ingredient to achieve the required characteristics for drug delivery in the body. These 2-component crystals are most commonly held together by hydrogen bonding between the two components and depending on

the interaction are defined as salts (proton transfer from donor to acceptor) or co-crystals (proton remains covalently bonded to the proton donor). Where a hydrogen bonding interaction is present, the exact nature of the hydrogen bond can have significant effects on the physicochemical properties of the crystal structure which can be vital in pharmaceutical applications, and the choice of second component is made to optimise these characteristics.[12–17] The important feature to characterise is the relative position of the proton within the hydrogen bond, which is dependent upon the shape of the proton potential energy surface between the two electronegative atoms, as shown in Figure 2.2.[6]

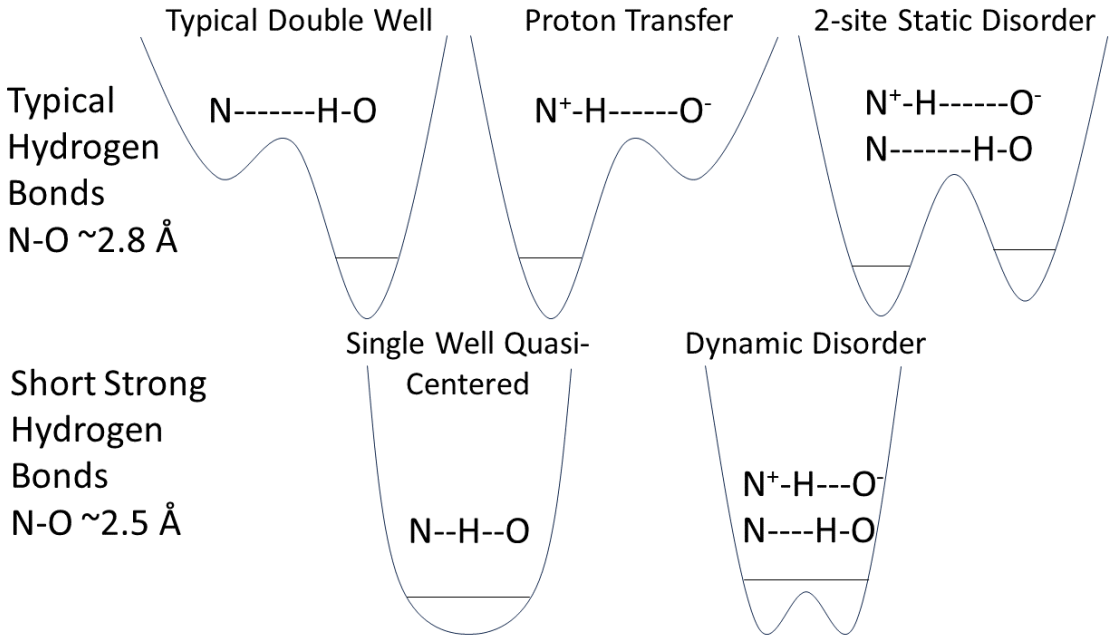



**Figure 2.2** Typical proton potential energy surfaces for hydrogen bonds. The most common situations are a typical double well and proton transfer, while the interesting characteristics of single well and dynamic disorder are observed in short strong hydrogen bonds.

The two most common cases consist of the typical hydrogen bond double well, and the proton transfer double well. The typical hydrogen bond double well potential is also referred to as the co-crystal structure, where the proton remains covalently bonded to the donor atom.[14,18] This results in a typical hydrogen bond interaction where the $\delta^+$ charge on the proton is attracted by the lone pair of electrons on the nitrogen acceptor. In the proton transfer case, the proton has transferred to the acceptor atom, and a negative charge left on the donor and

positive on the acceptor.[6] The resulting crystal is often referred to as a salt, due to the ionic nature of the interaction between positive and negative charges of the components. Where these two potential energy wells become close in energy, 2-site static disorder is possible where each of the salt and co-crystal structures are observed in a proportion dependent on the depth of the potential energy wells. However, in a standard strength hydrogen bond, the barrier between the potential wells is sufficient to prevent any dynamic processes and the proportion generally remains the same without venturing into extreme temperatures to overcome that barrier.[6]

In the short strong hydrogen bonding region where the donor acceptor distance is short there are two further possible situations. As the donor and acceptor atoms become closer, the barrier in a disordered system with two similar well depths becomes smaller, which at a certain point leads to dynamic disorder, whereby the proton has sufficient energy to hop between the two potential wells.[18] In this case, any change in chemical or physical environment will have an effect on the relative populations of each state meaning the properties of the crystal may change, leading to the temperature dependence often associated with SSHBs.[14,18,19]

The final possibility is that the proton potential energy surface approaches that of a single well. In a single well potential the proton is likely to be found towards the centre of the hydrogen bond, and is referred to as a quasi-centred hydrogen bond, although the potential well is often much broader in this case so a continuum of states around the centre of the potential well is also possible.[18,19] The nature of the bonding between the proton and both the donor and acceptor atoms is essentially the same in this case, with significant covalent character.[6] Here the proton often has a very low barrier to any motion, meaning environment dependent properties are frequently observed, normally associated with a change in location of the proton within the hydrogen bond again demonstrating temperature dependence.[19]

A term often quoted in literature is the salt co-crystal continuum, highlighting the continuous nature of the hydrogen bonding interaction dependent on the position of the proton, which specifies salts and co-crystals as the two extremes

with a proton residing towards the centre of a hydrogen bond forming a continuum between these two states. The continuous nature of this interaction has been demonstrated in measurements using various techniques.

### 1.1.2 Pharmaceuticals

Regulatory approval of new drugs involves a full characterisation, including the position of hydrogen atoms where this is important. Where the hydrogen bonding has an essential part to play in the drug molecule's properties, being able to accurately determine the hydrogen atom position will be vital for securing and defending patents.[9]

### 1.1.3 Traditional Methods of Characterisation

Crystallography is the umbrella term for methods used to determine a crystal structure. There are many techniques available to use if investigating the structure of organic crystals each of which can be used to access a variety of information about the crystal structure being investigated. The most commonly used methods in crystallography include X-ray diffraction, neutron diffraction and solid state nuclear magnetic resonance. As discussed below, each of these has advantages and disadvantages with regards to hydrogen bond characterisation and each provides access to different and important structural information.

#### 1.1.3.1 $\Delta pK_a$ approach

An easy way of predicting the character of hydrogen bonding is to determine and compare the acid dissociation constants of the donor and acceptor molecules at the interacting electronegative groups.[20] The shape of the potential energy surface of the hydrogen bond is directly related to the relative ability of the acid and base to associate with the hydrogen atom, with a strong acid (low pKa) relative to the base resulting in the formation of a salt and a relatively weak acid forming a cocrystal.[20] This relative difference in acid dissociation

constant therefore provides an empirical assessment of the likelihood of proton transfer from an acid to a base in an organic crystal structure.

$$\Delta pK_a = pK_a(\text{base}) - pK_a(\text{acid}) \quad (1)$$

Where $\Delta$pKa < 0 the structure will form a co-crystal (no proton transfer) and where $\Delta$pKa > 3 the structure will form a salt (proton transfer).[20,21] This empirical approach has been shown to be accurate for these cases almost universally. However, as with many empirical approaches to characterising a property, there is an intermediate range in which it is not possible to determine the nature of the hydrogen bonding using this method, where 0 < $\Delta$pKa < 3. When investigating organic crystals, this region is relatively common and an alternative approach is required to characterise these structures.[20,21]

#### 1.1.3.2 X-ray Diffraction

X-ray diffraction (XRD) is a method often carried out to determine crystal structures and determines the mean atomic positions in a sample. XRD can be carried out using a synchrotron radiation source for high resolution measurements, although lab sources are very widely available making the method very accessible as the standard approach to crystallography.[22] XRD makes use of X-ray diffraction patterns to determine atomic positions within a crystalline structure, making use of the repetitive nature of crystal structures to maximise the diffraction effect observed through constructive and destructive interference.[22] For the best diffraction patterns, a single crystal sample is required, which can be difficult to grow for some materials meaning powder diffraction is normally the simplest and easiest method.[22] The use of single crystal XRD will lead to a more accurate, higher resolution structure determined, and therefore it is important to account for the quality of the data in any XRD measurements, analysis and literature, particularly where hydrogen atoms are of interest.

As X-rays pass through a sample, they interact with the electron density within the structure, which cause constructive and destructive interference and the formation of a diffraction pattern unique to a particular electron density

distribution and unit cell.[22] In powder XRD measurements, this can lead to substantial peak overlap making structure determination difficult for molecules with large unit cells or low symmetry. For single crystal XRD measurements, the diffraction pattern is measured at a range of orientations, which removes the degeneracy of diffraction peaks making structure determination easier where single crystals can be grown. Using Fourier transforms, it is then possible to construct a three dimensional model of the electron density within the crystal, which in turn can be used to determine the mean atomic positions.

The quality of the diffraction data defines how accurate these mean atomic position will be. In addition to the data quality, a larger atom, such as C, with 6 electrons will have a much larger diffracting effect than a hydrogen atom. Therefore, the deviation from the mean is generally very small, and a good quality XRD structure is generally very close to the true structure for C, N, O.[22] However, due to being very small, relatively mobile and containing only a single electron, the sensitivity of X-ray diffraction data to hydrogen is low.[22] In many cases the exact hydrogen position is not very important; often there are only terminal hydrogen atoms present. However, in certain circumstances - hydrogen bonding in particular - the exact proton position is vital for the understanding of the structural and chemical properties of the crystal. Because of this limitation in the approach, various proton position refinement methods have been devised to ensure the proton is accurately described.[14,23] A common approach is by using specially designed refinement software to find the most likely proton position based on the measured electron density, and in the majority of cases this works well enough. However, a study comparing this refinement to accurate neutron diffraction measurements showed that the majority of the bond lengths were systematically underestimated, highlighting the fact that this model is working with data which is not directly sensitive to the hydrogen atom position. Therefore, while this approach may be sufficient to characterise a structure as salt or co-crystal, it is not sufficiently accurate or sensitive to hydrogen to be able to describe small but important changes in hydrogen bonding interactions.

The best practice method of hydrogen atom structure refinement for single crystal XRD is through the use of a difference density map. The experimental electron density map is compared to a theoretical model of the electron density map based on the structure.[24] The difference density map is determined by subtracting one from the other. The proton positions can then be fully refined to minimise the difference density map. As the proton electron density is low, there is still a level of uncertainty with this method, but is nonetheless widely used as best practice and generally agrees well with neutron diffraction measurements.[24–26]

#### 1.1.3.3 Neutron Diffraction

Neutron diffraction techniques are less commonly used than XRD, due primarily to the fact that neutron sources are less common and more difficult to access.[27] Neutron diffraction data is often used to provide complementary information to XRD studies.[26,28] Like XRD, neutron diffraction measures the diffraction pattern of a crystal structure when subject to a beam of neutrons. In XRD, the X-rays scatter off the electron density, whereas in neutron diffraction, the neutrons scatter directly off the atomic nuclei, resulting in effects due to different isotopes present, allowing selective measurements to be taken.[28] Due to their higher mass, neutrons generally work better than XRD on powder or non-crystalline structures due to the increased penetration depth, although require much larger samples.[28] The important feature of neutron diffraction for crystallography is the ability to directly scatter off hydrogen, meaning the location of the hydrogen atoms can be determined directly from the data, rather than by atom refinement as in XRD.[26,27] The flux of a neutron beam is generally low, meaning that the crystal size required is relatively large, at least $1mm^3$, and measurements of single crystals is difficult due to these size requirements.[27] However, it is possible to reduce the required crystal size by partially deuterating the sample to increase the interaction cross section.[28] On a neutron Fourier difference map, a hydrogen appears as a hole due to its negative scattering length, meaning it is very easily identified.[27]

The major disadvantage of neutron diffraction over alternative methods is the required sample preparation. In most cases, to make the interaction cross sections sufficiently large partial deuteration is required.[27] The preparation of a crystalline sample can be difficult with this condition depending on the structure of interest.

There have been many investigations of hydrogen bonding structure using neutron diffraction. One example is the investigation of short strong hydrogen bonding in pyridine 3,5-dicarboxylic acid with neutron diffraction.[29] Here proton migration was observed as a temperature dependent property with the N-H bond distance increasing from 1.21 Å to 1.31 Å over a temperature range of 15 K to 300 K, as shown in figure 1.3.[29]

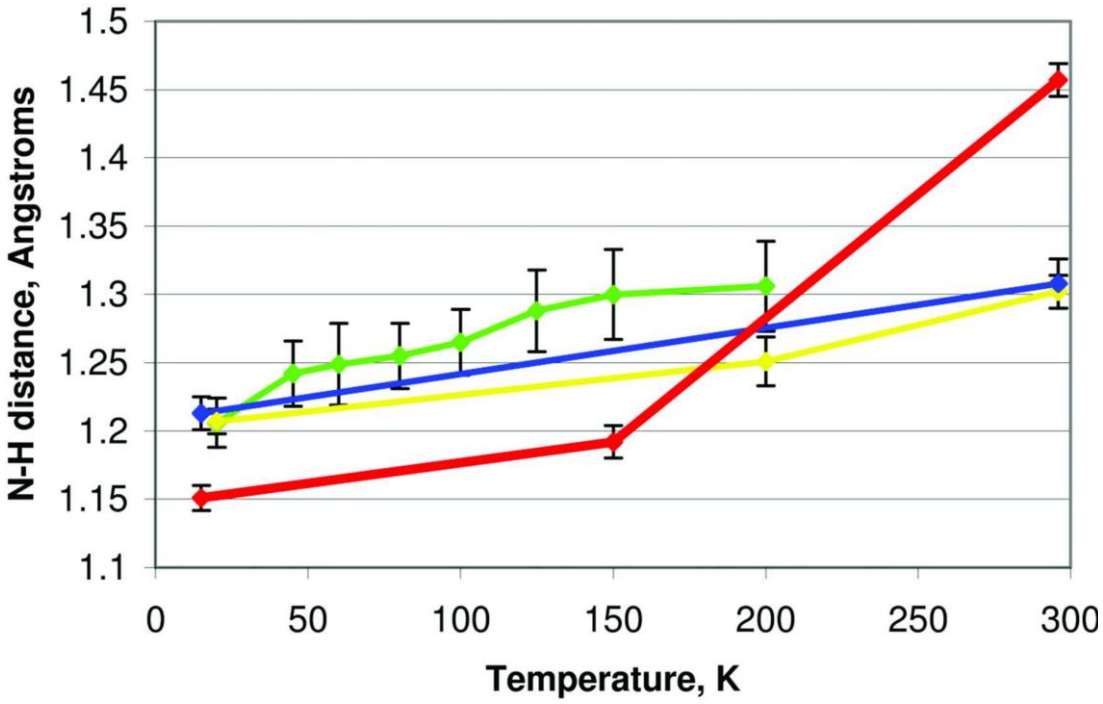



**Figure 1.3** Change in proton position investigated using neutron diffraction. Blue line is 3,5-dicarboxylic acid, red line is deuterated 3,5-dicarboxylic acid. Error bars on deuterate sample are smaller due to greater certainty of the position, but the data suggests that deuterating the sample has an effect om the structure so not directly comparable to the non-deuterated sample. Figure reproduced from ref. 29.

A fully deuterated version of the same molecule was also investigated to determine the effect of deuteration.[29] This was shown to increase the N-H distance within the hydrogen bond by 0.15 Å at 300 K, while reducing the distance by 0.06 Å as 15 K. This clearly has an effect on the hydrogen bonding interaction. However, the overall length of the hydrogen bond only increased by 0.015 Å at 15 K so the effects on the wider structure were minimal, but present. Therefore, as with other methods, there is still a significant trade-off between the most accurate data (the pure sample) and making the measurement feasible through approaches like deuteration.

Where used in conjunction with XRD, the neutron diffraction data is often used primarily to locate the hydrogen atoms, allowing the more accurate XRD data to locate the atoms of larger mass and electron density.[30] Overall, neutron diffraction is a more reliable method for determining proton positions in crystal structures than XRD, but may have reduced accuracy for other components where deuteration is used and as such the best studies combine these techniques to accurately determine the entire crystal structure.[26,30]

#### 1.1.3.4 Solid State Nuclear Magnetic Resonance (ss-NMR)

Solid state nuclear magnetic resonance (ss-NMR) is an alternative structural approach directly sensitive to the local nuclear environment, including the molecular, electronic and crystal structures.[31] Nuclear magnetic resonance gives rise to structural information through primarily nuclear properties, including magnetic shielding and spin-spin coupling. For these interactions to be visible in the spectrum, NMR relies upon the applied magnetic field to remove the degeneracy of the nuclear spin states.[31] These interactions then give rise to the key features of an NMR spectrum, chemical shifts and peak splitting, which are used to help determine the crystal structure.[31] The chemical shift arises from the interaction between a nuclear spin and the local electronic environment.[31] Using the chemical shift, functional groups can be identified. Spin-spin coupling originates from an interaction between two nuclear spins but also dependent on the local electron spins. This allows bonding to be determined, for example the bond order of a chemical bond between a pair of

C. Where the crystal structure is already known (for example through XRD for non-hydrogen atom positions), this makes peak assignment relatively easy.[31,32] However, where the crystal structure is unknown, or there are multiple conformations possible, peak assignment can be challenging and another method, either experimental or through modelling, is required. In these cases, the ss-NMR data is used as a constraint on the structure found via the alternative method.[12,31,32] The field known as NMR crystallography uses this approach by comparing experimental NMR data with calculated NMR resonances to validate, or otherwise, powder XRD data.[32] Where the calculated chemical shifts are in good agreement with experiment, this provides validation of the powder XRD structure. In these experiments, all relevant atoms are generally probed resulting in $^{1}H$, $^{13}C$ and $^{15}N$ data to compare and check for consistency.[32] When investigating SSHBs, the proton environment will exhibit a high frequency chemical shift in the $^{1}H$ NMR spectrum, due to the short nature of the hydrogen bond.

It has also been shown that the $^{1}H$ chemical shift of a proton involved in an oxygen donor nitrogen acceptor hydrogen bond can be used to determine the H-O bond distance in proteins.[33] Here the relationship was not found to be linear, but a third order polynomial.[33]

$$\Delta\delta = ad^{-3} + b \qquad (2)$$

where a and b are parameters to be fitted. Using values of a=19.2 and b=-2.3 a root mean square deviation of 0.2 Å was found in the experimental data.[33] This is not a particularly quantitative method, but the trend in the data is clear, increased chemical shift means shorter bond distance, and this applies to all hydrogen bonding.

Another example of work done using NMR in combination with other approaches is a study done on a photochromic anil using XRD and DFT to investigate the excited state causing the colour change.[12] Here it was confirmed that the key structural difference is in the hydrogen bond, where the proton transfers from the donor to the acceptor.[12] The method involved using the XRD structures as a starting point for DFT calculations, which were then optimised to ensure the theoretical NMR chemical shifts resembled the experimental $^{1}H$

NMR data, which is directly sensitive to the proton chemical environment. This was used to show that the actual structure was highly dynamic with fast rotations of various groups, information not available directly from XRD or DFT alone.[12] Again, this goes to show how useful combining multiple methods in the analysis of hydrogen bonding environments is and how it can lead to information which is otherwise not accessible. The use of $^{1}$H NMR alone to investigate crystal structures is relatively rare, but is a very useful tool to further constrain data obtained through other methods.[12,34,35]

## 1.2 Aims and Objectives

The primary aim of the research project was to examine the potential application of core level spectroscopy techniques to the field of crystal structure refinement as an alternative technique with desirable characteristics including ease of sample preparation, ease of measurement and direct sensitivity to hydrogen atom position. There are a number of issues highlighted above with the traditional crystal structure characterisation techniques which make each of them sub-optimal to use independently and expensive to combine multiple approaches to achieve the accuracy required in hydrogen bonded systems. Core level spectroscopy has been shown to exhibit significant sensitivity to atomic positions in the local chemical environment, along with the techniques being non-destructive, requiring a very small quantity of sample and measurements being rapid. The intention is to utilise this sensitivity to characterise organic pharmaceutical crystal structures through a combination of experimental spectroscopy with computational density functional theory calculations. The success or otherwise of this approach can easily be evaluated against the traditional methods of structure refinement for accuracy.

Utilising a series of test structures, firstly the method of structure refinement will be developed into a consistent and well defined process to ensure the information required can be obtained in a consistent manner. This work will build on previous investigations into the sensitivity of core level spectroscopy to hydrogen atom position and potential energy surfaces to identify the key

aspects of experimental spectra in allowing this analysis method to be developed.

An additional aim is to further understand and identify the underlying interactions at play within a hydrogen bonded system, in particular to identify the cause of spectral features and peak shifts with differing hydrogen atom positions, which would further improve our understanding of the hydrogen bond and its importance within these systems, while also explaining the sensitivity of core level spectroscopy techniques to this position.

The final aim of the research project was to challenge this new approach on a series of structures with inaccurate hydrogen atom positions and to refine these positions. This approach builds on previous work whereby the core level spectroscopy is being utilised to not only identify hydrogen bonding interactions, but to refine hydrogen atom positions in combination with density functional theory to show how sensitive he approach is and the ability to identify incorrect hydrogen atom position in molecular structures.

## 1.3 Experimental Techniques

The two experimental techniques detailed in this section are proposed as alternative methods of crystal structure refinement of the hydrogen atom positions. Each has been shown to exhibit significant and consistent sensitivity to the local chemical environment including visible shifts in spectra based on the position of a proton within a hydrogen bonding interaction.

### 1.3.1 XPS

X-ray Photoelectron Spectroscopy (XPS) is used to probe the core level electron binding energies of each of the components within a molecular structure.[36] The core level binding energy of an electron is the energy required to remove that electron from the molecule. While electrons in the core energy levels of atoms do not directly participate in any bonding interactions, the core level binding energies are highly sensitive to the local chemical and electrostatic

environment. The fundamental processes occurring in XPS involves the photo-emission of a core level electron due to the absorption of a photon, via the photo-electric effect.[36] This states that an electron with binding energy less than the energy of the incoming photon can interact and result in photoemission.

$$E_{kin} = h\nu - E_{BE} \quad (3)$$

In the gas phase, the kinetic energy ($E_{Kin}$) of the electron is equal to the difference in energy between the binding energy ($E_{BE}$) and the incoming photon ($h\nu$), as shown in equation 3. Equation 4 shows the modified expression for a solid, where the surface work function, φ, corresponds to the energy difference between the Fermi level, and the vacuum, resulting in additional energy loss when compared to gas phase samples. While it is important to factor the value of φ into any analysis of XP spectra, for any sample it will be a constant throughout the range of kinetic energies measured, so is effectively accounted for during the binding energy scale calibration of the spectrum.[37]

$$E_{kin} = h\nu - E_{BE} - \varphi \quad (4)$$

Photoelectrons therefore have a kinetic energy directly related to the binding energy of the core level from which it was emitted, with hv and $E_{kin}$ known and measured values respectively. The electrons being detected are from three primary sources, unscattered photoelectrons, Auger electrons and scattered photoelectrons.[36] The unscattered photoelectrons retain the kinetic energy they left the sample surface with by not colliding with anything, made possible with the relatively large mean free path for these electrons.[37] When detected, these electrons correspond to the photoelectron peaks in the spectrum. Inelastically scattered photoelectrons contribute towards the background in the spectrum, having lost an unknown amount of kinetic energy and contain no useful information.[37] These are generally those electrons which originate from deeper within the sample, so are more likely to scatter before reaching the sample surface. Auger electrons are also observed in XP spectra, and originate from the relaxation of the system. On photoemission, a core hole state is formed, such that the 1s orbital contains only a single electron. This is a very unstable state, which decays quickly through an electron transition from a higher energy

level. When a higher energy level electron fills the core hole state, there are two main processes which can occur to dissipate the excess energy, shown in Figure 1.4. The first involves emission of a photon through fluorescence.[36] The second process is an Auger emission, whereby the electron transfer into the core level causes the emission of another valence electron to ensure energy and momentum are conserved.[36,38] When this process occurs, Auger electrons are detected, and these have specific energies corresponding to the energy level from which the core hole state was filled. For low Z atoms, as found in organic crystals, the primary relaxation method is via Auger emission.

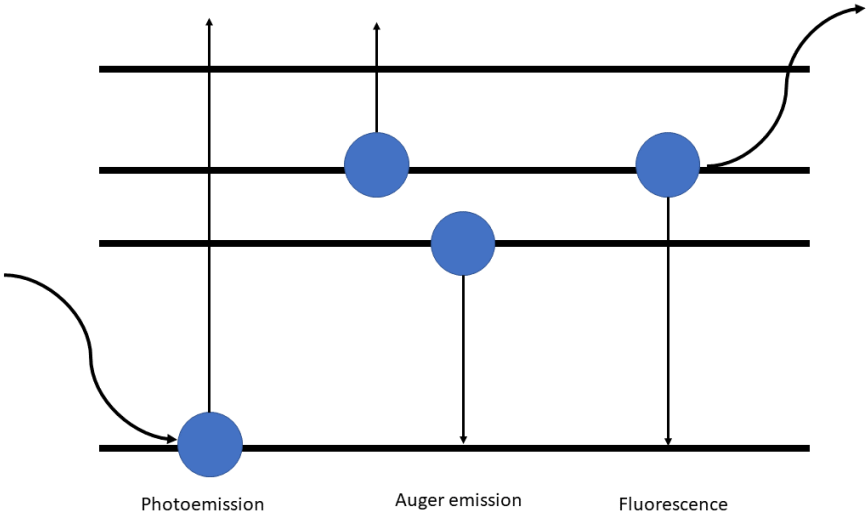



**Figure 1.4** Processes occurring in XPS. (a) Photoemission on absorption of a photon, (b) Auger emission to refill the core hole and (c) fluorescence to refill the core hole.

#### 1.3.1.1 Experimental set up

An XP spectrometer consists of many components. The most important of these are the X-ray source, the sample chamber (and any required vacuum pumps), hemispherical analyser and the detector.[37] An XP spectrometer uses a hemispherical analyser to determine the kinetic energy of any emitted electrons. A hemispherical analyser distinguishes between electrons with different kinetic energy by applying a known electric field to bend the path of the electrons between the two hemispherical electrodes, meaning the radius of the curved

path depends on the kinetic energy of the electron.[37] By varying the electric field, the full range of kinetic energies can be measured. The X-ray source can be a lab source, such as an Al or Mg $K_\alpha$ source, or can be from a synchrotron. For XPS measurements, only a single energy of X-rays is required, and must be a known constant to calculate the binding energy, as shown in equation 4.[37] This makes XPS an ideal use for lab sources, where only a single frequency is normally available from a single source. For the highest resolution spectra, synchrotron radiation is used as the X-ray beam can be tuned to the optimum energy for a specific sample, for example metallic samples often require a different incident X-ray energy to organic samples to maximise the signal to noise ratio.[37] Synchrotron radiation also provides a more focused and intense beam, which decreases the width of any observed peaks increasing the resolution of the spectra.[37] The sample chamber is generally kept under vacuum to maximise the mean free path of emitted electrons, maximising the measurement of useful unscattered electrons against those which have inelastically scattered.

X-rays have a certain penetration depth into the sample from which information can be obtained. This region is close to the surface due to the nature of X-ray energies used for XPS.[37] However, the penetration depth is normally sufficient to observe the properties of the bulk substance, with a few nanometers of depth probed, beyond the surface region and any surface structural effects. While surface structure analysis is the primary use of XPS, where surface properties are investigated, due to this penetration depth being on the nanometer scale, properties of the bulk sample can also be investigated, such as the bulk properties of organic crystals.[39]

An important consideration when making XPS measurements is charge compensation. The underlying processes involves the emission of electrons from a small areas of the sample. As the measurement progresses, this would lead to a build-up of positive charge at the surface.[40] This would then provide a greater attractive force to the remaining electrons in the sample, affecting their binding energy and shifting any peaks.[40] When this happens, a spectrum becomes very difficult to analyse and fit appropriately. This effect is then

exacerbated through differential charging. Differential charging occurs where the X-ray spot size is large and non-uniform, meaning that the centre of the spot charges at a different rate to the extremities.[40] This leads to a charge gradient within the measurement area leading to different charging effects at each point, making charging totally unpredictable.[40] These issues only occur in insulating samples, including organic crystals, and therefore an additional charge compensation must be included in the experimental setup to ensure these effects are minimised.

#### 1.3.1.2 Ultra High Vacuum XPS

Ultra-High Vacuum XPS (UHV-XPS) is the most common form of XPS measurement carried out, ensuring minimal background interference to the signal through ensuring the sample chamber is under ultra-high vacuum. This can provide some additional challenges, particularly for samples with a tendency to outgas under these extreme conditions. In UHV-XPS, charge compensation is provided by use of an electron flood gun. Throughout the measurement, low energy electrons are flooded into the measurement chamber, and are attracted to the positive charge.[41] This constant supply of low energy electrons therefore compensates for the electron emission by attempting to keep the sample neutrally charged.[41] For highly insulating samples, correctly charge compensating using a flood gun can be a challenge. This is because the number of electrons needs to be equal to the charging effect of the photoemission, otherwise the system becomes negatively charged which is as difficult to deconvolute.

#### 1.3.1.3 Near Ambient Pressure XPS

Near Ambient Pressure XPS (NAP-XPS) attempts to measure the same underlying XP spectrum as traditional UHV-XPS system, but under conditions much closer to everyday life meaning many more samples are suitable for measurement, including liquids and solutions. There are several advantages to this approach including the requirements for charge compensation.

Disadvantages to the NAP-XPS method include a lower signal to noise ratio due to the background balance gas in the sample chamber and interference of certain gasses with spectra making normalisation difficult where this occurs, although this can easily be avoided through use of an appropriate gas phase. Figure 1.5 provides a schematic diagram of the NAP-XPS experimental set up showing the key components. In NAP-XPS, charge compensation is provided through ionisation of the gas phase by the incoming X-ray beam and photoelectrons. As the beam ionises the gas phase, the free electrons are

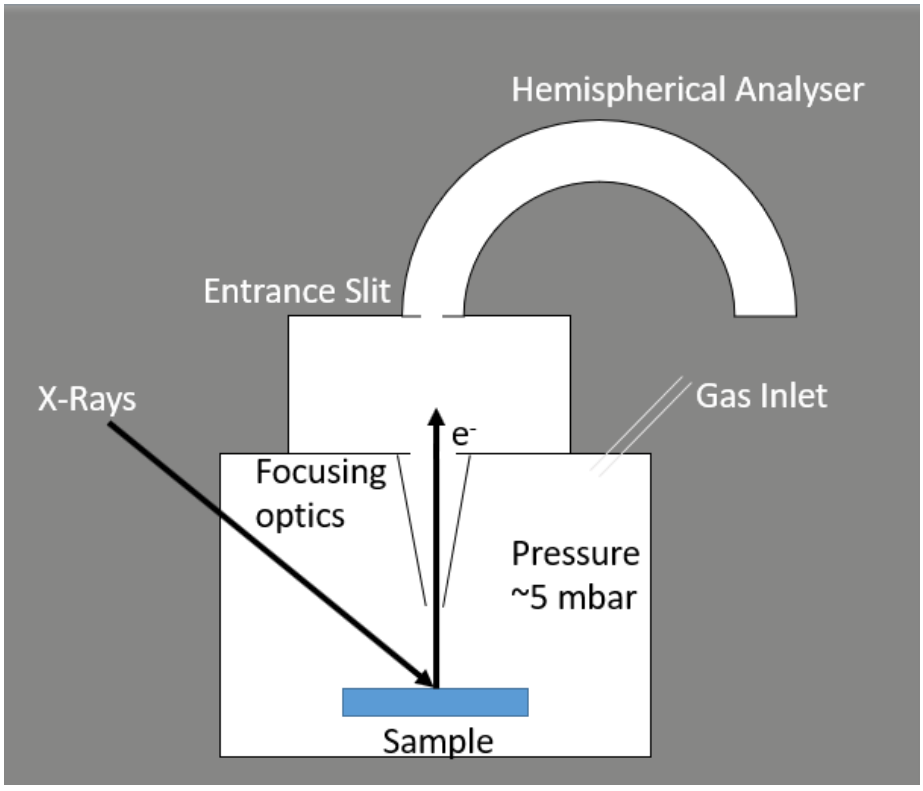



**Figure 1.5** Schematic diagram of a NAP-XPS experimental set up showing the key components. The choice of gas phase and the gas phase pressure are decided to avoid peak clash with the sample.

attracted to the charged surface of the sample, forming an electron cloud above the surface which provides compensation for the electron emissions to keep the sample neutrally charged. In general, this gas phase ionisation effect is easier to control as the free electrons form a cloud above the surface and do not directly affect the sample or any electron binding energies. The level of charge compensation provided by the gas phase is determined by the pressure of the gas phase in the measurement chamber; if there are more gas particles, then

more will be ionised per unit time. Overall, the choice of UHV vs NAP depends very much on the sample and the measurements of interest and whether quantitative or qualitative analysis is required.

#### 1.3.1.4 Spectrum processing and calibration

When carrying out XPS measurements, two different types of spectra are taken, a survey spectrum and then a series of high resolution spectra. Each of these measurements provide different but important information when analysing the data. The survey spectrum gives details about the relative abundances of the different atoms in the sample and can be used to identify any contamination on the surface, vital pieces of information for spectrum processing. The high resolution spectra are used for peak fitting of the individual atomic spectra to identify the binding energies of the different molecular environments. XPS is a fully quantitative technique. This means that the absolute values of the relative atomic abundances and peak positions can be determined entirely from an XP spectrum. Before peak fitting analysis can be attempted on XPS data, it is necessary to calibrate the spectrum correctly to ensure relative peak heights and binding energy scale are accurate to allow this quantitative analysis to take place. To do this, the spectrometer transmission function and a known binding energy scale reference are required. The transmission function essentially defines the efficiency of electron transmission of the specific instrument at a given measurement energy. This varies for every spectrometer and not in a linear fashion, although is often automatically applied by the analysis software.

In metallic and conducting samples, the Fermi level is easily obtained, and this has a fixed binding energy which is set to 0 eV allowing a confident calibration. Finding an appropriate binding energy scale reference is more challenging for insulating samples. Due to effects of the local chemical environment, it is not possible to consistently state that a certain atom type has a certain binding energy. There are several proposed methods in literature to calibrate the binding energy scale, and we utilise the most common approach of calibrating to the peak associated with adventitious carbon contamination on the sample surface.

All samples exposed to the air build up a layer of contamination on the surface, primarily consisting of carbon, which is known as adventitious carbon.[42] As this layer is not bonded to the sample, the binding energy is consistent across samples and is therefore used a binding energy scale reference. There are a range of values used in literature for the adventitious carbon binding energy, commonly used values are 284.8 eV and 285 eV.[42–44] When analysing XPS data, the absolute value of this calibration is not important, only that it is reported clearly to allow comparison and adjustment if necessary when analysing.

Using the survey spectrum, it is possible to calculate the relative abundance of each atom type in a sample. By comparing this with the expected stoichiometry, it is easy to determine the excess carbon attributable to adventitious carbon. Using this value, the carbon 1s high resolution spectrum can be fitted accordingly with the correct proportion of the signal attributed to adventitious carbon and calibrated to 285 eV, with the remaining signal fit as normal. Over a wide variety of samples, this approach has worked well and ensures consistency between the survey spectrum, high resolution spectra and the expected molecular structure.

#### 1.3.1.5 Peak Fitting and analysis

Using CasaXPS[45], Gaussian Lorentzian convolutions are fitted to the spectrum using the automatic peak fitting procedure.[45] The key component of peak fitting is ensuring consistency between all aspects of the data including the survey spectrum, the high resolution scans and the expected molecular structure. If all of these data are consistent then a good fit has been attained.[45] In general, the fitting parameters require a certain level of constraining to ensure the determined fit is realistic. The key features to ensure are consistent are the relative intensities of each peak relative to the stoichiometry of the crystal and that the full width half maximum (FWHM) of each of the peaks is similar. Unless there is a physical reason for the peak width to differ, all fitted peaks should be roughly the same width within the same spectrum.[45] By constraining the fit in these ways, we guarantee consistency with the survey spectrum and expected

stoichiometry. If a fit is not possible within these constraints, it implies there is something additional going on in the system which requires more investigation, which could be surface effects or a difference to the expected stoichiometry in the bulk sample.

The binding energy of core electrons is highly dependent on the chemical surroundings, particularly where electronegative atoms are present, and can result in BE shifts of up to 10 eV.[36,46] As the technique is accurate to ± 0.1 eV, this means the binding energies determined are very sensitive to small changes in the chemical environment, including hydrogen atoms.

#### 1.3.1.6 Literature using XPS to investigate hydrogen bonding

The most common uses for XPS are in surface analysis and structure determination making use of binding energy shifts to determine quantitatively any present functional groups and bonding environments to determine the molecule structure quantitatively.[34,47–50] A small amount of work has previously been done investigating hydrogen bonding using XPS, with some clear advantages over other methods. Figure 1.6 shows the relationship between N-H bond length and N 1s binding energy for a series of bipyridine complexes. While this is only a small sample of data, there is a very clear trend in the data, with a clear separation between the salt and co-crystal structures, with the salt structures sitting at around 401.5 eV and the co-crystal structures at around 399.5 eV. This shift of 2 eV can be attributed to the different types of hydrogen bonding present in the structures, with the proton residing closer to the donor (co-crystal) or acceptor (salt) and the electrostatic effect of the position of the positive proton on the core level binding energy of the nitrogen acceptor.[51]

Importantly, the NEXAFS (Near Edge X-ray Absorption Fine Structure) photon energy gives an almost identical trend with photon energies approximately 2 eV lower in each case, showing how these two methods can complement each other. This is the first indication that the XPS method is information rich compared to standard crystallography, with changes in the spectra due to change in hydrogen position present even though the H atoms are not directly probed. It has been shown that the XP spectra are dependent on the position of

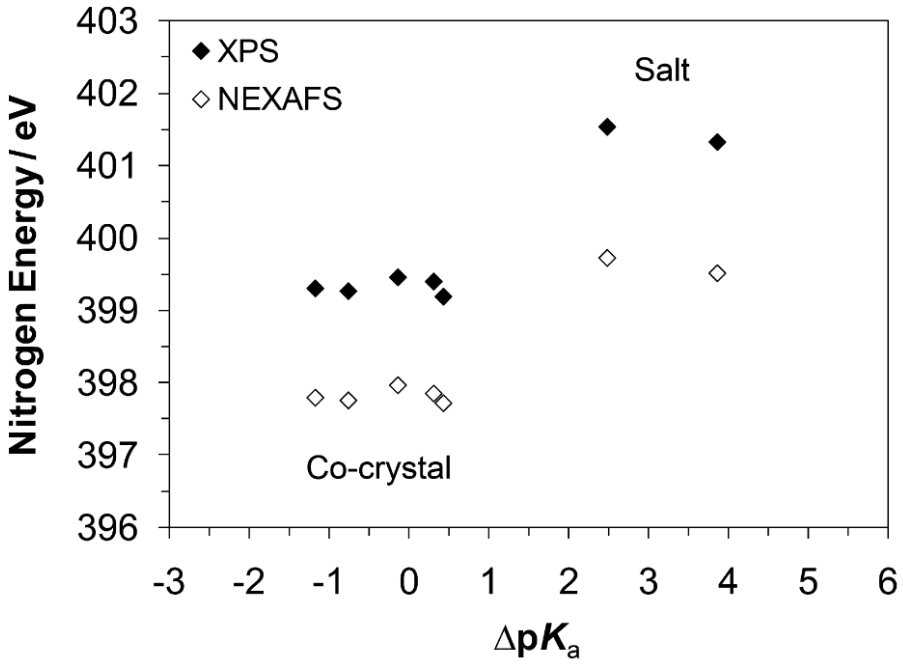



**Figure 1.6** N 1s Binding energy plotted against N--H distance for a range of crystal structures with salt and co-crystal interactions[51] . A clear distinction between salt and co-crystal is evident with a shift in binding energy. Figure reproduced from ref. 51.

a proton within a hydrogen bond, and can be observed in the spectra as either a shift and/or broadening of the peaks.[51–53] While this can give evidence of a change in the hydrogen bond structure from co-crystal to salt, it is difficult to predict with any certainty the precise position of the proton based entirely on XPS. [53]

A much larger set of 17 2-component organic crystal complexes was studied using XPS in conjunction with XRD and ss-NMR as additional sources of information.[35] Here the aim was to demonstrate that the use of XPS combined with ss-NMR provided the same information as alternatives using a simple approach. Of the 17 structures investigated, the most interesting are the isonicotinamide and theophylline complexes, showing significant spectral differences depending on the nature of the hydrogen bonding interaction. Figure 1.7 shows how much of an effect the hydrogen bonding has on the local environment of the adjacent atoms. For each of the N atoms not adjacent to the hydrogen bond, there is essentially no change in binding energy, even though the structures being investigated are different.[35]

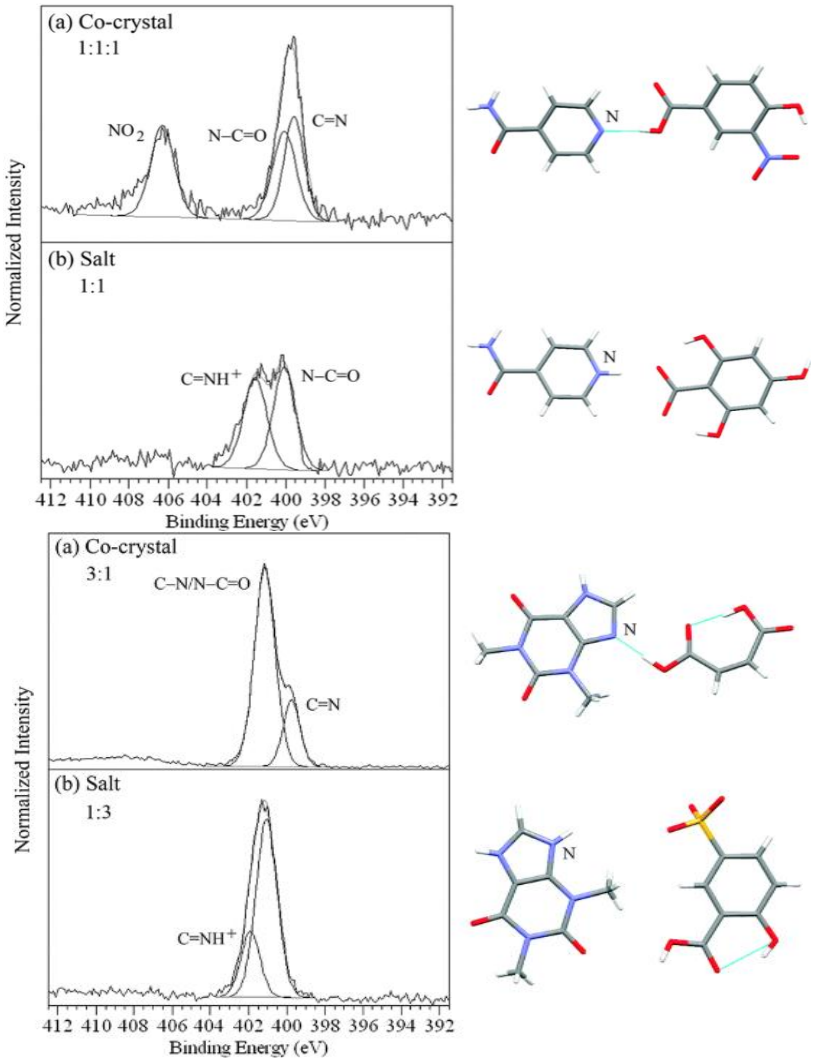



**Figure 1.7** N 1s XP spectra of isonicotinamide and theophylline complexes with different hydrogen bonding characteristics[35] Figure reproduced from ref. 35

However, for the atoms involved in the hydrogen bond, the salt structure shows an increase in binding energy for the relevant component of around 2 eV, again from around 399.5 eV to 401.5 eV, as seen previously.[35,51] The data obtained for this publication showed that the XPS data agreed directly with the ss-NMR/DFT data for all 17 structures, and correctly identified the nature of the hydrogen bonding present in each complex. [35]

An investigation into pyridine dicarboxylic acid 1-component crystals using XPS highlights the ability to observe changes in the salt cocrystal continuum in the quasi centred proton region. Three pyridine dicarboxylic acid complexes were investigated, each containing a different type of hydrogen bonding despite

having the same chemical formula.[19] The three complexes were measured using XPS to determine the effect of short strong hydrogen bonding on the N 1s emission line, in particular a quasi-centred type hydrogen bonding structure. Each of the three structures contain only a single nitrogen environment, making the analysis of the spectra relatively easy to interpret in terms of the hydrogen bonding occurring in the structure.

The three spectra shown in Figure 1.8 are clearly different, and show the same characteristics as all of the previous examples, with the salt structure sitting at around 401.5 eV and the co-crystal at 399.5 eV. In comparison to these peaks however, the third structure shows a completely different lineshape, and peak intensity.[19] From previous work done on this structure using neutron diffraction, it is known to exhibit some form of short strong hydrogen bonding behaviour due to the short N…O distance.[19,29] Of the two possibilities, dynamic disorder or quasi centred, dynamic disorder is easier to model, since there are still only two potential wells resulting in two main peaks, one for the co-crystal, another for

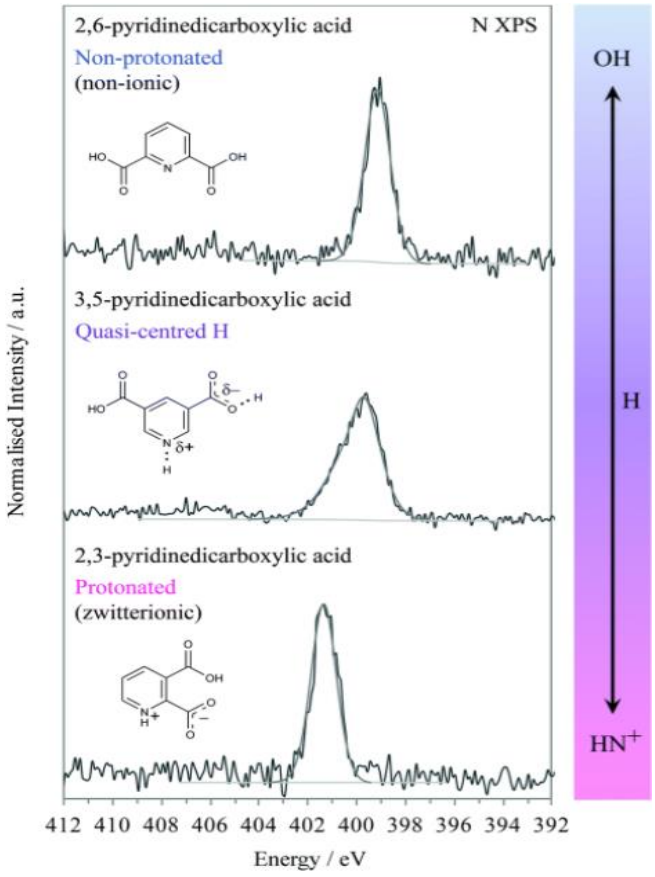



**Figure 1.8** N 1s XP spectra for the three PDCA structures displaying significant differences in both peak position and peak shape.[19] Hydrogen bonds are between the pyridine nitrogen and carboxylic acid group in each case. Figure reproduced from ref. 19.

the salt. Using this knowledge, an XP spectrum was modelled to illustrate the expected shape for the dynamic disordered system. This was a two component peak, with each component corresponding to the same binding energies as the pure salt and co-crystal structures, leading to a double peak.[19] The actual spectrum does not exhibit this behaviour, instead showing a distinct broadening compared to the other spectra. This suggests that a quasi-centred interaction is occurring, with the single well potential allowing the proton to reside over a range of positions around the bottom of the potential well.[19] Again, this returns to the relationship between the N 1s binding energy and the N---H distance. In this case it is even more interesting as an intermediate state has been found, confirming the presence of a continuum of states between the extremities of salt and cocrystal. Overall, this approach has allowed interpretation of the hydrogen bonding only previously possible to this level of accuracy through the use of a combination of techniques such as XRD and neutron scattering, in a much simpler way and potentially giving more information about the nature of the proton environment than other methods.

### 1.3.2 NEXAFS

Near edge X-ray absorption fine structure (NEXAFS) spectroscopy, also known as X-ray absorption near edge structure (XANES) spectroscopy, is a technique used to determine molecular structure through excitations between core energy levels and unoccupied molecular orbitals.[54]

NEXAFS spectroscopy measures the absorption coefficient over the region close to an absorption edge. An absorption edge occurs where the absorption coefficient increases in a step function at the binding energy of a core level electron.[54] As the photon energy is equal to or greater than the binding energy at this point, photoelectrons will be emitted resulting in an increase in absorption coefficient.[54] The main features of interest in NEXAFS and those related electron transitions between core level states and unoccupied molecular orbitals in the region above and below the absorption edge. As shown in quantum theory, for an electronic transition to occur, the energy of the incoming photon must be equal or greater than the energy difference between the two orbitals.

Therefore, as the energy of the incoming photon changes, different excitations will occur. The interaction cross section increases significantly where the photon energy is exactly equal to the energy of a particular transition resulting in resonance peaks corresponding to those transitions.[54] In samples containing π bonding or aromatic systems, the lowest energy of these transitions are 1s to π* which occur below the absorption edge. These generally form intense narrow peaks due to the long lifetime of such states before the core hole is filled through one of the decay processes.[54] This long lifetime leads to a small uncertainty in the energy, and therefore a narrow peak through the Heisenberg time-energy uncertainty relation. Rydberg energy levels then form a number of states up to the vacuum level. These also have a long lifetime and result in sharp peaks, although with lower intensity and a large number of resonances packed close to the ionisation potential.[54] The third transitions are those from the core level to σ* orbitals. σ* orbitals tend to have an energy above the vacuum level, despite them still being bound states. These have much shorter lifetimes and are less stable, resulting in significant broadening compared to other features, but in a similarly structural dependent manner.[54] While it is not possible to measure these processes individually, the return of the material to its ground state by filling the core hole is measurable resulting in the emission of energy in the form of photons or electrons depending on the core hole decay process. The two primary methods are photoemission and Auger emission. Fluorescence is the process whereby a photon is emitted with energy equal to the difference in energy levels. This will take place as a series of relaxations using spectroscopically allowed transitions.[54] Auger emission is a similar process, where another valence electron is emitted instead of a photon, with the electron gaining energy equal to the change in energy levels, ensuring conservation of energy and momentum.[36,38,54] Due to the energies involved, fluorescence is a much less likely process than Auger emission in organic crystals, and therefore the detection of electrons is most commonly used. The measurement of these electrons is then used to form the intensity in the spectrum, since the number of electrons detected is proportional to the number of core hole states forming at each photon energy.[54]

### 1.3.2.1 Experimental set up

A NEXAFS experimental set up consists of an X-ray source, a sample chamber and a selection of detectors depending on the measurement. In NEXAFS spectroscopy, the spectrum is plotted as a function of the incoming beam energy. This requires an X-ray source capable of providing a continuous spectrum of frequencies, monochromatically one at a time, scanning over the required near edge regions for the atoms present in the sample.[54] Normally, this requires a synchrotron X-ray source, although some NEXAFS lab sources are now available.[55] Unlike XPS, NEXAFS spectroscopy does not depend upon the measurement of the kinetic energy of emitted electrons, only on the number of electrons detected, and therefore does not use a hemispherical detector, but a detector at a specific angle relative to the incoming beam of X-rays to measure the drain current which is related to the total number of electrons.[56] A small voltage is applied across the sample to attract emitted electrons towards the detector. NEXAFS measurements are normally completed in UHV conditions. However, the new high throughput endstation on beamline B07-B at Diamond Light Source allows measurement to be completed in NAP conditions.[56] Using the near ambient pressure endstation on beamline B07-B, the sample chamber is filled with helium gas at 1 mbar for experimentation.[56] Helium is used as it does not have any absorption edges in the region of the basic organic components of crystals, C, N and O. This means that the He just forms a background which can be normalised out.[56] Figure 1.9 shows a schematic representation of the B07-B beamline.

### 1.3.2.2 Spectrum processing and calibration

Since the NEXAFS measurements described herein were being carried out for the first time with the NAP-NEXAFS setup on this new endstation, careful spectrum calibration was required to ensure the gas phase effect was normalised out of the data appropriately. Calibration of the photon energy scale requires comparison with some measured standards which were also measured alongside the samples under the same conditions.[57,58] The photon energies expected for these are very well known allowing a confident calibration. In

samples containing nitrogen, the calibration was carried out using gas phase nitrogen, of which there was a small quantity of contamination in the gas phase. This produces a distinctive set of peaks with well-known energy allowing a calibration to be taken.[57,58] Comparisons of this method with the standards confirmed an equivalent calibration was achieved.

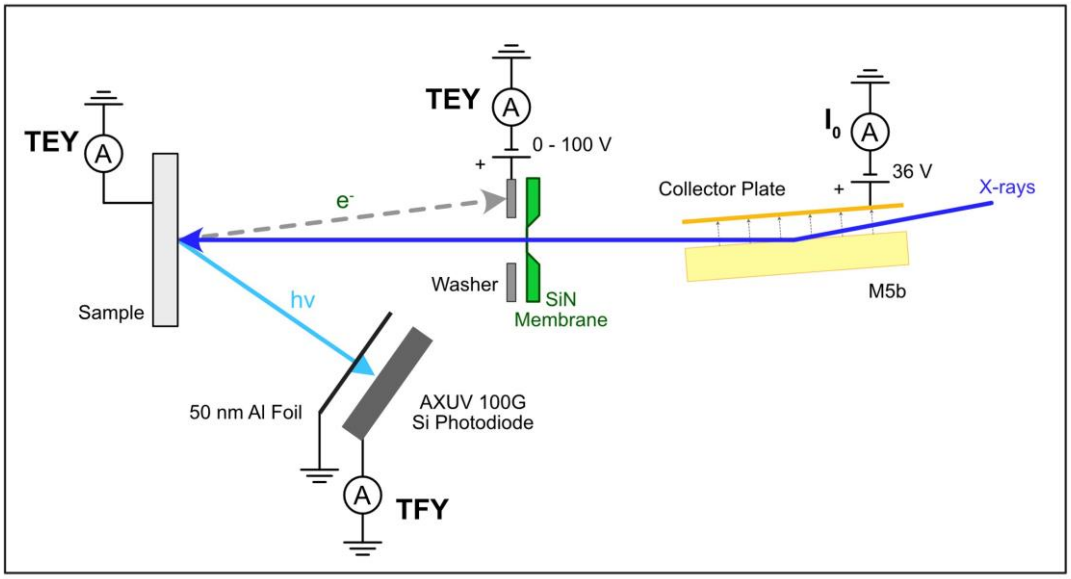



**Figure 1.9** Schematic diagram of the NAP-NEXAFS experimental set up at B07-B showing the key components and processes.[56] Everything to the left of the SiN membrane is within the sample chamber. Both measurement of electrons and fluorescence are possible through total fluorescence yield (TFY) and total electron yield (TEY) measurement.

A new processing procedure was designed for NAP-NEXAFS measurements due to the new beamline endstation in use.[56] This was to account for the effect of the gas phase and beamline on the measured spectrum by normalising them out. A background scan of the gas phase was utilised for this purpose which enabled all background signals to be measured and therefore normalised out of the measured spectrum. Some careful thought was required here to remove any path length effect of the gas phase vs sample measurements. The main feature to remove from the background was from the silicon nitride window into the sample chamber, causing a significant nitrogen signal loss over the range of photon energies at the absorption maximum. Without calibrating and normalising the spectra for this, they are non-physical with unexpected incursions below the base line which is not possible for an X-ray Absorption

Spectrum, highlighting how this is wrong. By completing the normalisation procedure, the spectra appear as expected, and match very well with published experimental data from other sources for the full range of amino acids.

#### 1.3.2.3 Peak fitting and analysis

Peak fitting analysis was carried out in conjunction with DFT calculations. By observing the peaks in the spectrum, it is possible to assign the transitions. For example, transitions in aromatic systems to the π* orbitals are always sharp, intense peaks, while transitions to σ* orbitals are much broader resonances, due to the large number of energy levels to excite into and their shorter lifetime.[54] This allowed each peak to be assigned to the relevant molecular orbital transitions and the orbital transition dipole moments.

There are two main types of feature in a NEXAFS spectrum of an organic crystal, π* and σ* resonances, and each originate from transitions from the core level into π* and σ* orbitals respectively.[54] The characteristic shapes of these resonances can then be used to identify structural properties, often through the use of a computational modelling method such as DFT. Given all of these additional variables being measured, NEXAFS produces a more unique molecular fingerprint than XPS, which similarly we can predict through density functional theory calculations.

#### 1.3.2.4 Literature using NEXAFS to investigate hydrogen bonding

While the NEXAFS technique is now relatively commonly used for surface studies, a less common use of NEXAFS is for the determination of structural information relating to hydrogen bonding, in a similar manner discussed for XPS. It has been shown that the NEXAFS method alone can be used to approximate bond lengths by measuring the difference in photon energy between the σ* shape resonance and ionisation potential for a particular bond.[59] This is shown for some simulated carbon-oxygen bond lengths in Figure 1.10 demonstrating how sensitive the technique is to changes in molecular structure, with changes in excess of 4 eV for a 0.1 Å change in bond length. A relatively

simple approach like this will never be accurate enough to determine a crystal structure to any level of accuracy, but could certainly be used to define the nature of a hydrogen bond as salt or co-crystal by determining one of the relevant bond distances.[59]

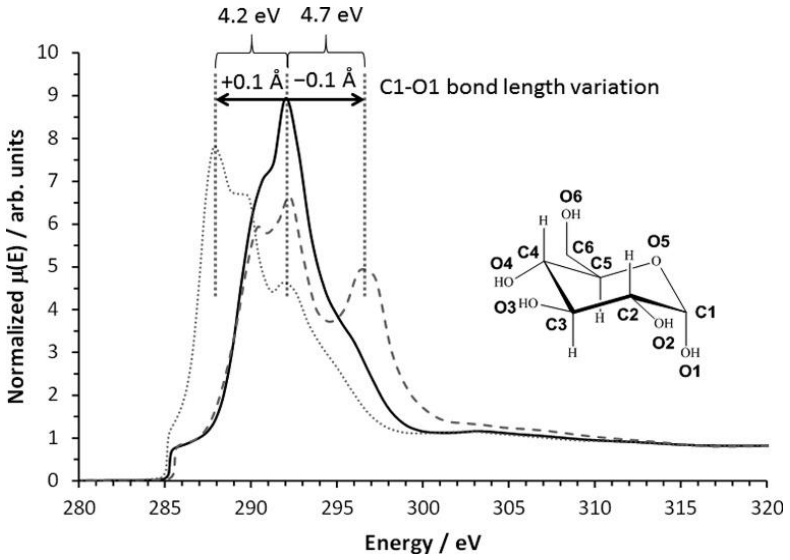



**Figure 1.10** Simulated NEXAFS spectrum showing a theoretical change in binding energy with bond length.[59] This demonstrates the theoretical sensitivity of the NEXAFS technique to individual bond lengths through a theoretical set of calculations highlighting the way in which NEXAFS can be used to determine very small changes in molecular structure. Reproduced from ref. 59.

By combining with computational modelling techniques, NEXAFS has also been used to directly observe changes in hydrogen bonding between two different polymorphs of para-amino benzoic acid.[44,60] In this study, both XPS and NEXAFS were used to identify changes in the hydrogen bonding structure between the α and β polymorphs, with reduced electron density on nitrogen for the β polymorph.[44] In the XP spectrum, it was possible to observe this as a very slight shift of 0.2 eV in the N 1s emission line.[44] Apart from this shift, no other changes in the spectrum were visible, meaning the interpretation of this on its own could be difficult. However, for the NEXAFS spectrum, significant changes between the polymorphs were observed due to changes in the unoccupied orbitals.[44] Figure 1.11 shows the two NEXAFS spectra labelled with the corresponding molecular orbitals. There are clear differences here, and visualisation of the molecular orbitals was required to explain the effects seen. For the 1s → 1π* resonance, the hydrogen bonding was found to not be directly

involved, with the molecular orbital covering much of the molecule and the changes occurring due to the carboxylic acid dimer. However, the 1s → 3π* resonance was shown to be a direct effect of the change in hydrogen bonding character. The 3π* molecular orbital was shown to be focused around the hydrogen bonded nitrogen, and was therefore very sensitive to any change in the hydrogen bond, including changes in electron density around the nitrogen due to the additional proton acceptor effect in the β polymorph.[44] The σ* resonance also shifted to a slightly lower energy in the β polymorph and this was shown to be due to a reduction in the C-N bond length.[44]

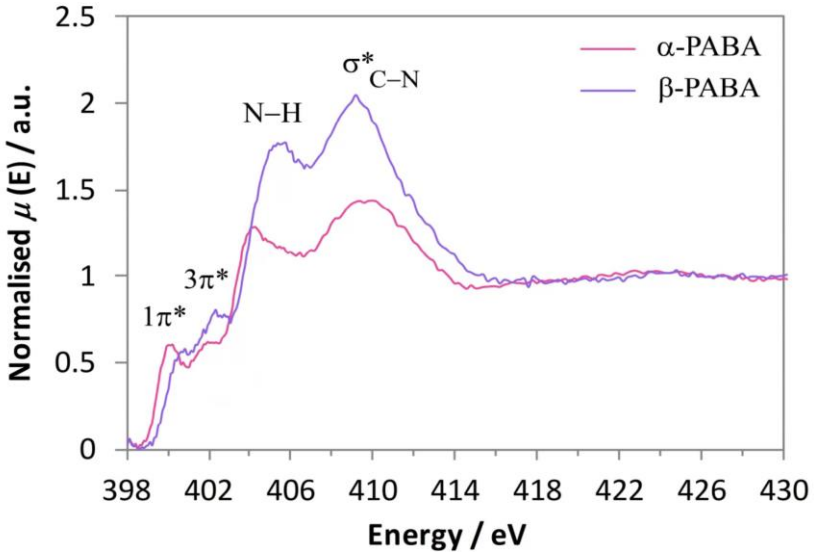



**Figure 1.11** N K-edge NEXAFS for both polymorphs of para-amino benzoic acid.[44] This shows how sensitive NEXAFS is to very minor changes in molecular structure, in particular where hydrogen bonding is present. Reproduced from ref. 44.

A more complicated study into the hydrogen bonding of water molecules at a liquid - vapour interface has been carried out and showed that water molecules at the interface exhibited acceptor only hydrogen bonding using NEXAFS and computational methods.[61] By studying the effect on NEXAFS spectra it was possible to interpret the shift in certain resonances as a change in the hydrogen bonding between the bulk and at the interface.[61] This feature is only visible on very short timescales, and alternative methods had not observed this when published, despite describing one of the most studied molecules.[61] Here the use of computational approaches allowed the NEXAFS spectra to be fully interpreted and improved the understanding of the nature of hydrogen bonding

interactions in water. A final example of the use of NEXAFS for the characterisation of hydrogen bonds in crystal structures is the same set of pyridine dicarboxylic acid complexes used in XPS.[19] In XPS, peak broadening was observed where quasi centred hydrogen bonding was present, and a similar effect was present in the NEXAFS spectrum. Here the σ* resonances have not been analysed as the information required is entirely within the 1s → π* resonance, as seen in Figure 1.12. In direct similarity to the XPS emission lines, there is a 2 eV shift between the co-crystal and salt structures, while the quasi centred proton hydrogen bond causes a broadening of the resonance and a decrease in intensity.[19]

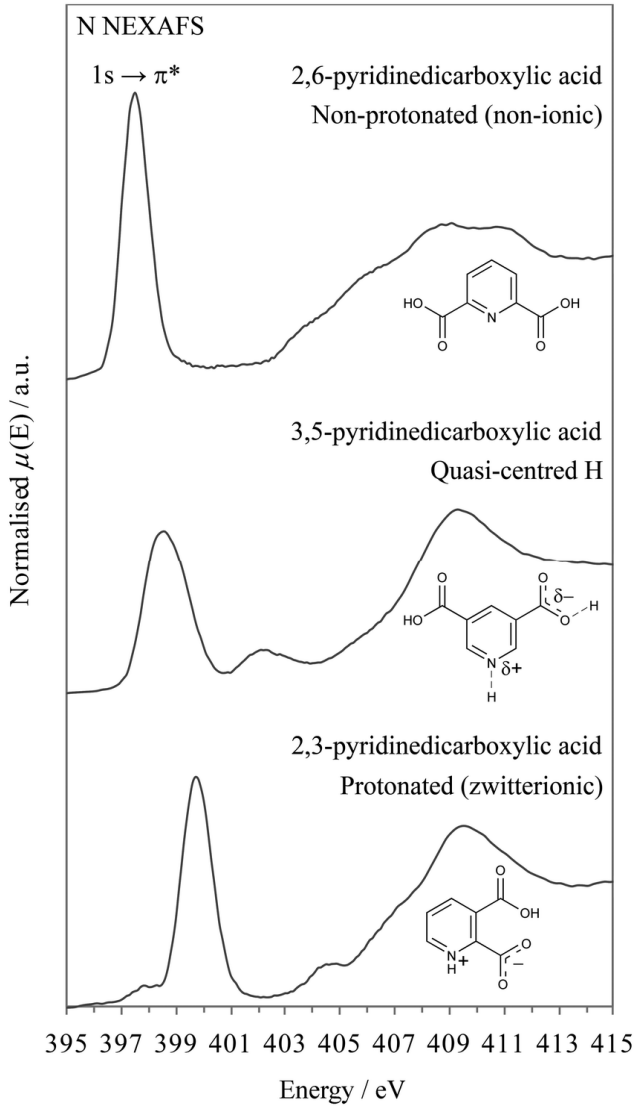



**Figure 1.12** N K-edge NEXAFS for the three pyridine dicarboxylic acid complexes.[19] Demonstrating the sensitivity of NEXAFS to proton position with broadened peak observed identifying the intermediate state. Figure reproduced from ref. 19.

There are clearly other effects visible in the σ* resonances with peaks shifting and similarities between the quasi centred structure and the salt structure. Since there is only a single nitrogen in the chemical structure, any differences can only be due to the different substitution locations, or the hydrogen bonding.[19] As stated for the equivalent XPS data, further work could be done to determine the origin of the peak broadening in an attempt to gain information about the exact nature of the protons within this hydrogen bond and their positions, or range of positions. While NEXAFS has been used as a tool for crystal structure refinement, very little work has been published regarding its use for the characterisation of hydrogen bonding in crystal structures, and even less on any quantitative analysis allowing hydrogen bonding to be fully described, including precise proton positions. While containing significant structural information, additional calculations are required to access this to its full extent by modelling crystal structures and molecular orbitals as accurately as possible, primarily through the use of density functional theory.

## 1.4 Computational Techniques

### 1.4.1 Density Functional Theory

Density Functional Theory (DFT) uses quantum theory to determine molecular structures and properties from first principles by solving Schrödinger's equation. Density functional theory is primarily implemented as a computational approach to solving quantum chemical problems including determining molecular structures, properties and interactions.[62] This approach fits perfectly with the spectroscopic techniques used, and the modelling of spectroscopic properties such as electron excitations is well understood in terms of time dependent DFT (TDDFT) calculations.[63] The fundamental underlying theory of DFT is exact, such that with the exact electron density ρ, and the exact 'functional' to apply to that electron density, any property of interest could in theory be calculated.[62] However, to make the calculations feasible, a series of approximations are made. The first is the Born-Oppenheimer approximation, made in almost all quantum chemistry calculations.[64] This assumes that the nuclear motion is significantly smaller than any electronic motion such that it can be ignored,

which in most cases is a good approximation and significantly simplifies calculations by retaining nuclear positions throughout.[64] The second simplification is performing the calculations within a ‘basis set’. This constrains any molecular orbitals to a linear combination of a series of basis functions. The exact value of any property will only be found by using a complete basis set (all possible functions) and it is well documented that larger basis sets give more accurate results, but also take longer to calculate so there is always a balance between accuracy and feasibility.[65] There are two main forms of basis set, Gaussian basis sets and plane wave basis sets. These two mathematical forms are chosen as they are very computationally efficient to work with. Gaussian basis sets are generally chosen for calculations on small systems and will take the form of a series of functions corresponding to s, p, d, f etc. type orbitals, a linear combination of which will describe the molecular orbitals in the system. Plane wave basis sets are generally used for larger systems with more than one repeating unit cell where the continuous nature of the plane wave makes the calculation very efficient and consistent across the entire molecule.

The final important approximation is the exchange correlation functional (Exc). The origin of the exchange-correlation energy comes from the principles of electron exchange and correlation in quantum theory. Exchange energy originates from the energy difference observed where exchange of identical electrons occurs, necessitating additional terms in the overall wave function due to interference in the various terms. The correlation component of the energy relates to the equivalent phenomenon in Hartree Fock theory, whereby an additional energy component is required to explain electron correlation to remove self-interaction error in particular systems. Both of these concepts are relatively easy to implement on an electron based wavefunction, but the form is not known for applying to the electron density, hence the requirement for approximate exchange correlation functionals. The form of the exact exchange correlation functional is unknown, there are many approximate Exc functionals designed for specific, or more general uses.[66–69] The total energy is calculated as the sum of the kinetic, potential, and exchange correlation contributions. The form of the kinetic and potential energy terms are known exactly, but the exchange correlation functional is always an approximation. A very commonly

used hybrid Exc is the B3LYP functional of Becke (1993).[67] This combines a proportion of “exact” Hartree Fock exchange energy with DFT Generalised Gradient Approximation (GGA) exchange energy. There is a significant body of literature showing that this exchange correlation functional provides reasonable accuracy across a wide range of ground state properties for small molecules and systems.[34,48,70]

#### 1.4.1.1 Experimental set up and details

There are many different types of DFT calculation for the calculation of different properties and a large variety of settings to optimise any calculation. The most basic DFT calculation is known as a single point energy calculation. This takes the molecular geometry of the input file and does a total energy calculation on that, which includes a calculation of the energies of each of the molecular orbitals in the system, along with a range of analytical properties. This calculation primarily involves minimising the energy of the system through the self-consistent field approach. This is an iterative process to find the minimum energy of the system by adjusting the electron density.

At the next level, a geometry optimisation calculation additionally attempts to find the minimum energy geometry of the system. This involves a further iterative process changing the nuclear positions to minimise the energy, followed by a self-consistent field approach for each geometry to optimise the electron density. This process continues until a minimum in energy is found (subject to a series of constraints). The choice of whether to complete a geometry optimisation depends on several factors, but primarily on whether the input .xyz coordinate file is based on experimental data or not. When a minimum is found, the optimisation calculation completes. However, the programme does not know if this is a local or global minimum in energy and an additional calculation must be completed to determine this. This is known as a frequency calculation, and calculates the UV-vis spectrum for the sample. If the structure is at the global minimum, all frequencies will be real numbers. If any are imaginary, then a local minimum is reached and an additional optimisation procedure is required.

Density Functional Theory calculations were completed using the *ORCA* quantum chemistry programme.[71,72] These were completed using the ARC4 high performance computing cluster at the University of Leeds. Making use of the abilities of the system, we have intended to model the appropriate observables in the most efficient but accurate way, but importantly fully consistent across all calculations done for any individual set of data. The purpose of this is to make all of the data fully comparable and as comparable to an experimental setup as possible. XPS and NEXAFS spectra can both be calculated and displayed based on a DFT calculation. These are determined from the underlying processes involved in each case, an electron emission from the core level and excitations from the core level to unoccupied molecular orbitals.

#### 1.4.1.2 Calculation of XP spectra

For the calculation of XP spectra, a simple ground state single point energy calculation is required. To determine the binding energy of an electron, the energy required to remove an electron from its core level orbital needs to be calculated. This is simply the energy of the appropriate molecular orbital. This will be the 1s orbitals for the core level spectra. This comes from Koopmans' theorem, which states that in closed shell HF theory, the ionisation energy, or binding energy, of an electron is equal to the negative of the orbital energy.[73] For comparison with experiment, the theoretical spectrum can be plotted by broadening each of the peaks with a Gaussian function. In the case of DFT, Koopmans theorem becomes an approximation due to the mixing of DFT and HF exchange in the hybrid exchange correlation functional, but nonetheless provides a suitably accurate result. In the case of calculating XP spectra, the choice of exchange correlation functional can significantly affect the absolute energies calculated, but so long as a suitable hybrid exchange correlation functional is chosen along with a suitably large basis set, the relative energies should be accurate, meaning the predicted spectrum should just be shifted compared to any experiment by calibrating the binding energy scale.[67]

A more thorough approach is to calculate the initial and final (core hole) states separately and determine the difference in overall energy. While this is the more correct method, *orca* struggles with the concept of a core hole and will attempt to fill it to minimise the energy, with no way to avoid this making calculation impossible. Some other quantum chemistry programmes are capable of forcing the core hole to remain.

As with all computationally produced results, the findings must be justified against experimental data, either through agreement and consistency, or through thorough explanation of any differences, including additional calculations to nullify these differences.

### 1.4.2 TDDFT

Electron transitions and excitations are widely calculated using DFT due to its applications in all forms of spectroscopy. The formalism for these calculations is known as Time Dependent Density Functional Theory (TDDFT), and although not strictly time dependent, can be used to determine the same information as the time dependent Schrödinger equation.

#### 1.4.2.1 Calculation of NEXAFS

There are two key values to calculate when looking at electron transitions between molecular orbitals, all of the possible transitions and their relative likelihood of occurring through the dipole transition moment (linked directly to selection rules and allowed transitions). A NEXAFS spectrum can then easily be constructed by using the energy of the transitions to define the position of a peak, and the dipole transition moment gives the relative intensity of each peak.[74]

TDDFT calculations are very commonly used to determine orbital transitions and these are used in the analysis of various spectroscopic techniques such as UV/Vis spectroscopy. These look at changes in the valence electron excitations.

To calculate NEXAFS spectra, the relevant processes need to be calculated, requiring the use of TDDFT to determine the excitation energies responsible for features in NEXAFS spectra. In TDDFT a further approximation is often made in order to simplify the calculation of excitation energies, in the form of the Tamm Dancoff approximation.[63] This has been shown to reduce the computational expense, while not significantly affecting the calculated excitation energies, and improving certain categories of excitation where triplet instabilities may be present. It also becomes more important to ensure a suitable exchange correlation functional is used to accurately model NEXAFS spectra. To calculate NEXAFS, a calculation of core → unoccupied excitations is carried out using TDDFT. This involves the calculation of the energy of multiple orbitals, and the differences between them, in addition to the transition dipole moments to define the intensity of each peak. Here the absolute energy of the orbitals becomes important, to ensure the core and unoccupied orbitals are calculated to the same level of accuracy, and a range separated hybrid functional is often required.[68,75,76]

A range separated hybrid functional varies the proportions of DFT exchange energy and Hartree Fock exchange energy depending on the distance of an interaction being calculated.[68] This means that in the short range for calculating core level orbitals, a high proportion of HF exchange can be included as required, while allowing roughly 20% HF exchange at larger distances for calculating the many unoccupied molecular orbitals at large distances. This means the molecular orbitals at both ranges are calculated to the same level of accuracy resulting in realistic excitation energies.

### 1.4.2.2 Literature using TDDFT to investigate hydrogen bonding

Figure 1.13 shows that the accuracy with which the excitation energies can be calculated using an appropriate exchange correlation functional, such as the Short Range Corrected-1 functional designed for calculating core level properties.[68] The butadiene and Si(CH3)3Cl spectra are simulated almost exactly compared to the experimental data, while the others have the correct

relative energies and would just require shifting by calibrating to one of the features.[68]

Clearly the relative intensities of the peaks are not accurate in all cases, which will be due to the uniform application of peaks with a full width half maximum (FWHM) of 0.3 eV across all excitations, whereas σ* resonances are much broader in experimental spectra. Overall, the requirement for an accurate exchange correlation functional is of vital importance in the calculation of NEXAFS spectra using TDDFT, due to the requirement to calculate the energies of two different molecular orbitals in very different environments.

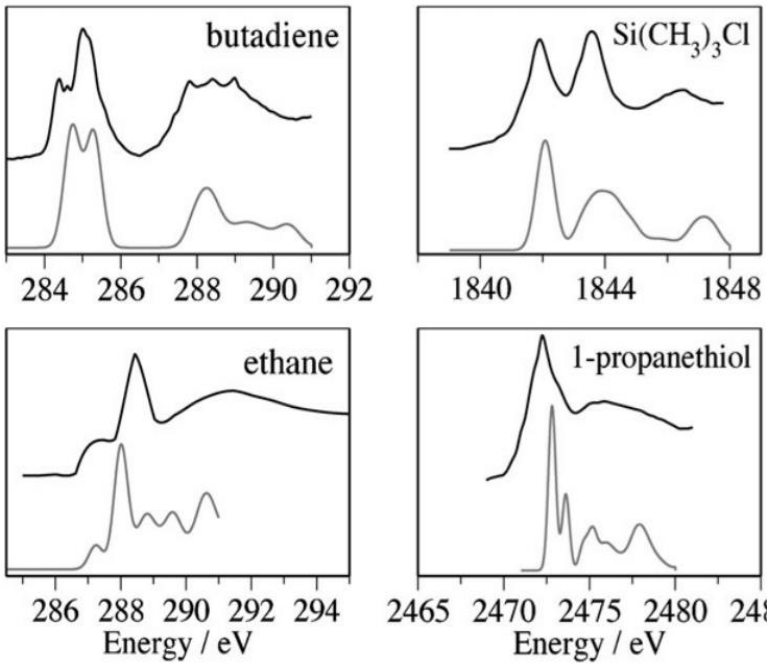



**Figure 1.13** Experimental spectra (upper line) and computed spectra (lower line) for four molecules using the SRC-1 range separated hybrid functional[68] It is clear that peak positions are in agreement with experiment. Reproduced from ref. 68.

As discussed briefly above, TDDFT has been used in many studies of crystal structures and surface analysis in conjunction with spectroscopic techniques. While perhaps not as common as its use with infrared, ultra violet or NMR spectroscopy, there is a significant amount of literature making use of the ability of TDDFT to calculate core level spectra from first principles. Investigating the para-amino benzoic acid molecule again, the experimental XPS and NEXAFS data has been refined by the use of DFT to elucidate the local interactions to

allow full peak assignment and characterisation of the effects of changes in the hydrogen bonding in the spectra.[52]

Figure 1.4 shows the calculated and experimental NEXAFS spectra and the peak assignments. Despite obviously corresponding to the same structure, it can be seen that there is a slight difference in the relative energies of the two main peaks, as well as the 3π* peak appearing further from the σ* feature in the experimental data. As discussed above, while this is not particularly an issue in this case, the use of an appropriate exchange correlation functional could have ensured the calculated photon energies matched those found experimentally.[68] Also shown is the visualisation of the calculated molecular orbitals.[52] This explains why the 2π* feature is missing from the spectrum, as the molecular orbital visualisation shows how the electron density of the 2π* orbital is focused on the ring with very little orbital overlap with the N 1s orbital, leading to a very low transition dipole moment not appearing in the N k-edge spectrum.

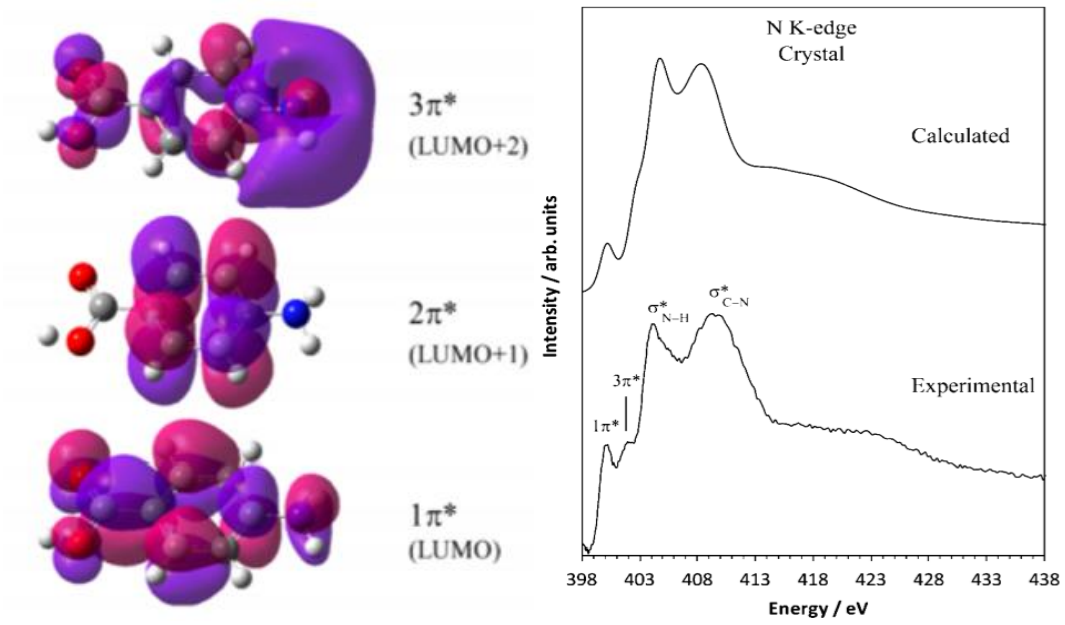



**Figure 1.14** Calculated molecular orbital visualisation determined using DFT (left) and experimental and calculated N K-edge NEXAFS (right) for para-amino benzoic acid[52] Reproduced from ref. 52

The agreement between experiment and theory for all of the spectra means that small changes in hydrogen bonding nature can be detected where changes in both experiment and theory match to confirm an effect, with the changes in hydrogen bonding between the α & β polymorphs also observed.[52] Finally, this approach can also be used as confirmation of the chemical structure, previously determined using XRD in a similar way to NMR crystallography.[52] Since the spectroscopic properties are correct and consistent, all of the chemical environments within the structure must be correct, as it has been shown how sensitive NEXAFS is to chemical environment.[54] Therefore, since the atomic positions for the non-proton atoms are very reliable from XRD data, this should be an appropriate method to refine the proton positions, as the proton positions in the structure should be the only major source of error.

## 1.5 Conclusions

The area of crystal structure refinement has long been dominated by techniques such as XRD and neutron diffraction. In most cases, these methods provide a good approach to determining the structure of molecules and crystals; however, each has its disadvantages in certain circumstances such as when requiring accurate hydrogen positions or having a structure not suitable for deuteration. An additional method sometimes used is ss-NMR as this tends to get around these issues by providing an additional constraint with sensitivity to the proton environments, but this is only of real use when combined with another method. An alternative approach which we have deployed here uses core level spectroscopies such as XPS and NEXAFS, where core electrons in atoms neighbouring a hydrogen bond are directly sensitive to changes in the nature of the hydrogen bonding. This has been shown to give access to significant amounts of structural information, particularly when describing the nature of hydrogen bonds within organic crystal structures and characterising these on the salt cocrystal continuum. There have been several empirical relationships discussed where various parameters can be approximated using XPS binding energies or NEXAFS photon energies. While this approach is still relatively new

in its use for crystallography, there is building evidence that significant contributions to our understanding of the hydrogen bonding in crystal structures can be obtained from these experiments, with plenty of structural information contained within the combination of spectrum and DFT calculated properties from the XRD structure. There is clearly further work to be done on the methods involved, including the potential for using the NEXAFS and XPS spectra as constraints on the DFT calculated structure meaning the atomic positions could be determined, in particular the proton positions in hydrogen bonds where other methods fail.

## 1.6 References


1 R. Tadmor, The London-van der Waals Interaction Energy Between Objects of Various Geometries, *Journal of Physics: Condensed Matter*, 2001, **13**, L195–L202.

2 S. Ota, Y. Tsutomu and T. Yasushi, Dipole-Dipole Interaction and its Concentration Dependence of Magnetic Fluid Evaluated by Alternating Current Hysteresis Measurement, *J Appl Phys*, 2015, **117**, 17713.

3 S. Emamian, T. Lu, H. Kruse and H. Emamian, Exploring Nature and Predicting Strength of Hydrogen Bonds: A Correlation Analysis Between Atoms-in-Molecules Descriptors, Binding Energies, and Energy Components of Symmetry-Adapted Perturbation Theory, *J Comput Chem*, 2019, **40**, 2868–2881.

4 A. C. Legon, Angular Geometries and other Properties of Hydrogen-bonded Dimers: A Simple Electrostatic Interpretation of the Success of the Electron-pair Model, Chem Soc Rev, 1987, **16**, 467-498.

5 L. Pauling, The Nature of the Chemical Bond IV. The Energy of Single Bonds and the Relative Electronegativity of Atoms , *J Am Chem Soc*, 1932, **54**, 3570–3582.

6 G. Gilli and P. Gilli, The Nature of the Hydrogen Bond: Outline of a Comprehensive Hydrogen Bond Theory, Oxford University Press, 2009.

7 M. Schmidtmann, M. J. Gutmann, D. S. Middlemiss and C. C. Wilson, Towards Proton Transfer in Hydrogen Bonded Molecular Complexes: Joint Experimental and Theoretical Modelling and an Energy Scale for Polymorphism, *CrystEngComm*, 2007, **9**, 743–745.

8 T. Steiner, I. Majerz and C. C. Wilson, First O-H-N hydrogen Bond with a Centered Proton Obtained by Thermally Induced Proton Migration, *Angewandte Chemie-International Edition*, 2001, **40**, 2651–2654.

9 J. Wouters and L. Quéré, Pharmaceutical Salts and Co-crystals, *The Royal Society of Chemistry Publishing*, 2012, 9-28.

10 C. B. Aakeröy, M. E. Fasulo and J. Desper, Cocrystal or salt: Does it really matter?, *Mol Pharm*, 2007, **4**, 317–322.

11 J. S. Stevens, S. J. Byard and S. L. M. Schroeder, Salt or Co-crystal? Determination of Protonation State by X-ray Photoelectron Spectroscopy (XPS), *J Pharm Sci*, 2010, **99**, 4453–4457.

12 B. L. Dale, N. R. Halcovitch, M. J. G. Peach and J. M. Griffin, Investigation of Structure and Dynamics in a Photochromic Molecular Crystal by NMR Crystallography, *Magnetic Resonance in Chemistry*, 2019, **57**, 230–242.

13 L. Saunders, H. Nowell, D. Allan and C. Wilson, Imaging H-atom Behaviour on I19, Diamond Light Source, *Acta Crystallographica a-Foundation and Advances*, 2017, **73**, C644–C644.

14 L. K. Saunders, H. Nowell, L. E. Hatcher, H. J. Shepherd, S. J. Teat, D. R. Allan, P. R. Raithby and C. C. Wilson, Exploring Short Strong Hydrogen Bonds Engineered in Organic Acid Molecular Crystals for Temperature Dependent Proton Migration Behaviour using Single Crystal Synchrotron X-ray Diffraction (SCSXRD), *CrystEngComm*, 2019, **21**, 5249–5260.

15 L. K. Saunders, H. Nowell, H. C. E. Spencer, L. E. Hatcher, H. J. Shepherd, L. H. Thomas, C. L. Jones, S. J. Teat, P. R. Raithby and C. C. Wilson, Tuning Charge-assisted and Weak Hydrogen Bonds in Molecular Complexes of the Proton Sponge DMAN by Acid Co-former Substitution, *CrystEngComm*, 2018, **20**, 3074–3083.

16 L. K. Saunders, H. Nowell, P. R. Raithby and C. C. Wilson, Crystal Engineering Urea Organic Acid Hydrogen Bonded Networks with Solvent Inclusion Properties, *CrystEngComm*, 2016, **18**, 5916–5929.

17 L. Saunders, C. Frampton, H. Nowell and D. Allan, Investigating Short Strong Hydrogen Bonds (SSHBs) with Potential Proton Migration Behaviour using Experimental Charge Density Analysis, *Acta Crystallographica a-Foundation and Advances*, 2018, **74**, E81–E81.

18 P. Gilli, V. Bertolasi, L. Pretto, V. Ferretti and G. Gilli, Covalent Versus Electrostatic Nature Of The Strong Hydrogen Bond: Discrimination Among Single, Double, And Asymmetric Single-Well Hydrogen Bonds By Variable-Temperature X-Ray Crystallographic Methods In Beta-Diketone Enol RAHB Systems, *J Am Chem Soc*, 2004, **126**, 3845–3855.

19 J. S. Stevens, S. Coultas, C. Jaye, D. A. Fischer and S. L. M. Schroeder, Core Level Spectroscopies Locate Hydrogen In The Proton Transfer Pathway-Identifying Quasi-Symmetrical Hydrogen Bonds In The Solid State, *Physical Chemistry Chemical Physics*, 2020, **22**, 4916–4923.

20 P. Gilli, L. Pretto, V. Bertolasi and G. Gilli, Predicting Hydrogen-Bond Strengths From Acid-Base Molecular Properties. The Pk(a) Slide Rule: Toward The Solution Of A Long-Lasting Problem, *Acc Chem Res*, 2009, **42**, 33–44.

21 J. G. Chen, M. A. McAllister, J. K. Lee and K. N. Houk, Short, Strong Hydrogen Bonds In The Gas Phase And In Solution: Theoretical Exploration Of Pka Matching And Environmental Effects On The Strengths Of Hydrogen Bonds And Their Potential Roles In Enzymatic Catalysis, *Journal of Organic Chemistry*, 1998, **63**, 4611–4619.

22 E. S. Ameh, A Review Of Basic Crystallography And X-Ray Diffraction Applications, *The International Journal of Advanced Manufacturing Technology*, 2019, **105**, 3289–3302.

23 S. C. Capelli, H.-B. Burgi, B. Dittrich, S. Grabowsky and D. Jayatilaka, Hirshfeld Atom Refinement, *IUCrJ*, 2014, **1**, 361–379.

24 E. Steiner, Density-Difference Maps In Quantum Chemistry, *Theor Chim Acta*, 1982, **60**, 561–572.

25 M. Woinska, S. Grabowsky, P. M. Dominiak, K. Wozniak and D. Jayatilaka, Hydrogen Atoms Can Be Located Accurately And Precisely By X-Ray Crystallography, *Sci Adv*, 2016, **2**, e1600192

26 A. Parkin, S. M. Harte, A. E. Goeta and C. C. Wilson, Imaging Proton Migration From X-Rays And Neutrons, *New Journal of Chemistry*, 2004, **28**, 718–721.

27 M. B. P. Paula, T. F. Koetzle and A. J. Schultz, Single Crystal Neutron Diffraction for the Inorganic Chemist – a Practical Guide, *Comments on Inorganic Chemistry*, **28**, 3–38.

28 M. Adachi and R. Kuroke, Complementary Information from Neutron Crystallography Studies in Structural Biology in Drug Discovery, Wiley Publishing, 2020, 277–293.

29 John. A. Cowan, J. A. K. Howard, J. McIntyre Garry, S. M.-F. Lo and I. D. Williams, Variable-Temperature Neutron Diffraction Studies Of The Short, Strong Hydrogen Bonds In The Crystal Structure Of Pyridine-3,5-Dicarboxylic Acid, *Acta Crystallographica Section B*, 2005, **61**, 724–730.

30 C. C. Wilson, K. Shankland and N. Shankland, Single-Crystal Neutron Diffraction Of Urea-Phosphoric Acid: Evidence For H-Atom Migration In A Short Hydrogen Bond Between 150 K And 350 K, *Zeitschrift Fur Kristallographie*, 2001, **216**, 303–306.

31 D. L. Bryce, NMR Crystallography: Structure And Properties Of Materials From Solid-State Nuclear Magnetic Resonance Observables, *IUCrJ*, 2017, **4**, 350–359.

32 K. D. M. Harris, NMR Crystallography as a Vital Tool in Assisting Crystal Structure Determination from Powder XRD Data, *Crystals (Basel)*, 2022, **12**, 1277.

33 G. Wagner, A. Pardi and K. Wuethrich, Hydrogen Bond Length And Proton NMR Chemical Shifts In Proteins, *J Am Chem Soc*, 1983, **105**, 5948–5949.

34 N. D. Feng, Q. Wang, A. M. Zheng, Z. F. Zhang, J. Fan, S. B. Liu, J. P. Amoureux and F. Deng, Understanding the High Photocatalytic Activity of (B, Ag)-Codoped TiO2 under Solar-Light Irradiation with XPS, Solid-State NMR, and DFT Calculations, *J Am Chem Soc*, 2013, **135**, 1607–1616.

35 J. S. Stevens, S. J. Byard, C. C. Seaton, G. Sadiq, R. J. Davey and S. L. M. Schroeder, Proton Transfer And Hydrogen Bonding In The Organic Solid State: A Combined XRD/XPS/Ssnmr Study Of 17 Organic Acid-Base Complexes, *Physical Chemistry Chemical Physics*, 2014, **16**, 1150–1160.

36 I. Michio, T. Masahiro, K. Hidetaka and K. Feiyu, Advanced_Materials_Science_and_Engineering, Elsevier Science, 2013.

37 G. Greczynski and L. Hultman, X-Ray Photoelectron Spectroscopy: Towards Reliable Binding Energy Referencing, *Prog Mater Sci*, 2020, **107**, 100591.

38 D. Price and K. Sköld, Neutron Scattering Methods in Experimental Physics, Acadmeic Press, 1986, vol 23.

39 X. N. Song, Y. Ma, C. K. Wang, P. M. Dietrich, W. E. S. Unger and Y. Luo, Effects of Protonation, Hydrogen Bonding, and Photodamaging on X-ray Spectroscopy of the Amine Terminal Group in Aminothiolate Monolayers, *Journal of Physical Chemistry C*, 2012, **116**, 12649–12654.

40 B. J. Tielsch and J. E. Fulghum, Differential Charging in XPS. Part I: Demonstration of Lateral Charging in a Bulk Insulator Using Imaging XPS, *Surface and Interface Analysis,* 1996, **24**, 422-427.

41 D. R. Baer, K. Artyushkova, H. Cohen, C. D. Easton, M. Engelhard, T. R. Gengenbach, G. Greczynski, P. Mack, D. J. Morgan and A. Roberts, XPS Guide: Charge Neutralization And Binding Energy Referencing For Insulating Samples, *Journal of Vacuum Science & Technology A*, 2020, **38**, 031204.

42 F. Mangolini, J. B. McClimon, F. Rose and R. W. Carpick, Accounting For Nanometer-Thick Adventitious Carbon Contamination In X-Ray Absorption Spectra Of Carbon-Based Materials, *Anal Chem*, 2014, **86**, 12258–12265.

43 J. S. Stevens, S. J. Byard, C. A. Muryn and S. L. Schroeder, Identification Of Protonation State By XPS, Solid-State NMR, And DFT: Characterization Of The Nature Of A New Theophylline Complex By Experimental And Computational Methods, *J Phys Chem B*, 2010, **114**, 13961–13969.

44 J. S. Stevens, A. Gainar, C. Jaye, D. A. Fischer, S. L. M. Schroeder and Iop, NEXAFS and XPS of p-Aminobenzoic Acid Polymorphs: The Influence of Local Environment in *16th International Conference on X-ray Absorption Fine Structure (XAFS)*, 2016, **712**, 012133.

45 N. Fairley and A. Carrick, *The Casa Cookbook: Recipes for XPS data processing*, Acolyte Science, 2011

46 N. Stojilovic, Why Can’t We See Hydrogen in X-ray Photoelectron Spectroscopy?, *J Chem Educ*, 2012, **89**, 1331–1332.

47 B. Bouchet-Fabre, K. Zellama, C. Godet, D. Ballutaud and T. Minea, Comparative Study Of The Structure Of A-Cnx And A-Cnx : H Films Using NEXAFS, XPS And FT-IR Analysis, *Thin Solid Films*, 2005, **482**, 156–166.

48 S. Carniato, J. J. Gallet, F. Rochet, G. Dufour, F. Bournel, S. Rangan, A. Verdini and L. Floreano, Characterization Of Hydroxyl Groups On Water-Reacted Si(001)-2x1 Using Synchrotron Radiation O 1s Core-Level Spectroscopies And Core-Excited State Density-Functional Calculations, *Phys Rev B*, 2007, **76**, 85321

49 K. Endo, S. Koizumi, T. Otsuka, M. Suhara, T. Morohasi, E. Z. Kurmaev and D. P. Chong, Analysis Of XPS And XES Of Diamond And Graphite By DFT Calculations Using Model Molecules, *J Comput Chem*, 2001, **22**, 102–108.

50 G. Iucci, V. Carravetta, P. Altamura, M. V Russo, G. Paolucci, A. Goldoni and G. Polzonetti, XPS, NEXAFS And Theoretical Study Of Phenylacetylene Adsorbed On Cu(100), *Chem Phys*, 2004, **302**, 43–52.

51 J. S. Stevens, L. K. Newton, C. Jaye, C. A. Muryn, D. A. Fischer and S. L. M. Schroeder, Proton Transfer, Hydrogen Bonding, And Disorder: Nitrogen Near-Edge X-Ray Absorption Fine Structure And X-Ray Photoelectron Spectroscopy Of Bipyridine-Acid Salts And Co-Crystals, *Cryst Growth Des*, 2015, **15**, 1776–1783.

52 J. S. Stevens, C. R. Seabourne, C. Jaye, D. A. Fischer, A. J. Scott and S. L. M. Schroeder, Incisive Probing of Intermolecular Interactions in Molecular Crystals: Core Level Spectroscopy Combined with Density Functional Theory, *Journal of Physical Chemistry B*, 2014, **118**, 12121–12129.

53 J. S. Stevens, S. J. Byard, C. C. Seaton, G. Sadiq, R. J. Davey and S. L. M. Schroeder, Crystallography Aided by Atomic Core-Level Binding Energies: Proton Transfer versus Hydrogen Bonding in Organic Crystal Structures, *Angewandte Chemie-International Edition*, 2011, **50**, 9916–9918.

54 J. Stohr, *NEXAFS spectroscopy*, Springer-Verlag, Berlin ; London, 1992.

55 M. Müller, M. Schellhorn and K. Mann, Laboratory-Scale Near-Edge X-Ray Absorption Fine Structure Spectroscopy With A Laser-Induced Plasma Source, *J Anal At Spectrom*, 2019, **34**, 1779.

56 G. Held, F. Venturini, D. C. Grinter, P. Ferrer, R. Arrigo, L. Deacon, W. Q. Garzon, K. Roy, A. Large, C. Stephens, A. Watts, P. Larkin, M. Hand, H. Wang, L. Pratt, J. J. Mudd, T. Richardson, S. Patel, M. Hillman and S. Scott, Ambient-Pressure Endstation Of The Versatile Soft X-Ray (Versox) Beamline At Diamond Light Source, *J Synchrotron Radiat*, 2020, **27**, 1153–1166.

57 O. Schwarzkopf, M. Borchert, F. Eggenstein, U. Flechsig, C. Kalus, H. Lammert, U. Menthel, M. Pietsch, G. Reichardt, P. Rotter, F. Senf, T. Zeschke and W. B. Peatman, The BESSY Constant Length Rowland Circle Monochromator, *Journal of Electron Spectroscopy and Related Phenomena*, 1999, **101**, 997-1001

58 M. R. Weiss, R. Follath, K. J. S. Sawhney and T. Zeschke, Absolute Energy Calibration For Plane Grating Monochromators, *Nuclear Instruments and Methods in Physics Research A*, 2001, **467**, 482-484

59 A. Gainar, J. S. Stevens, C. Jaye, D. A. Fischer and S. L. M. Schroeder, NEXAFS Sensitivity To Bond Lengths In Complex Molecular Materials: A Study Of Crystalline Saccharides, *Journal of Physical Chemistry B*, 2015, **119**, 14373–14381.

60 J. S. Stevens, A. Gainar, E. Suljoti, J. Xiao, R. Golnak, E. F. Aziz, S. L. M. Schroeder and Iop, NEXAFS Chemical State And Bond Lengths Of P-Aminobenzoic Acid In Solution And Solid State in *16th International Conference on X-ray Absorption Fine Structure (XAFS)*, 2016, **712**, 012136.

61 K. R. Wilson, M. Cavalleri, B. S. Rude, R. D. Schaller, A. Nilsson, L. G. M. Pettersson, N. Goldman, T. Catalano, J. D. Bozek and R. J. Saykally, Characterization Of Hydrogen Bond Acceptor Molecules At The Water Surface Using Near-Edge X-Ray Absorption Fine-Structure Spectroscopy And Density Functional Theory, *Journal of Physics-Condensed Matter*, 2002, **14**, L221–L226.

62 W. Kohn and L. J. Sham, Self-Consistent Equations Including Exchange And Correlation Effects, *Physical Review*, 1965, **140**, 1133.

63 S. Hirata and M. Head-Gordon, Time-Dependent Density Functional Theory Within The Tamm–Dancoff Approximation, *Chem Phys Lett*, 1999, **314**, 291–299.

64 M. Born and R. Oppenheimer, Zur Quantentheorie der Molekeln, *Ann Phys*, 1927, **389**, 457–484.

65 P. M. W. Gill, B. G. Johnson, J. A. Pople and M. J. Frisch, The Performance Of The Becke-Lee-Yang-Parr (B-LYP) Density Functional Theory With Various Basis Sets, Chem Phys Lett, 1992, **197**, 499-505

66 J. P. Perdew, K. Burke and Y. Wang, Generalized Gradient Approximation For The Exchange-Correlation Hole Of A Many-Electron System, *Phys Rev B*, 1996, **54**, 16533–16539.

67 A. D. Becke, Density-Functional Thermochemistry .3. The Role Of Exact Exchange, *Journal of Chemical Physics*, 1993, **98**, 5648–5652.

68 N. A. Besley, M. J. G. Peach and D. J. Tozer, Time-Dependent Density Functional Theory Calculations Of Near-Edge X-Ray Absorption Fine Structure With Short-Range Corrected Functionals, *Physical Chemistry Chemical Physics*, 2009, **11**, 10350–10358.

69 K. Burke and L. O. Wagner, DFT In A Nutshell, *Int J Quantum Chem*, 2013, **113**, 1601–1601.

70 K. Takaoka, T. Otsuka, K. Naka, A. Niwa, T. Suzuki, C. Bureau, S. Maeda, K. Hyodo, K. Endo and D. P. Chong, Analysis Of X-Ray Photoelectron Spectra Of Electrochemically Prepared Polyaniline By DFT Calculations Using Model Molecules, *J Mol Struct*, 2002, **608**, 175–182.

71 F. Neese, The ORCA Program System, *Wiley Interdiscip Rev Comput Mol Sci*, 2012, **2**, 73–78.

72 F. Neese, F. Wennmohs, U. Becker and C. Riplinger, The ORCA Quantum Chemistry Program Package, *Journal of Chemical Physics*, 2020, **152**, 224108.

73 Y. T. Koopmans, *Uber Die Zuordnung Von Wellenfunk-Tionen Und Eigenwerten Zu Den, Einzelnen Elektronen Eines Atoms*, *Physica,* 1934, **1**, 104-113.

74 N. A. Besley, A. T. B. Gilbert and P. M. W. Gill, Self-Consistent-Field Calculations Of Core Excited States, *Journal of Chemical Physics*, 2009, **130**, 124308.

75 M. J. G. Peach, P. Benfield, T. Helgaker and D. J. Tozer, Excitation Energies In Density Functional Theory: An Evaluation And A Diagnostic Test, *Journal of Chemical Physics*, 2008, **128**, 44118.

76 M. J. G. Peach, T. Helgaker, P. Sałek, T. W. Keal, O. B. Lutnæs, D. J. Tozer and N. C. Handy, Assessment Of A Coulomb-Attenuated Exchange-Correlation Energy Functional, *Physical Chemistry Chemical Physics*, 2006, **8**, 558–562.

# 2 Proton Transfer on the Edge of the Salt/Cocrystal Continuum: X-ray Photoelectron Spectroscopy (XPS) of Three Isonicotinamide Salts

*Paul T. Edwards,[1] Lucy K. Saunders,[2] Anuradha R. Pallipurath,[1] Andrew J. Britton,[1] Elizabeth A. Willneff,[3] Elizabeth J. Shotton,[2] Sven L. M. Schroeder[1,2]**

[1]School of Chemical and Process Engineering, University of Leeds, Leeds, LS2 9JT, UK
[2]Diamond Light Source Ltd, Harwell Science and Innovation Campus, Didcot, Oxfordshire
OX11 0DE, UK
[3]School of Design, University of Leeds, Leeds, LS2 9JT, UK
[4]Future Continuous Manufacturing and Advanced Crystallisation Hub, Research Complex at Harwell (RCaH), Rutherford Appleton Laboratory, Harwell, Didcot, Oxon, OX11 0FA, UK

* Sven L. M. Schroeder:

Email: s.l.m.schroeder@leeds.ac.uk

**Abstract**

X-ray photoelectron spectroscopy (XPS) has emerged as a technique that allows for characterisation and classification of hydrogen bonding and proton transfer interactions in organic crystal structures, in a way that is complementary to crystallography by X-ray or neutron diffraction. Here we analyse the nitrogen 1s core level binding energies of isonicotinamide systems with proton transfer between donor and acceptor groups at short distances. We show how a careful calibration of the binding energy scale places these salt systems correctly on the edge of the so-called salt-co-crystal continuum. We show how performing a fitting analysis of the data that is consistent with elemental analysis, expected stoichiometry, and quantification of adventitious carbon contamination facilitates the determination of absolute binding energies with accuracy and reproducibility within ±0.1 eV. The determined N 1s core level binding energies of the protonated isonicotinamide acceptors suggests that the local geometric arrangements of donor, acceptor and proton can influence the N 1s core level binding energy significantly.

## 2.1 Introduction

Creating crystals with an organic co-former is a widely used approach for tailoring the physical properties of pharmaceuticals and other functional organic compounds.[1] In the presence of proton donor-acceptor pairs, electrostatic interactions often drive the supramolecular assembly process, so the difference in $pK_a$ values of donor and acceptor groups is often used for physical property prediction.[2,3] For $\Delta pK_a$ differences > 3, proton transfer, and hence the formation of ionic salts is observed.[4–13] For negative $\Delta pK_a$ values, hydrogen bonding occurs and the product is classified as a co-crystal.[4–13] Between these two regimes, in the $\Delta pK_a$ range of 0…3, the nature of the donor-acceptor interaction is generally unpredictable and other properties, for example other non-covalent interactions, conformational variations and steric factors, can tip the balance from proton transfer to hydrogen bonding and vice versa. This region is often referred to as the salt co-crystal continuum.[7,11,14,15] It comprises not only examples of structures with proton-transfer or hydrogen-bonding, but also many cases that elude such classification, especially when donor-acceptor bond distances are short.[10] As a result, structures with proton transfer and short donor-acceptor distances have variously been referred to as salts,[4,5,13] charge-assisted short strong hydrogen bonds[16,17] or proton transferred interactions.[8,10,12,14,18] Co-crystals are usually described as hydrogen-bonded or containing neutral hydrogen bonds,[4,5,13,19] but short donor-acceptor distances can, for example, result in short-strong hydrogen bonds (SSHBs) in which the proton is partially transferred, but shared between donor and acceptor,[10] blurring the boundary between salt and co-crystal. Moreover, systems with multiple options for donor-acceptor pairing often exhibit two or more types of interactions within one crystal structure – the systems to be examined in this paper includes an example of such cases.[16]

Various studies[8–12] have demonstrated that the core level binding energy shifts detected at the acceptor by X-ray photoelectron spectroscopy (XPS) are sensitive to the position of the proton. Studies of a substantial range of salts and co-crystals with nitrogen acceptors have shown that the N 1s core level binding energy consistently discriminates Brønsted proton transfer (salts) from hydrogen bonding (co-crystal) through a chemical shift of approximately 2 eV.[7,10,11,14,15] Whether hydrogen bonding or proton transfer take place is determined by the relative depths of the double-well

potentials associated with the Brønsted interaction,[10] while the proton position is determined by the location of the potential minimum relative to donor and acceptor. The core level binding energy shifts in donor-acceptor systems reflect the local charge density variation induced by the proton. XPS thus highlights the electrostatic force exerted by the proton, at both the donor and the acceptor centres. Simply put, the negative charge at the donor site increases as the proton is more distant, while the positive charge at the acceptor increases as the proton comes closer. In this view, one has a continuum of electrostatic field effects on donor and acceptor, which is influenced by the distance between donor and acceptor centres as well as their individual distance to the proton. Systems in which the donor-acceptor distance is short (up to ~2.6 Å) allow strong potential well overlap, resulting in a low barrier or even a single well without any barrier.[10] Often referred to as short hydrogen bonds (SHBs) or short strong hydrogen bonds (SSHBs), their classification into salt and co-crystal is not clear-cut when the barrier is so low that the protons are able to populate a continuum of positions, and with a distribution that responds to temperature variations.[7,10,16,20] In line with this, XPS distinguished a quasi-centered population of hydrogens in a low-barrier SHB between an oxygen donor and a nitrogen acceptor unequivocally.[10] Compared to the N 1s emission from protonated and hydrogen-

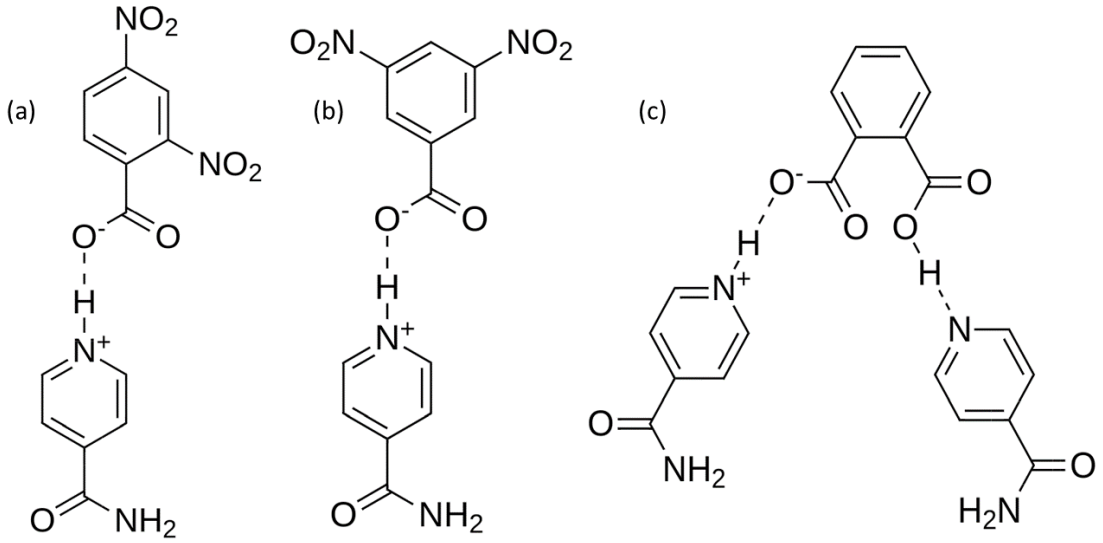



Figure 2.1. Chemical structures investigated comprising three two component isonicotinamide-acid (base-acid) systems with (a) 2,4-dinitrobenzoic acid (1:1), (b) 3,5-dinitrobenzoic acid (1:1) and (c) phthalic acid (2:1). N-H--O distances can be found in Appendix A, Table A1.

bonded functional groups, broadening of the N 1s emission line reflected the dynamic disorder in the quasi-single potential well, and the N 1s emission line centroid was at an intermediate binding energy (BE), neither characteristic for protonation nor for a hydrogen bond.[10] Recently, multi-component molecular crystals containing short strong hydrogen bonds (SSHBs) have been engineered by combining substituted organic acids with isonicotinamide (IN).[16] In the isonicotinamide 2,4-dinitrobenzoic acid (IN24DNBA) and isonicotinamide 3,5-dinitrobenzoic acid (IN35DNBA) 1:1 complexes (Figure 2.1) an interaction takes place between the carboxylic acid group OH donor and the pyridinic N acceptor of IN. In the case of isonicotinamide phthalic acid (INPA, Figure 2.1), the 2:1 IN:PA ratio of the components means two distinct interactions are present, between each of the carboxylic acid OH groups of phthalic acid (PA) and a pyridinic N of the two IN co-formers.[16] All three complexes are within the salt co-crystal continuum, i.e. in the range $0 < \Delta pK_a < 3$. Using X-ray crystallography, the hydrogen atom positions were determined using Fourier difference maps of the electron density.[16] It was found that IN formed charge-assisted hydrogen bonds with both 2,4- and 3,5-dinitrobenzoic acid, characterised by Brønsted proton transfer from the carboxylic acid to the pyridinic nitrogen group.[15,16] This suggests the existence of a sloping potential well, with its minimum located near the nitrogen acceptor. Phthalic acid in INPA had distinct intermolecular interactions with each of its two IN co-formers. One carboxylic acid group formed an SHB characterised by Brønsted protonation (as in IN24DNBA and IN35DNBA), while the other formed a more conventional hydrogen bond with the proton located nearer to the donor.[16] Our interest in applying XPS to these systems stems from their complementarity to our previous study of a quasi-centred SHB.[10] As in that study, O-N donor-acceptor distances in all three complexes are within the range of an SHB,[21] but in contrast to the quasi-centred SHB we have studied previously,[10,16] the Fourier difference maps indicated that the proton is located at the nitrogen acceptors as a result of proton transfer.[10,16] One of the reasons for examining these complexes further by XPS is that the relative proximity of the electrostatic potential from the donor may influence the charge at the acceptor sufficiently to result in a detectable N 1s core level shift relative to a non-SHB salt. If so, then XPS would distinguish protonation in a charge-assisted SHBs from

conventional Brønsted proton transfer with a longer donor-acceptor distance, allowing a more nuanced classification of salts. The effect on the N 1s BE was expected to be smaller than the BE shift associated with protonation, requiring a highly accurate determination of BE shifts. In addition to examining the mentioned effect on the N 1s core level binding energy, we therefore provide a detailed description of our data analysis procedure, to enable other researchers to follow this protocol and facilitate wider use of XPS for the characterisation of Brønsted interactions in organic crystals. The key to success of such studies is a reliable calibration of the core level electron BE scale with accuracy and precision (reproducibility) within ±0.1 eV for all investigated samples.[22] Throughout our work over the last decade we have found that such accuracy can be achieved by careful calibration of the BE scale through self-consistent analysis of the elemental composition and quantitative analysis of the C 1s emission line.[7–12,14,23] This facilitates accurate identification of the energy position of the aliphatic carbon contribution from adventitious carbon contamination present on the samples. This procedure overcomes the inaccuracies associated with the use of less elaborate BE scale calibrations based on C 1s emission from adventitious carbon.[24–26] We will demonstrate how the use of adventitious carbon as an internal BE standard can provide a very consistent calibration when a complete fitting analysis of the C 1s emission line is performed. We have reliably applied this procedure throughout all of our previous studies of organic systems, as it removes the uncertainties associated with taking the overall maximum of the C 1s emission line as the BE calibration peak. Using this method, the absolute N 1s binding energy involved in H-bonding is consistently found to be in the range 399 eV - 400 eV, while a salt is found between 401 eV - 402 eV.[7–12,14,23]

## 2.2 Experimental Methods

### 2.2.1 Materials

All three samples were prepared by evaporation from solution, as described in previous work.[16] The crystallizations were carried out in methanol (INPA) and ethanol (IN24DNBA and IN35DNBA) using a 2:1 stoichiometric ratio (INPA) or 1:1

ratio (IN24DNBA and IN35DNBA). Prior to measurement, crystalline samples were crushed to form a powder. In addition, pure isonicotinamide (IN, 99.9%) for XPS analysis was obtained from Flurochem.

### 2.2.2 X-ray Photoelectron Spectroscopy (XPS)

XP spectra were collected with a SPECS EnviroESCA NAP-XPS equipped with a monochromatic Al K-alpha X-ray source (1486.71 eV) operating at 42 W separated from the analysis chamber with a SiN window which illuminated the sample with a ~300 µm diameter beam footprint. Spectra were collected on powder samples pressed onto double sided adhesive carbon tape at ambient temperature in 7 mbar argon with a hemispherical Phoibus NAP 150 analyser operating in small-area mode with a source-analyser angle of 55° and a 1D delay-line detector. Under NAP conditions in the EnviroESCA instrument, insulating sample surfaces are inherently charge-neutralized through ionization of the gas phase by the X-ray beam, which provides positive ions and electrons to balance any surface charge.[26,27] Survey spectra were collected in one scan with a step size of 1 eV, pass energy of 100 eV and dwell time of 0.1 s. High resolution C 1s, N 1s, and O 1s core level spectra were collected with a step size of 0.1 eV, a dwell time of 100 ms per data point, and a pass energy of 50 eV. Spectra for C 1s and O 1s core levels were collected in 9 scans. Due to the lower atomic abundance of N within the sample, the N 1s core level spectrum was collected in 16 scans. The spectra were analysed using CasaXPS.[28] Shirley backgrounds were used and a GL(30) line shape (30% Lorentzian, 70% Gaussian) to fit the emission lines with appropriate components to describe the environment of each element in the sample.[29]

### 2.2.3 Quantitative Elemental Analysis

From early on in our XPS studies of organic materials we have found that quantitative elemental analysis of a wide-scan survey spectrum provides crucial boundary conditions for the correct interpretation of the C 1s emission line (Appendix A, Figure A1).[30] Identification of non-functionalised hydrocarbon contributions (‘adventitious carbon’) to the C 1s emission line is the most widely used binding energy calibration method for organic compounds (*vide infra*). Initially, the survey spectra were checked

to ensure the number and BEs of the emission lines were broadly consistent with the stoichiometry of the materials. Then it was ensured that the correct spectrometer transmission function was applied.[31,32] The transmission function is specific to the spectrometer and is influenced by various experimental parameters, such as the angles between X-ray source, sample and analyser entrance, operational detector and transfer lens settings, gas phase composition and pressure (for NAP instruments).[28,31,32] It must be re-determined when any changes to these settings are made. Having successfully carried out these intensity calibrations, the elemental composition analysis is performed in CasaXPS[28] to determine the detected atomic percentage of each element within the probed sample volume.[29] Comparison with the expected stoichiometry of the sample then allows the determination of the excess carbon quantity, $C_{adv}$, resulting from adventitious contamination. Subtraction of the excess carbon quantity then allows the calculation of a corrected elemental composition that can be compared to the expected stoichiometry (Table 2.1). This is calculated by weighting the effect of the adventitious carbon by the expected stoichiometry, giving equations 1 – 3.

$$[\mathrm{C}]_{\mathrm{corr}} = [\mathrm{C}] - [\mathrm{C}]_{\mathrm{adv}} \tag{1}$$

$$[\mathrm{N}]_{\mathrm{corr}} = [\mathrm{N}] + \frac{[\mathrm{N}]_0}{[\mathrm{N}]_0 + [\mathrm{O}]_0}[\mathrm{C}]_{\mathrm{adv}} \tag{2}$$

$$[\mathrm{O}]_{\mathrm{corr}} = [\mathrm{O}] + \frac{[\mathrm{O}]_0}{[\mathrm{N}]_0 + [\mathrm{O}]_0}[\mathrm{C}]_{\mathrm{adv}} \tag{3}$$

Here, $[N]_0$ and $[O]_0$ are the expected stoichiometric ratio and [C], [N] and [O] are the measured ratios. The results in Table 2.1 show that both N and O percentages are close to the expected stoichiometry of the sample.

**Table 2.1.** Element and adventitious carbon ($C_{adv}$) percentages compared to sample stoichiometries, including corrected values for N and O after removal of the $C_{adv}$ contributions.

| | % C | | | % N | | | % O | | |
|---|---|---|---|---|---|---|---|---|---|
| | Expected $[C]_0$ | Measured [C] | Corr. $[C]_{adv}$ | Expected $[N]_0$ | Measured [N] | Corr. $[N]_{corr}$ | Expected $[O]_0$ | Measured [O] | Corr. $[O]_{corr}$ |
| IN24DNBA | 54.2 | 70.5 | 16.3 | 16.6 | 12.8 | 18.7 | 29.2 | 16.7 | 27.1 |
| IN35DNBA | 54.2 | 64.2 | 10 | 16.6 | 15.0 | 18.6 | 29.2 | 20.8 | 27.2 |
| INPA | 66.7 | 72.6 | 5.9 | 13.3 | 10.2 | 12.6 | 20.0 | 17.2 | 20.7 |
| IN | 66.7 | 70.8 | 4.1 | 22.2 | 18.2 | 20.9 | 11.1 | 11.0 | 12.4 |

### 2.2.4 Binding Energy Scale Calibration

The importance of a coherent and consistent approach to the binding energy (BE) scale calibration in XPS analysis has recently been highlighted.[24,25] As already mentioned, for organic compounds the adventitious carbon contribution to the C 1s emission is often used as an internal standard for the BE, although sometimes emission from another characteristic group containing carbon has been used.[25,26,33,34] Consistency in the accounting for the adventitious carbon contribution in the elemental analysis and in the fitting analysis of the C 1s emission line improves the accuracy and precision of the binding energy calibration. We generally find that this method ensures not only consistency of the binding energy scales between measurements taken on different spectrometers, but it also enables the use of absolute binding energies with accuracy on the order of ±0.1 eV to identify the chemical state of a sample. In the present study, we will determine absolute N 1s binding energies in this manner to characterise and classify hydrogen bonds and protonated moieties.[7–14,29]

## 2.3 Results

### 2.3.1 C 1s Emission and Binding Energy Scale Calibration

Table 2.1 summarises the results of the elemental analysis of pure isonicotinamide (IN) and the three IN complexes examined in this paper. As measured, excess carbon ($C_{adv}$) is evident in all data, indicating significant adventitious carbon contamination that ranges from 4.1% in IN to 16.3% in IN24DNBA. Adventitious carbon contamination is unavoidable and stems from the small concentrations of surface-active contaminants that are present even in the purest solvents and chemicals, alongside ubiquitous environmental contaminants introduced by handling materials in a biological environment; even after the most thorough cleaning, glass vessels used in the laboratory have remaining surface contamination on the order of a single molecular overlayer.[33–35] After subtracting the adventitious carbon percentage from the elemental analysis corrected values for the % contributions of N and O are obtained, which are in good agreement with the expected N and O stoichiometry of the materials (Table 2.1). $C_{adv}$ also gives the fraction of C 1s signal

that must be attributed to adventitious carbon. Locating this intensity in the C 1s emission line profile allows a precise calibration of the BE scale to 285.0 eV at the centre of the $C_{adv}$ contribution. The expected functional group stoichiometry informs the relative intensities of the other C 1s component peaks, together with literature values for their BEs.[36] With these constraints in place, the C 1s fits presented in Figure 2.2 were obtained. The parameters of the components in these fits are summarised in Table 2.2.

**Table 2.2.** Best fit binding energies (eV), intensities (a.u.) and FWHMs (eV) for the fitted components to the C 1s emission shown in Figure 2.2.

| | | O−C=O | N−C=O | −(C−X)= | −CH=/$C_{adv}$ |
|---|---|---|---|---|---|
| IN24DNBA | BE / eV | 288.9 | 288.1 | 286.3 | 285.0 |
| | I / a.u. | 89.2 | 89.2 | 249.3 | 304.1 |
| | FWHM / eV | 1.45 | 1.45 | 1.45 | 1.67 |
| IN35DNBA | BE / eV | 289.0 | 288.2 | 286.2 | 285.0 |
| | I / a.u. | 101.0 | 101.0 | 267.7 | 285.8 |
| | FWHM / eV | 1.56 | 1.56 | 1.56 | 1.64 |
| INPA | BE / eV | 289.1 | 288.4 | 286.2 | 285.0 |
| | I / a.u. | 133.6 | 133.6 | 259.8 | 412.8 |
| | FWHM / eV | 1.45 | 1.45 | 1.45 | 1.60 |
| IN | BE / eV | | 287.9 | | 285.0 |
| | I / a.u. | n/a | 126.5 | n/a | 373.9 |
| | FWHM / eV | | 1.56 | | 1.74 |

Given that IN and its three complexes do not contain any aliphatic hydrocarbon moieties (only aromatic groups), one would perhaps expect that the adventitious contamination intensity appears on the high BE shoulder to the aromatic emission peak (−CH=), which for unsubstituted aromatic carbons is often found at 0.2-0.3 eV lower BE than the emission from aliphatic hydrocarbons.[36] However, there is no significant evidence for a separate adventitious hydrocarbon line in the data, probably because it coincides with the emission from IN, in which the electronegative pyridinic nitrogen appears to increase the C 1s BE of the unsubstituted aromatic ring carbons somewhat. For the three salts, the nitro and carboxylic acid substituents in the co-formers appear to have a similar effect on the unsubstituted aromatic carbons. The unsubstituted aromatic hydrocarbon and adventitious carbon contributions were therefore fitted as a single line that was used as the internal BE reference at 285.0 eV (note other values are sometimes used for calibration, e.g. 284.8 eV).[29] We would

like to emphasise that the fits summarised in Table 2.2 are in very good agreement both with the expected stoichiometry and the excess adventitious carbon signal indicated by the elemental analysis.

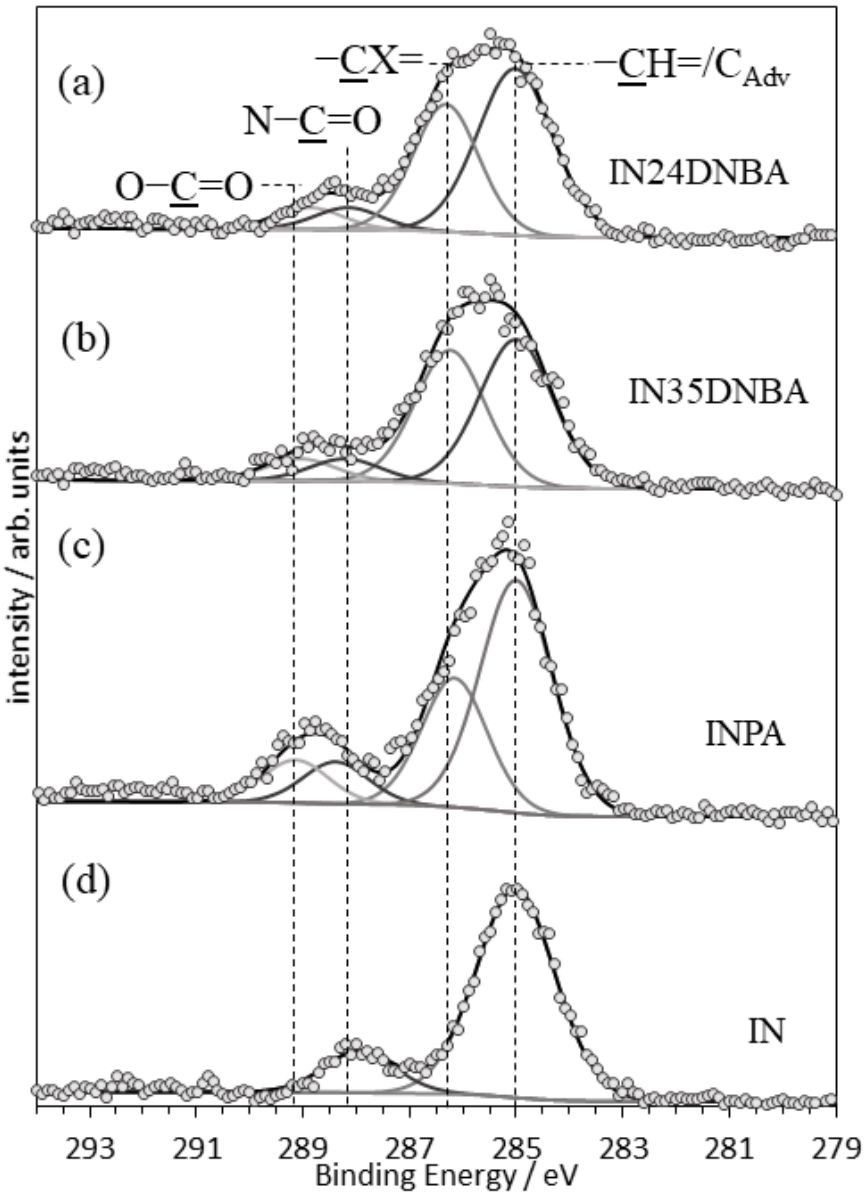



**Figure 2.2.** C 1s emission lines with best-fit components for the carboxylic acid (O−C=O) group, the amide (N−C=O) group from IN, aromatic carbons with electronegative substituents in DNBAs and PA (−(CX)=, with X = COOH, $NO_2$), unsubstituted aromatic carbons (−CH=), and adventitious carbon ($C_{Adv}$). (a) IN24DNBA, (b) IN35DNBA, (c) INPA and (d) pure isonicotinamide. The binding energy scale was calibrated at 285.0 eV with respect to the adventitious carbon contamination (Adv) emission, which is superimposed over the emission from the aromatic carbons (−C=).

We will show in the next section that this calibration of the BE results in N 1s BEs of 406.2 eV and 406.3 eV for the nitro groups in IN24DNBA and IN35DNBA, respectively (Figure 2.3), is in excellent agreement with all nitro N 1s BEs in our previous studies.[12,14] We are therefore confident that the presented C 1s analysis procedure represents best practice for a BE scale calibration based on the C 1s emission line.[24] This procedure was used in all our previous XPS studies of organic systems, which have produced a self-consistent set of binding energy values for various functional groups.[7–14]

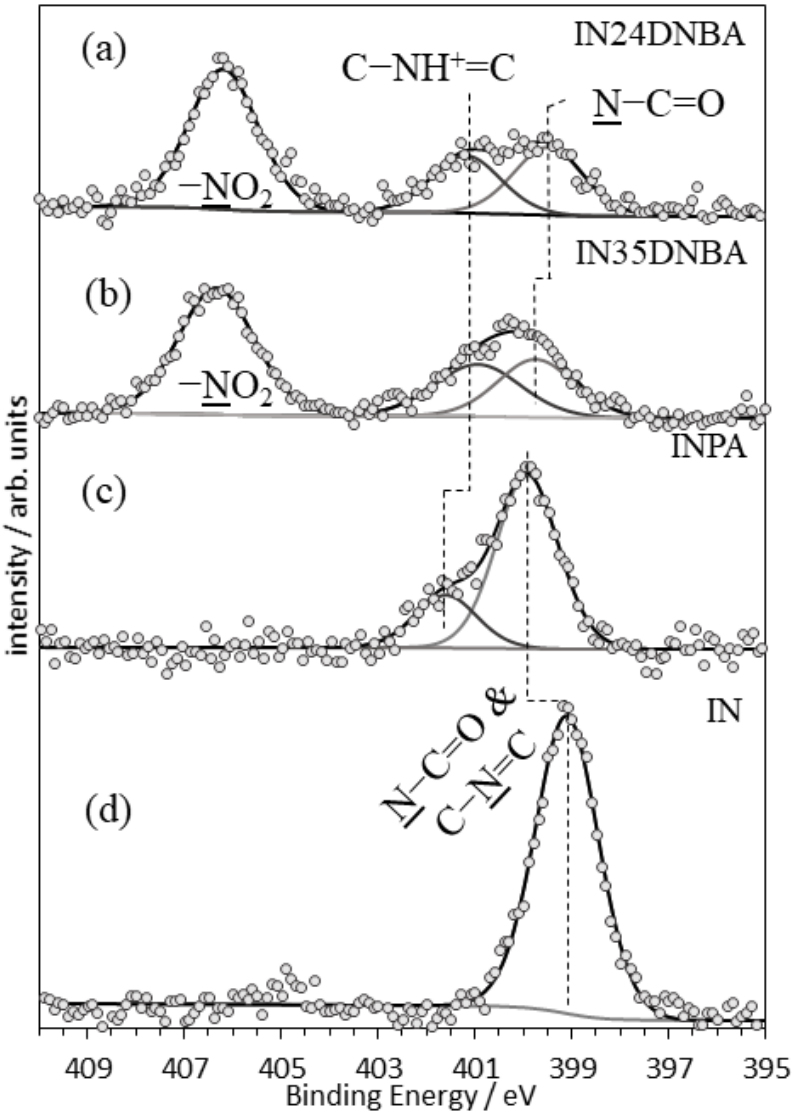



**Figure 2.3.** Nitrogen 1s XPS with fitted components for (a) IN24DNBA, (b) IN35DNBA, (c) INPA and (d) pure isonicotinamide. Binding energy scale calibration is taken with respect to the fitted adventitious carbon component peak.[29]

### 2.3.2 N 1s Emission

As in our previous studies of organic salts and co-crystals, we use the BE of the N 1s emission line of the proton acceptor group to classify the donor-acceptor interaction.[7–9,11,12] Figure 2.3 shows the fits to the N 1s emission lines for the three IN complexes alongside that of pure IN, and Table 2.3 contains the parameters used for the fits. We already mentioned that the C 1s BE calibration places the nitro group N 1s of IN24DNBA and IN35DNBA at BEs of 406.2 eV and 406.3 eV, respectively. These values are in excellent agreement with our previous results for nitro groups, suggesting accuracy and reproducibility on the order of ±0.1 eV for our BE calibration method, even across different XPS instruments.[12,14,22] In addition, during earlier measurements of these samples, we experienced some issues with sample charging and possible radiation damage (data available in the supplementary material, Figure A2 and Figure A3). Despite these issues, the same consistent fitting could be applied (with additional components) and the same binding energies determined.

The BEs of observed N 1s components for the four materials are reported in Table 2.3. Interestingly, although IN contains two chemically inequivalent nitrogen moieties (amide and pyridinic N) the BE of their N 1s emissions coincide within the spectral resolution achievable by experimental XPS, at the same BE of 399.1 eV. In line with this, the FWHM of the single emission line is, within the margins of error, identical to that of the individual N 1s components in the other N 1s spectra. The intensity of the combined emission line is consistent with the presence of two nitrogen atoms per formula unit (Table 2.1), as expected from the stoichiometry of IN. Our experimental N 1s spectrum for IN agrees very well with a recently published spectrum,[15] but the assignment of the amide and pyridinic N 1s BEs differs markedly. In the recently published spectrum, two separate emission lines with very low FWHM were forced (through peak fitting constraints) to model the single combined N 1s emission line without justification, resulting in values of 398.6 eV for C−N=C and 400.6 eV for N−C=O (note we have added 0.2 eV to account for the use of 284.8 eV as the BE scale reference point in the previous paper).[15] Given that the FWHM of the overall N 1s emission is similar to that of the single N 1s components reported in that paper,

the double-line fit is therefore difficult to justify, and we suggest adopting the value of 399.1 eV for the N 1s BE of both pyridinic and amide nitrogens.

**Table 2.3.** Best fit binding energies (eV), intensities (a.u.) and FWHMs (eV) for the fitted components to the N 1s emission shown in Figure 2.3. Single component fit describes both pyridinic and amide nitrogen indicated by *.

| | | N−C=O | C−N=C | C−NH⁺=C | −NO₂ |
|---|---|---|---|---|---|
| IN24DNBA | BE / eV<br>I / a.u.<br>FWHM / eV | 399.5<br>636.7<br>1.70 | n/a | 401.2<br>617.9<br>1.70 | 406.2<br>813.8<br>1.70 |
| IN35DNBA | BE / eV<br>I / a.u.<br>FWHM / eV | 399.8<br>1078.4<br>1.86 | n/a | 401.1<br>1038.6<br>2.10 | 406.3<br>1309.1<br>1.86 |
| INPA | BE / eV<br>I / a.u.<br>FWHM / eV | 399.9*<br>1093.1*<br>1.51* | 399.9*<br>1093.1*<br>1.51* | 401.6<br>821.6<br>1.51 | n/a |
| IN | BE / eV<br>I / a.u.<br>FWHM / eV | 399.1*<br>1184.9*<br>1.53* | 399.1*<br>1184.9*<br>1.53* | n/a | n/a |

For all three salts, the N 1s emission lines associated with IN in their crystal structures appear at higher BEs than for pure IN. For the amidic nitrogen (N−C=O), this is due to the interaction of the hydrogens with electronegative Brønsted acceptors in the crystal structure, resulting in withdrawal of electron density through N-H--O hydrogen bonds and a BE increase to 399.9 eV (INPA), 399.5 eV (IN24DNBA) and 399.8 eV (IN35DNBA). In the crystal structure of IN, a combination of N-H--O and weaker N-H--N hydrogen bonds are formed between amide groups, resulting in the lower BE. All three salts have a protonated pyridinic IN group (C−NH⁺=C), associated with N 1s BE shifts of 2.0 eV (IN35DNBA), 2.1 eV (IN24DNBA) and 2.5 eV (INPA) relative to pure IN, resulting in N 1s BEs of 401.1 eV, 401.2 eV and 401.6 eV, respectively. In INPA, half of the IN molecules have a hydrogen bond at the pyridinic nitrogen centre, characterised by a BE of 399.9 eV (C−N=C). All these results are fully consistent with previously observed BE ranges for hydrogen bonds in co-crystals (399 eV - 400 eV) and salts in which the nitrogen acceptor is protonated (401 eV – 402 eV).[7,8,11,12,14] Finally, we note that the intensities of all N 1s emission line components (Figure 2.3, Table 2.3) agree very

well with the stoichiometries expected from the crystal structures, indicating a consistent analysis.[28]

### 2.3.3 O 1s Emission

Due to the common presence of oxygen-containing contaminants on samples exposed to air (especially adsorbed $H_2O$ and hydroxyl groups) the O 1s emission lines are usually harder to interpret than the N 1s emission lines. The O 1s emission lines are included here to indicate that a fully consistent interpretation across all emission spectra can be obtained, and to provide reference data for future studies

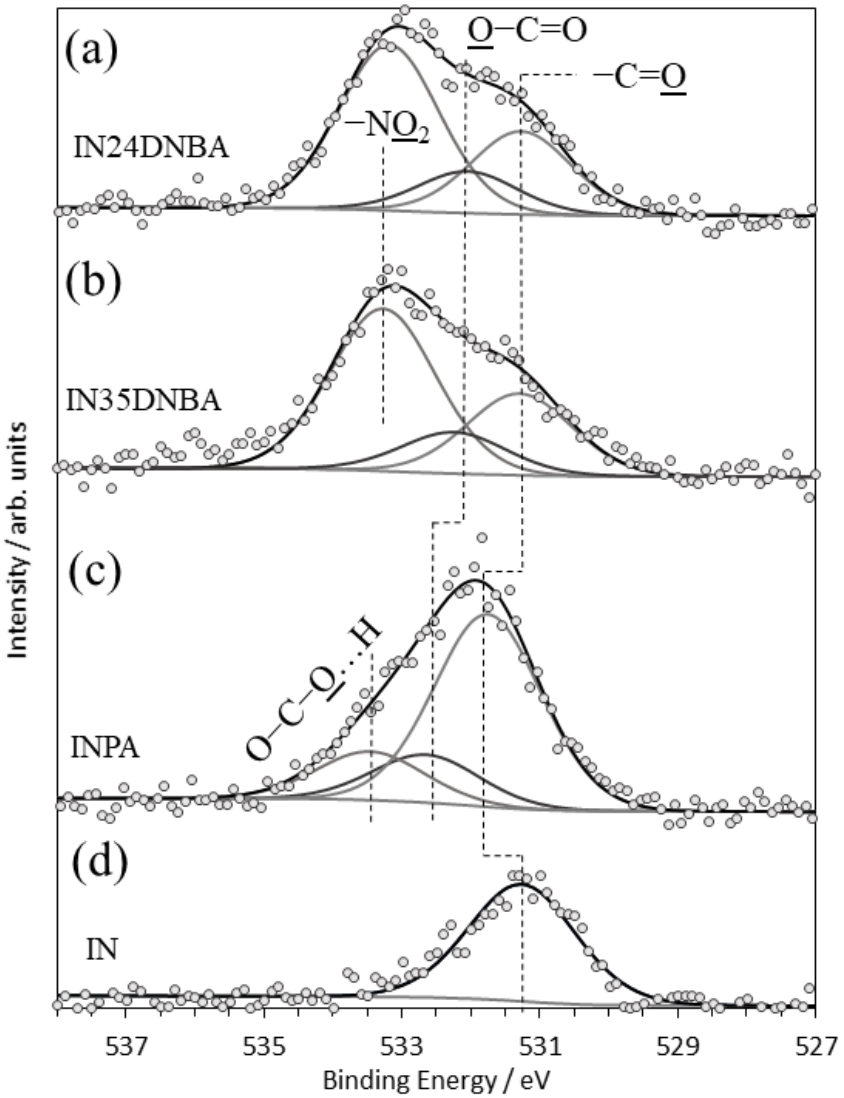



**Figure 2.4.** Oxygen 1s XPS with fitted components for (a) IN24DNBA, (b) IN35DNBA, (c) INPA and (d) pure isonicotinamide. Binding energy scale calibration is taken with respect to the fitted adventitious carbon peak.

of hydrogen bonding involving oxygen as donor and/or acceptor. Figure 2.4 and Table 2.4 give the data with relevant fits and the parameters obtained, respectively.

**Table 2.4.** Best fit binding energies (eV), intensities (a.u.) and FWHMs (eV) for the fitted components to the O 1s emission shown in Figure 2.4

| | | −C=O | O−C=O | $NO_2$ | O−C−O…H |
|---|---|---|---|---|---|
| IN24DNBA | BE / eV | 531.3 | 532.1 | 533.2 | |
| | I / a.u. | 240.2 | 194.8 | 337.9 | n/a |
| | FWHM / eV | 1.71 | 1.71 | 1.71 | |
| IN35DNBA | BE / eV | 531.3 | 532.3 | 533.3 | |
| | I / a.u. | 241.1 | 197.3 | 337.5 | n/a |
| | FWHM / eV | 1.80 | 1.80 | 1.80 | |
| INPA | BE / eV | 531.8 | 532.7 | n/a | 533.5 |
| | I / a.u. | 389.3 | 229.4 | | 233.3 |
| | FWHM / eV | 1.83 | 1.83 | | 1.83 |
| IN | BE / eV | 531.3 | | | |
| | I / a.u. | 280.1 | n/a | n/a | n/a |
| | FWHM / eV | 1.93 | | | |

## 2.4 Discussion

The N 1s BEs determined by XPS show that we can distinguish protonation from hydrogen bonding and classify the examined materials correctly as salts, in line with previously reported BEs for other systems. The XPS classification agrees with the crystallographically determined hydrogen positions in the structures.[16]

The examined salts fall within the $\Delta pK_a$ range (0…3) that is associated with the salt-co-crystal-continuum, for which it is difficult to predict whether hydrogen bonding or proton transfer occurs.[37] In addition, for short donor-acceptor distances the associated potential wells may be separated by a low barrier, so that the hydrogen population in the crystal may reflect a dynamic equilibrium of occupancy across both potential wells.[10] *In extremis*, a continuum of positions between donor and acceptor is possible.[10] Even though the salts examined here exhibit such short donor-acceptor (N---O) distances (see Table A1), there is no evidence in the N 1s emission lines to suggest a disordered or continuum occupancy: there are no additional shoulders or broadening of the emission line components associated with proton transfer (see FWHMs in Table 2.3). This is in line with the crystallographic analysis, which

unequivocally identified proton transfer to the pyridinic nitrogen acceptor of IN. In the same study, the complexes were also investigated with a view to temperature dependent proton migration.[16] Allowing the proton to re-occupy the donor site requires a shallow potential energy surface, which is often observed in structures with short Brønsted donor-acceptor distances,[10,16] However, the temperature-dependent analysis via Fourier difference electron density maps indicated only minor movement of the average hydrogen positions in the temperature range from 100 K to 350 K.[16] In fact, the analysis did not detect any significant proton migration for IN24DNBA and IN35DNBA.[16] For INPA, a slight shift of the average hydrogen positions towards the donor was detected, by about 0.05 Å.[16] This small shift is consistent with the pyridinic acceptor remaining protonated. This suggests that the potential well at the acceptors for all three salts is sufficiently deep to locate the proton, even at elevated temperatures. The weak temperature-dependent migration in INPA may be related to the local geometry of the donor-acceptor interaction: acceptor, hydrogen and donor positions are in a co-linear 180° arrangement. In contrast, they are arranged at an angle close to 160° in the DNBA complexes, which places the proton off the axis between donor and acceptor group.[16]

**Table 2.5.** $\Delta pK_a$ values compared with N-H, N--O distances (from XRD)[12,16,37,38] and N 1s Binding Energy

| | $\Delta pK_a$ | *D (N-H) / Å* | *D (N--O) / Å* | *N 1s BE / eV* |
|---|---|---|---|---|
| IN24DNBA | 2.18 | 1.05 | 2.60 | 401.2 |
| IN35DNBA | 0.94 | 1.08 | 2.54 | 401.1 |
| INPA (salt) | 0.72 | 1.10 | 2.55 | 401.6 |
| INPA (co-crystal) | -1.90 | 1.66 | 2.70 | 399.9 |

Due to the short distance between donor and acceptor, the electron density at the donor site may be expected to influence the electrostatic potential at the acceptor group, and hence its N 1s binding energy. Indeed, this appears to be borne out by the low N 1s BEs (401.1 eV / 401.2 eV) for the protonated pyridinic nitrogen in IN24DNBA and IN35DNBA (Figure 2.5). These BEs represent the lowest values we

have observed so far for any organic salt system (Figure 2.5). Therefore, while the barrier to the donor potential appears to be high enough to prevent, at least up to 350 K, significant re-occupation of the donor position, the proximity of the donor potential appears to reduce the depth of the potential well at the acceptor. This effect may be enhanced by the off-axis position of the proton. For INPA, the N 1s BE of the protonated pyridinic is higher (Figure 2.5), again in line with a potential well that is sufficiently deep to localize the transferred proton, but in this case with less influence of the electrostatic donor potential. The latter may be due to shielding of the donor potential by the electrostatic field of the co-linearly aligned proton. These considerations highlight the importance of the local geometry for classifying Brønsted interactions.

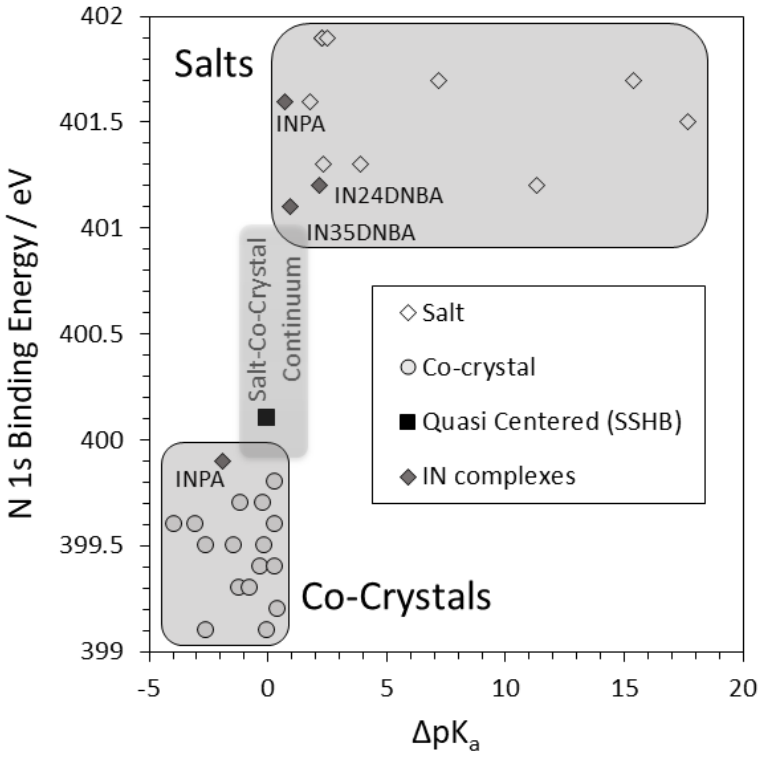



**Figure 2.5.** Relationship between the N 1s core level binding energies (XPS) with the $\Delta pK_a$ for a range of salt, co-crystal and short strong hydrogen bonded crystal structures, including the three complexes investigated here.[12]

It is interesting to examine the salt N 1s BEs in comparison with previously studied salt and co-crystal systems. First, we examine the BEs as a function of $\Delta pK_a$ (Figure 2.5), i.e., the difference between the acid dissociation constant ($pK_a$) of acid and

base [$\Delta pK_a = pK_a$ (base) – $pK_a$ (acid)], known as the $pK_a$ slide rule.[3] This empirical set of rules is generally defined such that a value of $\Delta pK_a > 3$ means a salt is formed, $\Delta pK_a < 0$ means a co-crystal is formed and $0 < \Delta pK_a < 3$ is not definitive.[3] Using the standard $pK_a$ values for the components, the $\Delta pK_a$ values for the three complexes are shown in Table 2.5.[19,38,39] The standard hydrogen bond between phthalic acid and half of the IN molecules in INPA is associated with a $\Delta pK_a$ of –1.90, so the empirical relationship correctly predicts the absence of proton transfer.[3,37] The three $\Delta pK_a$ values associated with proton transfer fall within the salt-co-crystal continuum range, where the empirical relationship is not sufficient to determine whether or not proton transfer occurs.[3] Figure 2.5 shows the relationship between the N 1s BE and the $\Delta pK_a$ value against a range of structures from the literature.[12] It is evident that the two clusters of N 1s BEs reliably classify systems into salt and co-crystals, including in the salt-co-crystal continuum range, $0 < \Delta pK_a < 3$. However, with the data from this paper included, there is a region of overlap around $\Delta pK_a = 0…1$, which we have not observed previously to the same extent.[12] These consolidated XPS results thus confirm that $\Delta pK_a$ in this range is not sufficient to predict which interaction will occur, whilst the binding energy provides a fully consistent approach to distinguishing between the two distinct interactions.

We note, however, that there is a clear separation of salt and co-crystal BEs when the XPS N 1s BE results are plotted as a function of the distance between nitrogen acceptor and proton, d(N-H). While this distance and the binding energy are impacted by many factors, including the surrounding electronic environment, donor-acceptor distance and the interaction angle, it provides an insight into the effect of the proton acceptor distance on the core electronic binding energy. It appears that protonation is associated with distances below 1.15 Å, whereas the shortest distance in a hydrogen bond is about 1.38 Å. Systems in the salt-co-crystal continuum would then appear to be associated with intermediate distances. So far, only one previously studied system can be confidently placed in this range,[10] and more systems should be evaluated to create a better understanding in this range of the diagram.

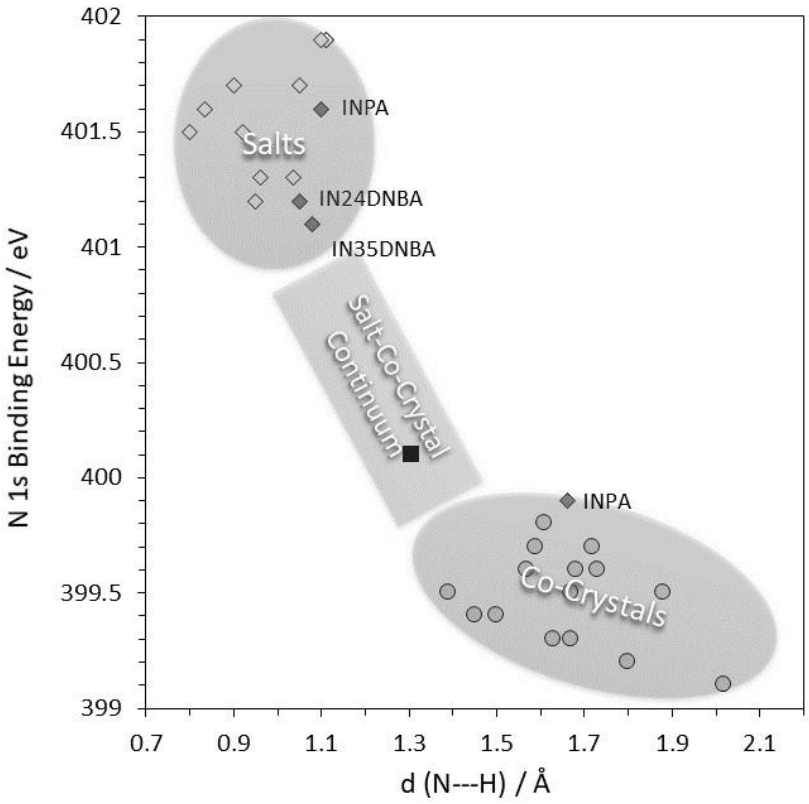



**Figure 2.6.** Relationship between the N 1s core level binding energies (XPS) with the distance between the Brønsted proton and the acceptor nitrogen in a range of crystal structures including the three complexes investigated here. All distances determined using X-ray diffraction and sourced from the Cambridge Structural Database.[40]

This classification is fully consistent with the physical process, whereby proton transfer leads to a shorter proton acceptor distance. The most important aspect of Figure 2.6 is the consistency with which salts and co-crystals are separated uniquely by the electron binding energy, with co-crystals between 399 eV – 400 eV, and salts between 401 eV – 402 eV.[7,12,14,30] This is the case despite the measurements being from a range of independent investigations carried out by our group over the last decade. Using the analysis procedure laid out in the present paper is key for achieving this consistency. Therefore, we suggest that the effect on the binding energy from the donor acceptor distance and interaction angle is sufficiently small such that the binding energy will always lie within the range we have consistently observed, with the variation within that range due to additional effects, including proton-acceptor distance, donor-acceptor distance as well as the angle between acceptor, proton and donor.

## 2.5 Conclusions

N 1s BEs determined by XPS correctly classified the nature of Brønsted interactions in three two-component isonicotinamide systems with short donor-acceptor distances. We demonstrated the data analysis procedure allowing full quantification of all data in a comprehensive XPS data set, including the quantitative fitting analysis of all (C, N, O) emission lines, the identification of the carbon contamination signal and the calibration of the binding energy scale through the C 1s emission line. For all three compounds, INPA, IN24DNBA and IN35DNBA, the N 1s BEs indicated the protonation of the pyridinic nitrogen of the isonicotinamide component. The results for IN24DNBA and IN35DNBA indicate that the transferred proton experiences significant electrostatic attraction from the donor group, resulting in low BEs that place these systems on the edge of the salt-co-crystal continuum. The transferred proton in INPA is aligned co-linearly with the donor and the acceptor and appears to shield the acceptor from the influence of the electron density at the donor. Summarising all XPS results we have obtained for other salt and co-crystal systems as a function of $\Delta pK_a$ values highlights the predictive failure of the $\Delta pK_a$ method in the $\Delta pK_a$ region from 0 to about 1. In this region, there appears to be higher predictive power in the distance between the acceptor group and the proton, suggesting a universal dependence of the observed N 1s binding energy on the distance to the proton.

**Acknowledgements**

XPS data were collected at the Versatile X-ray Spectroscopy Facility (VXSF), which is based in the Bragg Centre at the University of Leeds and funded by the Sir Henry Royce Institute (EPSRC Grants EP/R00661X/1, EP/S019367/1 and EP/P022464/1). PTE gratefully acknowledges a PhD studentship funded jointly by Diamond and through the EPSRC Doctoral Training Partnership (DTP) with the University of Leeds (EPSRC Grant EP/R513258/1). SLMS thanks the Royal Academy of Engineering, Diamond Light Source Ltd and Infineum UK for support of the Bragg Centenary Chair, and for financial support from the Future

Continuous Manufacturing and Advanced Crystallisation (CMAC) Hub (EPSRC Grant EP/P006965/1). All data supporting this study are provided either in the results section of this paper or in the ESI accompanying it.

**Supporting Information: Survey spectra of all samples, hydrogen bonding parameters and spectra showing the effects of radiation damage and sample charging.**

## 2.6 References

(1) Blagden, N.; de Matas, M.; Gavan, P. T.; York, P. Crystal Engineering of Active Pharmaceutical Ingredients to Improve Solubility and Dissolution Rates. *Advanced Drug Delivery Reviews*. July 30, **2007**, 617–630. DOI: 10.1016/j.addr.2007.05.011.

(2) Etter, M. C. Encoding and Decoding Hydrogen-Bond Patterns of Organic Compounds. *Accounts of Chemical Research* **1990**, *23* (4), 120–126.

(3) Gilli, P.; Pretto, L.; Bertolasi, V.; Gilli, G. Predicting Hydrogen-Bond Strengths from Acid-Base Molecular Properties. The pK(a) Slide Rule: Toward the Solution of a Long-Lasting Problem. *Accounts of Chemical Research* **2009**, *42* (1), 33–44. DOI: 10.1021/ar800001k.

(4) Aakeröy, C. B.; Fasulo, M. E.; Desper, J. Cocrystal or Salt: Does It Really Matter? *Molecular Pharmaceutics* **2007**, *4* (3), 317–322. DOI: 10.1021/mp060126o.

(5) Childs, S. L.; Stahly, G. P.; Park, A. The Salt-Cocrystal Continuum: The Influence of Crystal Structure on Ionization State. *Molecular Pharmaceutics* **2007**, *4* (3), 323–338. DOI: 10.1021/mp0601345.

(6) Wouters, J.; Quéré, L. *Pharmaceutical Salts and Co-Crystals*; Drug Discovery; The Royal Society of Chemistry, **2012**. DOI: 10.1039/9781849733502.

(7) Stevens, J. S.; Byard, S. J.; Schroeder, S. L. M. Salt or Co-Crystal? Determination of Protonation State by X-Ray Photoelectron Spectroscopy (XPS). *Journal of Pharmaceutical Sciences* **2010**, *99* (11), 4453–4457. DOI: 10.1002/jps.22164.

(8) Stevens, J. S.; Byard, S. J.; Schroeder, S. L. M. Characterization of Proton Transfer in Co-Crystals by X-Ray Photoelectron Spectroscopy (XPS). *Crystal Growth & Design* **2010**, *10* (3), 1435–1442. DOI: 10.1021/cg901481q.

(9) Stevens, J. S.; Byard, S. J.; Muryn, C. A.; Schroeder, S. L. Identification of Protonation State by XPS, Solid-State NMR, and DFT: Characterization of the Nature of a New Theophylline Complex by Experimental and Computational Methods. *J Phys Chem B* **2010**, *114* (44), 13961–13969. DOI: 10.1021/jp106465u.

(10) Stevens, J. S.; Coultas, S.; Jaye, C.; Fischer, D. A.; Schroeder, S. L. M. Core Level Spectroscopies Locate Hydrogen in the Proton Transfer Pathway-Identifying Quasi-Symmetrical Hydrogen Bonds in the Solid State. *Physical Chemistry Chemical Physics* **2020**, *22* (9), 4916–4923. DOI: 10.1039/c9cp05677g.

(11) Stevens, J. S.; Newton, L. K.; Jaye, C.; Muryn, C. A.; Fischer, D. A.; Schroeder, S. L. M. Proton Transfer, Hydrogen Bonding, and Disorder: Nitrogen Near-Edge X-Ray Absorption Fine Structure and X-Ray Photoelectron Spectroscopy of Bipyridine-Acid Salts and Co-Crystals. *Crystal Growth & Design* **2015**, *15* (4), 1776–1783. DOI: 10.1021/cg5018278.

(12) Stevens, J. S.; Byard, S. J.; Seaton, C. C.; Sadiq, G.; Davey, R. J.; Schroeder, S. L. M. Proton Transfer and Hydrogen Bonding in the Organic Solid State: A Combined XRD/XPS/SsNMR Study of 17 Organic Acid-Base Complexes. *Physical Chemistry Chemical Physics* **2014**, *16* (3), 1150–1160. DOI: 10.1039/c3cp53907e.

(13) Wang, T.; Stevens, J. S.; Vetter, T.; Whitehead, G. F. S.; Vitorica-Yrezabal, I. J.; Hao, H. X.; Cruz-Cabeza, A. J. Salts, Cocrystals, and Ionic Cocrystals of a "Simple" Tautomeric Compound. *Crystal Growth & Design* **2018**, *18* (11), 6973–6983. DOI: 10.1021/acs.cgd.8b01159.

(14) Stevens, J. S.; Byard, S. J.; Seaton, C. C.; Sadiq, G.; Davey, R. J.; Schroeder, S. L. M. Crystallography Aided by Atomic Core-Level Binding Energies: Proton Transfer versus Hydrogen Bonding in Organic Crystal Structures. *Angewandte Chemie-International Edition* **2011**, *50* (42), 9916–9918. DOI: 10.1002/anie.201103981.

(15) Tothadi, S.; Shaikh, T. R.; Gupta, S.; Dandela, R.; Vinod, C. P.; Nangia, A. K. Can We Identify the Salt-Cocrystal Continuum State Using XPS? *Crystal Growth and Design* **2021**, *21* (2), 735–747. DOI: 10.1021/acs.cgd.0c00661.

(16) Saunders, L. K.; Nowell, H.; Hatcher, L. E.; Shepherd, H. J.; Teat, S. J.; Allan, D. R.; Raithby, P. R.; Wilson, C. C. Exploring Short Strong Hydrogen Bonds Engineered in Organic Acid Molecular Crystals for Temperature Dependent Proton Migration Behaviour Using Single Crystal Synchrotron X-Ray Diffraction (SCSXRD). *CrystEngComm* **2019**, *21* (35), 5249–5260. DOI: 10.1039/C9CE00925F.

(17) Saunders, L. K.; Nowell, H.; Spencer, H. C. E.; Hatcher, L. E.; Shepherd, H. J.; Thomas, L. H.; Jones, C. L.; Teat, S. J.; Raithby, P. R.; Wilson, C. C. Tuning Charge-Assisted and Weak Hydrogen Bonds in Molecular Complexes of the Proton Sponge DMAN by Acid Co-Former Substitution. *Crystengcomm* **2018**, *20* (22), 3074–3083. DOI: 10.1039/c8ce00443a.

(18) Schmidtmann, M.; Gutmann, M. J.; Middlemiss, D. S.; Wilson, C. C. Towards Proton Transfer in Hydrogen Bonded Molecular Complexes: Joint Experimental and Theoretical Modelling and an Energy Scale for Polymorphism. *Crystengcomm* **2007**, *9* (9), 743–745. DOI: 10.1039/b709136m.

(19) Sánchez-Férez, F.; Ejarque, D.; Calvet, T.; Font-Bardia, M.; Pons, J. Isonicotinamide-Based Compounds: From Cocrystal to Polymer. *Molecules* **2019**, *24* (22). DOI: 10.3390/molecules24224169.

(20) Saunders, L. K.; Pallipurath, A. R.; Gutmann, M. J.; Nowell, H.; Zhang, N.; Allan, D. R. A Quantum Crystallographic Approach to Short Hydrogen Bonds. *CrystEngComm* **2021**. DOI: 10.1039/d1ce00355k.

(21) Steiner, T.; Majerz, I.; Wilson, C. C. First O-H-N Hydrogen Fond with a Centered Proton Obtained by Thermally Induced Proton Migration. *Angewandte Chemie-International Edition* **2001**, *40* (14), 2651–2654 DOI: 10.1002/1521-3773(20010716)40:14<2651::aid-anie2651>3.0.co;2-2.

(22) Deshkovskaya, A.; Shchukarev, A. XPS Study of Fluorine Implantation Unduced Centres in Fused Silica; *10th International Conference "Interaction of Radiation with Solids"* 2013.

(23) Stevens, J. S.; Gainar, A.; Jaye, C.; Fischer, D. A.; Schroeder, S. L. M.; Iop. NEXAFS and XPS of P-Aminobenzoic Acid Polymorphs: The Influence of Local Environment. In *16th International Conference on X-ray Absorption Fine Structure (XAFS)*; Journal of Physics Conference Series; **2016**; Vol. 712 DOI: 10.1088/1742-6596/712/1/012133.

(24) Greczynski, G.; Hultman, L. Compromising Science by Ignorant Instrument Calibration—Need to Revisit Half a Century of Published XPS Data. *Angewandte Chemie* **2020**, *132* (13), 5034–5038. DOI: 10.1002/ange.201916000.

(25) Greczynski, G.; Hultman, L. X-Ray Photoelectron Spectroscopy: Towards Reliable Binding Energy Referencing. *Progress in Materials Science*. Elsevier Ltd January 1, 2020. DOI: 10.1016/j.pmatsci.2019.100591.

(26) Baer, D. R.; Artyushkova, K.; Cohen, H.; Easton, C. D.; Engelhard, M.; Gengenbach, T. R.; Greczynski, G.; Mack, P.; Morgan, D. J.; Roberts, A. XPS Guide: Charge Neutralization and Binding Energy Referencing for Insulating Samples. *Journal of Vacuum Science & Technology A* **2020**, *38* (3), 031204. DOI: 10.1116/6.0000057.

(27) Dietrich, P. M.; Bahr, S.; Yamamoto, T.; Meyer, M.; Thissen, A. Chemical Surface Analysis on Materials and Devices under Functional Conditions – Environmental Photoelectron Spectroscopy as Non-Destructive Tool for Routine Characterization. *Journal of Electron Spectroscopy and Related Phenomena* **2019**, *231*, 118–126. DOI: 10.1016/j.elspec.2017.12.007.

(28) Fairley, N.; Carrick, A. *The Casa Cookbook: Recipes for XPS Data Processing*; Casa cookbook; Acolyte Science: Knutsford, Cheshire, 2005.

(29) Shard, A. G. Practical Guides for X-Ray Photoelectron Spectroscopy: Quantitative XPS. *Journal of Vacuum Science & Technology A* **2020**, *38* (4), 041201. DOI: 10.1116/1.5141395.

(30) Stevens, J. S.; Schroeder, S. L. M. Quantitative Analysis of Saccharides by X-Ray Photoelectron Spectroscopy. *Surface and Interface Analysis* **2009**, *41* (6), 453–462. DOI: 10.1002/sia.3047.

(31) Reed, B. P.; Cant, D. J. H.; Spencer, S. J.; Carmona-Carmona, A. J.; Bushell, A.; Herrera-Gómez, A.; Kurokawa, A.; Thissen, A.; Thomas, A. G.; Britton, A. J.; Bernasik, A.; Fuchs, A.; Baddorf, A. P.; Bock, B.; Theilacker, B.; Cheng, B.; Castner, D. G.; Morgan, D. J.; Valley, D.; Willneff, E. A.; Smith, E. F.; Nolot, E.; Xie, F.; Zorn, G.; Smith, G. C.; Yasufuku, H.; Fenton, J. L.; Chen, J.; Counsell, J. D. P.; Radnik, J.; Gaskell, K. J.; Artyushkova, K.; Yang, L.; Zhang, L.; Eguchi, M.; Walker, M.; Hajdyła, M.; Marzec, M. M.; Linford, M. R.; Kubota, N.; Cortazar-Martínez, O.; Dietrich, P.; Satoh, R.; Schroeder, S. L. M.; Avval, T. G.; Nagatomi, T.; Fernandez, V.; Lake, W.; Azuma, Y.; Yoshikawa, Y.; Shard, A. G. Versailles Project on Advanced Materials and Standards Interlaboratory Study on Intensity Calibration for X-Ray Photoelectron Spectroscopy Instruments Using Low-Density Polyethylene. *Journal of Vacuum Science & Technology A* **2020**, *38* (6), 063208. DOI: 10.1116/6.0000577.

(32) Zommer, L. Determination of the Spectrometer Transmission Function for XPS Quantitative Analysis. *Vacuum* **1995**, *46* (5–6), 617–620. DOI: 10.1016/0042-207X(94)00143-X.

(33) ISO 15472:2010. Surface Chemical Analysis X-Ray Photoelectron Spectrometers Calibration of Energy Scales. ISO: Geneva 2010.
(34) ASTM E1523-15. Standard Guide to Charge Control and Charge Referencing Techniques in X-Ray Photoelectron Spectroscopy (ASTM International, West Conshohocken, PA, 2015)
(35) Mangolini, F.; McClimon, J. B.; Rose, F.; Carpick, R. W. Accounting for Nanometer-Thick Adventitious Carbon Contamination in X-Ray Absorption Spectra of Carbon-Based Materials. *Analytical Chemistry* **2014**, *86* (24), 12258–12265. DOI: 10.1021/ac503409c.
(36) Naumkin, A. v.; Kraut-Vass, A.; Gaarenstroom, S. W.; Powell, C. J. NIST X-ray Photoelectron Spectroscopy Reference Database 20 Version 4.1, DOI: 10.18434/T4T88K
(37) Chen, J. G.; McAllister, M. A.; Lee, J. K.; Houk, K. N. Short, Strong Hydrogen Bonds in the Gas Phase and in Solution: Theoretical Exploration of PKa Matching and Environmental Effects on the Strengths of Hydrogen Bonds and Their Potential Roles in Enzymatic Catalysis. *Journal of Organic Chemistry* **1998**, *63* (14), 4611–4619. DOI: 10.1021/jo972262y.
(38) Williams R. *PKa Data Compiled by R. Williams,* 2011 https://organicchemistrydata.org/hansreich/resources/pka/pka_data/pka-compilationwilliams.pdf [Online accessed on 20/01/2021].
(39) Harding, A. P. The Prediction of Mutagenicity and pKa for Pharmaceutically Relevant Compounds Using “Quantum Chemical Topology” Descriptors; PhD Thesis, The University of Manchester, 2010.
(40) Groom, C. R.; Bruno, I. J.; Lightfoot, M. P.; Ward, S. C. The Cambridge Structural Database. *Acta Crystallographica Section B: Structural Science, Crystal Engineering and Materials* **2016**, *72* (2), 171–179. DOI: 10.1107/S2052520616003954.

# 3 Determination of H-atom Positions in Organic Crystal Structures by NEXAFS Combined with Density Functional Theory: A Study of Two-Component Systems Containing Isonicotinamide

*Paul T. Edwards,[1, 2] Lucy K. Saunders,[2] David C. Grinter,[2] Pilar Ferrer,[2] Georg Held,[2] Elizabeth J. Shotton,[2]* Sven L. M. Schroeder[1,2,3]**

[1]School of Chemical and Process Engineering, University of Leeds, Leeds, LS2 9JT, UK
[2]Diamond Light Source, Harwell Science & Innovation Campus, Didcot, OX11 0DE, UK
[3]Future Continuous Manufacturing and Advanced Crystallisation Hub, Research Complex at Harwell (RCaH), Rutherford Appleton Laboratory, Didcot, OX11 0FA, UK

* Sven L. M. Schroeder:
Email: s.l.m.schroeder@leeds.ac.uk

* Elizabeth J. Shotton:
Email: elizabeth.shotton@diamond.ac.uk

**Abstract**

It is important to be able to identify the precise position of H-atoms in hydrogen bonding interactions to fully understand the effects on the structure and properties of organic crystals. Using a combination of Near Edge X-ray Absorption Fine Structure (NEXAFS) spectroscopy and Density Functional Theory (DFT) quantum chemistry calculations we demonstrate the sensitivity of core level X-ray spectroscopy to the precise H-atom position within a donor – proton – acceptor system. Exploiting this sensitivity, we then combine the predictive power of DFT with the experimental NEXAFS, confirming the H-atom position identified using single crystal X-ray Diffraction (XRD) techniques more easily than using other H-atom sensitive techniques, such as neutron diffraction. This proof of principle experiment confirms the H-atom positions in structures obtained from XRD, providing evidence for the potential use of NEXAFS as a more accurate and easier method of locating H-atoms within organic crystals.

## 3.1 Introduction

Hydrogen bonding and Brønsted proton transfer interactions are fundamental to the formation of organic crystal structures.[1] The location of the H-atom between the donor and acceptor atoms in these interactions has an important effect on the structural properties of the crystal; understanding is particularly relevant for regulatory bodies identifying structures for patent protection of pharmaceutical compounds.[2–4] Hydrogen bonding in co-crystals, and proton transfer in salts, form the extremities of the salt-cocrystal continuum describing the two states where the hydrogen atom is either located close to the proton donor (co-crystal) or the proton acceptor (salt).[1,4,5] Our previous studies have consistently demonstrated how core level spectroscopy in the form of X-ray Photoelectron Spectroscopy (XPS) is sensitive to the position of the H-atom through the 1s core level binding energy (BE) shift at the proton acceptor, as shown in Figure 3.1.[3,6–9] This effect is also present in near-edge X-ray absorption fine structure (NEXAFS) spectroscopy through a shift dependent upon both the core level BE and the energy of the final state unoccupied molecular orbitals.[3,7–17] Due to the large number of unoccupied π* (where π bonding is present) and σ* orbitals, which are both highly sensitive to the molecular and electronic structure,[17] the technique has the potential to enable a structure refinement directly sensitive to the H-atom. The origin of the features in a NEXAFS spectrum are outlined in Figure 3.1, with the energy of the N1s → π* transition indicated by the length of the arrow. The shift in energy from the pure isonicotinamide to the hydrogen bonded complex can be attributed to the effect of the complex lowering the energy of the 1s orbital, resulting in a shift in the photon energy of ΔE.

Previous studies have focused on the direct effects on the experimental NEXAFS spectra for identifying bond lengths. The so-called “Bond length with a ruler” method has been proposed [12,17–20] identifying trends between the energy position of the ionization potential and σ* shape resonances. In many cases, this provides a useful tool to identify differences between similar compounds. However, the approach is not sufficiently sensitive to conformational variations and other non-covalent interactions to act as a more general structure refinement tool.[12]

Complementing experimental NEXAFS, time dependent DFT (TDDFT) calculations can be conducted to model the theoretical NEXAFS spectrum based upon the calculated excitation energies and transition dipole moments.[21] The approximations made in the formalism of DFT have been shown to not adversely affect the properties

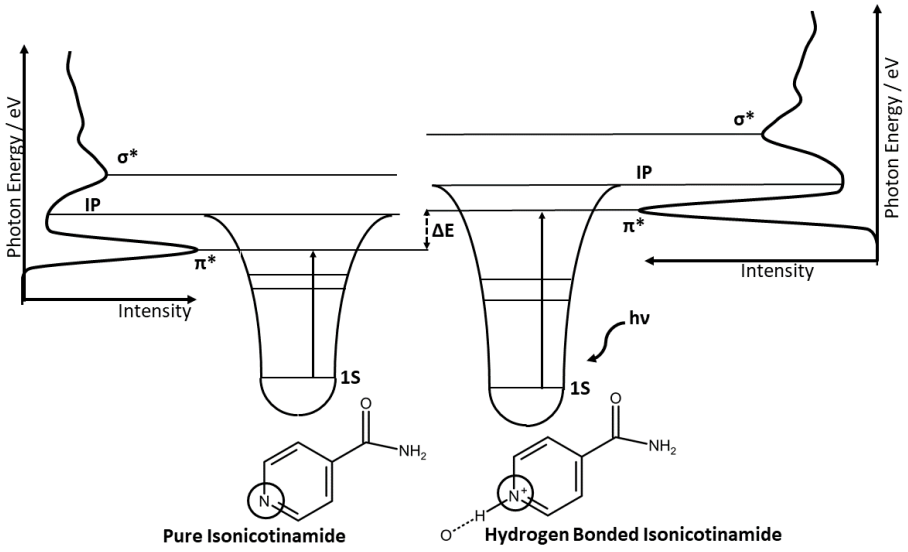



**Figure 3.1.** Schematic of the principal transitions for the pyridinium nitrogen showing the energy levels of the 1s, π* and σ* orbitals, and the ionization potential. The effect of hydrogen bonding and proton transfer increasing the photon energy of the N 1s → π* transition is shown due to stabilizing of the N 1s orbital in the hydrogen bonded complex.

calculated.[21–23] One of the most common exchange correlation functionals used for DFT calculations is the B3LYP functional,[24] and there have been studies where this is shown to underestimate excitation energies compared to experiment.[22,23,25] However, relative excitation energies are found to be correct, and therefore a uniform shift is generally applied to allow comparison to experimental data.[11,26] Using this approach, it is possible to model the effect of H-atom position on the experimental NEXAFS spectrum, with the aim of identifying how sensitive the technique is to the precise location of the H-atom.

An alternative commonly used method for finding such a precise H-atom location is neutron diffraction, although sample preparation is more onerous and data acquisition more laborious.[27,28] The difficulty of locating H-atoms using X-ray diffraction (XRD) stems from the fundamental insensitivity of the technique to H-atom positions. The diffraction of X-rays relies on electron density at the scattering sites,

and by their nature hydrogen atoms have low electron density.[29] Since this electron is occupied in a bond, this electron density is also biased towards the donor atoms to which the H-atom is bonded.[27] This leads to a general lack of sensitivity to the exact H-atom location, and a requirement for additional refinement methods for H-atom positions.[30,31] Here we aim to complement and improve on this theoretical refinement modelling with NEXAFS spectroscopy and DFT calculations exploring the determination of the location of H-atoms within a hydrogen bond for systems for which their location has previously been determined.

In a recent study we examined three isonicotinamide systems using XPS, in which isonicotinamide forms 1:1 salts with 2,4-dinitrobenzoic acid and 3,5-dinitrobenzoic acid, and a 2:1 structure with phthalic acid containing two distinct interactions between the components: one Brønsted proton transfer interaction and one standard hydrogen bond.[6,31] For these isonicotinamide 2,4-dinitrobenzoic acid (IN24DNBA), isonicotinamide 3,5-dinitrobenzoic acid (IN35DNBA) and isonicotinamide phthalic acid (INPA) systems the expected chemical shift of the N 1s BE was observed, confirming that the core level BEs are consistent with the crystallographically determined interactions. Using NEXAFS, we have now investigated these complexes further to confirm the XRD-based H-atom location procedure experimentally (the XRD-derived hydrogen bond distances can be found in Table B1 of the supporting information in appendix B).[31] In effect, this forms the basis of a “proof of principle” experiment showing how NEXAFS could be used to complement structure refinement by XRD.

## 3.2 Modelling and Experimental Methods

### 3.2.1 Density Functional Theory Calculations

DFT calculations were completed using the *ORCA* software package version 5.0.1, utilizing the University of Leeds high performance computing facilities.[32,33] Example *ORCA* input file and .xyz files for calculations are available in the supporting information. For molecular geometry optimization calculations, the B3LYP[24] exchange correlation functional was used along with the def2-TZVP basis set[34], def2/J auxiliary basis set and the RIJCOSX approximation.[35] For the ground state

geometry optimization, this is a sufficient level of theory allowing accurate molecular geometries to be calculated in reasonable timescales. Additionally, geometry optimization of larger clusters of molecules were calculated using the same approximations. Starting structures were obtained from the Cambridge Structural Database (CSD).[36] Geometry optimization parameters were all left at the default values in *ORCA*. Hydrogen position optimization was carried out on all structures (such that the other atoms remained in the XRD measured positions), in addition to a full optimization of every atom position in the molecule. Where a full optimization was done, a vibrational frequency calculation was also completed to ensure the structure reached a global minimum.

For the modelling of NEXAFS spectra, TDDFT calculations were run on the various crystal and DFT optimized structures. TDDFT is used to calculate excited states and can be used to model an X-ray absorption spectrum based on the transition probabilities between occupied and unoccupied molecular orbitals and the energies calculated.[21] By allowing only transitions from core level orbitals of specific atoms, we can calculate a theoretical NEXAFS spectrum accounting for all of the electronic transitions within a certain energy range.[21] The same B3LYP[24] exchange correlation functional and def2-TZVP basis set were used in all calculations.[34] By default, the Tamm-Dancoff approximation[37] was applied which simplifies the calculation of excited states without significantly affecting the energies calculated, and improving them for some cases involving triplet excited states. Due to the consistent underestimate of excitation energies using this exchange correlation functional, a uniform shift of typically +12.4 eV was applied to all calculated spectra.

### 3.2.2 Materials

All three samples were prepared by evaporation from solution, as described in previous work.[31] The crystallizations were carried out in methanol (INPA) and ethanol (IN24DNBA and IN35DNBA) using a 2:1 stoichiometric ratio (INPA) or 1:1 ratio (IN24DNBA and IN35DNBA).[31] The individual components for the crystallizations (isonicotinamide (99.9%), 2,4-dinitrobenzoic acid (98%), 3,5-

dinitrobenzoic acid (99%), and phthalic acid (99.5%)) were obtained from Sigma-Aldrich.

### 3.2.3 Near Edge X-ray Absorption Fine Structure (NEXAFS)

NEXAFS characterization of the powder samples was carried out at Diamond Light Source beamline VerSoX B07-B utilizing the near ambient pressure NEXAFS end station enabling fast sample transfer.[38] Measurements were carried out using total electron yield (TEY) detection, measuring the drain current through the sample mounting plate with the soft X-ray beam at normal incidence and at room temperature. Samples were mounted on metal sample holders with carbon tape on copper foil, with the crystalline powders pressed onto the carbon tape. The photon energy scale was calibrated for beamline effects and surface charging using a small trace of nitrogen gas in the sample chamber which is therefore present in the gas phase spectra. This is well known to form a peak at 400.8 eV [39,40], to which the photon energy scale was calibrated by applying a uniform shift to each spectrum. The sample chamber was operated with He gas at 1 mbar. He was chosen as there are no absorption peaks in the region around the C, N, O K-edges, and because it minimizes the electron yield signal of fluorescence gas phase absorption. The incident X-ray beam intensity $I_0$ is required for determining the absorption spectrum (see supplementary information) and was acquired by measuring the He gas phase spectrum when the sample plate was removed from the beam path. There is a difference in the gas phase beam path from which electrons are emitted when the sample is removed in this way, and this needs to be accounted for during the spectrum normalization process (see supplementary information). Athena XAS analysis software[41] was used to model the spectra through curve fitting with error functions to model the absorption edge steps associated with ionization at the photoemission thresholds and Gaussian bell curves to model the absorption lines associated with core excitations to unoccupied bound states. On fitted spectra, these edge steps are indicated by the ionisation potential (IP) and absorption peaks indicated by the relevant transition. The energy of the ionisation potentials was determined using the binding energy of the relevant group taken from XPS measurement.

## 3.3 Results

DFT modelled spectra were calculated using the orca_mapspc program within the *ORCA* software suite[33], using a Gaussian line broadening of 0.6 eV to compute the X-ray absorption spectra, as this best matches the experimental spectra for the N 1s → π* transitions. The photon energy scale is shifted by approximately 12 eV such that the main peak matches with the experimentally observed value. This is necessary due to the consistent underestimation of excitation energies calculated using the B3LYP exchange correlation functional.[22] An alternative exchange correlation functional, designed for the calculation of core level properties[42] was tested as well, see Figure B5. While the absolute values were generally in agreement with the experiment, the relative offset of the peak positions compared poorly to experiment, particularly for the amide groups. We also calculated the spectra using a range of larger basis sets (including additional polarization functions, additional

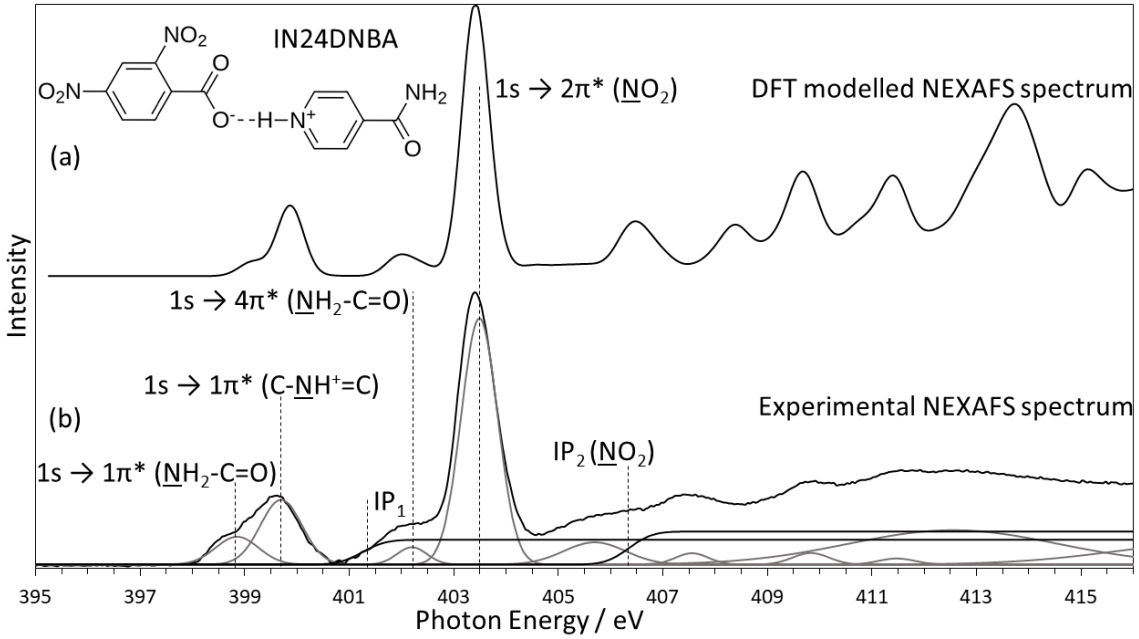



**Figure 3.2.** Modelled and experimental N K-edge NEXAFS spectrum of IN24DNBA showing excellent agreement of π* resonance peak positions. Experimental NEXAFS spectrum is fitted with Gaussian peaks identifying transitions and error functions ($IP_1$ and $IP_2$) at the core electron binding energies determined using XPS. N 1s → π* transitions are identified based on the origin of the core level electron and the π* orbital related to the relevant ring structure. Gaussian peaks at photon energies above 405 eV are to allow data fitting only, and we extract no physical meaning from these.

diffuse functions) to investigate the basis set convergence. We observed that additional diffuse basis functions lower the energies of the outermost orbitals reducing the excitation energies of electrons into these molecular orbitals. However, the peaks of interest, in particular the N 1s → π* transitions, remained at the same energy, leading to the decision to use the lower computational cost of the smaller def2-TZVP basis set.[34] The calculations using alternative exchange correlation functional and basis sets are available in the supporting information (Appendix B, Figures B4 and B5).

### 3.3.1 Isonicotinamide 2,4-dinitrobenzoic acid (IN24DNBA)

According to X-ray crystallography, the IN24DNBA complex forms a 1:1 salt structure.[31] Using XPS, we have confirmed this through an N 1s binding energy shift at the proton acceptor of +2 eV compared to a co-crystal, with relative N 1s emission intensities for all nitrogen moieties in line with the bulk stoichiometry, confirming that the composition within the near-surface region probed by XPS is the same as the bulk.[6] Figure 3.2 shows the modelled and experimental nitrogen K-edge NEXAFS spectra for this salt. It is immediately obvious that the modelled structure very closely resembles the experimental NEXAFS data. The photon energy scale of the modelled spectrum was shifted by 12.15 eV from the originally calculated value, using the nitro group 1s → 2π* transition for photon energy scale calibration. Peak assignments are aided through the use of DFT modelling, since the electronic transitions responsible for each peak can be directly identified, as shown in Figure 3.2(b). The three nitrogen environments (amide ($NH_2$-C=O), pyridinium ring (C-$NH^+$=C) and nitro (-$NO_2$)) are each associated with a distinct N 1s → π* resonance, consistent with our previous XPS analysis.[6] The nitro group is the most intense and visible at a photon energy of 403.5 eV. The protonated nitrogen of the pyridinium results in the peak at 399.7 eV, while the amide 1s → 1π* transition appears as a low energy shoulder to the protonated nitrogen, the peak of that shoulder being at 398.9 eV. Figure 3.2(b) shows the experimental spectrum including Gaussian peak fitting and step functions for the ionization potentials, using the N 1s binding energies determined from XPS. Higher energy peaks in the electron backscattering regime are also fitted with Gaussians, as this facilitates more accurate fits in the NEXAFS range, but we stress that Gaussians are physically meaningless in this region, which reflects oscillatory

extended X-ray absorption fine-structure (EXAFS) variations of the absorption coefficient. The low energy shoulder to the nitro peak at 402.2 eV can easily be attributed to the amide N 1s → 4π* transitions, while higher energy transitions are primarily due to transitions to σ* orbitals (shape resonances). The intensities of the transitions associated with the three nitrogen species are influenced by the nature of the initial and final state. This is unlike XPS, where the peak area is proportional to the excitation cross section of the atomic initial state, and therefore proportional to the concentration of the excited moiety within the probed volume. The absorption intensity in NEXAFS is additionally influenced by the net transition rate from the excited core level to the unoccupied molecular orbitals, resulting in the relatively weak amide absorption band. In previous work, we have not observed this effect to the same extent,[12,13,19,43] but encouragingly the DFT calculated spectrum predicts the same relative intensities, confirming this analysis.

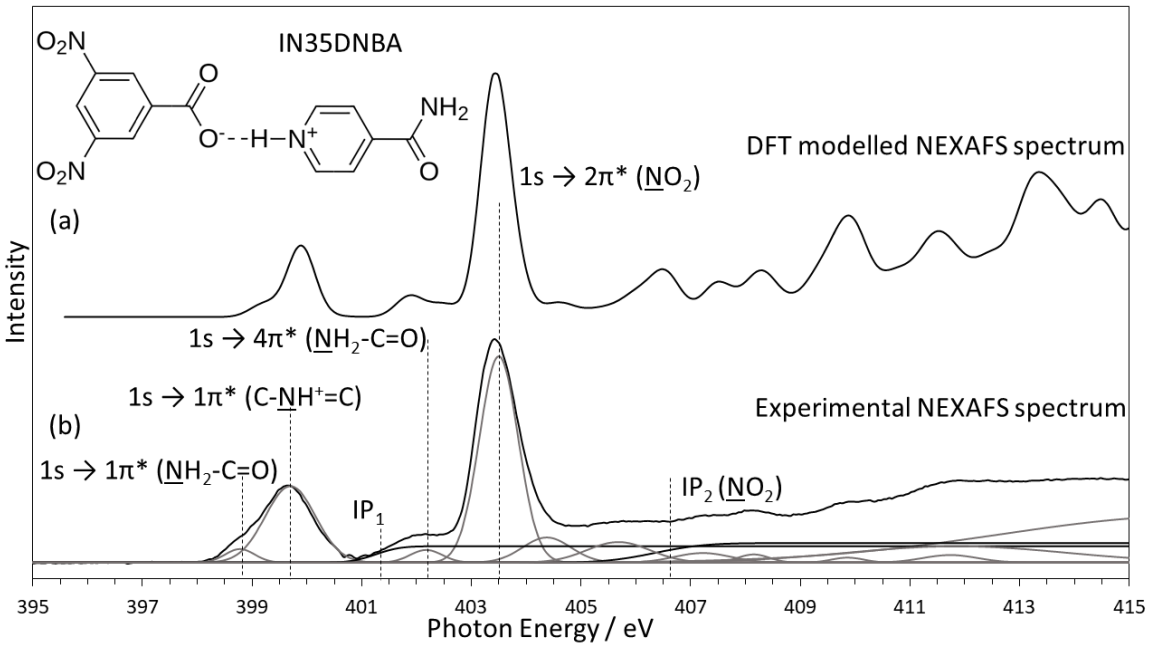



**Figure 3.3.** Modelled and experimental N K-edge NEXAFS spectrum of IN35DNBA showing excellent agreement of π* resonance peak positions. IP's taken from XPS measurement. Overall peak shapes also fully consistent with salt structure with low energy amide peak (398.8 eV) and higher energy protonated nitrogen peak (399.7 eV). Additional feature compared with IN24DNBA acid above nitro peak (404.4 eV) present in both spectra visible as a slight broadening.

### 3.3.2 Isonicotinamide 3,5-dinitrobenzoic acid (IN35DNBA)

The IN35DNBA complex forms a 1:1 salt structure.[31] Figure 3.3 shows the experimental and modelled NEXAFS spectra for this sample. It is immediately clear that the overall peak shapes form the same pattern as IN24DNBA, with an additional shoulder peak above the nitro peak, and different σ* resonances, again with the modelled spectrum suggesting larger intensities than observed experimentally. This is primarily due to a much larger width of the σ* resonances in the experimental spectrum leading to lower maximum intensities. The differences between the spectrum for IN24DNBA and IN35DNBA demonstrates the sensitivity of the NEXAFS technique compared with other core level spectroscopies. When these samples were measured using XPS, the resultant spectra are identical, and only the slight differences in the unoccupied molecular orbitals allows us to distinguish between these samples using NEXAFS, as we know the core level binding energy is essentially the same. This is an important result, indicating just how sensitive NEXAFS is to the chemical environment, and highlighting the importance of the entire NEXAFS region of XAS spectra, not only the most prominent π* resonances. Table 3.1 summarizes the photon energies of the N 1s → π* transitions. We clearly observe that the energies for these transitions are almost identical between IN24DNBA and IN35DNBA – in agreement with XPS. It is the additional information and the higher energy σ* resonances in the spectra which allow us to distinguish between the two similar complexes using NEXAFS. For the majority of the σ* resonances in both IN24DNBA and IN35DNBA, there are corresponding peaks in both the experimental and modelled spectrum, suggesting the differences between the complexes is real. The fit for IN35DNBA follows the same principles as IN24DNBA and is shown in Figure 3.3(b). The peak intensity for the amide component forms a smaller proportion of the peak areas than for IN24DNBA. This could be due to the slightly different molecular orientations in this complex and the differing positions of the nitro groups. Additionally, the peak fitted just above the nitro group 1s → 2π* fits with the slight additional shoulder peak observed theoretically in the DFT calculated spectra. The fit is consistent with previous work on the sample, including XRD and XPS studies.

### 3.3.3 Isonicotinamide phthalic acid (INPA)

The INPA complex forms a 2:1 structure, containing one hydrogen bond interaction and one proton transfer interaction between the phthalic acid and isonicotinamide components.[31] In XPS measurements, this structure exhibits a shift of + 1.7 eV in the N 1s binding energy related to one of the two nitrogen acceptors, confirming this analysis.[6] The experimental NEXAFS spectrum shown in Figure 3.4(b) appears to confirm this structure, with obvious increase in photon energy of the protonated nitrogen acceptor N 1s → 2π* peak, leading to the distinctive double peak. In contrast to the other two complexes, the DFT calculated NEXAFS spectrum (Figure 3.4(a)) shows some differences in the relative intensities of the peaks, and an additional feature at 403.3 eV which is absent in the experimental spectrum. Using DFT modelling we were able to assign the π* resonances to transitions from the core levels of each of the nitrogen atoms in the structure. Figure 3.4(b) shows our assignment of the N 1s → π* transitions based on knowledge of the XP spectra and relative intensities of the peaks observed. We have already seen that the interaction cross section for the amide groups is significantly lower than for the pyridine ring nitrogen. Therefore, both the amide and unprotonated pyridine ring nitrogen energies

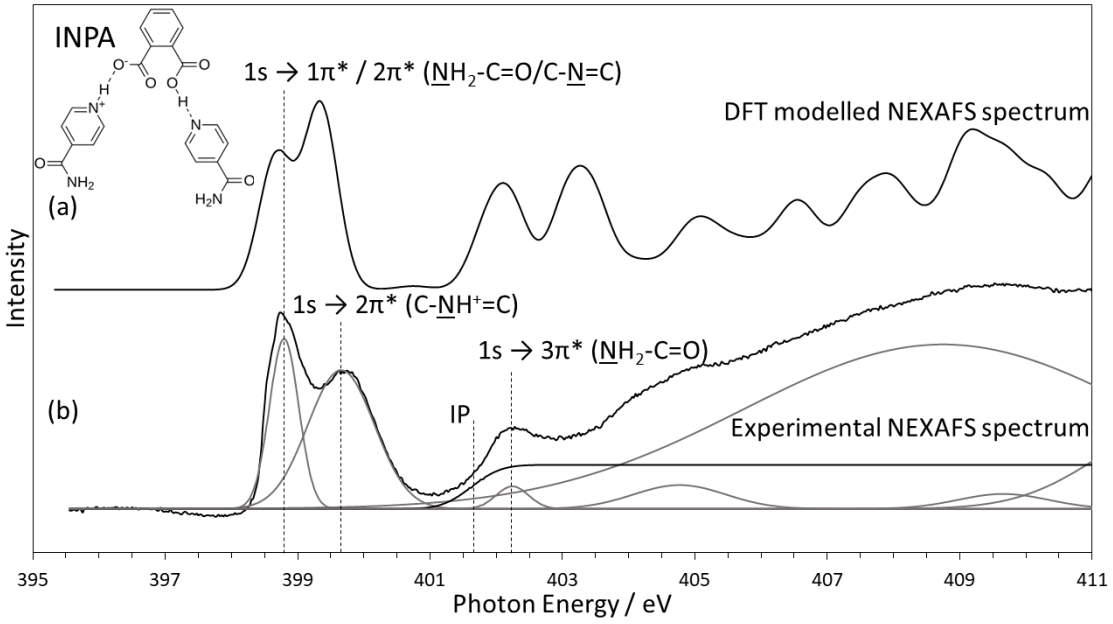



**Figure 3.4.** Modelled and experimental N K-edge NEXAFS spectrum of INPA showing excellent agreement of π* resonance peak positions. However, the DFT calculated spectrum has altered peak intensities for the peaks at 398.8 eV and 399.7 eV, suggesting a difference in structure.

overlap, leading to the additional intensity in the lower energy peak at 398.8 eV. This photon energy is in excellent agreement with the other two complexes, as shown in Table 3.1. Interestingly, the higher energy protonated nitrogen N 1s → 2π* peak at 399.7 eV exhibits broadening compared to the lower peak. There is no reason for this based on the instrumentation used, so this feature may suggest the presence of disorder in the H-atom position, potentially in agreement with previous work.[31] Our

**Table 3.1.** Best fit photon energies (eV) for the N 1s → π* transitions. DFT energies after uniform shift applied taken from the calculated excitation energies. *Additional changes to the molecular structure to reproduce peak shifts and intensities from the experimental spectrum. **Molecular Cluster calculation to account for longer range interactions.

| | | N−C=O | C−N=C | C−NH⁺=C | −NO₂ |
|---|---|---|---|---|---|
| IN24DNBA | Energy / eV<br>DFT energy / eV | 398.9<br>399.2 | n/a | 399.7<br>399.9 | 403.5<br>403.5 |
| IN35DNBA | Energy / eV<br>DFT energy / eV | 398.8<br>399.2 | n/a | 399.7<br>399.9 | 403.5<br>403.5 |
| INPA | Energy / eV<br>DFT energy / eV<br>DFT energy* / eV<br>DFT Energy** /eV | 398.8<br>399.5<br>398.9<br>398.8 | 398.8<br>398.8<br>398.8<br>398.7 | 399.7<br>399.6<br>399.6<br>399.6 | n/a |

analysis using the DFT calculation indicates that the energy of the amide nitrogen shifts to contribute to the higher of the two peaks, explaining the difference in relative intensities observed. These $NH_2$ groups form weak hydrogen bonds to oxygen at neighbouring amide groups, and carboxylic acid groups in phthalic acid. Additionally, unlike the other two complexes, the higher energy σ* peaks do not align well with the experimental data, with several peaks aligning with troughs. It is likely that longer range interactions between multiple components influence the energies of the amide N 1s → π* and σ* peaks which is not reproduced in calculations involving only an isolated complex.

## 3.4 Discussion

### 3.4.1 Isonicotinamide phthalic acid peak intensities

To assess the sensitivity of NEXAFS to the H-atom location we have carried out a calculation scanning across a range of N--H distances at the amide nitrogen to

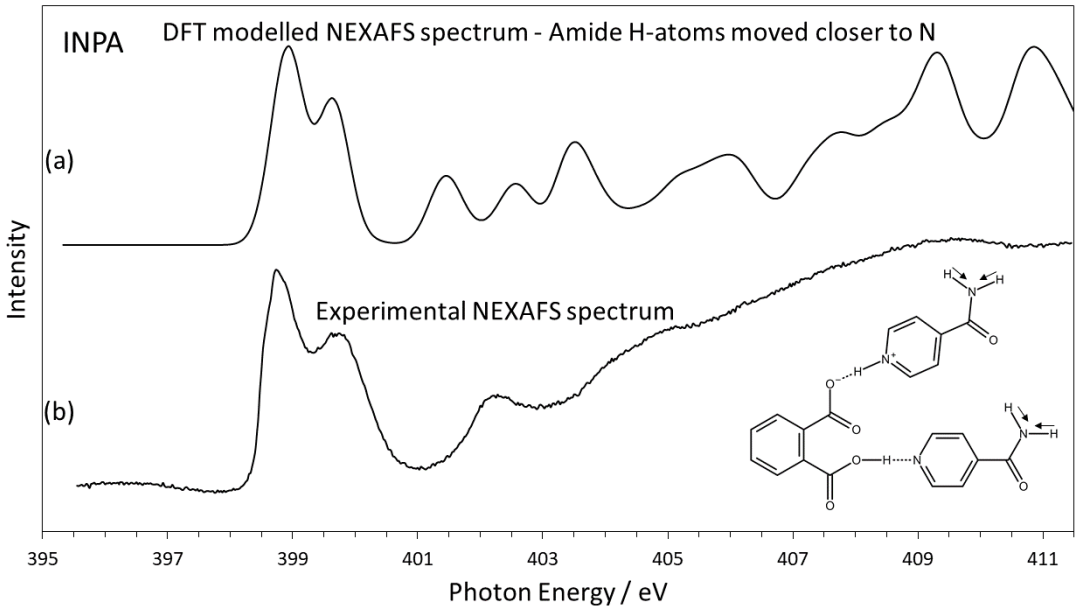



**Figure 3.5.** Modelled and experimental N K-edge NEXAFS spectrum of INPA showing excellent agreement of π* resonance peak positions and peak intensities. N--H distances at the amide groups have been reduced by 0.15 Å compared to the original database geometry to correctly match the experimental spectrum.

investigate the effect on the NEXAFS spectrum of INPA. By moving these H-atoms closer to the amide nitrogen by 0.15 Å, while leaving all other atomic positions as reported in the CSD untouched, we find the energy of the 1s → π* resonance decreases to that observed experimentally (Figure 3.5 and Table 3.1). Figure 3.5 shows this modelled spectrum compared to the experimental spectrum, and the relative intensities of the double peak appear to be corrected compared to the CSD geometry (Figure 3.4). This result is a demonstration of the high sensitivity of NEXAFS to the H-atom location. It also underlines our idea of a refinement procedure combining NEXAFS with DFT to locate H-atoms more accurately within hydrogen bonds. The simplistic approach taken here, movement of the H-atom while leaving the remaining structure static, is not meant to imply that the best fit H-atom

location is a correct representation of the crystal structure. Indeed, the resulting N-H distance of 0.77 Å is too short to be realistic. But it does give insight into the sensitivity of the NEXAFS technique, with a 0.15 Å change in N-H distance leading to a chemical shift of -0.7 eV.

### 3.4.2 DFT of cluster structures – isonicotinamide phthalic acid

We have also calculated an INPA spectrum using a larger cluster of molecules, at the CSD geometry, to investigate the effects of longer-range interactions on the predicted spectrum. The details for IN24DNBA and IN35DNBA are shown in Figure B2 in the supporting information, as the longer-range interactions have much less of an effect in these complexes. For INPA, the importance of longer-range interactions on the energy of the amide nitrogen resonances has become clear, with Figure 3.6 showing the cluster model spectrum compared with the experimental data. The relative peak intensities are predicted correctly in this case, the only difference between this and the initial calculation being additional molecules in the cluster. To avoid any effect of the peripheral molecules in the structure, the theoretical spectrum

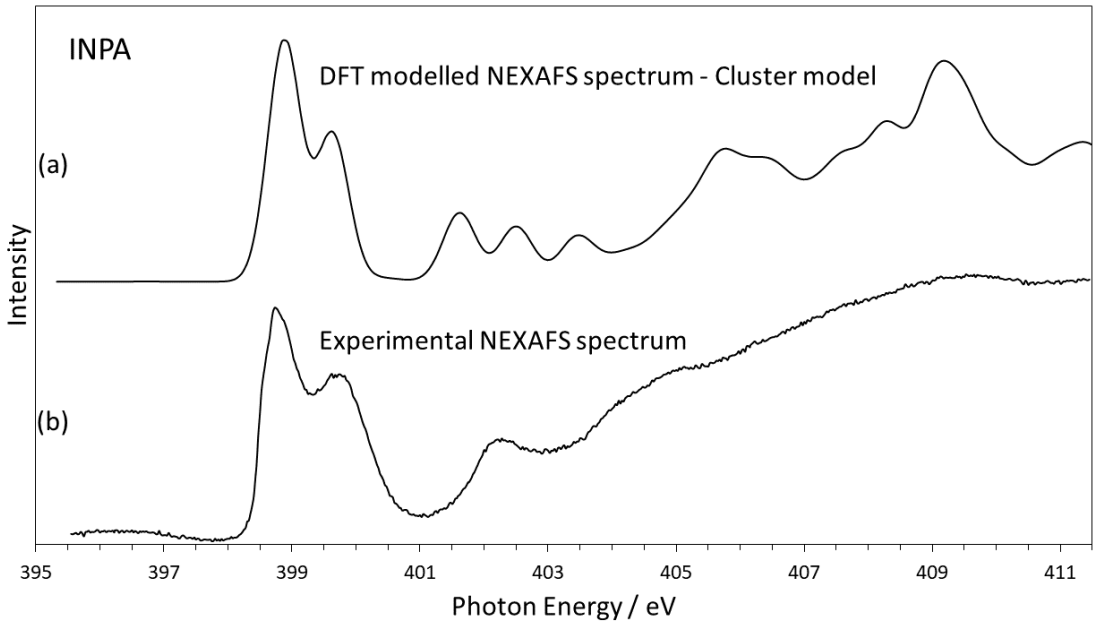



**Figure 3.6.** Modelled and experimental N K-edge NEXAFS spectrum of INPA showing excellent agreement of π* resonance peak positions and peak intensities. Calculation based on a cluster of molecules to simulate longer range interactions involved in the crystal structure.

was formed based on transitions from the central molecule only. This improvement in the energy of the amide peak shows how vital a correct molecular geometry is when calculating a theoretical spectrum. The σ* resonances are slightly improved over the initial calculation, with the higher energy peaks aligning with the experimental data and the unexplained peak at 403.3 eV being significantly reduced in intensity. Another effect not visible in the DFT calculation is the broadening of the protonated nitrogen peak. This further suggests the broadening is due to some disorder of the H-atom location, with the DFT calculation focusing on just one conformation. To simulate this effect, an average over a range of H-atom positions (at the protonated nitrogen site) has to be taken.

### 3.4.3 The effect of H-atom position on peak position.

In order to quantify how sensitive the spectra are to the location of the H-atom, an important test was to observe the theoretical effect of moving the H-atom on the calculated NEXAFS spectrum. An initial test of this was to calculate the spectra for the three samples for a range of N--H distances from the CSD structure to the optimized hydrogen structure in 0.1 Å steps, changing no other parameters of the geometry. Figure 3.7 shows the effect on the IN24DNBA nitrogen K-edge spectrum. As expected, both the nitro peak (403.5 eV) and the amide peaks (399.2 eV and 402.2 eV) remain in the same position (except for the optimized hydrogen structure where all H-atom positions are changed). We observe a continuous shift in the photon energy calculated for the nitrogen acceptor, showing that, at least theoretically, the core level excitation energies are sufficiently sensitive to the H-atom location to identify a difference of 0.1 Å. The inset plot in Figure 3.7 shows the relationship between photon energy calculated and the N--H bond distance. Using this roughly linear relationship, the bond length can be determined based on the experimental photon energy. Further to this, the CSD structure shows excellent agreement with the experimental spectrum (see Figure 3.2), including relative intensities and peak positions, indicating that the H-atom location is accurately refined in the XRD analysis in these relatively simple samples. This theoretical approach implies a direct sensitivity of the nitrogen acceptor π* resonance energy

to the position of the H-atom and opens up a new way to crystal structure refinement utilizing this rapid and simple technique.

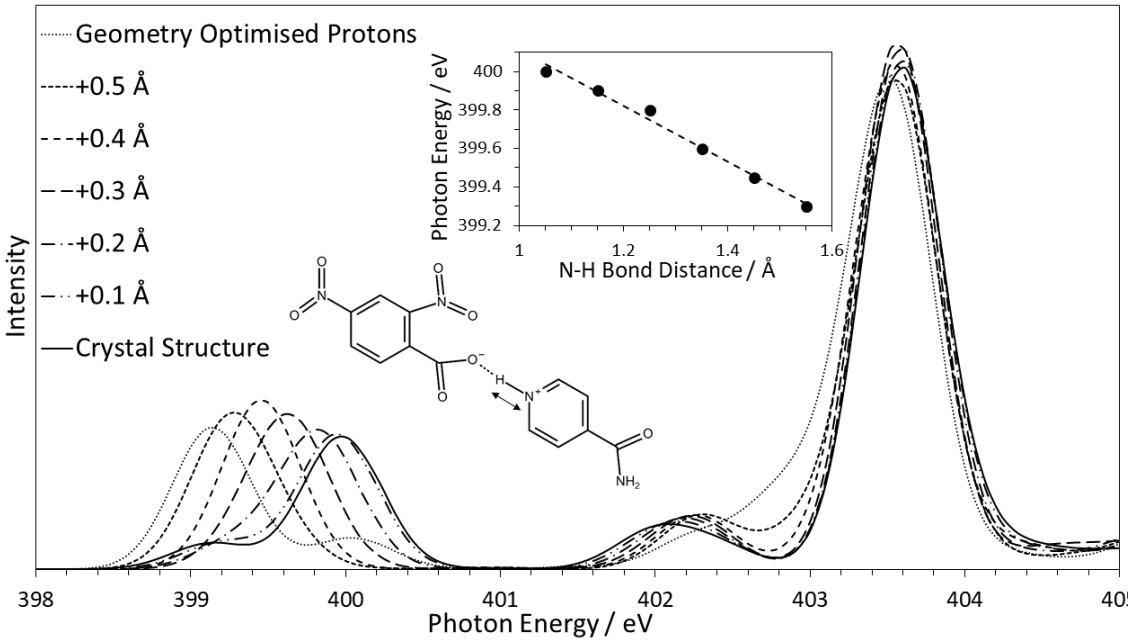



**Figure 3.7.** Series of DFT calculated NEXAFS spectra moving the H-atom across the hydrogen bond from the structure in the crystal database to the "optimized hydrogen" structure. The peak positions for the amide nitrogen is different for the "optimized hydrogen" structure calculation since all hydrogen atom positions were optimized.

### 3.4.4 Difference Spectra – complexes compared to individual components

To visualize the effect of hydrogen bonding and proton transfer interactions on the spectrum, difference spectra for the three complexes were calculated using experimental data for the individual components of the complexes. The difference spectra between the individual components and the complexes were calculated by subtracting the sum of the components in the correct stoichiometric ratio from the experimental spectrum of the complex. This is shown in Figure 3.8. Care was taken to normalize each of the spectra with respect to the high-energy baseline (which in effect is proportional to the N 1s XPS intensity) to ensure the relative signals from each of the components were correct. Here, we can clearly observe the N 1s → 1π* peaks shift from an energy of 398.5 eV in pure isonicotinamide to higher energy in

the crystal complex, and a broadening of the peak, as a result of the splitting of the peak into the amide (399.2 eV) and pyridinium (399.7 eV) nitrogens. This is particularly obvious in INPA because of the double peak due to the additional unprotonated pyridine nitrogen, leading to an even more pronounced broadening in the difference spectrum. The trough (401.8 eV) and peak (402.3 eV) corresponding to the amide N 1s → 4π* transition shows an equivalent increase in energy due to the forming of the complex. The N 1s → 2π* peak at 403.5 eV is due to the measured intensity being higher in the complexes than the pure dinitrobenzoic acid components, even with full normalization of the spectra. We also observe some additional features at higher photon energies due to changes in the unoccupied molecular orbitals when the complex is formed.

We have previously shown that the core level binding energy of electrons (determined using XPS) is directly affected by the distance between the H-atom and the proton acceptor.[6] However, the magnitude of the binding energy shifts also depends on the surround electrons and atoms and therefore there is no universal

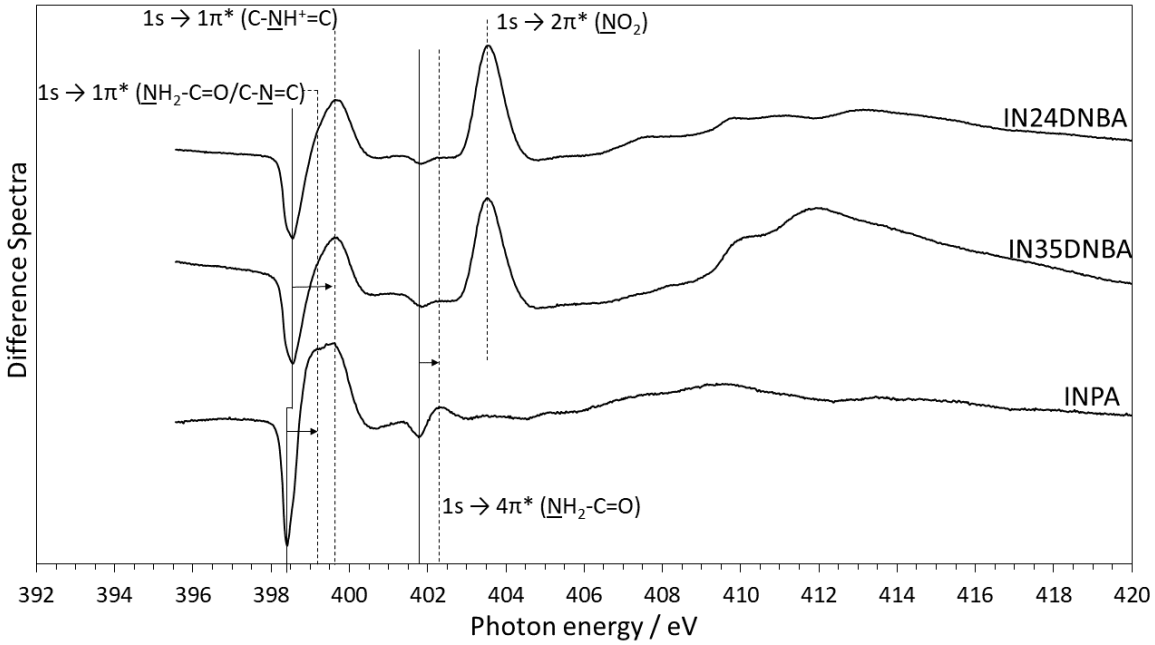



**Figure 3.8.** Difference spectra between experimental spectra and sum of the individual component's experimental spectra (See Figure B3). All three complexes exhibit a clear shift of both the amide and pyridine ring nitrogen species from the pure components to the crystal complex. Negative peaks (solid lines) correspond to the peak position in the pure compounds and arrows indicate the peak shift from pure compound to crystal complex.

dependency.[6] This means we have reached the limit of usefulness of XPS in characterization of organic hydrogen bonded crystals. The use of NEXAFS utilizes the increased sensitivity and structural dependence of the spectra on both the core N 1s level and the unoccupied π* orbitals to determine more precisely the location of the H-atom within the hydrogen bond, by combining the predictive power of DFT with NEXAFS.

## 3.5 Conclusions

The combination of experimental and theoretical nitrogen K-edge NEXAFS spectra has been used to accurately determine the structure of three two-component organic crystals by optimising the H-atom locations and comparing to H-atom location refinement procedures used in XRD. The observed chemical shifts and relative intensities of nitrogen acceptor atoms are used to combine the predictions of DFT calculations with NEXAFS experiment to obtain a fully consistent picture of the crystal structure and the important interactions. For all three complexes, the chemical shift and relative intensities calculated for the protonated nitrogen match exactly with experiment based on the CSD molecular geometries, indicating that the XRD proton location refinement in these structures is correct. In INPA, this peak is broadened, suggesting some degree of disorder, as mentioned in a previous XRD-based paper.[31] Ignoring longer range interactions between neighbouring molecular units leads to incorrect peak positioning of amide nitrogen N 1s → 1π* resonances in DFT calculations, indicating the importance of these interactions on the observed crystal structure. In addition, the combination of DFT with NEXAFS for refinement of H-atom positions has been established through a theoretical demonstration of the effect of H-atom position on the measured spectra, with changes in peak position above the detection limit for changes in H-atom position of < 0.1 Å. The advantage of NEXAFS over XRD is its fundamental direct sensitivity to the H-atom locations. This approach is also more practical than alternatives such as neutron diffraction, with measurements of the nitrogen K-edge typically taking on the order of 15 minutes, allowing rapid sample analysis. We intend to use this technique to refine the H-atom location in some more challenging situations, particularly where there

is uncertainty in the H-atom location within a hydrogen bonded system, such as quasi-centered short strong hydrogen bonds.

**Supporting Information Available:** Chemical structure and hydrogen bonding information, Experimental NEXAFS normalization procedure. Additional fitted spectra and fitting parameters, Cluster model for IN24DNBA and IN35DNBA and spectra of individual components. DFT calculation input files and geometry files with structural data, including comparison of exchange correlation functional and basis sets.

**Acknowledgements**

We gratefully acknowledge funding from the EPSRC and Diamond Light Source to cover a PhD studentship (EPSRC Grant EP/R513258/1). We thank Diamond Light Source for access to beamline B07-B (SI29334) during the early stages of comissioning that contributed to the results presented here. SLMS thanks the Future Continuous Manufacturing and Advanced Crystallisation (CMAC) Hub for financial support (EPSRC Grant EP/P006965/1). Our thanks to Anna Kroner and Nathan Hennessy for their assistance with the beamtime. The Density Functional Theory calculations were undertaken using ARC4, part of the High Performance Computing facilities at the University of Leeds, UK. All data supporting this study are provided either in the results section of this paper or in the ESI accompanying it.

## 3.6 References

1 Childs, S. L.; Stahly, G. P.; Park, A. The Salt-Cocrystal Continuum: The Influence of Crystal Structure on Ionization State, *Mol. Pharmaceutics* **2007**, 4, 323–338.

2 Wouters, J; Quéré, L. *Pharmaceutical Salts and Co-crystals*, The Royal Society of Chemistry, 2012.

3 Stevens, J. S.; Byard, S. J.; Schroeder S. L. M. Salt or Co-crystal? Determination of Protonation State by X-ray Photoelectron Spectroscopy (XPS), *J. Pharm. Sci.* **2010**, 99, 4453–4457.

4 Aakeröy, C. B.; Fasulo, M. E.; Desper, J. Cocrystal or Salt: Does it Really Matter?, *Mol. Pharmaceutics* **2007**, 4, 317–322.

5 Wang, T.; Stevens, J. S.; Vetter, T.; Whitehead, G. F. S.; Vitorica-Yrezabal, I. J.; Hao, H. X.; Cruz-Cabeza, A. J. Salts, Cocrystals, and Ionic Cocrystals of a "Simple" Tautomeric Compound, *Cryst. Growth Des.* **2018**, 18, 6973–6983.

6 Edwards, P. T.; Saunders, L. K.; Pallipurath, A. R.; Britton, A. J.; Willneff, E. A.; Shotton, E. J.; Schroeder, S. L. M. Proton Transfer on the Edge of the Salt/Cocrystal Continuum: X-Ray Photoelectron Spectroscopy of Three Isonicotinamide Salts, *Cryst. Growth Des.* **2021**, 21, 6334-6340.

7 Stevens, J. S.; Schroeder, S. L. M. Quantitative Analysis of Saccharides by X-ray Photoelectron Spectroscopy, *Surf. Interface Anal.* **2009**, 41, 453–462.

8 Stevens, J. S.; Coultas, S.; Jaye, C.; Fischer, D. A.; Schroeder, S. L. M. Core Level Spectroscopies Locate Hydrogen in the Proton Transfer Pathway-Identifying Quasi-Symmetrical Hydrogen Bonds in the Solid State, *Phys. Chem. Chem. Phys.* **2020**, 22, 4916–4923.

9 Stevens, J. S.; Byard, S. J.; Seaton, C. C.; Sadiq, G.; Davey, R. J.; Schroeder, S. L. M. Proton Transfer and Hydrogen Bonding in the Organic Solid State: a Combined XRD/XPS/ssNMR Study of 17 Organic Acid-Base Complexes, *Phys. Chem. Chem. Phys.* **2014**, 16, 1150–1160.

10 Stevens, J. S.; Byard, S. J; Schroeder, S. L. M. Characterization of Proton Transfer in Co-Crystals by X-ray Photoelectron Spectroscopy (XPS), *Cryst. Growth Des.* **2010**, 10, 1435–1442.

11 Stevens, J. S.; Byard, S. J.; Muryn, C. A.; Schroeder, S. L. M. Identification of Protonation State by XPS, Solid-State NMR, and DFT: Characterization of the Nature of a new Theophylline Complex by Experimental and Computational Methods, *J. Phys. Chem. B* **2010**, 114, 13961–13969.

12 Gainar, A.; Stevens, J. S.; Jaye, C.; Fischer, D. A.; Schroeder, S. L. M. NEXAFS Sensitivity to Bond Lengths in Complex Molecular Materials: A Study of Crystalline Saccharides, *J. Phys. Chem. B* **2015**, 119, 14373–14381.

13 Stevens, J. S.; Gainar, A.; Jaye, C.; Fischer, D. A.; Schroeder, S. L. M. NEXAFS and XPS of p-Aminobenzoic Acid Polymorphs: The Influence of Local Environment, *J. Phys.: Conf. Ser.* **2016**, 712, 012133.

14 Stevens, J. S.; Newton, L. K.; Jaye, C. ; Muryn, C. A.; Fischer, D. A.; Schroeder, S. L. M. Proton Transfer, Hydrogen Bonding, and Disorder: Nitrogen Near-Edge X-ray Absorption Fine Structure and X-ray Photoelectron Spectroscopy of Bipyridine-Acid Salts and Co-crystals, *Cryst. Growth Des* **2015**, 15, 1776–1783.

15 Stevens, J. S.; Seabourne, C. R.; Jaye, C.; Fischer, D. A.; Scott, A. J.; Schroeder, S. L. M. Incisive Probing of Intermolecular Interactions in Molecular Crystals: Core Level Spectroscopy Combined with Density Functional Theory, *J. Phys. Chem. B*, **2014**, 118, 12121–12129.

16 Stevens, J. S.; Byard, S. J.; Seaton, C. C.; Sadiq, G.; Davey, R. J.; Schroeder, S. L. M. Crystallography Aided by Atomic Core-Level Binding Energies: Proton Transfer versus Hydrogen Bonding in Organic Crystal Structures, *Angew. Chem. Int. Ed.* **2011**, 50, 9916–9918.

17 Stohr, J. *NEXAFS spectroscopy*, Springer-Verlag, Berlin, 1992.

18 Bianconi, A.; Dell'Ariccia, M.; Gargano, A.; Natoli, C. R.; Bianconi, A. *Bond Length Determination Using XANES;* Incoccia, L., Stipcich, S., Eds; Springer Berlin Heidelberg, Berlin, Heidelberg, 1983, pp. 57–61.

19 Stevens, J. S.; Gainar, A.; Suljoti, E.; Xiao, J.; Golnak, R.; Aziz, E. F.; Schroeder, S. L. M. NEXAFS Chemical State and Bond Lengths of p-Aminobenzoic Acid in Solution and Solid State, J. Phys.: Conf. Ser. **2016,** 712, 012136.

20 Stohr, J.; Sette, F.; Johnson, A. L. Near-Edge X-Ray-Absorption Fine-Structure Studies of Chemisorbed Hydrocarbons: Bond Lengths with a Ruler, Phys. Rev. Lett. **1984**, 53, 1684-1687.

21 Runge E.; Gross, E. K. U. *Density-Functional Theory for Time-Dependent Systems*, Phys. Rev. Lett. **1984**, 52, 997-1000.

22 Peach, M. J. G.; Benfield, P.; Helgaker, T.; Tozer, D. J. Excitation Energies in Density Functional Theory: An Evaluation and a Diagnostic Test, *J. Chem. Phys.* **2008**, 128, 44118.

23 Besley, N. A.; Gilbert, A. T. B.; Gill, P. M. W. Self-Consistent-Field Calculations of Core Excited States, *J. Chem. Phys.* **2009**, 130, 124308.

24 Becke, A. D. Density-Functional Thermochemistry 3. The Role of Exact Exchange, *J. Chem. Phys.* **1993**, 98, 5648–5652.

25 Peach, M. J. G.; Helgaker, T.; Sałek, P.; Keal, T. W.; Lutnæs, O. B.; Tozer, D. J.; Handy, N. C. Assessment of a Coulomb-Attenuated Exchange-Correlation Energy Functional, *Phys. Chem. Chem. Phys.* **2006**, 8, 558–562.

26 Endo, K.; Koizumi, S.; Otsuka, T.; Suhara, M.; Morohasi, T.; Kurmaev, E. Z.; Chong, D. P. Analysis of XPS and XES of Diamond and Graphite by DFT Calculations using Model Molecules, *J. Comput. Chem.* **2001**, 22, 102–108.

27 Parkin, A.; Harte, S. M.; Goeta, A. E.; Wilson, C. C. Imaging Proton Migration from X-rays and Neutrons, *New J. Chem.* **2004**, 28, 718–721.

28 Paula, M. B. P.; Koetzle, T. F.; Schultz, A. J. Single Crystal Neutron Diffraction for the Inorganic Chemist – a Practical Guide, *Comments Inorg. Chem.* ***2007,*** 28, 3–38.

29 Warren, B. E. *X-ray Diffraction*, Courier Corporation, 1990.

30 Sheldrick, G. M. Crystal Structure Refinement with SHELXL, *Acta Crystallogr., Sect. C: Struct. Chem.* **2015**, 71, 3-8.

31 Saunders, L. K.; Nowell, H.; Hatcher, L. E.; Shepherd, H. J.; Teat, S. J.; Allan, D. R.; Raithby, P. R.; Wilson, C. C. Exploring Short Strong Hydrogen Bonds Engineered in Organic Acid Molecular Crystals for Temperature Dependent Proton Migration Behaviour using Single Crystal Synchrotron X-ray Diffraction (SCSXRD), *CrystEngComm.* **2019**, 21, 5249–5260.

32 Neese, F. The ORCA Program System, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, 2, 73–78.

33 Neese, F.; Wennmohs, F.; Becker, U.; Riplinger, C. The ORCA Quantum Chemistry Program Package, *J. Chem. Phys.* **2020**, 152, 224108.

34 Schäfer, A.; Huber, C; Ahlrichs, R. Fully Optimized Contracted Gaussian Basis Sets of Triple Zeta Valence Quality for Atoms Li to Kr, *J. Chem. Phys.* **1994**, 100, 5829–5835.

35 Neese, F.; Wennmohs, F.; Hansen, A.; Becker, U. Efficient, Approximate and Parallel Hartree-Fock and Hybrid DFT Calculations. A "Chain-of-Spheres" Algorithm for the Hartree-Fock Exchange, *Chem. Phys.* **2009**, 356, 98–109.

36 Groom, C. R.; Bruno, I. J.; Lightfoot, M. P.; Ward, S. C. The Cambridge Structural Database, *Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater.* **2016**, 72, 171–179.

37 Hirata S.; Head-Gordon, M. Time-Dependent Density Functional Theory within the Tamm–Dancoff Approximation, *Chem. Phys. Lett.* **1999**, 314, 291–299.

38 Held, G.; Venturini, F.; Grinter, D. C.; Ferrer, P.; Arrigo, R.; Deacon, L.; Garzon, W. Q.; Roy, K.; Large, A.; Stephens, C.; et. al. Ambient-Pressure Endstation of the Versatile Soft X-ray (VerSoX) Beamline at Diamond Light Source, *J. Synchrotron Radiat.* **2020**, 27, 1153–1166.

39 Schwarzkopf, O.; Borchert, M.; Eggenstein, F.; Flechsig, U.; Kalus, C.; Lammert, H.; Menthel, U.; Pietsch, M.; Reichardt, G.; Rotter, P.; et. al. The BESSY constant length Rowland circle monochromator, *J. Electron Spectrosc. Relat. Phenom.* **1999**, 101, 997-1001.

40 Weiss, M. R.; Follath, R.; Sawhney, K. J. S.; Zeschke, T. Absolute energy calibration for plane grating monochromators, *Nucl. Instrum. Methods Phys. Res.* **2001**, 482-484.

41 Ravel, B.; Newville, M. Athena, Artemis, Hephaestus: Data Analysis for X-ray Absorption Spectroscopy using IFEFFIT, *J. Synchrotron Radiat.* **2005**, 12, 537–541.

42 Besley, N. A.; Peach, M. J. G.; Tozer, D. J. Time-Dependent Density Functional Theory Calculations of Near-Edge X-ray Absorption Fine Structure with Short-Range Corrected Functionals, *Phys. Chem. Chem. Phys.* **2009**, 11, 10350–10358.

43 Thomason, M. J.; Seabourne, C. R.; Sattelle, B. M.; Hembury, G. A.; Stevens, J. S.; Scott, A. J.; Aziz, E. F.; Schroeder, S. L. M. Self-Association of Organic Solutes in Solution: a NEXAFS Study of Aqueous Imidazole, *Faraday Discuss.* **2015**, 179, 269–289.

# 4 Hydrogen Bonds in Molecular Crystals: Strong Correlation between Proton Positions and 1s Core Level Binding Energies

Paul T. Edwards
*School of Chemical and Process Engineering,*
*University of Leeds, Leeds, LS2 9JT, UK and*
*Diamond Light Source, Harwell Science & Innovation Campus, Didcot, OX11 0DE, UK*

Elizabeth J. Shotton
*Diamond Light Source, Harwell Science & Innovation Campus, Didcot, OX11 0DE, UK*

Sven L. M. Schroeder*
*School of Chemical and Process Engineering, University of Leeds, Leeds, LS2 9JT, UK*
*Diamond Light Source, Harwell Science & Innovation Campus, Didcot, OX11 0DE, UK and*
*Future Continuous Manufacturing and Advanced Crystallisation Hub,*
*Research Complex at Harwell (RCaH), Rutherford Appleton Laboratory, Didcot, OX11 0FA, UK*
(Dated: December 2, 2023)

Hydrogen bonding and proton transfer, often referred to as Brønsted interactions, represent the strongest local cohesive forces in the organic solid state. Core level spectroscopies are sensitive probes for characterizing the electronic structure of these electrostatic bonds. We present a simple model for relating experimental 1s core level binding energy shifts at nitrogen proton acceptors with the proton distance. The nitrogen atom is modelled as a central positive point charge surrounded by an inner and an outer shell containing the elementary charges of core and valence electrons, respectively. A positive elementary point charge, counterbalanced by a negative charge at the location of the proton donor, is used to calculate the dependence of the potential energy at the location of the core electrons on the distance between nitrogen and proton. The results agree well with experimental and theoretical core level binding energy shifts predicted by density functional theory. Our analysis underlines that the whole spectrum of Brønsted interactions in the organic solid state is a continuum of geometry-dependent proton-acceptor Coulomb interactions, providing a universal principle for classifying these interactions.

## 4.1 INTRODUCTION

Brønsted acid-base interactions are the strongest forces that contribute to the cohesive energy in crystalline phases of small organic molecules [1, 2]. They occur when a functional group carrying a hydrogen atom (a proton donor) is in proximity to an atom with high local electron density (a proton acceptor) that attracts the proton from the donor group [2]. In a classical hydrogen bond, the proton is somewhat displaced towards the acceptor group, but remains in the potential well of the donor atom because of a potential energy barrier to the acceptor site. When the proton encounters no significant potential energy barrier between the donor and the acceptor then it can migrate to the acceptor and proton transfer takes place [2]. Proton transfer creates an ionic bond and the resulting phase is conventionally called a salt, referencing salt formation in the classic Brønsted acid/base interaction [1, 3]. The nature of the intermolecular Brønsted interactions in organic phases with more than one component is commonly used to classify such phases [4–12]: it distinguishes the ionic salt systems from cocrystals, which are characterised by classical hydrogen bonds. This dichotomic system does not capture systems containing short hydrogen bonds (SHBs), in which the distance between acceptor and donor is shorter than in a conventional hydrogen bond. The resulting overlap of the atomic potentials reduces the potneital barrier so much that protons dynamically populate both donor and acceptor sites [3]. Such interactions create hydrogen bonds with a fractional ionic charge that eludes the salt/cocrystal classification [6, 13, 14].

Brønsted interactions are commonly thought of as dominated by Coulomb electrostatics [15, 16], although simple models to describe them as such have had limited success [17–19]. However, X-ray photoelectron spectroscopy (XPS) studies have consistently indicated that the 1s core level binding energy of nitrogen proton acceptors in N-H-O hydrogen bonds is approximately 2 eV lower than in the protonated state [6, 20, 21]. Here we will

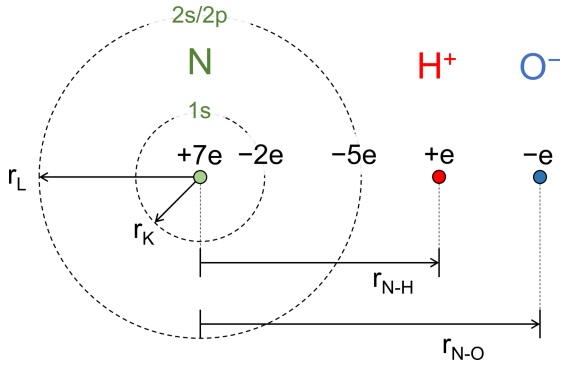



FIG. 4.1. Schematic representation of the primary interactions accounted for by the simple electrostatic model. These consist of interactions between the core level electron and i) the atomic nucleus ii) the valence electrons iii) the hydrogen atom and iv) the donor oxygen.

show that this binding energy shift, and variations in the N 1s binding energies in hydrogen-bonded systems in general, can be rationalized almost quantitatively by simple point charge electrostatics based on the distance of the proton to the acceptor.

### 4.2 ELECTROSTATIC MODEL

To estimate the 1s core level binding energy shift of a nitrogen acceptor hydrogen-bonded to an oxygen donor we solve Poisson's equation for the potential energy in a model (figure 4.1) comprising:

(i) the nitrogen nucleus as a central point charge of seven positive elementary charges,

(ii) the nitrogen 1s core electrons as a shell of two negative elementary charges located at a radial distance $r_K$ of 0.08 Å to the nucleus, corresponding to the maximum in the 1s orbital radial distribution function in a nitrogen atom [22],

(iii) an outer shell of five negative elementary charges modelling the remaining nitrogen 2s and 2p electrons, located at a radial distance $r_L$ corresponding to the nitrogen atomic radius (0.71 Å) [23].

(vi) a proton ($H^+$) modelled as a positive point charge at a given N-H distance, $r_{N-H}$, to the nitrogen nucleus

(v) a single negative charge maintaining overall charge neutrality, located at the distance $r_{N-O}$ to the oxygen donor (O) nucleus.

The total electric potential energy at $r_K$ in the system of electrostatic charges in figure 1 then becomes the sum of the relevant Coulomb interactions,

$$V = \sum_i \frac{q_i}{4\pi\varepsilon_r(r_K - r_i)} \tag{1}$$

where $q_i$ is the charge of the $i^{th}$ component of the potential, $\varepsilon_r$ is the relative permittivity of the medium, $r_K - r_i$ is the distance from the $i^{th}$ component of the potential to the K shell. By modelling the electron shells as insulating spherical shells, the electrostatic potential inside the valence shell is a constant average that depends on the radius of the shell. The contribution to the overall potential is

$$V = \frac{q_{valence}}{4\pi\varepsilon_r r_L} \tag{2}$$

The four contributions to the model N 1s core level energies are the Coulomb interactions with the acceptor nucleus, the valence electrons, the external proton and the donor. We model the region for $r < r_L$ as a homogeneous electron gas with $\varepsilon_r \approx 3\varepsilon_0$ [24]. For $r > r_L$ (outside of the nitrogen atom) we have a vacuum with $\varepsilon_r = \varepsilon_0$, the permittivity of free space [24]. In effect, interactions between the N 1s core electron and neighboring atoms (except the hydrogen and oxygen donor atoms) in the real organic solid state are thereby modelled as a background electrostatic potential that is independent of the chemical composition of the solid. This is likely to be a good approximation because in the real organic solid state electron density variations by neighbouring low-Z atoms around the donor-acceptor bonds are relatively small, exerting only a secondary contribution to the potential energy of the core level N 1s electron. By the same reasoning, core hole relaxation and other final state contributions to the core level binding energy are neglected as well. The neglected interactions affecting the electrostatic potential would include the self-interaction of the core level electrons, intraatomic interactions with other electrons, interactions with electron density in the surrounding molecular structure, and any external fields.

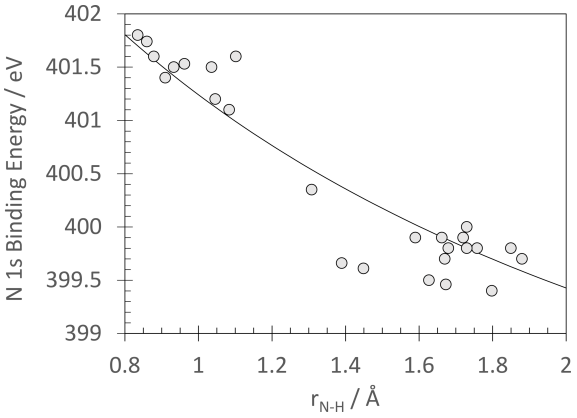



FIG. 4.2. Core level binding energies predicted by electrostatics modelling (full line) for a range of N-H distances, comprising salts (N-H distance $\sim$ 1.0 Å) and cocrystals (N-H distance $\sim$ 1.8 Å). The datapoints are experimental N 1s core level binding energies for salts and cocrystals ([6] and references therein).

### 4.3 RESULTS AND DISCUSSION

In figure 4.2 we present the calculated potential energy variation experienced by the model 1s electron charges as a function of the distance $r_{N-H}$, with the $r_{N-O}$ fixed at 2.8 Å. Best agreement with available experimental core level binding energies was obtained by applying a constant offset of 9.1 eV, which was added to the potential energy. This added constant accounts for the net influence of the model simplifcations, which include the omitted interactions in the organic solid state (*vide supra*). We note that its magnitude is consistent with the assumption that the effect of these interactions on the core level binding energy shift is negligible, as 9.1 eV is less than 3% of the absolute potential energy of the N 1s electrons.

The modelled binding energy continuum in Figure 2 features an inverse dependence on the proton distance $r_{N-H}$, as expected for a radial electrostatic problem. The key feature is the magnitude of the change in electrostatic potential over the range of proton distances from salts ($\sim$ 1.0 Å) to cocrystals ($\sim$ 1.8 Å). An upward shift of +2 eV is observed as the proton is positioned closer to the acceptor, establishing a continuum of binding energy shifts between salts and cocrystals. This shift agrees well with the experimentally observed values ([6] and references therein) between cocrystal and salt structures. A principal limit to the applicability of this simple model arises in the region where covalent bonds form, $r_{N-H} <\sim 0.8$ Å, which are not just driven by electrostatic interactions. Being able to identify the cause of core level binding energy shifts in terms of simple Coulomb interactions is not only useful for understanding and modelling hydrogen bonds in crystalline structures, but also validation of the previously noted correlation [6, 13] between the acceptor N 1s core level binding energy shift and the acceptor-proton distance in a hydrogen bond.

To validate the results obtained by electrostatics modelling, density functional theory (DFT) calculations were carried out using the ORCA package, version 5.0.2 [25, 26]. Calculations were performed for a range of salt and cocrystal structures to examine how the functional form and/or position of the potential energy curve depended on the type of dominant interaction. For various salts, good agreement between the core level binding energies variations predicted by DFT and the electrostatics model calculation was found. The isonicotinamide 3,5-dinitrobenzoic acid salt system [6] was chosen as a representative system for examining the influence of the acceptor-proton distance, $r_{N-H}$. Single point energy calculations were carried out using the B3LYP [27] exchange correlation functional and def2-TZVP basis set [28] providing a suitable level of accuracy for calculating ground state properties with realistic relative energy differences between the structures with changing proton position. An additional shift in energy was applied to the DFT calculated values, which compensates for the consistent underestimate of energies by 12 eV using the B3LYP exchange correlation functional. To calculate the N 1s binding energies in a way comparable to the electrostatic model, the atomic positions (i.e., the crystal structure) was not permitted to relax, meaning that the only change from one calculation to the next was the position of the proton. A potential energy curve was then constructed from the energies of the N 1s core levels, displaying the effect of moving the proton on the binding energy of the core level electron (figure 4.3).

It is instructive to compare the DFT and electrostatic modelling results. The DFT calculations account for all significant interactions in the solid state. Like the core level shifts predicted by electrostatics, DFT correctly models the decrease of the N 1s core level binding energy above

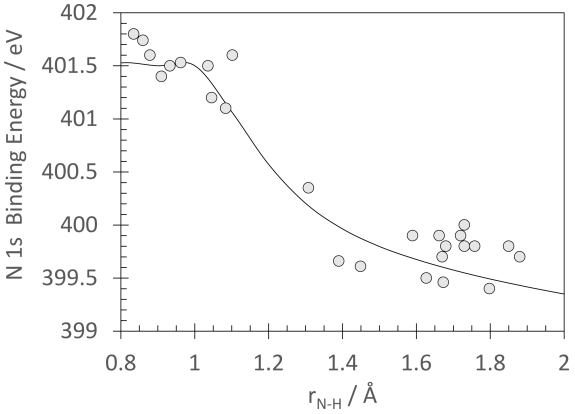



FIG. 4.3. DFT-calculated core level electron binding energy shifts (line) for a range of N-H distances compared to experimental results ([6] and references therein) for salts (N-H distance $\sim$ 1.0 Å) and cocrystals (N-H distance $\sim$ 1.8 Å).

$r_{N-H} =\sim$ 1.0 Å as the N-H distance is increased. The decrease is somewhat steeper than in the electrostatics model, suggesting the presence of interactions not captured by simple electrostatics modelling. However, apart from this slight difference in gradient, the overall shape and magnitude of the curve is similar, most importantly the magnitude of the predicted binding energy difference between salts and cocrystals, which is still $\sim +2$ eV. This underlines that the dominant interaction in hydrogen bonding is the electrostatic potential variation induced by the proton. Compared to the electostatic modelling there is also a flattening of the curve below an acceptor-proton distance of $r_{N-H} =\sim$ 1.0 Å that can be attributed to the formation of covalent bonds in which significant electron density is localized between acceptor and proton. This shields the core level electrons from further electrostatic effects.

However, outside of the proton transfer range both modelling approaches indicate that the N 1s core level binding energies follow a continuous universal curve, suggesting that all systems, including what are often considered distinct fractional charge systems between salts and co-crystals, are reasonably described as a single electrostatic continuum of hydrogen bonds. The intermediate systems, e.g. short strong hydrogen bonds (SSHBs) and charge-assisted hydrogen bonds, occur due specific local geometries in the solid state, which are associated with short donor-acceptor distances (short hydrogen bonds, SHBs). Being able to estimate the distance between acceptor and proton from the electrostatically induced core level binding energy shift at the acceptor site goes a long way to identifying such intermediate cases. Since the two potential wells are separated by only a shallow or no barrier, proton disorder can occur in SHBs, which can become evident in core level spectra through a broadening of the emission line [3], which aids their identification further.

To test the predictive strength of the electrostatics modelling we calculated the N 1s binding energies for 27 systems with known experimental N 1s core level binding energies [3, 6, 11, 20, 21, 29, 30] from crystallographic [31] values for $r_{N-H}$ and $r_{N-O}$. The results for all 27 systems are given in Table C1 in appendix C. Figure 4.4 presents a correlation plot summarising the predicted N 1s binding energies vis-a-vis the experimental values determined by XPS. Data above the line indicate an overestimate of the binding energy, while data below represent an underestimate. While the model overall has high predictive power, especially considering the approximations made, there is an overall trend towards an overestimate of binding energies in cocrystals and an underestimate of binding energies in salts. Across all systems, the standard deviation between predicted and experimental values is approximately 0.3 eV (see supporting information), with the maximum observed deviation being 0.74 eV. In all cases, the absolute deviations between experiment and model were much smaller than the 2 eV difference between ionic bond formation (salt) and hydrogen bonding (cocrystal). Indeed, these results indicate that a measurement of the N 1s core level binding energy can be used to identify short hydrogen bonds (SHBs) in the continuum between salts and cocrystals with high confidence. A case in point is the 3,5-pyridinedicarboxylic acid SHB system (blue diamond in the centre of figure 4.4) that would be correctly identified through its N 1s binding energy just above 400 eV [3], which is higher than the binding energies in the conventional hydrogen bonds. It should also be noted that the main outliers in the cocrystal region (Figure 4.4) are of molecular structures involving 4,4'-bipyridine as the proton acceptor [11], suggesting that for these systems unidentified additional inter- and/or intramolecular interactions may introduce a systematic error. The crystallographic origin of the unusually low N 1s binding energy for the isonicotinamide-phthalic acid sys-

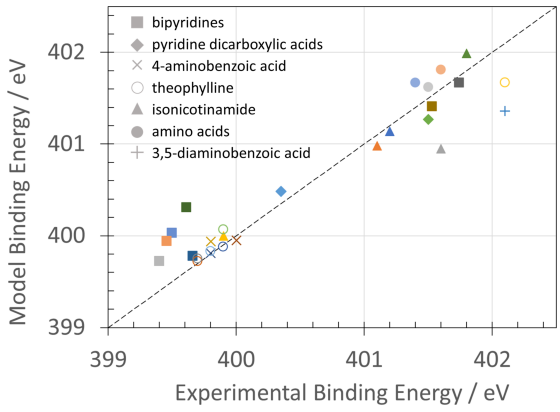



FIG. 4.4. Comparison of N 1s core level binding energies predicted by the electrostatics model with experimentally observed values. The model calculations were performed with crystallographically determined $r_{N-H}$ and $r_{N-O}$ distances (see text).

tem, the most extreme outlier among the salts, has recently been discussed [6]. Overall, the results indicate that simple electrostatics modelling provides a foundation for determining proton positions in a hydrogen bond from the core level binding energy of the acceptor.

### 4.4 CONCLUSIONS

Simple point charge electrostatics reveal how variations in the distance to the proton determine the 1s core level binding energy shifts of nitrogen Brønsted acceptor in the organic solid state. They reproduce experimentally observed binding energy shifts as well as predictions from DFT modelling. The nitrogen 1s binding energies of 27 systems were predicted from donor, acceptor and proton positions in published crystal structures, with a standard deviation between predicted and experimental nitrogen 1s bindng energies of 0.3 eV. The absolute error never exceeded 0.75 eV. These results indicate that proton-acceptor distances in Brønsted systems can be estimated from the acceptor core level binding energies. Our analysis also highlights that all types of Brønsted interactions, including proton transfer, classic hydrogen bonding and intermediate interactions associated with short hydrogen bonds, should be considered as one continuum of electrostatic interactions that depends on the local geometric arrangement of acceptor, donor and proton.

PTE gratefully acknowledges a PhD studentship funded jointly by Diamond and through the EPSRC Doctoral Training Partnership (DTP) with the University of Leeds (EPSRC Grant EP/R513258/1). SLMS thanks the Royal Academy of Engineering, Diamond Light Source Ltd and Infineum UK for support of the Bragg Centenary Chair, and for financial support from the Future Continuous Manufacturing and Advanced Crystallisation (CMAC) Hub (EPSRC Grant EP/P006965/1). The Density Functional Theory calculations were undertaken using ARC4, part of the High Performance Computing facilities at the University of Leeds, UK. All data supporting this study are provided either in the results section of this paper or in the electronic supporting information accompanying it.

* s.l.m.schroeder@leeds.ac.uk

[1] J. G. Sławomir, *Hydrogen Bonding—New Insights* (Springer Netherlands, 2006).
[2] T. Steiner, Angew. Chem. Int. Ed. **41**, 48 (2002).
[3] J. S. Stevens, S. Coultas, C. Jaye, D. A. Fischer, and S. L. M. Schroeder, Phys. Chem. Chem. Phys. **22**, 4916 (2020).
[4] S. Tothadi, T. R. Shaikh, S. Gupta, R. Dandela, C. P. Vinod, and A. K. Nangia, Cryst. Growth Des. **21**, 735 (2021).
[5] S. L. Childs, G. P. Stahly, and A. Park, Mol. Pharm. **4**, 323 (2007).
[6] P. T. Edwards, L. K. Saunders, A. R. Pallipurath, A. J. Britton, E. A. Willneff, E. J. Shotton, and S. L. M. Schroeder, Cryst. Growth Des. **21**, 6332 (2021).
[7] C. B. Aakeröy, M. E. Fasulo, and J. Desper, Mol. Pharm. **4**, 317 (2007).
[8] T. Wang, J. S. Stevens, T. Vetter, G. F. S. Whitehead, I. J. Vitorica-Yrezabal, H. X. Hao, and A. J. Cruz-Cabeza, Cryst. Growth Des. **18**, 6973 (2018).
[9] J. Wouters and L. Quéré, *Pharmaceutical Salts and Co-crystals* (The Royal Society of Chemistry, 2012).
[10] J. S. Stevens, S. J. Byard, and S. L. M. Schroeder, J. Pharm. Sci. **99**, 4453 (2010).
[11] J. S. Stevens, L. K. Newton, C. Jaye, C. A. Muryn, D. A. Fischer, and S. L. M. Schroeder, Cryst. Growth Des. **15**, 1776 (2015).
[12] P. Gilli, L. Pretto, V. Bertolasi, and G. Gilli, Acc. Chem. Res. **42**, 33 (2009).
[13] P. T. Edwards, L. K. Saunders, D. C. Grinter, P. Ferrer, G. Held, E. J. Shotton, and S. L. M. Schroeder, J. Phys. Chem. A **126**, 2889 (2022).
[14] L. K. Saunders, H. Nowell, L. E. Hatcher, H. J. Shepherd, S. J. Teat, D. R. Allan, P. R. Raithby,

and C. C. Wilson, CrystEngComm **21**, 5249 (2019).
[15] P. Gilli, V. Bertolasi, L. Pretto, V. Ferretti, and G. Gilli, J. Am. Chem. Soc. **126**, 3845 (2004).
[16] S. C. van der Lubbe and C. F. Guerra, Chem. Asian J. **14**, 2760 (2019).
[17] J. J. Dannenberg, L. Haskamp, and A. Masunov, J. Phys. Chem. A **103**, 7083 (1999).
[18] M. J. O'Meara, A. Leaver-Fay, M. D. Tyka, A. Stein, K. Houlihan, F. Dimaio, P. Bradley, T. Kortemme, D. Baker, J. Snoeyink, and B. Kuhlman, J. Chem. Theory Comput. **11**, 609 (2015).
[19] L. C. Allen, J. Am. Chem. Soc. **97**, 6921 (1975).
[20] J. S. Stevens, A. Gainar, C. Jaye, D. A. Fischer, and S. L. M. Schroeder, J. Phys. Conf. Series **712**, 012133 (2016).
[21] J. Stevens, S. Byard, C. Seaton, G. Sadiq, R. Davey, and S. L. M. Schroeder, Angew. Chem. Int. Ed. **123**, 10090 (2011).
[22] J. T. Waber and D. T. Cromer, J. Chem. Phys. **42**, 4116 (1965).
[23] J. Emsley, *The Elements*, 3rd ed. (Clarendon Press, 1998).
[24] F. Brosens, J. T. Devreese, and L. F. Lemmens, Phys. Rev. B **21**, 1363 (1980).
[25] F. Neese, Wiley Interdiscip. Rev. Comput. Mol. Sci. **2**, 73 (2012).
[26] F. Neese, F. Wennmohs, U. Becker, and C. Riplinger, J. Chem. Phys. **152**, 224108 (2020).
[27] A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).
[28] A. Schäfer, C. Huber, and R. Ahlrichs, J. Chem. Phys. **100**, 5829 (1994).
[29] J. S. Stevens, S. J. Byard, C. A. Muryn, and S. L. M. Schroeder, J. Phys. Chem. B **114**, 13961 (2010).
[30] J. S. Stevens, S. J. Byard, C. C. Seaton, G. Sadiq, R. J. Davey, and S. L. M. Schroeder, Phys. Chem. Chem. Phys. **16**, 1150 (2014).
[31] C. R. Groom, I. J. Bruno, M. P. Lightfoot, and S. C. Ward, Acta Cryst. B **72**, 171 (2016).

# 5 Crystallographic Refinement of Proton Positions and Site Occupancies by Combining Core Level Spectroscopies with Density Functional Theory

*Paul T. Edwards,[1, 2] Elizabeth J. Shotton,[2]* Sven L. M. Schroeder[1,2,3]**

[1]School of Chemical and Process Engineering, University of Leeds, Leeds, LS2 9JT, UK
[2]Diamond Light Source, Harwell Science & Innovation Campus, Didcot, OX11 0DE, UK
[3]Future Continuous Manufacturing and Advanced Crystallisation Hub, Research Complex at Harwell (RCaH), Rutherford Appleton Laboratory, Didcot, OX11 0FA, UK

* Sven L. M. Schroeder: Email: s.l.m.schroeder@leeds.ac.uk

* Elizabeth J. Shotton: Email: elizabeth.shotton@diamond.ac.uk

**Abstract**

Pinpointing precise proton positions within a hydrogen bond is of enormous practical importance for property prediction and development of new molecular materials. Here we show how combining core level spectroscopies with density functional theory (DFT) calculations generates a quantitative description of dynamic proton occupancies in the organic solid state. We analyse five different crystal structures, selected by a Cambridge Structural Database (CSD) search for systems with potential for short strong hydrogen bonds or proton disorder. The analysis strategy is to achieve consistency between results of two core level spectroscopies, X-ray photoelectron spectroscopy (XPS) and near-edge X-ray absorption fine-structure (NEXAFS), and DFT calculations of the orbital energies by refinement of proton positions within the CSD structures. For one system, no refinement was necessary as DFT, XPS and NEXAFS data were fully consistent with the CSD structure. The other four structures required refinement of proton positions to achieve the same level of consistency. Our study strongly underlines the complementarity of XPS and NEXAFS, which explicitly include quantitative information about proton disorder due to their ultrafast nature, in improving crystallographic structures where diffraction methods are ambiguous.

## 5.1 Introduction

Hydrogen bonding interactions are fundamental to the formation of many organic crystalline structures.[1] The location of the H-atom shared between the donor and acceptor atoms in these bonds has an important effect on the properties of the crystal; understanding this local bonding is even relevant for regulatory bodies identifying structures for patent protection of pharmaceutical compounds.[2–4] Hydrogen bonding in co-crystals and proton transfer in salts form the extremities of what is often called the salt-co-crystal continuum of Brønsted interactions. They describe the states where the hydrogen atom is either located close to the proton donor (co-crystal) or at the proton acceptor (salt).[1,4,5] In certain cases, the proton resides approximately equidistant between the donor and acceptor atoms; this can occur where the potential energy well holding the proton is sufficiently broad to allow a central position to be occupied. This happens for short donor acceptor distances, leading to the formation of weak covalent-like bonds between the proton and both the donor and acceptor.[6] This instance of a hydrogen bond is known as a Short Strong Hydrogen Bond (SSHB). Our previous studies have demonstrated how core level spectroscopy in the form of X-ray Photoelectron Spectroscopy (XPS) and near-edge X-ray absorption fine structure (NEXAFS) spectroscopy can be used to locate the position of the H-atom quite accurately through the 1s core level binding energy and 1s → π* transition energy of the proton acceptor.[3,7–17]

A commonly used method for determining precise H-atom positions is neutron diffraction, although preparation of a suitable crystal can be onerous and data acquisition is more laborious.[18,19] Locating H-atoms by X-ray crystallography, specifically X-ray diffraction (XRD), requires overcoming the fundamental insensitivity of the technique to H-atoms, as the cross sections for X-ray scattering scales with the electron density, and hydrogen is the element with the lowest local electron density.[20] The electron of a hydrogen atom also tends to be localized near the atom to which hydrogen is bonded.[21] This leads to a general insensitivity to the exact location of its nucleus (the proton) and a requirement for additional refinement methods for H-atom positions.[22]

In two recent studies we examined three isonicotinamide systems using XPS and NEXAFS, in which isonicotinamide was present as 1:1 salts with 2,4-dinitrobenzoic acid and 3,5-dinitrobenzoic acid, and a 2:1 structure with phthalic acid that contained two distinct interactions between the components: one Brønsted proton transfer interaction and one standard hydrogen bond.[7,11] For these systems, the expected chemical shift of the N 1s binding energy (BE) was observed, confirming that the core level BEs are consistent with the crystallographically determined interactions. Using NEXAFS, we have previously investigated these complexes further to confirm the XRD-based H-atom location procedure experimentally.[11,22] In effect, this formed the basis of a “proof of principle” experiment showing how NEXAFS and XPS could be used to complement structure refinement by XRD.[11]

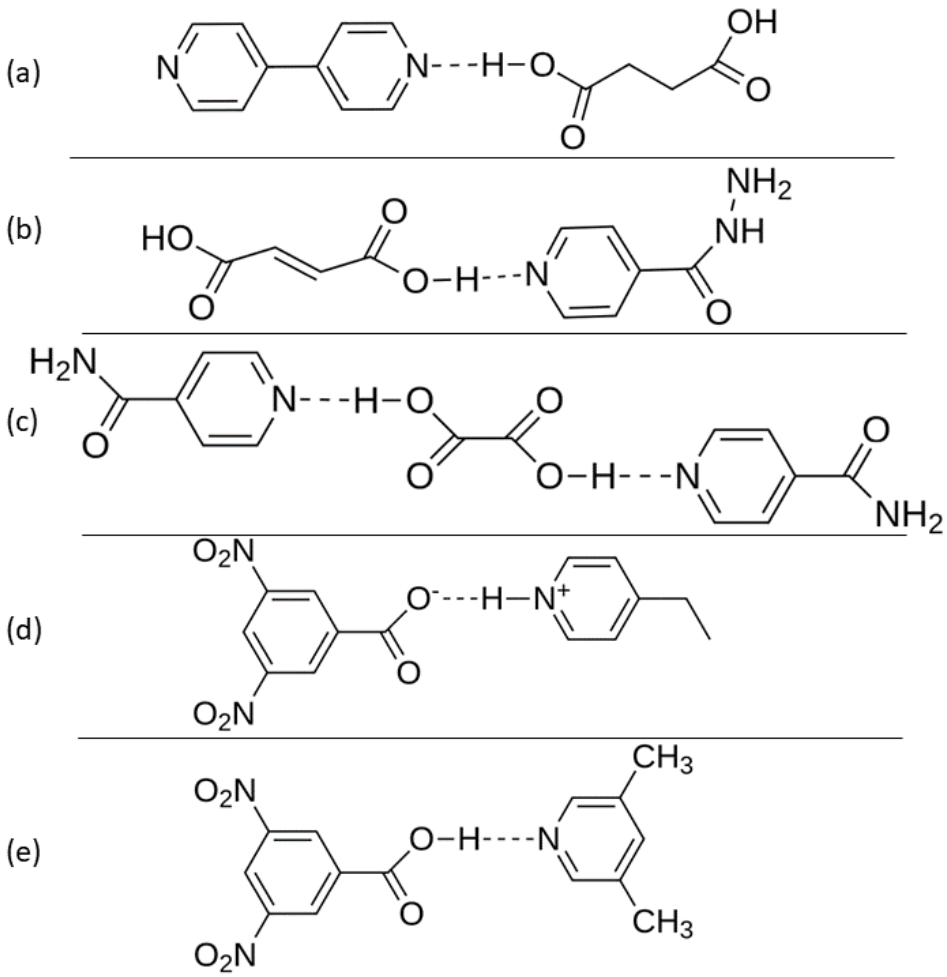



**Figure 5.1.** Molecular structures of the 5 samples investigated as found in the crystallography database. (a) 44BPYSA, (b) INHFA, (c) INOA, (d) 4EP35DNBA and (e) 35L35DNBA.

Having successfully demonstrated the potential of the method, we now develop this method of proton position refinement further by investigating five 2-component crystal structures with short N--H-O interactions, combining XPS, NEXAFS and density functional theory (DFT) to attempt to localize the hydrogen atoms and explain all features of the spectra in a consistent manner. It is this level of self-consistency between the methods that gives high confidence in the analysis. Figure 5.1 shows the chemical structure of the five samples investigated, while table 5.1 outlines the characteristics of these samples as found in the Cambridge Structural Database (CSD).[23] These samples were selected specifically to challenge the approach with a selection of samples with proton positions that could result in either salt, cocrystal or intermediate states, They were identified through a CSD search using appropriate parameters to find structures exhibiting characteristics of short strong hydrogen bonds with a short N…O distance, and with a $\Delta pK_a$ in the region between 0 and 3 where the formation of a salt or cocrystal is possible.[23,24] Each of these structures

**Table 5.1. Samples and their properties from the Cambridge Structural Database at room temperature.[23] Molecular structures are shown in figure 5.1.**

| Sample Name | Sample ID | Salt/Cocrystal | N--H Distance / Å | N..O Distance / Å | CSD Reference |
|---|---|---|---|---|---|
| 4,4'-Bipyridine Succinic Acid | 44BPYSA | Cocrystal | 1.33 | 2.62 | GOKBAL |
| Isonicotinic Acid Hydrazide Fumaric Acid | INHFA | Cocrystal | 1.33 | 2.54 | LATTAD |
| Isonicotinamide Oxalic Acid | INOA | Cocrystal | 1.39 | 2.56 | ULAWAF |
| 4-Ethylpyridine 3,5-Dinitrobenzoic Acid | 4EP35DNBA | Salt | 1.20 | 2.56 | BATYAZ01 |
| 3,5-Lutidine 3,5-Dinitrobenzoic Acid | 35L35DNBA | Cocrystal | 1.37 | 2.54 | PUHROZ09 |

contain N--H-O hydrogen bonds between a pyridine ring nitrogen acceptor and a carboxylic acid group oxygen donor. 4EP35DNBA and 35L35DNBA contain highly electronegative dinitro benzoic acid groups, similarly to previously investigated samples, while the other three structures contain one pyridine ring interacting with a simple aliphatic dicarboxylic acid. According to the database, only one of the five (4EP35DNBA) exhibits proton transfer behaviour while the other four samples are simply cocrystal. However, based on the reported bond distances, the relatively short N…O distance leaves open the opportunity for these structures to exhibit the characteristics of short strong hydrogen bonding, and therefore fulfil the requirements of the structural database search.

## 5.2 Materials and Methods

*Materials*

All five samples were prepared by evaporation from solution, as described in previous works.[25–29] The crystallizations were carried out in solvent using the correct stoichiometric ratios.[25–29] Methanol was used for 44BPYSA, INHFA and 4EP35DNBA. A 1:1 water ethanol mixture was used for INOA and Acetonitrile was used for 35L335DNBA. The individual components for the crystallizations (isonicotinamide (99.9%), 3,5-dinitrobenzoic acid (99%), oxalic acid (99%), 4,4'-bipyridine (99%), succinic acid (99%), 3,5 Lutidine (98%), 4-ethylpyridine (98%), isonicotinic acid hydrazide (99%), fumaric acid (99%) were obtained from Merck Life Sciences UK.

*Density Functional Theory Calculations*

DFT calculations were completed using the *ORCA* software package version 5.0.1, utilizing the University of Leeds high performance computing facilities.[30,31] For molecular geometry optimization calculations, the B3LYP[32] exchange correlation functional was used along with the def2-TZVP basis set[33], def2/J auxiliary basis set and the RIJCOSX approximation.[34] Starting structures were obtained from the Cambridge Structural Database (CSD).[23] For the modelling of NEXAFS spectra, time-dependent DFT (TDDFT) calculations were run on the various crystal and DFT

optimized structures. TDDFT is used to calculate excited states and can be used to model an X-ray absorption spectrum based on the transition probabilities between occupied and unoccupied molecular orbitals and the energies calculated.[35] By allowing only transitions from core level orbitals of specific atoms, we can calculate a theoretical NEXAFS spectrum accounting for all of the electronic transitions within a certain energy range.[35] The same B3LYP[32] exchange correlation functional and def2-TZVP basis set were used in all calculations.[33] By default, the Tamm-Dancoff approximation[36] was applied which simplifies the calculation of excited states without significantly affecting the energies calculated, and improving them for some cases involving triplet excited states. Due to the consistent underestimate of excitation energies using this exchange correlation functional, a shift of +12.4 eV was applied to all calculated spectra to calibrate the photon energy scale.

*X-ray Photoelectron Spectroscopy (XPS)*

XPS characterization of the powder samples was carried out at the University of Leeds Versatile Soft X-ray Facility (VXSF). XPS spectra were collected with a SPECS EnviroESCA NAP-XPS equipped with a monochromatic Al $K_\alpha$ X-ray source (1486.71 eV) operating at 42 W, which is separated from the analysis chamber by a SiN window and illuminates the sample with a ~300 μm-diameter beam footprint. Spectra were collected for powder samples pressed onto double-sided adhesive carbon tape at ambient temperature in 7 mbar argon, with a hemispherical Phoibos NAP 150 analyzer operating in a small-area mode with a source–analyzer angle of 55° and a 1D delay-line detector. Under NAP conditions in the EnviroESCA instrument, electrically insulating sample surfaces are charge-neutralized through ionization of gas phase atoms by energetic electrons emitted from the sample and, to a lesser extent, the X-ray beam itself. This provides positive ions and electrons to balance any surface charge. Survey spectra were collected in one scan with a step size of 1 eV, a pass energy of 100 eV, and a dwell time of 0.1 s. High-resolution C 1s, N 1s, and O 1s core-level spectra were collected with a step size of 0.1 eV, a dwell time of 100 ms per data point, and a pass energy of 50 eV. Spectra for C 1s and O 1s core levels were collected in nine scans. Because of the lower atomic abundance of N within the sample, the N 1s core-level spectrum was collected in 16

scans. The spectra were analyzed using CasaXPS.[37] Shirley backgrounds and a GL(30) line shape (30% Lorentzian, 70% Gaussian) were used to fit the emission lines with appropriate components to describe the environment of each element in the sample.

*Near Edge X-ray Absorption Fine Structure (NEXAFS)*

NEXAFS characterization of the powder samples was carried out at Diamond Light Source beamline VerSoX B07-B utilizing the near ambient pressure NEXAFS end station enabling fast sample transfer.[38] Measurements were carried out using total electron yield (TEY) detection, measuring the drain current through the sample mounting plate with the soft X-ray beam at normal incidence and at room temperature. Samples were mounted on metal sample holders with carbon tape on copper foil, with the crystalline powders pressed onto the carbon tape. The photon energy scale was calibrated for beamline effects and surface charging using a small trace of nitrogen gas in the sample chamber which is therefore present in the gas phase spectra. This is well known to form a peak at 400.8 eV, [39,40] to which the photon energy scale was calibrated by applying a uniform shift to each spectrum. The sample chamber was operated with He gas at 1 mbar. He was chosen as there are no absorption peaks in the region around the C, N, O K-edges, and because it minimizes the electron yield signal of fluorescence gas phase absorption. The incident X-ray beam intensity $I_0$ is required for determining the absorption spectrum (see Appendix B) and was acquired by measuring the He gas phase spectrum when the sample plate was removed from the beam path. There is a difference in the gas phase beam path from which electrons are emitted when the sample is removed in this way, and this needs to be accounted for during the spectrum normalization process (see Appendix B).

*Crystallography*

The structures of the prepared crystalline compounds were confirmed through single crystal X-ray Diffraction measurement at Diamond Light Source beamline I19. Measurements were made at room temperature using radiation with a wavelength of $\lambda$=0.6889 Å.

## 5.3 Results

### *5.3.1 XPS*

Spectra were processed as described in detail in our previous paper on self-consistent XPS analysis of small molecule organic materials.[7] This ensures that the determined fits are all consistent with the expected sample structure and relative

**Table 5.2. Experimental Binding Energies. C 1s. Adventitious carbon proportion determined from the survey spectra. The C-X component includes all aromatic carbon covalently bonded to any other electron withdrawing group including nitro, amide, carboxylic acid. Expected % are calculated from the anticipated molecular structure, accounting for the known quantity of adventitious carbon. Very good agreement across all five samples.**

| | | C-C= + Adv | C−X | C−ONH | C-OOH |
|---|---|---|---|---|---|
| 44BPYSA | Energy / eV | 285.0 | 285.6 | | 288.8 |
| | Relative % | 41.7 | 47.6 | n/a | 10.7 |
| | Expected % | 40.8 | 47.3 | | 11.8 |
| INHFA | Energy / eV | 285.0 | 285.8 | 288.0 | 289.1 |
| | Relative % | 38.7 | 37.9 | 11.7 | 11.7 |
| | Expected % | 40.8 | 39.4 | 9.9 | 9.9 |
| INOA | Energy / eV | 285.0 | 286.0 | 288.0 | 288.6 |
| | Relative % | 52.9 | 23.6 | 11.8 | 11.8 |
| | Expected % | 52.8 | 23.6 | 11.8 | 11.8 |
| 4EP35DNBA | Energy / eV | 285.0 | 286.2 | | 288.7 |
| | Relative % | 69.8 | 24.1 | n/a | 6.0 |
| | Expected % | 69.7 | 24.2 | | 6.0 |
| 35L35DNBA | Energy / eV | 285.0 | 286.1 | | 288.4 |
| | Relative % | 66.7 | 25.9 | n/a | 7.5 |
| | Expected % | 67.5 | 26.0 | | 6.5 |

atomic abundances. The binding energy scale calibration was carried out relative with respect to the C 1s emission component attributed to adventitious carbon on the sample surface, setting its BE to 285.0 eV. Adventitious carbon contributions were determined using the known elemental composition and the elemental analysis of the survey spectrum to quantify the amount of excess carbon. Excess carbon contents ranged from 8 to 21 mol%, across the samples we examined, comparing well to levels identified in multiple previous studies of organic materials.[7]

The multicomponent fits to the C 1s spectra were all consistent with the expected elemental composition, with relative abundances within 2% of the ideal values as

**Table 5.3. Experimental N 1s Binding Energies. BE Calibration with respect to the adventitious carbon peak. Expected % calculated from the anticipated molecular structure. Good agreement for 44BPYSA and INHFA samples, but unexpected splitting of hydrogen bonded nitrogen peak observed in INOA, 4EP35DNBA and 35L35DNBA.**

| | | $N\text{-}H_2$ | N-C | $N\text{-}H^+$ | $NO_2$ |
|---|---|---|---|---|---|
| 44BPYSA | Energy / eV | | 399.4 | | |
| | Relative % | n/a | 100 | n/a | n/a |
| | Expected % | | 100 | | |
| INHFA | Energy / eV | 399.9 | 399.9 | 400.7 | |
| | Relative % | 33.3 | 33.3 | 33.3 | n/a |
| | Expected % | 33.3 | 33.3 | 33.3 | |
| INOA | Energy / eV | 399.3 | 399.3 | 400.9 | |
| | Relative % | 50.0 | 19.2 | 30.8 | n/a |
| | Expected % | 50.0 | 50.0 | 0 | |
| 4EP35DNBA | Energy / eV | | 400.0 | 401.3 | 405.9 |
| | Relative % | n/a | 6.0 | 15.9 | 78.1 |
| | Expected % | | 0 | 33.3 | 66.7 |
| 35L35DNBA | Energy / eV | | 399.4 | 401.0 | 405.8 |
| | Relative % | n/a | 3.8 | 18.8 | 77.4 |
| | Expected % | | 33.3 | 0 | 66.7 |

shown in Table 5.2. Table 5.3 summarizes the fits to the experimental data at the N 1s K-edge. Experimental spectra are shown in Figure 5.2. The first two samples agree perfectly with the expected elemental composition and crystallographic data, exhibiting the expected cocrystal structure. In the case of INHFA, the -NH binding energy is not involved in hydrogen bonding, and the higher binding energy is entirely due to the relatively high electronegativity of other functional groups in its local environment. The binding energy of the hydrogen bonded nitrogen is well within the co-crystal range, and therefore the potentially intermediate hydrogen position needs further investigation. For each of the last three samples relative atomic abundances do not match with the expected values. For the dinitrobenzoic acid complexes, the nitro group appears to dominate the relative abundance more than it should by

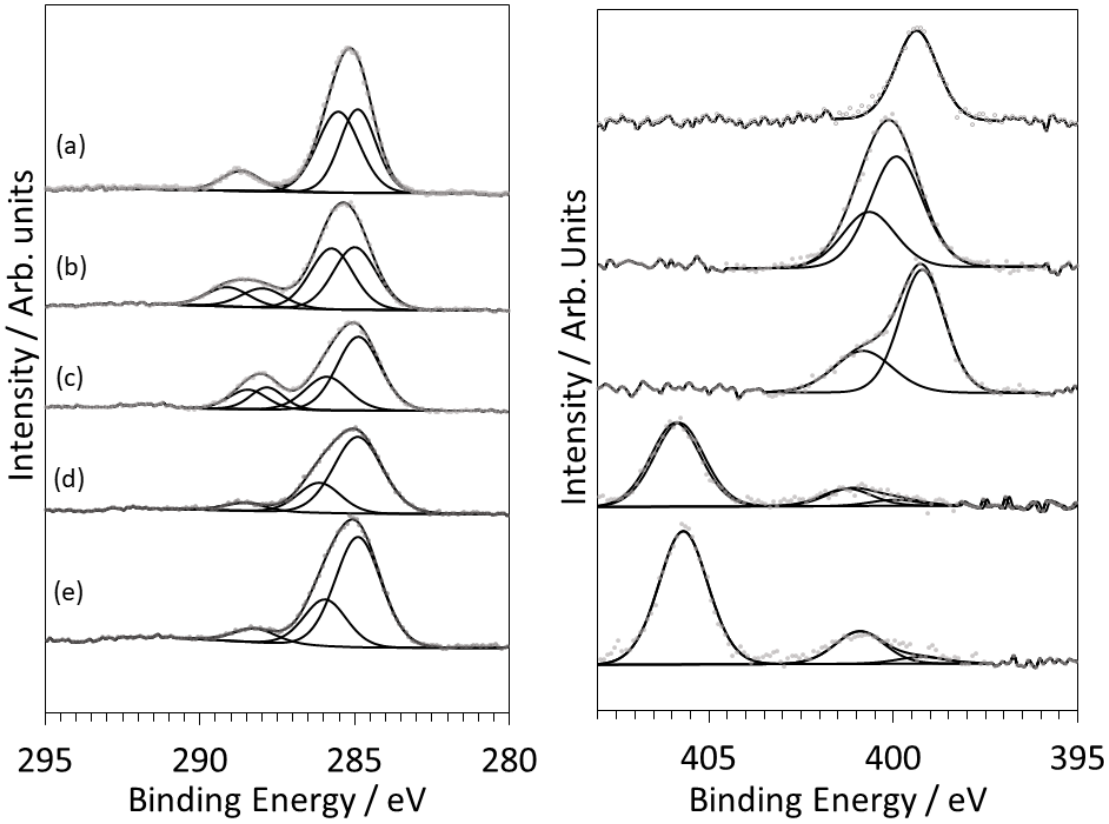



**Figure 5.2.** Fitted Experimental XPS spectra of the 5 samples, left C 1s and right N 1s. (a) 44BPYSA, (b) INHFA, (c) INOA, (d) 4EP35DNBA and (e) 35L35DNBA. Peak positions and relative intensities are reported in table 5.2 (C 1s) and table 5.3 (N 1s). Peak fits consistent with fixed FWHM and peak intensities to match expected stoichiometry.

around 10%. This is likely to be due to a surface effect, possibly a non-stoichiometric excess of the dinitrobenzoic acid coformer at the surface, as observed in a previous study.[3] It does not affect the analysis of the hydrogen bonding due to the clear binding energy separation between the amino and nitro N 1s emission components. From the crystallographic data, each of these three samples are expected to form either a salt (4EP35DNBA) or cocrystal (35L35DNBA and INOA). In each case, the peak relating to the proton acceptor nitrogen is split between two distinct binding energies with apparent disorder across the donor and acceptor groups. For 35L35DNBA and INOA, despite expecting a co-crystal, it would appear the majority of interactions involve at least a level of proton transfer. For 35L35DNBA, there appears to be full proton transfer with a binding energy of 401.3 eV, while in INOA

this could be partial proton transfer due to the lower binding energy of 400.9 eV. For 4EP35DNBA, both binding energy positions are towards the salt structure, although the majority of interactions appear to fall in the region below full proton transfer with binding energy of 400.6 eV. We can be reasonably confident in the peak assignments made here due to the consistent method of analysis utilized, and the discrepancies from published structures need investigating further using NEXAFS and DFT to identify the cause of these discrepancies.

### *5.3.2 NEXAFS*

NEXAFS spectra were processed in accordance with our previously published method for calibrating and normalizing NAP-NEXAFS spectra.[11] There are two effects we normalize out of the spectra during this process: the effect of the gas phase He in the X-ray beam path in the sample chamber and the effect of the silicon nitride membrane at the entrance to the sample chamber. The processed and calibrated spectra are shown in Figure 5.3. NEXAFS spectra are not quantitative in the same way as XPS, meaning that the peak intensity is not directly a measure of the relative abundance of a particular nitrogen environment. Instead, the peak intensity depends upon the oscillator strength of the electronic transition between the core level initial state and the valence orbital final state of the electron. The orbital overlap between these states gives rise to the transition electric dipole moment, the likelihood of a particular electron transition taking place, which is directly related to the peak intensity observed in NEXAFS. As can be seen in the spectra, the nitro ($NO_2$) groups have the highest transition dipole moment, due to the large orbital overlap between the nitro groups and the 1π* state on the aromatic ring. In 44BPYSA, there is only a single nitrogen environment resulting in a single 1s → 1π* peak, followed by 1s → 2π* and 1 s→ 3π* resonances. The 1s → π* resonances in each of the remaining samples is a little more complex due to the overlap of various nitrogen environments, requiring DFT calculations to determine the origin of each peak. However, similarly to the XPS data for INOA, 35L35DNBA and 4EP35DNBA, there is a split in the peak corresponding to the hydrogen bonded nitrogen environment suggesting more than one proton position in the structure is present at room temperature.

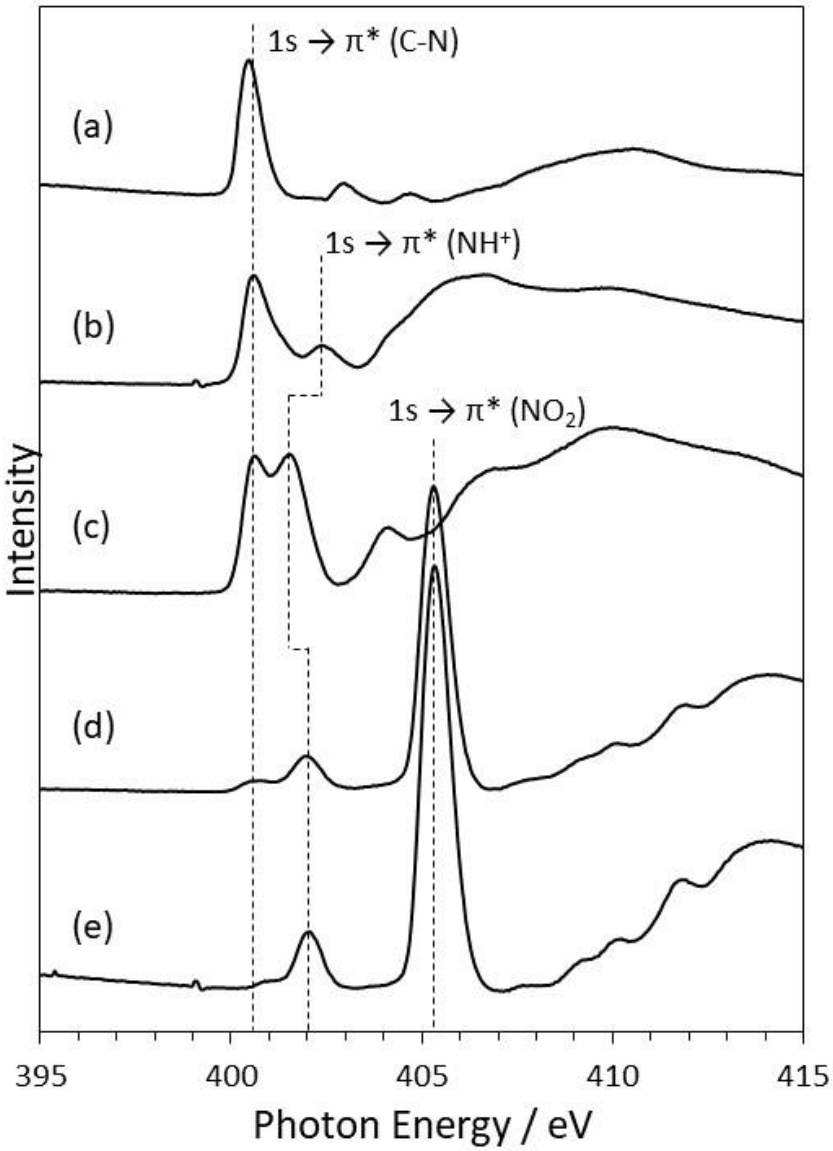



**Figure 5.3.** Experimental NEXAFS spectra of the 5 samples. (a) 44BPYSA, (b) INHFA, (c) INOA, (d) 4EP35DNBA and (e) 35L35DNBA. Important peak positions labelled showing equivalent peaks over the five structures. Non proton transferred peaks (C-N and $NO_2$) very consistent in energy while $NH^+$ shows some variation between samples.

### *5.3.3 DFT*

DFT modelled spectra were calculated using the orca_mapspc program within the *ORCA* software suite,[31] using a Gaussian line broadening of 0.6 eV to compute the X-ray absorption spectra, as this best matches the experimental spectra for the N 1s → π* transitions. The photon energy scale is shifted by approximately 12 eV such that the main 1s→π* peak matches with the experimentally observed value. This is necessary due to the consistent underestimation of excitation energies calculated

using the B3LYP exchange correlation functional.[41] Initially, the literature structures of the samples were used to simulate the expected spectra for comparison to the experimental NEXAFS. These can be seen in Figure 5.4. In each of the spectra, the main features of the experimental data are present (Figure 5.3). For 44BPYSA, each of the three 1s → π* are present at the correct photon energies. In INHFA, the primary 1s → π* peak is split correctly with the expected relative intensities, although the photon energies are slightly off. For the three remaining samples, the errors in the simulated spectra are much more obvious, with a missing peak/should peak in

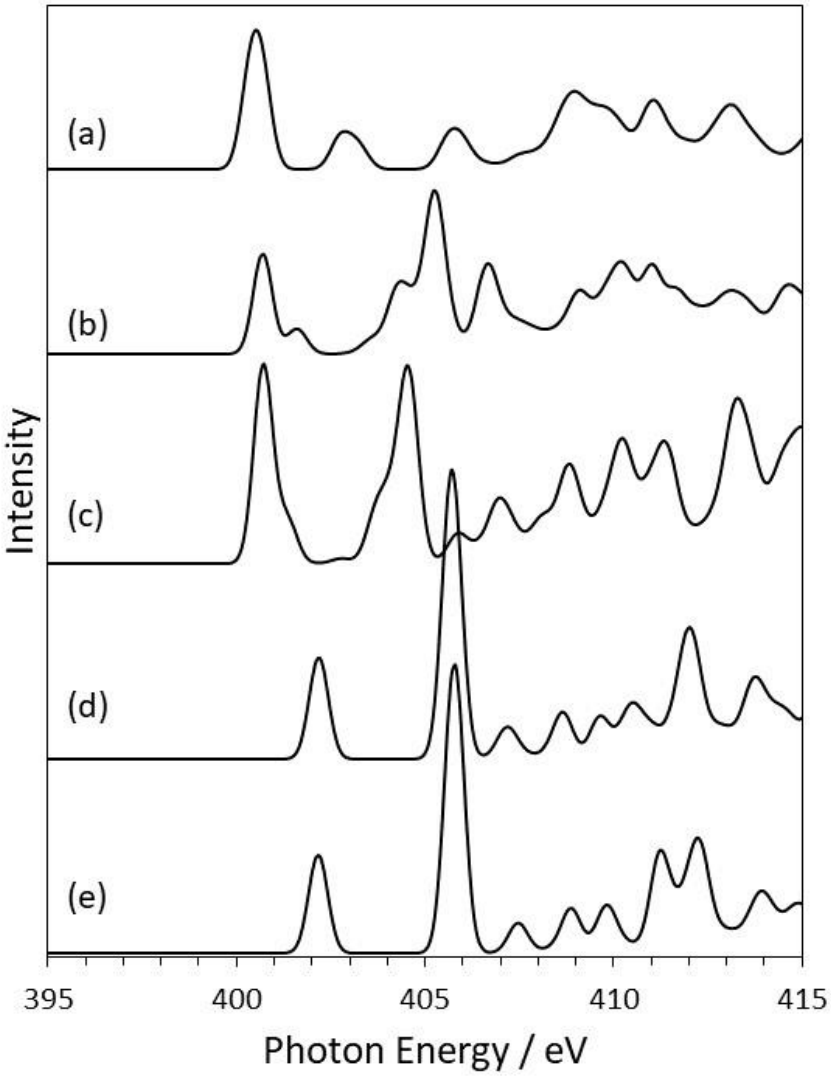



**Figure 5.4.** Simulated NEXAFS spectra of the 5 samples based on the literature geometry. (a) 44BPYSA, (b) INHFA, (c) INOA, (d) 4EP35DNBA and (e) 35L35DNBA. Some significant differences in peak shapes and positions compared to experimental data suggesting incorrect input structures for the calculations.

all three indicating that the structure being simulated is not a good reflection of the real system.

## 5.4 Discussion

We have investigated the five samples using three methods: XPS, NEXAFS and DFT. Each of these approaches provides different information about the systems, but all also provide complementary information where the results are consistent with each other. In the initial results shown in Table 5.3 as well as Figures 5.3 and 5.4 there are clearly some interesting features to investigate further and potentially suggest the reported crystal structures of these materials may need further attention.

### *5.4.1 4,4'-Bipyridine Succinic Acid*

According to X-ray crystallography, the 4,4'-bipyridine complex forms a 1:1 co-crystal structure, although the precise location of the hydrogen atom is ambiguous and could fall in the region known as the salt-co-crystal continuum. Using XPS, NEXAFS and DFT, we have identified the proton position through a fully consistent analysis comparing all aspects of these three techniques. In all three, a single 1s → 1π* peak or binding energy is observed confirming there is only a single stable state at room temperature with the proton in the same position in all molecules in a single nitrogen environment. The position of this peak is fully consistent with a co-crystal structure with a BE of 399.3 eV. Additionally, the NEXAFS experimental data are equally convincing as to the co-crystal nature of the structure, with this sample providing the best agreement between experimental and simulated NEXAFS spectra. The three resonances are found at 400.5 eV, 403 eV and 404.9 eV in experimental data and 400.5 eV, 403 eV and 405.4 eV in the calculated spectra. These are in good agreement and consistent with each other, suggesting that the structure of 44BPYSA is accurate in the structural database, with N…H distance of 1.325 Å and an N…O distance of 2.624 Å.

### *5.4.2 INHFA*

The isoniazid fumaric acid complex appears to form a co-crystal structure at room temperature according to the structure available in the CSD. However, the N…H bond distance is shown to be 1.332 Å which is very close to the centre of the hydrogen bond. On first sight, neither the XPS nor NEXAFS data are consistent with this structure, with binding and photon energies in the region expected for a sample which is a co-crystal. The DFT calculation based on the defined structure confirms there is a difference between the experimental data and the simulated spectra with a smaller difference in photon energy between the two main 1s → π* features (400.7 eV and 402.5 eV experimentally but 400.7 eV and 401.6 eV for the database structure). In addition, the calculated core level orbital energies (essentially equivalent to BE in XPS) suggest that the hydrogen bonded nitrogen acceptor would have a BE similar to the -NH group, rather than similar to the $NH_2$ group, which is clearly not the case based upon the measured XP spectrum.

Following our procedure set out previously,[11] a set of calculations were run moving only the proton involved in hydrogen bonding over a range of positions between the proton donor and acceptor. Based on the issues outlined so far, we would expect to find an improved fit with experimental data, with the proton closer to the oxygen donor, making it more like a typical cocrystal. This is in fact the case, with the optimal structure found with the proton shifted 0.4 Å towards the oxygen position in the CSD structure. The simulated NEXAFS spectrum for this structure is shown in Figure 5.5 with the two 1s → π* features separated by the correct photon energy and the σ* resonances starting at the correct photon energy of 405 eV. Compared to the experimental spectrum, the larger 1s → π* resonance does not exhibit a high energy shoulder to the same extent. The origin of this shoulder is the 1s → π* ($NH_2$) resonance, which due to the low orbital overlap with the 1π* orbital has a small contribution to the NEAXFS spectrum. This feature is present in the simulation and at the correct photon energy (1 eV above the proton acceptor nitrogen resonance), the interaction cross section just appears to be slightly larger experimentally than

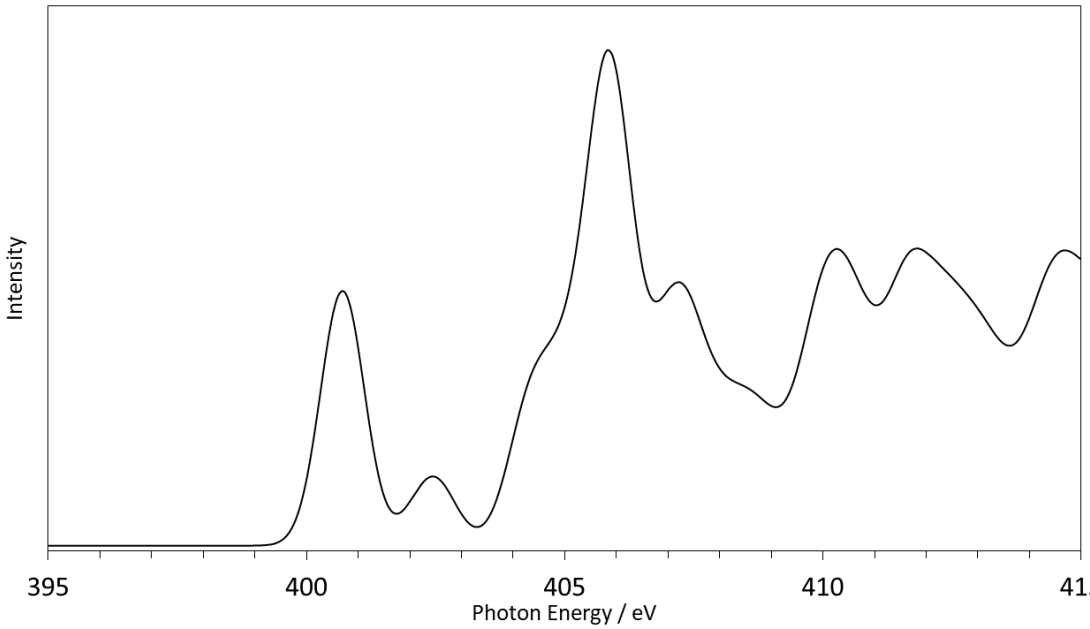



**Figure 5.5.** Simulated NEXAFS spectrum of INHFA with the hydrogen bond proton moved 0.4 Å towards the proton donor showing improved agreement to experimental data.

predicted. Therefore, with this new structure, the experimental and theoretical NEXAFS are fully consistent and consistent with a co-crystal structure.

The improved structure was compared with the experimental XPS to ensure full consistency, as shown in table 5.4. The effect of moving the proton primarily affects the orbital energy of the proton acceptor nitrogen atom, reducing it to be within 0.1

**Table 5.4. INHFA summary of results showing the change in proton position and the effect on the calculated binding energies and electron transition energies to better agree with experiment.**

| | N…H /Å | N…O / Å | N 1s → π* / eV | XPS N1s BE / eV |
|---|---|---|---|---|
| Literature (DFT) | 1.33 | 2.54 | 400.7<br>401.7 | 399.9<br>401.0<br>401.1 |
| Experimental | n/a | 2.54 | 400.7<br>402.5 | 399.9<br>399.9<br>400.7 |
| Improved structure (DFT) | 1.73 | 2.54 | 400.7<br>402.5 | 399.9<br>399.9<br>400.9 |

eV of the $NH_2$ group, as observed experimentally. This additional test confirms that our experimental measurements correspond to this altered structure with the proton 0.4 Å closer to the oxygen atom than reported.

### *5.4.3 INOA*

The isonicotinamide oxalic acid complex forms a 2:1 co-crystal structure according to the CSD. Again, this was identified as being close to the region where quasi-centred hydrogen bonds are found. It can clearly be identified from figures 5.3 and 5.4 that there is a significant difference between the experimental spectrum and the simulated spectrum. The simulated spectrum contains a single 1s → π* peak, consisting of a main resonance and a shoulder to high energy. The main peak corresponds to the two proton acceptor nitrogen atoms in the unit cell, and the shoulder originates from the amide groups. As we have observed previously, these amide groups are not directly attached to the aromatic ring, and therefore the orbital overlap with the 1π* orbital is reduced leading to a low transition cross section. In the experimental data, there is clearly an approximate 1:1 split between two separate peaks, indicating that a certain proportion of the proton acceptor nitrogen atoms are in a different state. This analysis of the NEXAFS data is consistent with XPS, where the acceptor nitrogen appeared to split into two separate components. Therefore, this suggests that the database structure of INOA is correct for some of the hydrogen bonding interactions, but at room temperature there appears to be a population of sites in a different state. To find this state, a scan over one of the two hydrogen bonding proton positions was calculated to model a 1:1 split of the two states, and the optimal structure was found, as shown in figure 5.6, with one of the protons moved 0.4 Å towards the nitrogen acceptor. A split such as this has been observed before in a 2:1 complex with one salt and one co-crystal interaction.[7,11] Table 5.5 shows a summary of the results for the improved structure. In the simulated NEXAFS spectrum, the peak separation of the 1s → π* resonance is correct, along with the higher energy peak exhibiting a slight broadening due to the amide groups (with low interaction cross section) having a slightly lower 1s → 1π* resonance energy.

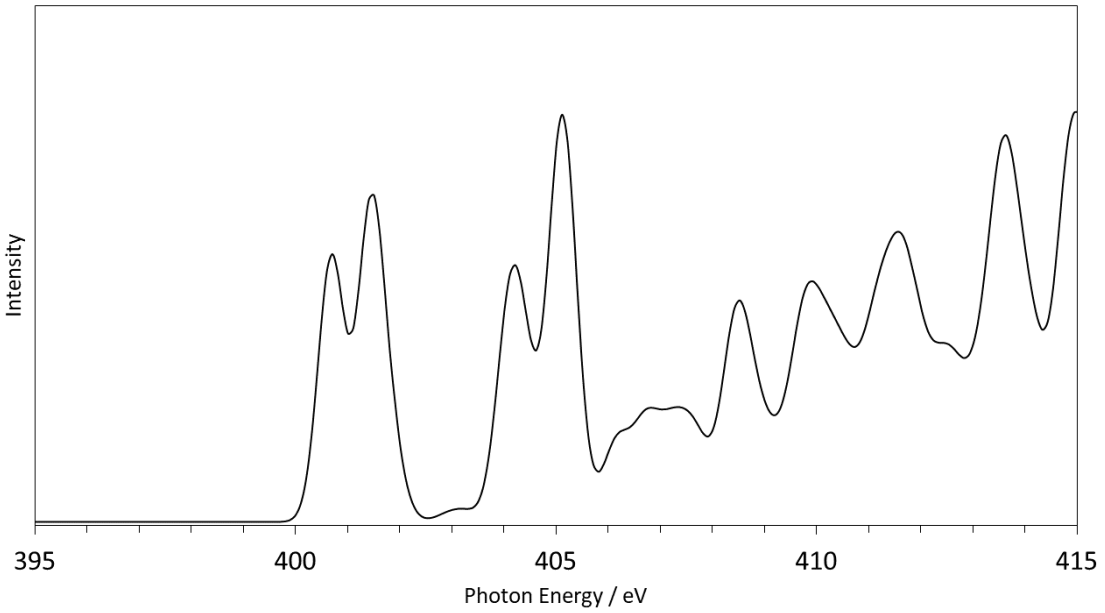



**Figure 5.6.** Simulated NEXAFS spectrum of the INOA with hydrogen bond proton moved 0.4 Å towards the proton acceptor, demonstrating improved agreement with experiment.

The second set of π* peaks in the simulated spectrum are 1s → 2π*, visible in the experimental spectrum at 404.3 eV although less pronounced. The amide groups also explain the increased peak height of the second peak compared to the experimental data where the additional intensity is due to the amide components. This implies that the split of nitrogen acceptor states must be slightly skewed towards

**Table 5.5. INOA (2:1) summary of results showing the change in half of the hydrogen bonded proton positions and the effect on the calculated binding energies and electron transition energies to better agree with experiment.**

| | N…H /Å | N…O / Å | N 1s → π* / eV | XPS N1s BE / eV |
|---|---|---|---|---|
| Literature (DFT) | 1.39 | 2.56 | 400.7<br>400.7<br>401.4<br>401.4 | 399.3<br>399.3<br>399.5<br>399.5 |
| Experimental | n/a | 2.56 | 400.7<br>401.6<br>401.6<br>401.6 | 399.3<br>399.3<br>399.3<br>400.9 |
| Improved structure (DFT) | 0.99&1.39 | 2.56 | 400.7<br>401.4<br>401.4<br>401.5 | 399.3<br>399.5<br>399.5<br>401.2 |

the lower photon energy co-crystal state to result in the two peaks having equal intensity as shown experimentally. This is also in agreement with the experimental XPS where a roughly 60:40 split was identified between the salt and co-crystal interactions as shown in table 5.3. Again, the binding energies determined from the DFT calculation are in good agreement with experiment for the improved structure, with values within 0.3 eV for all four nitrogen environments in the complex. Since all of these techniques are consistent, this suggests strongly that the sample at room temperature may behave as described with two distinct hydrogen bond interactions, one behaving as a salt with full proton transfer to the nitrogen acceptor and the other behaving as a cocrystal. The distribution of these states appears to favor the cocrystal (as found in the CSD) by a ratio of 6:4 to give the correct peak intensities in XPS and NEXAFS. In this case, these differences between simulated spectrum based on the literature structure and the experimental data cannot be explained by a deficiency in the computational method – the experimental data clearly indicate the inconsistency and the simulation just confirms it.

### *5.4.4 4EP35DNBA*

The 4-ethylpyridine 3,5-dinitrobenzoic acid complex forms a salt structure close to the salt co-crystal continuum according to the crystal structure database. From both the experimental XPS and NEXAFS, we have observed a splitting of the peak corresponding to the hydrogen bonded nitrogen environment, which is not present in the simulated NEXAFS, nor do the binding energies of this splitting make sense for a purely salt structure. As there is only a single nitrogen environment (excluding the nitro groups which are easily assigned) there must be some degree of proton disorder in the system such that a certain proportion of protons reside in a different position, suggesting that the proton potential energy surface may have two relatively stable positions. Since the structure from the CSD agrees well with the positions of the nitro peak and the high binding energy of the remaining peaks, this forms a useful starting point to model the situation.

To obtain a binding energy lower than the simulated peak, a structure involving the proton residing closer to the proton donor oxygen is required, so a series of scans were completed to identify the optimum structure for this second conformation. Using

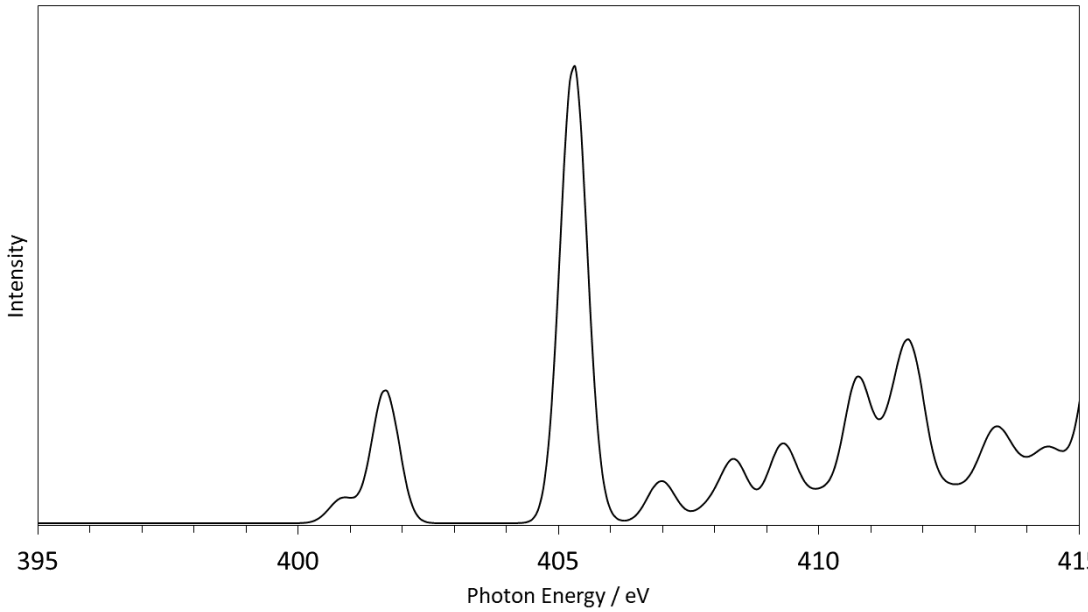



**Figure 5.7.** Simulated NEXAFS spectrum of the 4EP35DNBA with hydrogen bond proton moved 0.5 Å towards the proton donor in ¼ of molecules. This proportion was determined from the relative intensities of XPS peaks. Agreement with experimental spectrum is improved.

the quantitative XPS measurement, we can determine the anticipated proportion of each hydrogen bonding interaction.

Figure 5.7 shows the calculated spectrum for the combination of crystal structure with the amended proton position in a 3:1 ratio. This shows a good agreement with experimental data, with the intense $NO_2$ peak at 405.5 eV used for energy scale calibration. The peak at 402 eV is in the same place as experimentally, while the new peak from the improved structure is within 0.3 eV of the experimental value, as shown in table 5.6. The simulated XPS binding energies are also in line with the experimentally determine values all within 0.5 eV of the experimental values. While the accuracy of this is reduced compared to the previous samples, it is still an improvement on the original crystal structure by way of the observed additional broadening of the emission line.

### *5.4.5 35L35DNBA*

According to the CSD, the 3,5-lutidine 3,5-dinitrobenzoic acid complex forms a co-crystal structure close to the salt cocrystal continuum region with an N…H distance roughly equal to the O…H distance. As shown in figures 5.3 and 5.4, the simulated

**Table 5.6. 4EP35DNBA summary of results showing the change in a quarter of the hydrogen bonded proton positions and the effect on the calculated binding energies and electron transition energies to better agree with experiment.**

| | N…H /Å | N…O / Å | N 1s → π* / eV | XPS N1s BE / eV |
|---|---|---|---|---|
| Literature (DFT) | 1.20 | 2.56 | 402.0<br>402.0<br>405.5 | 401.8<br>401.8<br>405.9 |
| Experimental | n/a | 2.56 | 400.8<br>402.0<br>405.5 | 400.0<br>401.3<br>405.9 |
| Improved structure (DFT) | 1.20&1.70 | 2.56 | 401.1<br>402.0<br>405.5 | 399.6<br>401.8<br>405.9 |

NEXAFS spectrum does not match the experimental data, missing a slight shoulder peak to the 1s → π* transition for the nitrogen acceptor. Since there is only a single nitrogen environment in this structure (other than the nitro groups), there must either be two stable states, each corresponding to a different photon energy, or the peak is broadened due to proton disorder around a centred position. Based on the NEXAFS spectrum, it would appear the first option is most likely due to two distinct peaks being present, similarly to 4EP35DNBA. From the XPS experimental data, it would appear that the majority of hydrogen bonding interactions involve proton transfer with a binding energy of 401 eV, and a minority of co-crystal interactions. Therefore, to model this using DFT, a combination of two calculations is required, one with each proton position and then combined in the appropriate proportions. In the structural database, there is a second structure of 35L35DNBA with a much longer N…H distance, which was used as for the second proton position. Using these structures, the spectrum in figure 5.8 was simulated. The relative contribution of each component was determined from the XPS fits, resulting a 6:1 ratio of salt to co-crystal contribution to the total. Figure 5.8 shows a good agreement with the experimental data, with the equivalent shoulder peak now present in the spectrum. Additionally, the $NO_2$ peak appears at the same photon energy in both calculations, confirming the effect is limited to the proton acceptor nitrogen. Table 5.7 shows how well the NEXAFS data for the improved structure agree with experiment, with a maximum energy difference between peak positions of 0.2 eV.

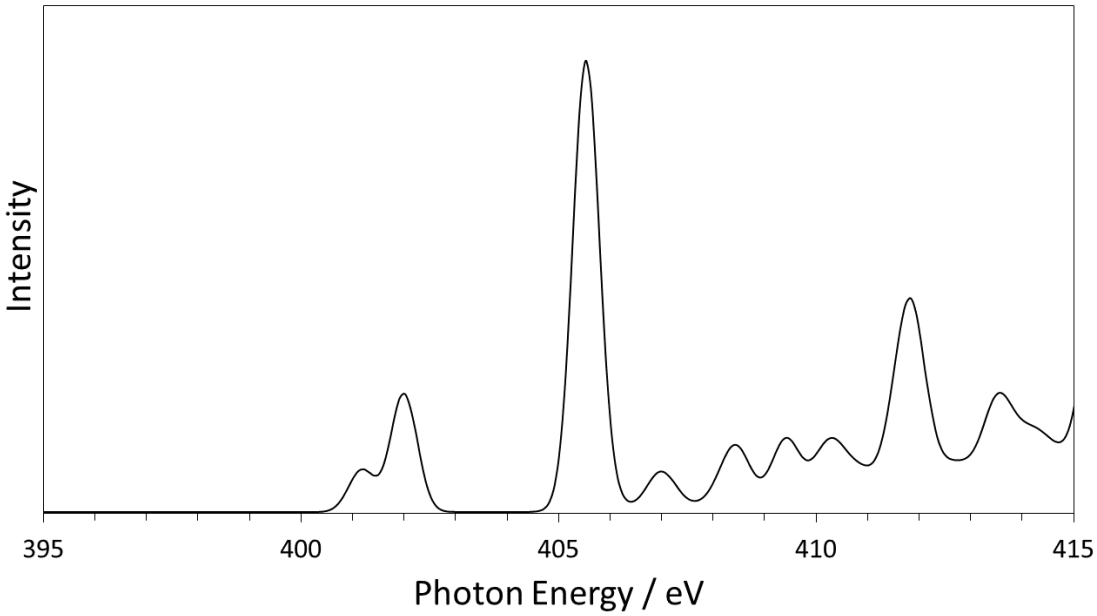



**Figure 5.8.** Simulated NEXAFS spectra of the 35L35DNBA combining the two structures found in the database to model the peak shape observed experimentally.

However, the simulated XPS data appears less convincing for this structure. The nitro group is used for energy scale calibration of the DFT calculation, so is always consistent. The improved structure appears to give the correct binding energy for the

**Table 5.7. 35L35DNBA summary of results showing the change in 1/7 of the hydrogen bonded proton positions and the effect on the calculated binding energies and electron transition energies to better agree with experiment. The final structure was required to ensure both the transitions and binding energies agreed with experiment.**

| Structure / method | N…H /Å | N…O / Å | N 1s → π* / eV | XPS N1s BE / eV |
|---|---|---|---|---|
| Literature (DFT) | 1.37 | 2.54 | 401.9<br>401.9<br>405.4 | 401.8<br>401.8<br>405.8 |
| Experimental | n/a | 2.54 | 401.1<br>402.1<br>405.4 | 399.4<br>401.0<br>405.8 |
| Improved structure (DFT) | 1.37&1.66 | 2.54&2.56 | 401.0<br>401.9<br>405.4 | 399.5<br>401.8<br>405.8 |
| Final Structure (DFT) | 1.26&1.66 | 2.56 | 401.0<br>401.7<br>405.4 | 399.5<br>401.2<br>405.8 |

nitrogen acceptors in the co-crystal state, so the issue appears to be with the original crystal structure found in the database. Therefore, an additional scan was completed, based upon the second structure to obtain a new salt structure, identified as the final structure in table 5.7. This shows a much-improved set of binding energies calculated with only a slight decrease in the relevant 1s → π* photon energy.

## 5.5 Conclusions

This study highlights the level of geometric and electronic structure details accessible by combining experimental NEXAFS and XPS with computational DFT to generate a self-consistent description of organic small molecule crystal structures. Five two-component complexes which were chosen specifically to challenge the XPS/NEXAFS/DFT approach to proton refinement. The procedure followed predicts new proton positions for disordered hydrogen bonding interactions and providing an improved level of accuracy compared to the structures found in the CSD. Of the five samples investigated, only one resulted in consistency with the crystallographic structure found in the CSD. The remaining four samples all required further refinement of the proton position, with either peak broadening or splitting, suggesting disorder across two different proton positions throughout the sample. Our structure refinement based on the CSD structures was completed through variations in proton positions until the DFT predictions of orbital energies were consistent with the experimental data. The disordered states found for three of the crystal structures, where the proton resides in two different positions in an appartently fixed ratio, is interesting and should be investigated further, for example by temperature-dependent measurements that may shift the population of each state.

## Acknowledgements

We gratefully acknowledge funding from the EPSRC and Diamond Light Source to cover a PhD studentship (EPSRC Grant EP/R513258/1). We thank Diamond Light Source for access to beamline B07-B (SI29334) during the early stages of comissioning that contributed to the results presented here. SLMS thanks the

Future Continuous Manufacturing and Advanced Crystallisation (CMAC) Hub for financial support (EPSRC Grant EP/P006965/1). Our thanks to Anna Kroner and Nathan Hennessy for their assistance with the beamtime. The Density Functional Theory calculations were undertaken using ARC4, part of the High Performance Computing facilities at the University of Leeds, UK. All data supporting this study are provided either in the results section of this paper or in the ESI accompanying it.

## 5.6 References

1 S. L. Childs, G. P. Stahly and A. Park, The Salt-Cocrystal Continuum: The Influence Of Crystal Structure On Ionization State, *Mol Pharm*, 2007, **4**, 323–338.

2 J. Wouters and L. Quéré, *Pharmaceutical Salts and Co-crystals*, The Royal Society of Chemistry Publishing, 2012, 9-28.

3 J. S. Stevens, S. J. Byard and S. L. M. Schroeder, Salt Or Co-Crystal? Determination Of Protonation State By X-Ray Photoelectron Spectroscopy (XPS), *J Pharm Sci*, 2010, **99**, 4453–4457.

4 C. B. Aakeröy, M. E. Fasulo and J. Desper, Cocrystal Or Salt: Does It Really Matter?, *Mol Pharm*, 2007, **4**, 317–322.

5 T. Wang, J. S. Stevens, T. Vetter, G. F. S. Whitehead, I. J. Vitorica-Yrezabal, H. X. Hao and A. J. Cruz-Cabeza, Salts, Cocrystals, And Ionic Cocrystals Of A ‘Simple’ Tautomeric Compound, *Cryst Growth Des*, 2018, **18**, 6973–6983.

6 C. L. Perrin, Are Short, Low-Barrier Hydrogen Bonds Unusually Strong?, *Acc Chem Res*, 2010, **43**, 1550–1557.

7 P. T. Edwards, L. K. Saunders, A. R. Pallipurath, A. J. Britton, E. A. Willneff, E. J. Shotton and S. L. M. Schroeder, Proton Transfer on the Edge of the Salt/Cocrystal Continuum: X-Ray Photoelectron Spectroscopy of Three Isonicotinamide Salts, *Cryst Growth Des*, 2021, **21**, 6332–6340.

8 J. S. Stevens and S. L. M. Schroeder, Quantitative Analysis Of Saccharides By X-Ray Photoelectron Spectroscopy, *Surface and Interface Analysis*, 2009, **41**, 453–462.

9 J. S. Stevens, S. Coultas, C. Jaye, D. A. Fischer and S. L. M. Schroeder, Core Level Spectroscopies Locate Hydrogen In The Proton Transfer Pathway-Identifying Quasi-Symmetrical Hydrogen Bonds In The Solid State, *Physical Chemistry Chemical Physics*, 2020, **22**, 4916–4923.

10 J. S. Stevens, S. J. Byard, C. C. Seaton, G. Sadiq, R. J. Davey and S. L. M. Schroeder, Proton Transfer And Hydrogen Bonding In The Organic Solid State: A Combined XRD/XPS/SSNMR Study Of 17 Organic Acid-Base Complexes, *Physical Chemistry Chemical Physics*, 2014, **16**, 1150–1160.

11 P. T. Edwards, L. K. Saunders, D. C. Grinter, P. Ferrer, G. Held, E. J. Shotton and S. L. M. Schroeder, Determination of H-Atom Positions in Organic Crystal Structures by NEXAFS Combined with Density Functional Theory: a Study of Two-Component Systems Containing Isonicotinamide, *J Phys Chem A*, 2022, **126**, 2889–2898.

12 A. Gainar, J. S. Stevens, C. Jaye, D. A. Fischer and S. L. M. Schroeder, NEXAFS Sensitivity to Bond Lengths in Complex Molecular Materials: A Study of Crystalline Saccharides, *Journal of Physical Chemistry B*, 2015, **119**, 14373–14381.

13 J. S. Stevens, S. J. Byard and S. L. M. Schroeder, Characterization of Proton Transfer in Co-Crystals by X-ray Photoelectron Spectroscopy (XPS), *Cryst Growth Des*, 2010, **10**, 1435–1442.

14 J. S. Stevens, S. J. Byard, C. A. Muryn and S. L. Schroeder, Identification Of Protonation State By XPS, Solid-State NMR, And DFT: Characterization Of The Nature Of A New Theophylline Complex By Experimental And Computational Methods, *J Phys Chem B*, 2010, **114**, 13961–13969.

15 J. S. Stevens, A. Gainar, C. Jaye, D. A. Fischer, S. L. M. Schroeder and Iop, NEXAFS and XPS of p-Aminobenzoic Acid Polymorphs: The Influence of Local Environment in *16th International Conference on X-ray Absorption Fine Structure (XAFS)*, 2016, **712**, 012133.

16 J. S. Stevens, L. K. Newton, C. Jaye, C. A. Muryn, D. A. Fischer and S. L. M. Schroeder, Proton Transfer, Hydrogen Bonding, and Disorder: Nitrogen Near-Edge X-ray Absorption Fine Structure and X-ray Photoelectron Spectroscopy of Bipyridine-Acid Salts and Co-crystals, *Cryst Growth Des*, 2015, **15**, 1776–1783.

17 J. S. Stevens, A. Gainar, E. Suljoti, J. Xiao, R. Golnak, E. F. Aziz, S. L. M. Schroeder and Iop, NEXAFS Chemical State And Bond Lengths Of P-Aminobenzoic Acid In Solution And Solid State in *16th International Conference on X-ray Absorption Fine Structure (XAFS),* 2016, **712**, 012136.

18 A. Parkin, S. M. Harte, A. E. Goeta and C. C. Wilson, Imaging Proton Migration From X-Rays And Neutrons, *New Journal of Chemistry*, 2004, **28**, 718–721.

19 M. B. P. Paula, T. F. Koetzle and A. J. Schultz, Single Crystal Neutron Diffraction For The Inorganic Chemist – A Practical Guide, *Comments on Inorganic Chemistry*, **28**, 3–38.

20 B. E. Warren, X-ray Diffraction, *Courier Corporation*, 1990.

21 A. Parkin, S. M. Harte, A. E. Goeta and C. C. Wilson, Imaging Proton Migration From X-Rays And Neutrons, *New Journal of Chemistry*, 2004, **28**, 718–721.

22 L. K. Saunders, H. Nowell, L. E. Hatcher, H. J. Shepherd, S. J. Teat, D. R. Allan, P. R. Raithby and C. C. Wilson, Exploring Short Strong Hydrogen Bonds Engineered In Organic Acid Molecular Crystals For Temperature Dependent Proton Migration Behaviour Using Single Crystal Synchrotron X-Ray Diffraction (SCSXRD), *CrystEngComm*, 2019, **21**, 5249–5260.

23 C. R. Groom, I. J. Bruno, M. P. Lightfoot and S. C. Ward, The Cambridge Structural Database, *Acta Crystallogr B Struct Sci Cryst Eng Mater*, 2016, **72**, 171–179.

24 P. Gilli, L. Pretto, V. Bertolasi and G. Gilli, Predicting Hydrogen-Bond Strengths from Acid-Base Molecular Properties. The pK(a) Slide Rule: Toward the Solution of a Long-Lasting Problem, *Acc Chem Res*, 2009, **42**, 33–44.

25 S. Cherukuvada and T. N. Guru Row, Comprehending The Formation Of Eutectics And Cocrystals In Terms Of Design And Their Structural Interrelationships, *Cryst Growth Des*, 2014, **14**, 4187–4198.

26 L. Rajput, M. Banik, J. R. Yarava, S. Joseph, M. K. Pandey, Y. Nishiyama and G. R. Desiraju, Exploring The Salt-Cocrystal Continuum With Solid-State NMR Using Natural-Abundance Samples: Implications For Crystal Engineering, *IUCrJ*, 2017, **4**, 466–475.

27 S. Cherukuvada and A. Nangia, Fast Dissolving Eutectic Compositions Of Two Anti-Tubercular Drugs, *CrystEngComm*, 2012, **14**, 2579–2588.

28 I. Majerz and M. J. Gutmann, Mechanism Of Proton Transfer In The Strong OHN Intermolecular Hydrogen Bond, *RSC Adv*, 2011, **1**, 219–228.

29 P. Vishweshwar, A. Nangia and V. M. Lynch, Molecular Complexes Of Homologous Alkanedicarboxylic Acids With Isonicotinamide: X-Ray Crystal Structures, Hydrogen Bond Synthons, And Melting Point Alternation, *Cryst Growth Des*, 2003, **3**, 783–790.

30 F. Neese, The ORCA Program System, *Wiley Interdiscip Rev Comput Mol Sci*, 2012, **2**, 73–78.

31 F. Neese, F. Wennmohs, U. Becker and C. Riplinger, The ORCA Quantum Chemistry Program Package, *Journal of Chemical Physics*, 2020, **152**, 224108.

32 A. D. Becke, Density-Functional Thermochemistry 3: The Role Of Exact Exchange *Journal of Chemical Physics*, 1993, **98**, 5648–5652.

33 A. Schäfer, C. Huber and R. Ahlrichs, Fully Optimized Contracted Gaussian Basis Sets Of Triple Zeta Valence Quality For Atoms Li To Kr, *J Chem Phys*, 1994, **100**, 5829–5835.

34 F. Neese, F. Wennmohs, A. Hansen and U. Becker, Efficient, Approximate And Parallel Hartree-Fock And Hybrid DFT Calculations. A 'Chain-Of-Spheres' Algorithm For The Hartree-Fock Exchange, *Chem Phys*, 2009, **356**, 98–109.

35 E. Runge and E. K. U. Gross, Density-Functional Theory for Time-Dependent Systems, *Physical Review Letters,* 1984, vol. 52.

36 S. Hirata and M. Head-Gordon, Time-Dependent Density Functional Theory Within The Tamm–Dancoff Approximation, *Chem Phys Lett*, 1999, **314**, 291–299.

37 N. Fairley and A. Carrick, *The Casa Cookbook: Recipes for XPS data processing*, Acolyte Science, 2012

38 G. Held, F. Venturini, D. C. Grinter, P. Ferrer, R. Arrigo, L. Deacon, W. Q. Garzon, K. Roy, A. Large, C. Stephens, A. Watts, P. Larkin, M. Hand, H. Wang, L. Pratt, J. J. Mudd, T. Richardson, S. Patel, M. Hillman and S. Scott, Ambient-pressure endstation of the Versatile Soft X-ray (VerSoX) beamline at Diamond Light Source, *J Synchrotron Radiat*, 2020, **27**, 1153–1166.

39 O. Schwarzkopf, M. Borchert, F. Eggenstein, U. Flechsig, C. Kalus, H. Lammert, U. Menthel, M. Pietsch, G. Reichardt, P. Rotter, F. Senf, T. Zeschke and W. B. Peatman, The BESSY Constant Length Rowland Circle Monochromator, *Journal of Electron Spectroscopy and Related Phenomena*, 1999, **101**, 997-1001

40 M. R. Weiss, R. Follath, K. J. S. Sawhney and T. Zeschke, Absolute Energy Calibration For Plane Grating Monochromators, *Nuclear Instruments and Methods in Physics Research A*, 2001, **467**, 482-484.

41 M. J. G. Peach, P. Benfield, T. Helgaker and D. J. Tozer, Excitation Energies In Density Functional Theory: An Evaluation And A Diagnostic Test, *Journal of Chemical Physics*, 2008, **128**, 44118.

# 6 Discussion, Additional Results and Analysis

This section will discuss the four papers and their results and implications in further detail to combine the findings into a coherent and cohesive story demonstrating the consistency with which core level spectroscopy combined with density functional theory and other theoretical approaches is able to refine crystal structures, in addition to the deeper understanding of the underlying interactions occurring in systems where more than a single conformation is present without the difficulties of techniques such as the gold standard approach of neutron diffraction. [1]

## 6.1 Overview of the papers

Four papers have been written, with the first two peer reviewed and published.[2,3] The first, published in September 2021, was on the topic of salt co-crystal characterisation using XPS to demonstrate the universality of the effect on core level binding energy, in addition to highlighting the consistency of our binding energy scale calibration technique.[3] The second paper, published in May 2022, built on the topic of core level spectroscopy for salt co-crystal characterisation by using NEXAFS to characterise samples on the salt co-crystal continuum. This study also combined the use of DFT to confirm the crystal structures found in literature compared with theoretical and experimental NEXAFS spectra.[2] The third paper, submitted for peer review, focuses on the theoretical background and a new approach to modelling the universal effect of proton position on the core level binding energy. This shows that we can successfully describe the effect through the use of classical electrostatics applied to the core level electron of the proton acceptor, as we probe in XPS measurements. A fourth paper is currently in draft and combines the results of the published papers to show how the use of core level spectroscopies including XPS and NEXAFS combined with DFT can be used to refine the proton positions within an unknown or unrefined crystal structure for which the atomic positions (except hydrogen) are known from XRD experiments.

Combined, these four pieces of work demonstrate the potential benefits to the pharmaceutical industry of utilising these alternative approaches to crystal structure characterisation with sufficiently accurate structures determined with rapid measurements and simple calculations. Each paper builds on the previous data and samples investigated and the end result is a demonstration of the potential uses of the technique as hypothesised at the beginning of the research project.

### 6.1.1 Proton Transfer on the Edge of the Salt/Co-crystal Continuum: X-Ray Photoelectron Spectroscopy of Three Isonicotinamide Salts

The intention of the Crystal Growth & Design paper (chapter 2) was to communicate two key messages. Firstly, that our method of consistent binding energy scale calibration using adventitious carbon accurately determines absolute binding energies. Over a large range of samples and studies using multiple spectrometers by people in the research group, we have consistently found that our binding energies following energy scale calibration result in the expected binding energies compared to literature. The purpose of this was to provide a counter argument following a series of recently published papers claiming that the approach using adventitious carbon was inconclusive and subject to large errors, making any quantitative analysis meaningless.[4,5] Whilst there are a number of published papers where careless binding energy scale calibration using the adventitious carbon component has resulted in incorrect peak assignment, chapter 2 sets out our approach in a step by step fashion to clarify that the issues seen in various literature are through an incorrect or inconsistent binding energy scale calibration and not a fundamental issue with the approach. A key feature of this approach is ensuring consistency between all parts of the XPS measurement, including the survey spectrum, high resolution scans and the fits of each of the absorption edges. Without this holistic approach, it is possible for inconsistencies to creep in, and any such discrepancies need explaining when published. An important feature of XPS, which was highlighted through comparison on XPS data taken experimentally

compared with literature on the same sample, is the ability of Gaussian line fitting to result in multiple good fits with totally different conclusions and resulting binding energies. In the case identified, there were two fits which were possible, ours was a single peak defining all of the nitrogen environments in the sample (isonicotinamide), while the literature had defined this as two peaks.[6] While there is no physical reason it should not be represented as two peaks, as there are two distinct nitrogen environments, the peak fits proposed did not match the consistent approach we use as outlined in chapter 2. In the published data, these two peaks were of significantly different width, with no explanation as to why this would be the case, and no obvious explanation, while a single peak fit to the data results in a very good fit as shown in figure 2.3(d). We are therefore confident in our approach to XPS data analysis, calibration and peak fitting and have demonstrated over a range of experimental data that the correct absolute binding energies can be determined.[3,7–13]

Secondly, and of more relevance to the title of the PhD project, chapter 2 includes an initial investigation into the use of XPS for the characterisation of hydrogen bonding characteristics based on the proton acceptor binding energy. There is a growing pool of experimental XPS data from the group and in the literature more widely identifying shifts in the proton acceptor binding energy and relating this directly to the hydrogen bonding interaction present in the sample and its characterisation as either a salt or co-crystal. In figures 2.5 and 2.6 we demonstrate that over a large number of samples measured, there is complete consistency in how XPS characterises salts and co-crystals, with a much clearer distinction between the two states than the $\Delta pK_a$ approach simply by interpreting the binding energy of the nitrogen acceptor atom. We observe a + 2 eV shift in the binding energy of the nitrogen acceptor between co-crystal and salt structures, which is well within the margins of error for XPS which are less than ±0.1 eV for a high quality measurement with correct calibration. Throughout, we demonstrate how using our approach we end up with consistency in our binding energies of equivalent groups in a range of samples, see figures 2.2, 2.3 and 2.4, thereby highlighting our confidence in the calibration technique applied. A key conclusion from this aspect of chapter 2 is ensuring a good fit to the data to allow correct calibration using the adventitious

carbon peak. Many of the issues highlighted in the recent papers are due to this full procedure not being consistently followed, with either the amount of expected adventitious carbon not explicitly determined or incorrect and illogical fitting of the adventitious carbon and other peaks with inconsistent peak widths (without explanation). Using this approach we have therefore suggested an alternative fit for the isonicotinamide N k edge which meets the criteria for consistency. We have also demonstrated the ability of XPS in the field of crystal structure characterisation, both through the new experimental data and comparisons to literature XPS binding energies, to confirm that the approach of using core level spectroscopy as a complementary technique when characterising crystal structures is a useful additional constraint on locating proton positions. Following our consistent approach to the analysis procedure, we showed how we can observe very small changes in the core level binding energy of proton acceptor groups to identify the hydrogen bonding interaction within the salt co-crystal continuum, including in an intermediate case. This ability to confirm structural properties in the bulk sample was highlighted and discussed as to its future uses potentially in pharmaceutical characterisation. In each of the three example samples, we expected to find a salt interaction, with the third sample also exhibiting a second distinct hydrogen bonding interaction in half of the hydrogen bonds, which is characterised as a neutral co-crystal interaction. This was confirmed with binding energies indicating full proton transfer in the three cases and no proton transfer in the fourth. By using three samples with known hydrogen bonding interactions, this was a useful test and proof of principle experiment showing how the analysis procedure worked and how dependent the core level binding energy is on the hydrogen bonding interaction present adjacent to the nitrogen acceptor.

### 6.1.2 Determination of H-Atom Positions in Organic Crystal Structures by NEXAFS Combined with Density Functional Theory: a Study of Two-Component Systems Containing Isonicotinamide.

The second paper (chapter 3) published investigated the same samples as the first, but using the complementary technique of NEXAFS, which also probes core level states, although additionally through excitations of electrons to unoccupied valence molecular orbitals, making the spectrum dependent on both the core level and valence level, resulting in an even more structure dependent spectrum. In this complementary second paper, the intention was to demonstrate how NEXAFS can equally be used in a similar approach to XPS for the characterisation of hydrogen bonding interactions in organic crystals, and further our understanding of these interactions, in addition to then using the method to confirm the positioning of protons in the literature structures. For the experimental section of this work, the new NAP-NEXAFS high throughput endstation at beamline B07-B at Diamond was utilised in what was the first user commissioning experiment.[14] Therefore, there was a need to develop a data processing workflow including a calibration to remove any effects of the gas phase. This was required in particular as the background measurement (for background subtraction) was a measurement of the gas phase helium with a different measurement path length to when the sample was present. The steps taken to remove any dependence on that different path length are included in the supplementary information (appendix B), where we described our method of consistent analysis, equivalent to the XPS procedure outlined in chapter 2 to ensure consistency in all aspects. The purpose of chapter 3 was therefore to demonstrate as a proof of principle that NEXAFS can also be used to determine the nature of hydrogen bonding interactions, with an additional level of sensitivity to the proton position such that the proton position can be determined through comparison with DFT calculations. As a proof of principle experiment, we deliberately decided to use the same samples as previously investigated using XPS, and we therefore anticipated that the NEXAFS and DFT results would all be fully consistent with literature and structures found in the

Cambridge Structural Database. Experimental NEXAFS exhibited the expected behaviour, with the clear shift in photon energies equivalent to XPS observed in the salt structures compared to co-crystal. In the first two samples which form simple salt structures, DFT calculations simulating the NEXAFS spectra agreed well with experimental data, using the structure found in the CSD. By this point, we have now demonstrated consistency across three different analytical techniques, with results from XPS, NEXAFS and DFT all concluding the same structure as predicted from XRD measurements done previously. This gives us confidence that these techniques are able to describe the hydrogen bonding interaction based on the core level binding energy of the proton acceptor group.

However, to identify the level with which NEXAFS is sensitive to the precise proton location, the third sample was required which showed a slight discrepancy between the experimental and simulated NEXAFS spectra, with differences in the relative intensities of the 1s $\rightarrow \pi^*$ peaks, such that we were could not explain these discrepancies without further investigation as to the cause of the discrepancy using the structure found in the Cambridge Structural Database, through an investigation of the DFT calculation. To ensure this effect was not simply a deficiency in the DFT method used, a series of equivalent calculations using alternative exchange correlation functionals were completed. These all resulted in similar discrepancies, although also introduced further issues compared to experimental data, thus confirming the suitability of our choice of functional. We then conducted a series of tests to isolate the parameter affecting the simulated spectrum. This firstly involved visualising the relevant molecular orbitals involved in the excitations where the energies were offset from experiment. By completing this analysis of the core level molecular orbitals calculated during the simulation, it was possible to determine that the issue lay not with the hydrogen bonded nitrogen, but with the peripheral amide group. Two options were investigated to resolve the issue: 1. Change in the proton position of these peripheral amide groups to increase the energy or 2. Model a larger system containing a number of unit cells to ensure all intermolecular interactions are accounted for. In chapter 3 each of these results are shown, with both possibilities giving appropriately improved simulated spectra. The difficulty then lies in determining the correct solution, as discussed

in chapter 3. In addition to the 1s → π* transitions which we focus on, a comparison of the higher energy 1s → σ* resonances can be used to refine the correct interpretation. Figures 3.5 and 3.6 highlight how these σ* resonances have been used to distinguish between two very similar 1s → π* transition peaks through comparison of the σ* resonance with experiment. All of these additional analyses which are available with the NEXAFS technique allow for this much more robust and complete analysis and gives the ability to question the legitimacy of the structures in cases where the literature molecular structure has not fully refined the positions of the protons. Therefore, while both XPS and NEXAFS techniques can be confidently used to characterise systems as salt or co-crystal, the additional dependencies of the NEXAFS technique make it even more sensitive to the position of the proton in these systems, whether the proton is involved in hydrogen bonding or not, resulting in the ability to identify the proton positions through comparison with DFT calculated spectra.

### 6.1.3 Classical Electrostatics Predicts the Core Level Binding Energy Shifts at Brønsted Acceptor Sites

The third paper (chapter 4) discusses the causes of the binding energy shift observed in XPS and outlines a simple model capable of describing the observed behaviour using basic electrostatics. The intention was to identify the primary cause of this binding energy shift and attempt to model the effect. The basis for the model was founded on various literature claiming that hydrogen bonding is primarily an electrostatic interaction between the proton donor, the proton and the electronegative proton acceptor. While this is a relatively commonly stated fact, there are no simple models in which pure electrostatics can model the bulk of the behaviour observed in hydrogen bonds and little evidence to prove that electrostatic interactions could describe the behaviour we observe in these two component hydrogen bond interactions including the 2 eV shift in binding energies between salt and co-crystal as demonstrated in chapter 2. In chapter 4, we demonstrate how this simple model not only models this 2 eV shift, suggesting it originates primarily from the electrostatic interaction between the proton and proton acceptor, but also predicts binding energies for

the proton acceptor based upon the electrostatics of the simple system consisting of the proton donor, acceptor and proton plus the relevant electrons to within 1 eV. All other interactions are assumed to remain constant and an optimised constant is included in the calculation based on a range of test structures minimising the error. Over a range of structures for which there is experimental XPS data, this produced a mean error of 0.3 eV, which is less than an order of magnitude greater than the expected error in the best XPS measurements (0.1 eV) demonstrating clearly the electrostatic dependence of the system. Although electrostatics clearly describes the largest component of the system well, there are clearly other effects outside the scope of the model. The most obvious of these is the constant used to describe all the external electrostatic interactions with the system from the surrounding atoms and molecules. As the distance from the hydrogen bonding system increases, the effect of the electrostatic force decreases with a $1/r^2$ dependence. Therefore, only the immediate neighbouring atoms to the system, such as the carboxylic acid group which the oxygen donor is part of and a nitrogen substituted ring acting as the acceptor, will have any noticeable effect on the binding energy of the core level electron. This background electrostatic interaction then forms the basis of the constant term in the model, and will be essentially the same for all systems. The close neighbouring atoms then make up the remainder of the constant term. This part will vary depending on the make-up of the molecules present, with a highly electronegative component close by could affect the binding energy, and we observed this throughout our comparisons with experimental data. Table 6.1 shows some of the errors calculated alongside the nearest group to highlight the effect we observe. It is clear that where highly electronegative groups are present, such as in 3,5-dinitrobenzoic acid and 5-sulphosalicylic acid, the electrostatic model underestimates the binding energy, whereas samples with a strong base, such as theophylline, coupled with a weak acid are well described. Samples containing the weak base 4,4’-bipyridine result in an overestimate of the binding energy when coupled with a weak acid. This highlights the consistency of these errors in the electrostatic model, indicating that it is not the underlying electrostatic principles causing the errors calculated, but the lack of any dependence on the surroundings where highly

electronegative, or a lack of electronegative groups are in the structure close to the hydrogen bond.

**Table 6.1** Comparison of errors from electrostatic model grouped by closest neighbour group. This shows that the error observed in the electrostatic model is not random and there is a clear trend between the neighbouring group and the error observed.

| **Dominant group** | **Mean error in the Electrostatic model Binding Energy / eV** |
|---|---|
| 3,5-dinitrobenzoic acid | -0.43 |
| 5-sulphosalicylic acid | -0.28 |
| Theophylline | 0.05 |
| 4,4’-bipyridine | 0.43 |

Other causes of error, although likely a smaller magnitude of error than the approximation of the surrounding atoms, include self-interaction errors between electrons, dipole-dipole/van der Waals forces and any non-classical quantum effects due to relativistic electrons. However, in almost all cases, these will be a constant for a particular structure, independent of the proton position and therefore also form part of the constant term. The fact that there is such a small mean error despite all of these potential causes of error isolates the electrostatic effect as the overriding cause of the behaviour we observe in XPS and NEXAFS by investigating the core level electrons. This development of a model to understand the underlying interactions occurring in hydrogen bonding interactions highlights the usefulness of simple models in furthering our understanding of complex systems. A similar model would be non-trivial for NEXAFS modelling, due to the dependence on both the core level (which we have shown can be modelled reasonably well using electrostatics) and the unoccupied molecular orbitals which would be significantly more challenging to model with electrostatics due to the large number of influences on the system and non-localised molecular orbitals. For this reason we have not attempted this

and allowed DFT calculations to provide the theoretical modelling for comparison to experimental data. As discussed in the introduction, modelling of XPS with DFT is slightly more difficult and the additional theoretical approach improves the confidence in any DFT modelling we carry out.

For the comparison of the electrostatic model with the equivalent DFT, the following calculations were carried out using a test molecule of isonicotinamide 3,5 dinitrobenzoic acid. In order to model an equivalent “electrostatic curve” as determined from the electrostatic model, a series of calculations where only the proton position moved were carried out, providing a scan over the same range as the electrostatic model from typical salt to co-crystal, the aim being to model the equivalent situation to determine any effects in the energies calculated which are not picked up through classical electrostatics. The key aspect of this calculation was the exclusion of any neighbouring atom effects by restricting any change of atom coordinates to only the proton. While this results in a series of structures which are not at the energy minimum, since the test structure used is a salt at room temperature, the purpose of the calculation was to identify any differences between the calculation and the electrostatic model. This involved only modelling the change in the core level binding energy (taken as the core level molecular orbital energy) when altering only the position of the proton. Therefore, in a similar way to the electrostatic model, the interactions with neighbouring atoms remain constant (at least in terms of first order interactions) resulting in a direct comparison and ignoring the wider molecular geometry. Prior to choosing this particular structure for the DFT calculations, a number of test calculations were completed on structures which form salts and co-crystals at room temperature to ensure a suitable structure was chosen. In all cases, the shape of the curve remained the same. As discussed, in the region between and including salts and co-crystals, both theoretical approaches result in a very similar curve with equivalent difference in binding energy from salt to co-crystal. Outside of this region, we observe some differences, particularly where the proton moves particularly close to the proton acceptor and other effects dominate the system, resulting in an abrupt end to the smooth $1/r^2$ curve. This demonstrates that in hydrogen bonding interactions, the primary source of the

interaction energies are from electrostatics, but where the distances reduce such that a covalent bond is formed, other factors take over.

Short strong hydrogen bonds, or charge assisted hydrogen bonds, are often referred to as exhibiting partial covalent character, such that there is a partial formation of covalent bonds between the proton and it's donor and/or the acceptor. In addition to the above test involving a selection of salt and co-crystal structures to ensure there was no significant difference in the shape of the potential energy curve calculated using DFT, one structure containing a very short donor acceptor and roughly centred proton was modelled. Where this partial formation of covalent bonds is observed and the donor acceptor distance is sufficiently short such that this partial covalent bonding region is reached, we observe that the levelling off of the binding energy occurs at a lower energy, indicating that the formation of a covalent bond occurs at lower energy, potentially explaining the partial covalent character observed even when the proton is in the centred position. Further work needs to be done to determine the reason that a covalent bond would result in this flattening and how this affects the electrostatics of the system when a covalent bond is formed.

### 6.1.4 Core Level Spectroscopy to Determine H-atom Positions in Organic Crystal Structures Combined With Density Functional Theory: Short Strong Hydrogen Bonds and Proton Disorder

The intention of the fourth paper (chapter 5) was to showcase the use of NEXAFS, XPS and DFT in crystal structure refinement and to highlight the advantages over traditional methods on a set of structures chosen to challenge the approach in a culmination of the three previous investigations. The basis for chapter 5 was to identify some crystal structures from the CSD where the proton positions had not been fully refined and as such left a certain level of uncertainty in their position. In addition, the structures were also constrained to any displaying characteristics of Short Strong Hydrogen Bonding, with donor acceptor distances of less than 2.6 Å. These constraints were chosen to identify a set of structures to follow on from the SSHB isonicotinamide complexes on which we tested the approach. In addition, this set of structures was chosen

deliberately to challenge the approach of core level spectroscopy on some less common systems with unknown and non-standard hydrogen bonding interactions in an attempt to further our understanding of hydrogen bonding interactions and their effect on crystal structural properties. Five structures were identified from this search of the database and were prepared in house using the method of evaporation from solution, with the appropriate components and solvents as described in the literature, referenced in chapter 5. The solutions were left to evaporate for a week, following which the sample crystals had formed. Following this, the samples were measured using XRD to obtain the unit cell and ensure our product was of the crystals intended through comparison with the unit cells from the CSD. In all five cases, the crystals had formed as expected into the following unit cells as described in literature.

This then gave an expected stoichiometry, and a structure on which to perform DFT calculations with known atomic positions, except for the hydrogen atoms. Initially, the structure in the database was used to perform a baseline calculation to determine the expected NEXAFS spectra and binding energies equivalent to XPS peaks based upon the structure found in the database. In four of the five samples, this was found to insufficiently describe the structure through comparison with experimental data such that there were significant differences in peak positions and intensities. As shown, a series of optimisations were then completed to refine the structure to produce spectra which are consistent with both the NEXAFS and XPS experimental data. Similarly to the outcomes presented in chapters 2 and 3, we are confident in this optimisation process determining the structure due to the consistency between multiple methods. Taking the starting structure from XRD measurements from the CSD, we are already confident of the non-H positions and as such leave those as found. Therefore, the only changes made are to the proton position, actually maintaining additional consistency with the accurate components of the XRD measurements too. Further to this, the sensitivity of the NEXAFS resonance energies were highlighted through visible shifts of 2 eV between co-crystal and salt, a difference in N--H distance of only up to 0.8 Å. In all cases, the same approach was taken to ensure all possibilities were explored. Firstly, an orbital analysis was carried out to identify the relevant

molecular orbitals, particularly for those transitions where the CSD structure does not fully agree with the experimental data. The key orbitals to probe are the core level orbitals (1s) as these allow us to identify the relevant atom causing any discrepancies in the spectrum. In all four cases, it was the hydrogen bonded nitrogen causing the issue, and therefore the analysis shown in the chapter 5 was carried out. This is in contrast with the situation in chapter 3 where the issue was caused by the energy positioning of the non-hydrogen bonded amide hydrogen and its interactions with surrounding molecules. Using the experimental data as the constraint, a series of calculations are run over a range of possible structural changes to identify the requirements to resemble the experimental data. Once this has been narrowed down, an in depth analysis to identify the correct situation is completed. This perfect structure is optimised to both the NEXAFS and XPS experimental data, both of which change independently in the DFT calculation due to the NEXAFS dependence on additional energy levels. This means there are two degrees of freedom on which to optimise the structure, leading to only one optimal structure in each case.

#### 6.1.4.1 Applying the Electrostatics model to these additional systems investigated.

In addition to the use of NEXAFS, XPS and DFT to refine the structures of the five organic co-crystals investigated as shown in chapter 5, we also compare the refined structures (obtained through our DFT optimisation process) to the experimental XPS data through the use of the electrostatic model to confirm the placement of protons with respect to the binding energies observed. In each case, this is again in remarkably good agreement with experiment. The results are shown in table 6.2 below.

**Table 6.2** The electrostatic model applied to the five systems in chapter 5. Errors are in line with those found in the original work with a mean of 0.15 eV.

| Sample | Experimental Binding Energy / eV | Electrostatics Binding energy / eV |
|---|---|---|
| 44BPYSA | 399.4 | 400.2 |
| INHFA | 399.9 | 399.7 |
| INOA | 399.3 & 400.9 | 400.1 & 401.2 |
| 4EP35DNBA | 400.0 & 401.3 | 399.8 & 400.8 |
| 35L35DNBA | 399.4 & 401.0 | 399.8 & 400.7 |

In each case, the errors determined are within the ranges previously identified for the larger set of structures, with a mean error of 0.15 eV and a maximum error of 0.8 eV. This error is sufficiently small to consistently distinguish between salt and co-crystal. Similarly to the previous electrostatic data, the errors clearly depend on the electronegativity of the components, with binding energies underestimated for those with a high electronegativity. This provides some additional evidence and adds to the consistency of the results from the approach taken. This ability to maintain consistency across a range of techniques taking data from a variety of sources enhances the confidence with which the proton position has been accurately refined.

Along with the accuracy of the experimental techniques, we have shown how theory can result in a similar level of accuracy, even when only taking a basic level of theory in account. Over a range of structures for which we have XPS data, the electrostatics model resulted in a mean error of less than 0.5 eV, which is a very small margin of error for a simple electrostatics problem in the real world, and ignoring any quantum effects or external crystal structural effects, which will cause a shift in cases where highly electronegative groups are near the region of the hydrogen bond. These components to the binding

energy should be relatively small, but the small level of error suggests that actually they are vanishingly small in most crystals. While we came to this conclusion in chapter 4, it is very clear from this analysis of the data that the overwhelming effect at play on the hydrogen bonding interaction (at least the difference due to the proton position) is entirely electrostatic.

#### 6.1.4.2 Checking for consistency between XPS and NEXAFS for the isonicotinamide samples investigated in Chapters 2 and 3

In a further test and confirmation of the refinement procedure utilised in the chapter 5, we now apply the same approach to the isonicotinamide systems investigated previously. In the two separate investigations, both XPS and NEXAFS were used to characterise the structure and confirm the nature of the hydrogen bonding interaction, however the two techniques were not combined in the same way to reach a conclusion as to the correct interpretation, particularly for the isonicotinamide phthalic acid complex for which there were still two possible explanations for the differences in the NEXAFS.

**Table 6.3** Comparison of DFT calculated binding energies with experimental XPS for the three isonicotinamide complexes INPA, IN35DNBA and IN24DNBA. Agreement of binding energies is good, except for an underestimate of the nitro group binding energy by 2 eV.

| Sample | Experimental Binding Energies / eV | DFT Calculated Binding Energies / eV |
|---|---|---|
| IN24DNBA | 399.5, 401.2, 406.2 | 399.6, 401.1, 404.2 |
| IN35DNBA | 399.8, 401.1, 406.3 | 399.7, 401.2, 404.0 |
| INPA (amide proton) | 399.9, 401.6 | 399.8, 401.6 |
| INPA (cluster model) | 399.9, 401.6 | 399.8, 404.4 |

In chapter 3, we have already achieved a level of consistency between the NEXAFS data and the DFT calculation, so it is the additional consistency with

XPS that is necessary here. Table 6.3 shows that both IN24DNBA and IN35DNBA agree well with the experimental XPS, with the $NO_2$ binding energy underestimated in each case. This confirms that the DFT calculation represents the hydrogen bonding interaction well by agreeing with both experimental NEXAFS and XPS and we can be confident that the crystal structure is that found in the database. However, for INPA two different molecular structures resulted in correct NEXAFS spectra being simulated, and this highlights the importance of additionally constraining to XPS experimental data. In the table 6.3, it becomes very clear which molecular structure more accurately represents the experimental XPS data, with the change in hydrogen position on the amide group significantly improving the calculated binding energies over the cluster model, where the proton transfer binding energy is significantly higher than observed experimentally.

### 6.1.5 Accuracy of XPS and NEXAFS

One large advantage of the core level spectroscopy techniques is the relatively small errors during measurement. It is generally accepted that a Binding Energy from XPS can be accurate to the nearest 0.1 eV, and our data is consistent with this, reproducing binding energies to within 0.1 eV. This is why we can identify the different hydrogen bonding characteristics through the relatively small 2 eV shift in core level binding energy. Similarly, NEXAFS spectra taken on the near ambient pressure endstation are also accurate to within 0.1 eV, giving two similar, but sufficiently different as regards the information obtainable, techniques. In both cases, the accuracies stated are only achieved through the analysis procedures outlined in the chapters 2 and 3 and ensuring any inconsistencies are explored and corrected or explained before finalising the analysis. As discussed previously, there has recently been some debate in the literature regarding the accuracy of binding energy scale calibration in non-conducting samples; conducting samples are easy to calibrate correctly through identification of the Fermi level. However, throughout we have shown that we can accurately perform this calibration using the adventitious carbon contribution to the C 1s emission line. Nevertheless, throughout the work

carried out, there were multiple occasions in the analysis of the results where consistency could not be achieved due to issues with experimental data which are discussed below.

#### 6.1.5.1 Experimental issues with UHV-XPS

During the measurement of the samples, initially using the UHV-XPS instrument, we encountered problems with the charge compensation of the samples, with a significant shoulder peak appearing at low binding energy at each emission line. The main cause of inconsistent spectra was due to an issue with our non-conducting samples in the UHV-XPS spectrometer such that the charge compensation was insufficient and resulted in distorted peak shapes, making peak fitting difficult. Initially, this effect was not necessarily obvious, as carbon 1s emission lines contain a variety of components meaning overall peak shapes can vary. However, in these cases, the key feature which caused the inconsistency was in the fitting of the adventitious carbon component, for which the intensity is calculated from the survey spectrum. In almost all organic crystals, this should form part of the lowest binding energy component, as it is not bonded to the sample. However, in these cases, a shoulder peak at low binding energy meant that this could not be fit with the appropriate intensity, suggesting an issue. This was identified as an issue due to the requirement for consistency between the survey spectrum (in which this effect would not be visible) and the C 1s emission line. The cause of this effect is uneven charging of the sample surface resulting in differential charging. This means that different areas of the sample within the beam boundaries charge at different rates, causing a range of shifts in the binding energies determine due to the charges. In the cases we identified, the cause was found to be the flood gun used on the UHV-XPS instrument, and remeasurement using the NAP-XPS instrument solved all of the charging issues. In these cases, the importance of the consistent analysis is clear, as it allows this poor quality data to be picked up and discarded, increasing the reliability of data analysed in this way as it should have filtered out any poor data like this. Additionally, a forced fit to the distorted data was possible and would lead to incorrect binding energy scale calibration,

with potential issues with analysis of those binding energies. Following the discovery of these differential charging effects, the samples were remeasured using the NAP-XPS instrument where charge compensation is provided by the ionised gas phase which appears to result in better charge compensation for these samples.

#### 6.1.5.2 Beam Damage and Vacuum conditions

We have also experienced the effect of X-ray exposure in both NEXAFS and XPS on samples through beam damage and found that certain types of sample are particularly vulnerable to beam damage, which result in a sharp nitrogen free radical peak at around 404.5 eV in samples containing nitrogen, as shown in the supporting information to chapter 2 (appendix A). This occurs where a sample is exposed for a long time at high energy. Therefore this was not observed in NEXAFS to the same extent as XPS, due to the lower photon energies used. However, utilising these measurements, we have carried out a full analysis by accounting for this additional peak in the elemental analysis used to determine the elemental composition and the same results were achieved as from the follow up measurement without beam damage features. Following remeasurement, without beam damage, analysis was completed on both sets of results and it was concluded that it is still possible to conduct a fully consistent analysis with beam damaged data, simply by identifying and excluding the relevant peaks. While beam damaged data is not publishable, this does demonstrate how consistent core level spectroscopy is, even where the experimental conditions do not fully suit the sample. This highlights again a key feature of our analysis methods resulting in the reproducibility of any results and this requires a thorough and consistent analysis procedure. This highlights the flexibility of the approach and also the universality of the effects meaning they are distinguishable from contaminants and beam damage.

### 6.1.5.3 Commissioning of the new B07-B high throughput NEXAFS endstation

For the measurement of our NEXAFS spectra, the new high throughput endstation at beamline B07-B, Diamond Light Source was used. Since our experiment was the first user commissioning experiment to take place, proving the accuracy and consistency of the measured spectra was an important aspect of the beamtime. In order to ensure accurate spectra were being recorded, a series of NEXAFS measurements were completed on a set of amino acids. These were chosen as part of the commissioning beamtime as a set of samples which are well known with experimental spectra in the literature from other instruments. The initial purpose of the measurements was: (1) to ensure the new NEXAFS sample chamber and measurement worked as expected and (2) that the normalisation procedure allowed us to reproduce the expected peak shapes, positions and intensities accurately compared to known experimental data from other instruments.

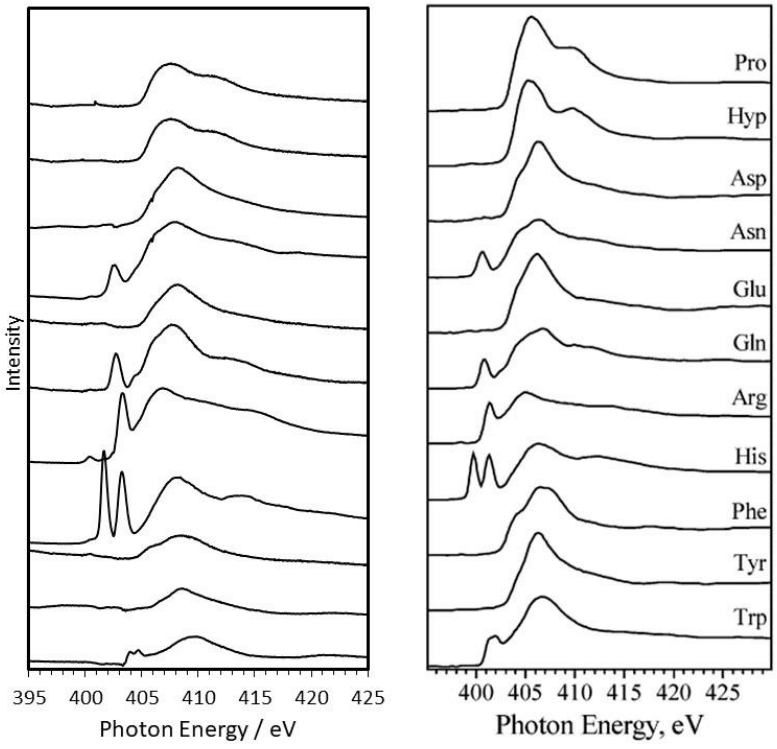



**Figure 6.1** N k-edge NEXAFS spectra of half of the amino acid samples. Our experimental data (left) compared to a literature set of data[15] (right) exhibiting excellent agreement in peak position and shape.

As can be seen in figure 6.1 for a subset of the samples, the experimental spectra are essentially indistinguishable from the literature, and this was the case for all 20 samples. The normalisation and calibration procedure is completed in two parts, firstly the photon energy scale is calibrated. This is done using a known standard, measured at a specific absorption edge to identify the offset of the photon energy from the true value. This standard is measured several times throughout the experiment to ensure it does not change. Once the photon energy scale is calibrated, the intensities must be normalised against the background He gas spectrum, which otherwise distorts peak intensities due to its varying effect over the range of photon energies and the different path length of the X-ray beam in the sample chamber when measuring the background. Along with the consistency with which we can analyse the data, the agreement between experimental and literature spectra for the set of amino acids confirms the accuracy of the new endstation and our NEXAFS data.

### 6.1.6 Accuracy and Reliability of DFT calculations

Density Functional theory has been utilised as an additional constraint on the experimental data as described in the four papers. To ensure this constraint is realistic and as accurate as possible, some optimization tests were completed on the DFT calculations to ensure reasonable structures and results were obtained over a range of different structures and calculation types.

As with the DFT calculations used throughout the experimental analysis, this additional check on the results obtained through the computational approach further highlights the power of first principles calculations in the field of crystallography and crystal structure analysis. As discussed in each of the papers, the decision regarding the parameters of such a first principle calculation are vitally important to ensure maximum reliability and reproducibility for comparison with experimental data. Throughout, a range of exchange correlation functionals and basis sets were trialled to ensure the best matches to experimental data were obtained. It was anticipated that a different exchange correlation functional and basis set may be required for TDDFT calculations of excitation energies from the core level orbitals to unoccupied valence molecular

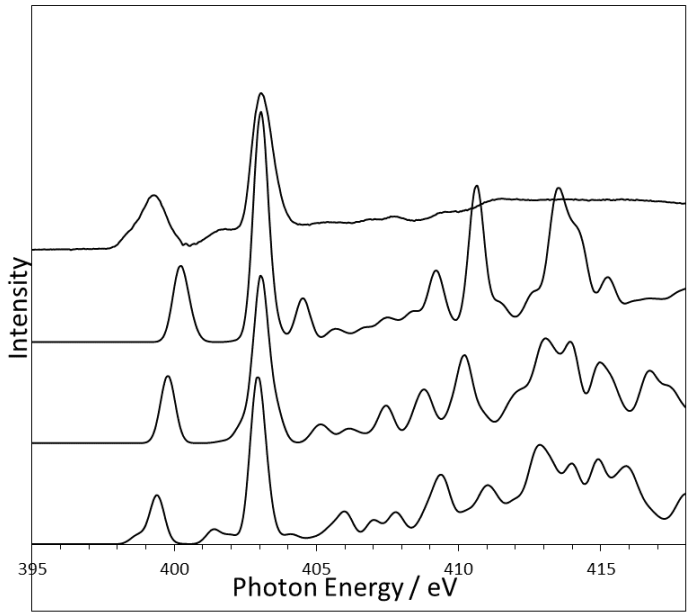
orbitals compared to those required for single point energy calculations in the ground state required in the calculation of XPS.

#### 6.1.6.1 Choice of Exchange Correlation Functional and Basis set

The B3LYP[16] functional is one of the most widely used in quantum chemistry and works in a wide variety of applications including ground state and excited state properties and was therefore the first option. The reason for its applicability to a wide range of applications is the combination of Hartree Fock and DFT theories to more accurately model the exchange component of the total energy. It is this mixing of components which has allowed this and similar hybrid exchange correlation functionals to become the standard. For calculations involving excited states, range separated hybrid functionals are often used, and these are designed to assign the correct proportion of HF and DFT exchange at each interaction distance, with an increased proportion of HF exchange generally required at larger distances to correctly predict excited state properties. There is also evidence that a larger proportion is required at short length scales and therefore these range separated hybrid functionals are not generally suitable for core level calculations, since the HF exchange is at a minimum at short distance. Consequently, functionals have been designed around the use case of core level spectroscopy, specifically optimised to a set of NEXAFS experimental data, such as the Short Range Corrected Functionals (SRC-1 and SRC-2).[17] The parameters were determined through an optimisation process iteratively improving the calculated energies on a test set of excitation energies. This resulted in the two published functionals, one of which is implemented easily in all quantum chemistry packages, with a single error function similarly to range separated hybrid functionals, the second requiring two such parameters allowing the shape of the exchange correlation partitioning function (between HF and DFT exchange) to be tuned more accurately to the test set of structures, and therefore more difficult to implement computationally

The simpler of these exchange corelation functionals (SRC-1) was tested and compared to the commonly used B3LYP hybrid and CAM-B3LYP range

separated hybrid functionals using the same def2-TZVP basis set to ensure the most appropriate and accurate was used for all calculations.[18] A requirement for this was that the chosen method was as accurate across the full range of calculations such that analysis could compare between calculations in a direct fashion without the need to account for any changes in functional in addition to the uncertainties present in DFT calculations anyway. This comparison is again the key feature in enabling a consistent set of analysis across all calculations and techniques experimentally and computationally.




| Exchange Correlation Functional | C-$\underline{N}H^+$ / eV | C-$\underline{N}H_2$ / eV |
|---|---|---|
| Experimental | 399.4 | 398.5 |
| SRC-1 | 400.3 | 400.3 |
| CAM-B3LYP | 399.8 | 399.8 |
| B3LYP | 399.5 | 398.7 |

**Figure 6.2** Comparison of calculated NEXAFS spectra from three different exchange correlation functionals compared with experimental data. Photon energy scale calibrated to the $NO_2$ peak. 0.2 eV shift for SRC-1, 12.3 eV for CAM-B3LYP and 14.2 eV for B3LYP. The B3LYP functional clearly best represents the experimental data.

In each of the tests, the same conclusion was reached; that the B3LYP functional with def2-TZVP basis set consistently resulted in the most correct relative energies across a wide range of samples and calculations. As is often the case with DFT calculations, the absolute energies calculated are offset due to deficiencies in the total energy calculation, with experimental energies typically 10-15 eV below the experimental value. While the SRC-1 functional predicted correct absolute energies for a number of samples, the inaccuracies

in the relative binding energies and excitation energies makes any subtle changes difficult to identify, as shown in figure 6.2. Therefore, the choice of the B3LYP functional was made, using the most intense peak in the spectrum to calibrate the energy scale to allow the absolute binding energies to be modelled for a full comparison with experiment. This calibration removes any effect of the exchange correlation functional from the absolute energy determinations, with the DFT calculation being used to calculate the relative differences in binding and excitation energies.

These tests identified a series of interesting characteristics of the three functionals. Firstly, the B3LYP functionals appeared to perform equally well across the full range of calculations, with a consistent underestimate of energies which was accounted for through a uniform shift to the calculated energies. This is the reason that this functional was chosen. The CAM-B3LYP functional performed well for excited state properties, but less well for the core level properties of interest here, primarily due to the lack of HF exchange at the short range. The SRC-1 functional performed the best for core level properties correctly modelling the absolute values for core level binding energies in a number of cases. However this general improvement in the absolute energies calculated resulted in the relative energies of some groups to deteriorate, to the extent that it was no longer comparable to experimental data, particularly evident in figure 6.2 where the relative energies of the peaks are up to 1 eV off the experimental data. This appeared to particularly affect the regions at the edge of the molecules, with these having higher energies than observed experimentally and therefore completely changing any analysis completed, while the B3LYP functional predicted these relative energies correctly.

## 6.2 Additional complementary investigations

In addition to the main aspects of research carried out as discussed, a number of additional projects were undertaken alongside which complement the main focus of proton position refinement providing some additional data which improves our understanding of the nature of hydrogen bonding in organic crystals, particularly where short strong hydrogen bonding interactions are

present and intermediate states are observed. A classic example of this is identified in chapter 1, where a broadened XPS and NEXAFS peak is identified in 3,5-pyridine dicarboxylic acid where only a single nitrogen environment should be present. It is clear from of a visual qualitative analysis of the broadened peak that there are at least two main components contributing to the peak and that from the asymmetrical nature of the peak that there is no equivalence between the states, with the occupancies potentially dependent on experimental conditions such as temperature. Using the process of Fourier deconvolution, we have investigated the cause of this and the underlying hydrogen bonding interactions occurring.

### 6.2.1 Deconvolution of broadened XPS peaks

Some work carried out was investigating the possibility of deconvoluting broadened XPS peaks using Fourier methods. This followed on from some literature in which a peak corresponding to a single nitrogen environment was shown to broaden, with the expectation that this was caused by a variety of proton positions across the sample a Boltzmann distribution dependent upon the proton potential energy surface. A Fourier deconvolution procedure was developed to deconvolute the underlying Gaussian line shape of the emission line, the shape we would expect without any broadening, from the observed spectrum to attempt to understand the underlying cause of the broadening through what we call the “broadening function”. In theory, the experimental spectrum including the broadened peak is a convolution of the Gaussian peak shape with a broadening function specific to the sample. The aim is to deconvolute this broadening function to investigate the underlying cause of the broadening. The process involves the experimental spectrum along and an ideal Gaussian peak. To each of these, a Fourier transform is performed. In this space, the experimental spectrum becomes the product of the Gaussian and the broadening function, so it is trivial to obtain the broadening function. The broadening function in real space is then obtained by carrying out an inverse Fourier transform. In order to obtain realistic broadening functions, a cut off frequency is applied; otherwise the complete broadening function results in an

oscillating function with little physical meaning. Applying this cut off frequency does not affect the meaning of the broadening function, but any attempt to convolute this with an appropriate Gaussian function will not replicate the experimental data since the higher order frequencies have been lost.

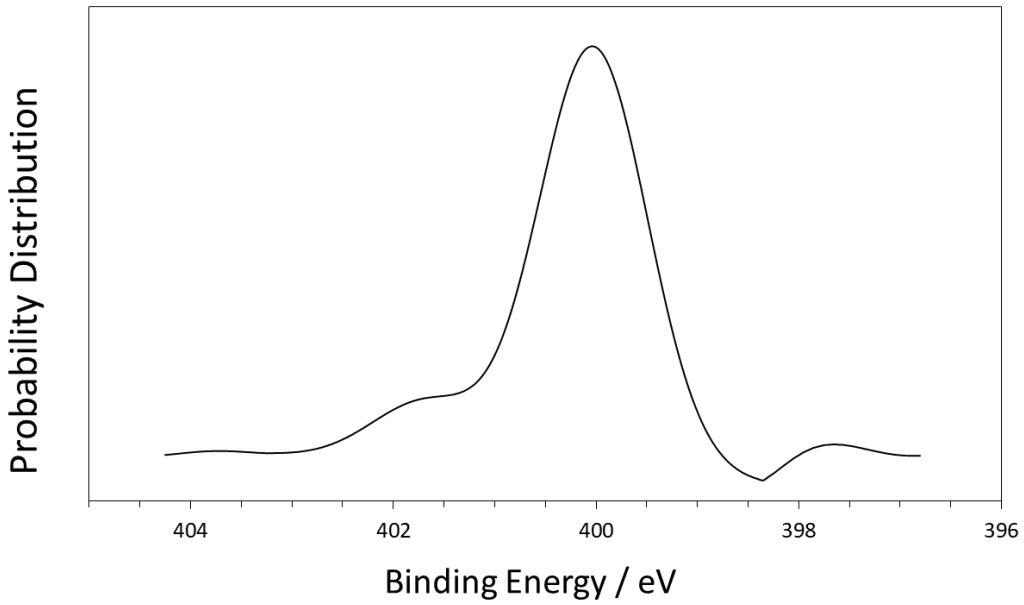



**Figure 6.3** Broadening function for broadened N 1s emission line observed in 3,5-pyridine dicarboxylic acid calculated using Fourier Deconvolution technique. This appears to relate to the population density of states.

Figure 6.3 shows a calculated broadening function, which appears to define the population density of different states, describing the population of varying proton positions within the proton potential energy surface. Using this we can observe how there is a probability that the proton resides in a large range of positions as measured by XPS. This confirms that the hydrogen bonding interaction can be classified as a very low barrier two site disordered system, with a non-zero probability of the proton residing between these two minima at room temperature. That fact that the distribution does not reach zero between the two peaks indicates that there is a continuum of possible states, some of which are Boltzmann suppressed by the potential energy surface. This additional technique of XPS analysis is further evidence of the additional information which is available through the use of core level spectroscopy, with this approach giving useful insights into the nature of hydrogen bonding in this short

strong hydrogen bonded system. A future investigation is required to identify the split in energies, and this could easily be calculated from a temperature dependent series of measurements.

### 6.2.2 Histidine angle dependent NEXAFS

During our measurements of the amino acids, an additional feature of the C k-edge spectrum of Histidine was identified such that the peak intensities varied based on the crystallite being measured, and therefore by the orientation of the crystallite. As these measurements were being carried out to ensure the accuracy and reproducibility of the new NEXAFS endstation, multiple measurements were taken of each sample at each absorption edge, allowing this effect to be observed. It is known that NEXAFS spectra are dependent on the relative orientations of the molecular orbitals in non-centrosymmetric structures, but we are not aware of this phenomenon previously being identified in structures as simple as amino acids. By measuring a series of spectra covering different crystallites, we postulated that it may be possible to reconstruct the crystallite orientation from the varying peak intensities. To investigate this we conducted a series of DFT calculations to identify the importance of orientation in L-histidine. The DFT calculated “x”, “y” and “z” components of the spectra are identified in figure 6.4 for the optimised structure of histidine.

This was optimised using DFT, with a starting structure from the crystal structure database. These components at an initial visual check appear to correspond to the three features observed in the experimental spectra. It is clear that the lower experimental spectra correspond primarily to the z orientation, with the middle peak (288 eV) displaying the doublet structure. The upper experimental spectrum appears more like the x orientation, with the intermediates then combining these two. Further analysis of the experimental and calculated data shows that we can identify the changes seen in the experimental spectra through variations in the peak intensities corresponding to the x, y and z components. From this we can identify the appropriate proportions of each component to the experimental spectra, and therefore

postulate a rough idea of the orientation(s) present in the particular area of crystal irradiated with X-rays. This effect stems from the lack of symmetry in the crystal structure of histidine making the spectra highly orientation dependent relative to the π bonding in the aromatic ring structure. Due to the lower abundance of nitrogen and oxygen in the structure, the same effect is not observed, although in theory the same approach can be applied to observe the theoretical differences based on the DFT calculations carried out. This is just another way in which the NEXAFS technique can be used to gain a greater insight into the systems investigated.

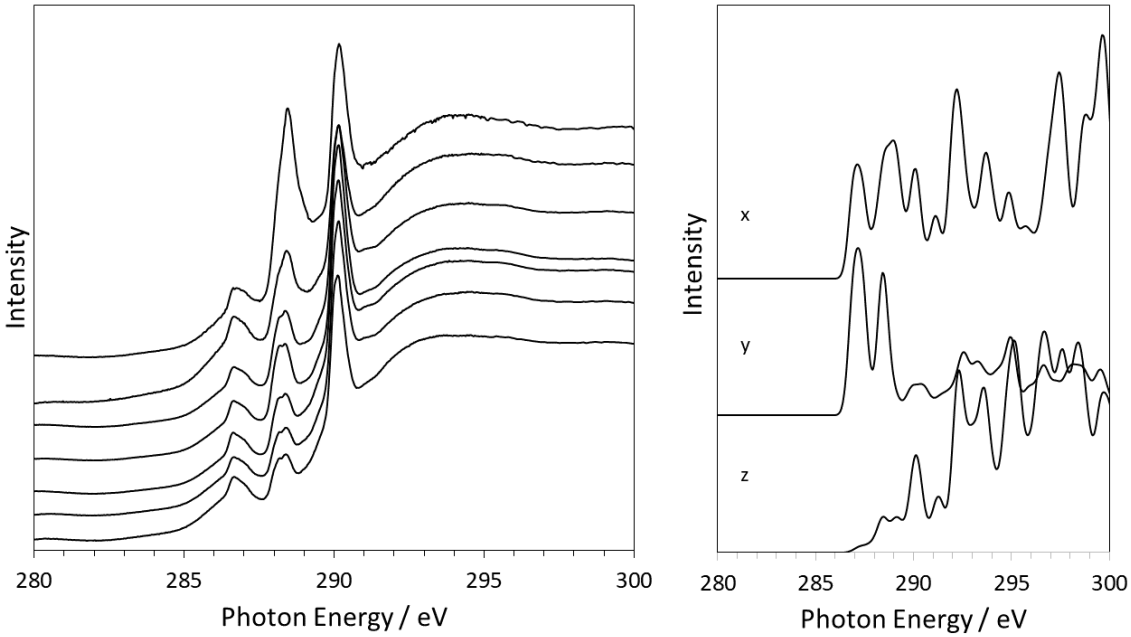



**Figure 6.4** Experimental C k-edge NEXAFS of histidine showing dependence on orientation (left) and x, y and z components of DFT calculated NEXAFS spectrum of Histidine (right)

## 6.3 Conclusions and Future work

As demonstrated above, the four papers address some of the issues faced when attempting to classify salts and co-crystals by formulating a new approach based upon methods or theoretical approaches which are by definition directly sensitive to the proton position in organic structures. This direct sensitivity gives an increased level of accuracy when characterising samples and could be of

use in multiple industries including pharmaceuticals where co-crystallisation is a method used to formulate new drugs with desirable properties. This increased accuracy comes about firstly from the direct sensitivity to the proton position through the electrostatic nature of the interactions as shown in chapter 4, in addition to the level of consistency demonstrated throughout the four investigations whereby the characterisation of the samples were correctly identified in all cases identified. In a number of samples, there was additional information over and above that available through alternative techniques and also identified cases where the traditional approaches had not fully refined the proton positions such that the proton was incorrectly assigned. In addition to these improvements in accuracy and additional information, both core level spectroscopy techniques involve rapid data acquisition with good quality spectra taking in the order of 15 minutes to collect, while DFT calculations on these relatively small molecules can be completed in a number of hours.

### 6.3.1 Relevance of findings to industry

While the pure science behind these findings is obviously of interest, the relevance to industry is of even greater importance to the usefulness of the techniques and theory behind it. By showing that the theory and experimental evidence is sound, we hope to provide the evidence required to show that this type of crystal structure refinement is possible and desirable. This additional evidence has been procured such that we are confident in the resulting analysis and results due to the fully consistent results across a range of techniques and theoretical approaches. There are very few techniques in which this level of self-consistency can be achieved across such a wide range of crystal structures and therefore the confidence we have that this process can be used as an improved structure refinement method.

The pharmaceutical industry relies on the ability to characterise and understand the products with which they work and produce. There are a wide variety of analytical techniques used to achieve this, each with its own set of disadvantages in either cost or difficulty. The aim of this approach was to develop a technique without the time, difficulty, cost or accuracy disadvantages

to demonstrate alternative ways in which we can analyse crystalline samples. Throughout the four papers we have shown how this is the case and demonstrated how we can use this approach to further refine crystalline structures without the difficulty and expense of neutron diffraction experiments.

### 6.3.2 Future Work

There are a number of future avenues in which to build on the research covered. Firstly, using the techniques outlined and the procedures for consistent analysis demonstrated, a further investigation could be completed into SSHB's, to further understand some of the unique properties observed in these crystal structures, which is not fully understood in terms of the hydrogen bonding interactions. One of these which is reasonably common is temperature dependence of the hydrogen bonding interaction whereby the hydrogen atom position is dependent on the temperature. This occurs in SSHBs with a proton potential energy surface approaching a single well since there is minimal barrier to movement of the proton. Using XPS this temperature dependence has already been observed in the literature through a shift in the binding energy as anticipated. However, we believe it would be possible to observe these intermediate states over a temperature dependent set of XPS and NEXAFS data through the shift in the core level binding energy, which we could then refine to definitive hydrogen atom positions demonstrably changing with temperature. In addition to this, there are a number of such samples where a colour change, among other physical properties, takes place over the temperature range 150 K – 300 K which is also directly dependent on the hydrogen bonding. It should also be possible to determine the hydrogen atom positions and show the differences in these samples. At the time of our experiments, the cooling equipment had not been installed; however this is now in place allowing such an experiment to take place going forward.

Another area in which this research could be investigated further is through further refinement of the electrostatic model, in particular in an attempt to remove the ambiguity surrounding the dependence of the local environment of the hydrogen bond and on particular functional groups. An initial thought would be to include a term in the calculation to account for an "external" electrostatic

field which would be calculated to have a dependence on the local chemical environment. A second factor affecting the hydrogen bond is the hydrogen bond angle from donor to acceptor. It would also be possible to account for any additional shielding effects of the angle on the net electrostatic field at the core level electron. A further enhancement to finalise the simple model would be to generalise the system to any proton acceptor (currently based on a nitrogen acceptor) to show the universality of the effect in all hydrogen bonded systems.

A third area in which the research covered here can be worked on going forwards is to continue improvement and refinement of the analysis procedures required to carry out the level of crystallographic refinement required of these methods. At present, the process of data analysis of the core level spectroscopy along with potentially a large number of DFT calculations is a very manual process, much of which could potentially be automated. Now that the proof of principle has been completed, the additional research required is therefore to determine the possibility of automation of the DFT part of the analysis through the definition of appropriate parameters to carry out the comparison with experimental data in an equivalently consistent manner when carried out manually. The resulting system would then be capable of carrying out the structure refinement process without any specialist knowledge of the techniques used, which could open up the method to a much wider number of users. This would also allow a large test set of structures to be measured and refined, to create a database in the literature demonstrating the ability of the approach on a wider range of structures.

Finally, further comparative analysis on a large number of structures needs to be completed compared to primary data from “gold standard” methods such as neutron diffraction using identical samples, which would then act as confirmation that the technique works consistently over a wide range of differing structures. Each of these proposals for future work will further embed the technique of structure refinement using core level spectroscopy as a potential approach alongside the traditional methods.

### 6.3.3 Conclusions

The use of core level spectroscopy combined with computational *ab initio* density functional theory has been demonstrated in the field of crystallography refining hydrogen atom positions in organic crystal structures. Throughout the work, the processes required for this level of hydrogen position refinement have been identified, thoroughly tested and optimised to result in a fully consistent analysis procedure utilising multiple techniques. Over the series of four papers, it has been demonstrated that this consistency can be used to improve our knowledge and understanding of organic crystalline materials, their structures, and the hydrogen bonding interactions present. Firstly utilizing the definitions of salt and co-crystal structures, we investigated whether XPS and NEXAFS provide suitable techniques to distinguish between these states. It turns out that this approach using the quick and easy techniques of XPS and NEXAFS results in a method consistently able to characterise samples as salt or co-crystal, subject to a thorough and robust calibration and data analysis procedure laid out in the chapters 2, 3 and 5. Throughout, the key aspect of the analysis of core level X-ray spectroscopy is to obtain consistency between all of the various techniques. As with all peak fitting analysis, it is possible to fit the emission line in a number of convincing ways depending upon the expected outcome. It is also possible to interpret NEXAFS data in multiple ways by assigning peaks to transitions. These two techniques then also need to be consistent with the Density Functional Theory calculations, which can provide the absolute orbital transitions for NEXAFS and then confirm whether the experimental NEXAFS and XPS spectra are consistent, or if they require some additional fitting constraints, or explanations in terms of the molecular structure, to ensure consistency between techniques. Across a range of samples in the literature and investigated first hand, the results show that salts and co-crystals are identified with a level of accuracy approaching that of neutron diffraction.

To obtain further information about the nature of hydrogen bonding and explain binding energy shifts observed in XPS, a model system based on electrostatics was devised. This was used to demonstrate the electrostatic nature of hydrogen bonding and discover the underlying cause of the + 2 eV shift in binding energy

between salt and co-crystal, which we can now attribute directly to the electrostatic potential of the hydrogen atom on the core level electron.

Finally, the three techniques of XPS, NEXAFS and DFT were combined to showcase the sensitivity to the proton position within a hydrogen bonding interaction by refining the position based on unrefined XRD data. In the five samples, additional information was obtained and the proton position optimised to ensure full consistency between the three methods. In three of the samples, it was discovered that there was a relatively low barrier in the proton potential energy surface, with two distinct peaks appearing where only a single nitrogen environment is present. The combination of DFT with NEXAFS was used to ensure the peak positions were correct, and allowed the two different populations to be determined. As an additional constraint, XPS binding energies were compared to the calculated core level molecular orbitals which further refined the correct molecular structure from DFT by constraining an additional degree of freedom.

Overall, this technique has been developed to accurately and informatively refine the position hydrogen atoms in hydrogen bonding interactions in organic crystals using techniques which have been shown to be directly sensitive to that position. Additionally, the theory has confirmed that the properties probed in core level spectroscopy including binding energies and excitations from the core level are directly dependent on the position of the hydrogen atom through the electrostatic potential difference based on the proton acceptor distance. Using this direct sensitivity to the proton, hydrogen atom positions have been refined in a number of structures showing improvement to the structures found in the crystallography database through a consistent analysis combining the three techniques.

## 6.4 References

1 M. B. P. Paula, T. F. Koetzle and A. J. Schultz, Single Crystal Neutron Diffraction for the Inorganic Chemist – a Practical Guide, *Comments on Inorganic Chemistry*, **28**, 3–38.

2 P. T. Edwards, L. K. Saunders, D. C. Grinter, P. Ferrer, G. Held, E. J. Shotton and S. L. M. Schroeder, Determination of H-Atom Positions in Organic Crystal Structures by NEXAFS Combined with Density Functional Theory: a Study of Two-Component Systems Containing Isonicotinamide, *J Phys Chem A*, 2022, **126**, 2889–2898.

3 P. T. Edwards, L. K. Saunders, A. R. Pallipurath, A. J. Britton, E. A. Willneff, E. J. Shotton and S. L. M. Schroeder, Proton Transfer on the Edge of the Salt/Co-crystal Continuum: X-Ray Photoelectron Spectroscopy of Three Isonicotinamide Salts, *Cryst Growth Des*, 2021, **21**, 6332–6340.

4 G. Greczynski and L. Hultman, Compromising Science by Ignorant Instrument Calibration—Need to Revisit Half a Century of Published XPS Data, *Angewandte Chemie*, 2020, **132**, 5034–5038.

5 G. Greczynski and L. Hultman, X-Ray Photoelectron Spectroscopy: Towards Reliable Binding Energy Referencing, *Prog. Mater. Sci.*, 2020, **107**

6 S. Tothadi, T. R. Shaikh, S. Gupta, R. Dandela, C. P. Vinod and A. K. Nangia, Can We Identify The Salt-Co-Crystal Continuum State Using XPS?, *Cryst Growth Des*, 2021, **21**, 735–747.

7 J. S. Stevens, S. J. Byard, C. A. Muryn and S. L. Schroeder, Identification Of Protonation State By XPS, Solid-State NMR, And DFT: Characterization Of The Nature Of A New Theophylline Complex By Experimental And Computational Methods, *J Phys Chem B*, 2010, **114**, 13961–13969.

8 J. S. Stevens, A. Gainar, C. Jaye, D. A. Fischer, S. L. M. Schroeder and Iop, NEXAFS and XPS of p-Aminobenzoic Acid Polymorphs: The Influence of Local Environment in *16th International Conference on X-ray Absorption Fine Structure (XAFS)*, 2016, **712**, 012133.

9 J. S. Stevens, S. J. Byard and S. L. M. Schroeder, Salt Or Co-Crystal? Determination Of Protonation State By X-Ray Photoelectron Spectroscopy (XPS), *J Pharm Sci*, 2010, **99**, 4453–4457.

10 J. S. Stevens, S. J. Byard and S. L. M. Schroeder, Characterization of Proton Transfer in Co-Crystals by X-ray Photoelectron Spectroscopy (XPS), *Cryst Growth Des*, 2010, **10**, 1435–1442.

11 J. S. Stevens, L. K. Newton, C. Jaye, C. A. Muryn, D. A. Fischer and S. L. M. Schroeder, Proton Transfer, Hydrogen Bonding, and Disorder: Nitrogen Near-Edge X-ray Absorption Fine Structure and X-ray Photoelectron Spectroscopy of Bipyridine-Acid Salts and Co-crystals, *Cryst Growth Des*, 2015, **15**, 1776–1783.

12 J. S. Stevens, S. J. Byard, C. C. Seaton, G. Sadiq, R. J. Davey and S. L. M. Schroeder, Proton Transfer And Hydrogen Bonding In The Organic Solid State: A Combined XRD/XPS/SSNMR Study Of 17 Organic Acid-Base Complexes, *Physical Chemistry Chemical Physics*, 2014, **16**, 1150–1160.

13 J. S. Stevens, S. J. Byard, C. C. Seaton, G. Sadiq, R. J. Davey and S. L. M. Schroeder, Crystallography Aided by Atomic Core-Level Binding Energies: Proton Transfer versus Hydrogen Bonding in Organic Crystal Structures, *Angewandte Chemie-International Edition*, 2011, **50**, 9916–9918.

14 G. Held, F. Venturini, D. C. Grinter, P. Ferrer, R. Arrigo, L. Deacon, W. Q. Garzon, K. Roy, A. Large, C. Stephens, A. Watts, P. Larkin, M. Hand, H. Wang, L. Pratt, J. J. Mudd, T. Richardson, S. Patel, M. Hillman and S. Scott, Ambient-Pressure Endstation Of The Versatile Soft X-Ray (VerSoX) Beamline At Diamond Light Source, *J Synchrotron Radiat*, 2020, **27**, 1153–1166.

15 Y. Zubavichus, A. Shaporenko, M. Grunze and M. Zharnikov, Innershell Absorption Spectroscopy Of Amino Acids At All Relevant Absorption Edges, *Journal of Physical Chemistry A*, 2005, **109**, 6998–7000.

16 A. D. Becke, Density-Functional Thermochemistry 3: The Role Of Exact Exchange, *Journal of Chemical Physics*, 1993, **98**, 5648–5652.

17 N. A. Besley, M. J. G. Peach and D. J. Tozer, Time-Dependent Density Functional Theory Calculations Of Near-Edge X-Ray Absorption Fine Structure With Short-Range Corrected Functionals, *Physical Chemistry Chemical Physics*, 2009, **11**, 10350–10358.

18 A. Schäfer, C. Huber and R. Ahlrichs, Fully Optimized Contracted Gaussian Basis Sets Of Triple Zeta Valence Quality For Atoms Li To Kr, *J Chem Phys*, 1994, **100**, 5829–5835.

# Appendix A: Proton Transfer on the Edge of the Salt/Co-Crystal Continuum: X-ray Photoelectron Spectroscopy (XPS) of Three Isonicotinamide Salts – Supporting Information

*Paul T. Edwards,[1] Lucy K. Saunders,[2] Anuradha R. Pallipurath,[1] Andrew J. Britton,[1] Elizabeth A. Willneff,[3] Elizabeth J. Shotton,[2] Sven L. M. Schroeder[1,2]**

[1]School of Chemical and Process Engineering, University of Leeds, Leeds, LS2 9JT, UK

[2]Diamond Light Source Ltd, Harwell Science and Innovation Campus, Didcot, Oxfordshire

OX11 0DE, UK

[3]School of Design, University of Leeds, Leeds, LS2 9JT, UK

[4]Future Continuous Manufacturing and Advanced Crystallisation Hub, Research Complex at Harwell (RCaH), Rutherford Appleton Laboratory, Harwell, Didcot, Oxon, OX11 0FA, UK

* Sven L. M. Schroeder:

Email: s.l.m.schroeder@leeds.ac.uk

## Survey Spectra for the calculation of elemental composition and adventitious carbon

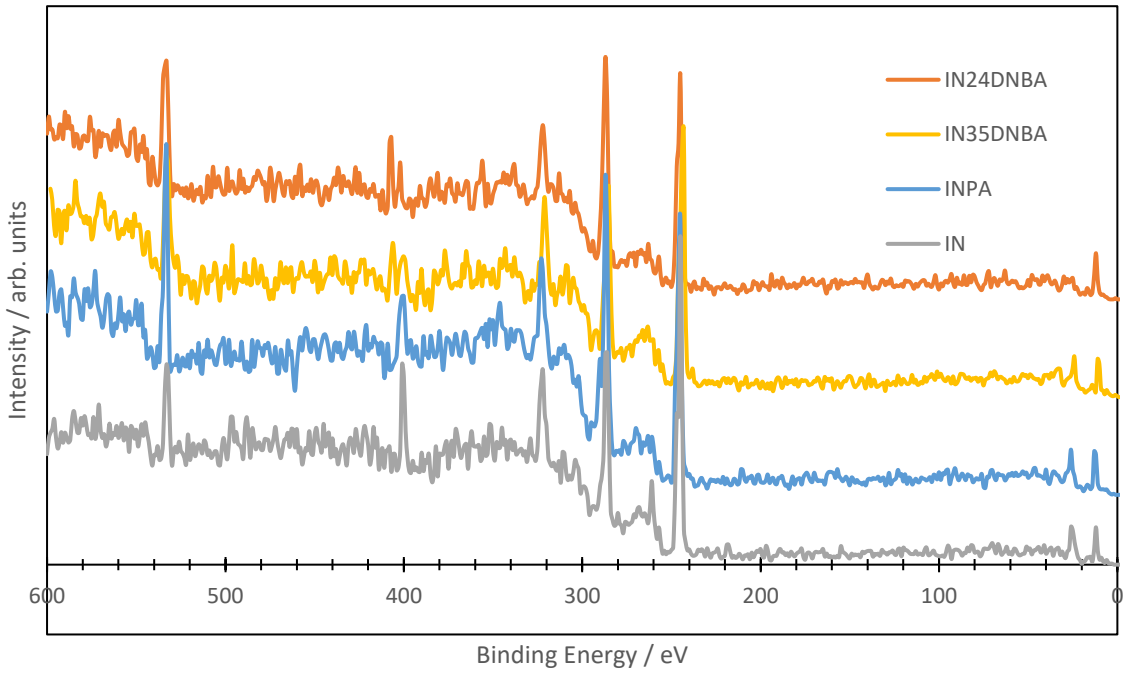



**Figure A1**: Survey Spectra for the four samples measured using the NAP-XPS instrument.

The four survey spectra shown in Figure A1 show those obtained when measuring the final data used in the paper. The C 1s, N 1s and O 1s emission lines are obviously present at roughly 290 eV, 400 eV and 530 eV as expected. The additional features at 245 eV and 320 eV are due to the Argon 2p and 2s emissions from the Argon in the chamber.

**Table A1: Hydrogen bond interaction lengths and angles for the complexes from XRD. [1]**

| IN24DNBA | □□H / Å | H···O / Å | N···O / Å | □□H···O / ° |
|---|---|---|---|---|
| N□H···O | 1.046 | 1.572 | 2.600 | 166.2 |

| IN35DNBA | □□H / Å | H···O / Å | N···O / Å | □□H···O / ° |
|---|---|---|---|---|
| N□H···O | 1.084 | 1.484 | 2.535 | 161.4 |

| INPA | □□H/ N···H / Å | H□O/H···O / Å | N···O / Å | □□H···O / ° |
|---|---|---|---|---|
| N□H···O | 1.102 | 1.449 | 2.550 | 176.7 |
| N···H□O | 1.662 | 1.042 | 2.696 | 179.2 |

**Initial XPS measurements showing radiation damage and charging effects**

Early measurements shown in Figure A2 were taken using a UHV-XPS instrument which under the chosen measurement conditions produced spectra including a low binding energy component often associated with sample charging. Despite these challenges, the conclusions from the N 1s spectra remain the same as for the published data, with the difficulties in binding energy scale calibration and component peak fitting overcome by fitting all of the appropriate charging effects and using the same consistent method of binding energy scale calibration using the adventitious carbon.

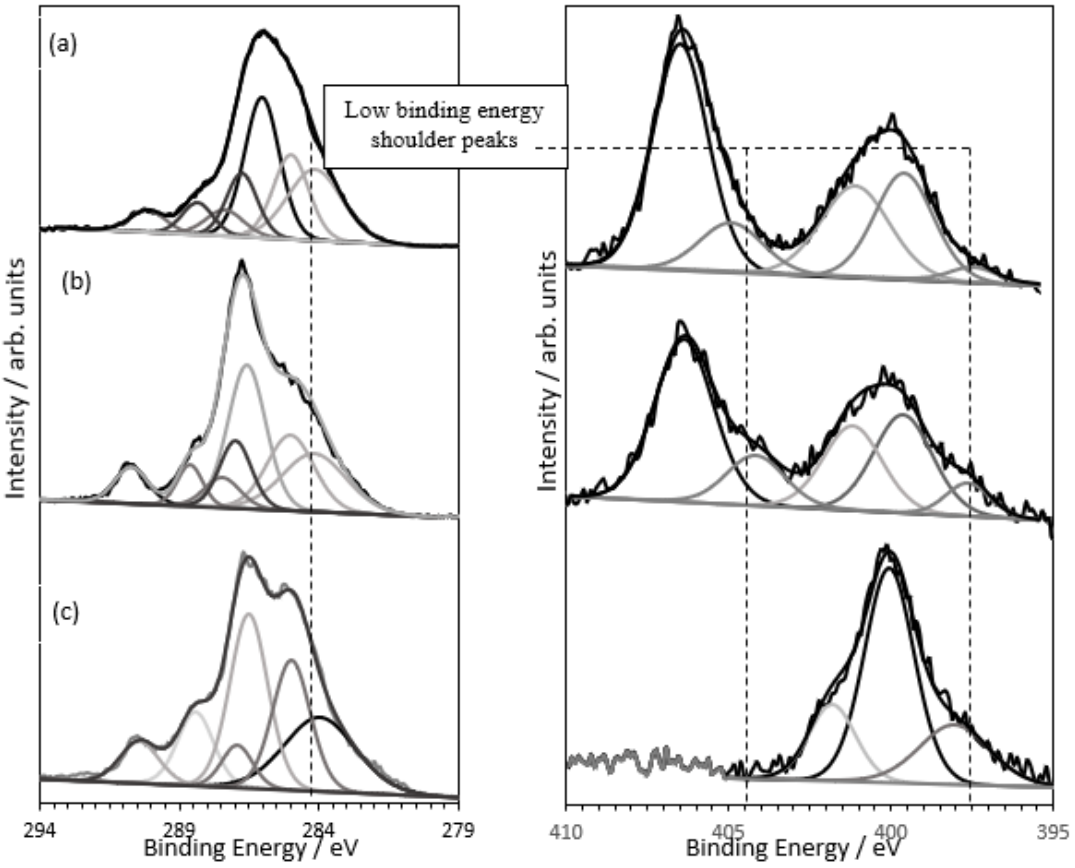



**Figure A2**: Measurements of (left) C 1s and (right) N 1s emission lines using ultra high vacuum XPS (UHV-XPS) for (a) IN24DNBA, (b) IN35DNBA and (c) INPA. In each case a low binding energy shoulder peak is present due to differential charging.

Figure A3 shows the initial measurements of the four samples using the NAP-XPS instrument, with an additional component evident at ~404.5 eV which is believed to be attributed to radiation damage and an uncontrolled amount of residual air in the analysis chamber. The exact chemical origins of this additional feature are uncertain but could be associated with the formation of azo or azido groups. However, the remaining peaks are unchanged compared to the published data. This implies that the method using XPS works even where spectral features not directly associated with sample chemistry are observed, so long as careful and consistent analysis identifies and accounts for these features, and the same self-consistent method of binding energy scale calibration is carried out.

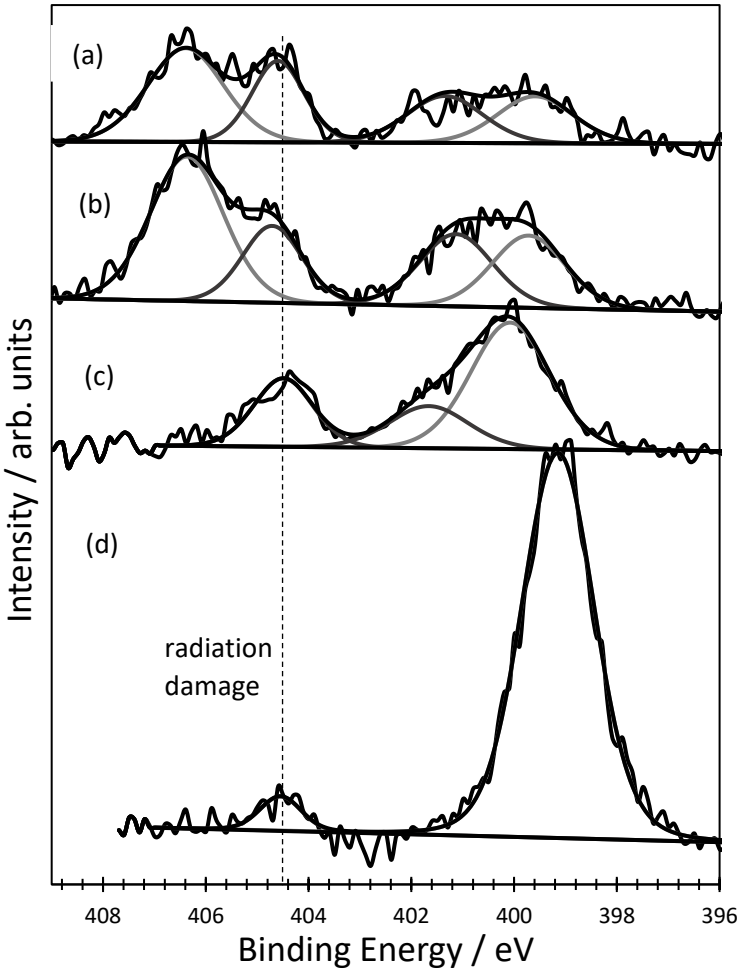



**Figure A3**: Measurements of N 1s emission lines of (a) IN24DNBA, (b) IN35DNBA, (c) INPA and (d) IN using NAP-XPS with radiation damage peaks identified by dashed line. All remaining peak locations are consistent with final results due to consistent method of binding energy scale calibration.

## References

[1] L. K. Saunders, H. Nowell, L. E. Hatcher, H. J. Shespherd, S. J. Teat, D. R. Allan, P. R. Raithby and C. C. Wilson, *Cryst. Eng. Comm.,* vol. 21, no. 35, pp. 5249-5260, 2019.

# Appendix B: Determination of H-atom Positions in Organic Crystal Structures by NEXAFS Combined with Density Functional Theory: A study of Two-Component Systems Containing Isonicotinamide – Supporting information

*Paul T. Edwards,[1, 2] Lucy K. Saunders,[2] David C. Grinter,[2] Pilar Ferrer,[2] Georg Held,[2] Elizabeth J. Shotton,[2]* Sven L. M. Schroeder[1,2,3]**

[1]School of Chemical and Process Engineering, University of Leeds, Leeds, LS2 9JT, UK
[2]Diamond Light Source, Harwell Science & Innovation Campus, Didcot, OX11 0DE, UK
[3]Future Continuous Manufacturing and Advanced Crystallisation Hub, Research Complex at Harwell (RCaH), Rutherford Appleton Laboratory, Didcot, OX11 0FA, UK

* Sven L. M. Schroeder: Email: s.l.m.schroeder@leeds.ac.uk

* Elizabeth J. Shotton: Email: elizabeth.shotton@diamond.ac.uk

## The three crystal structures investigated including hydrogen bonding information and hydrogen bond distances

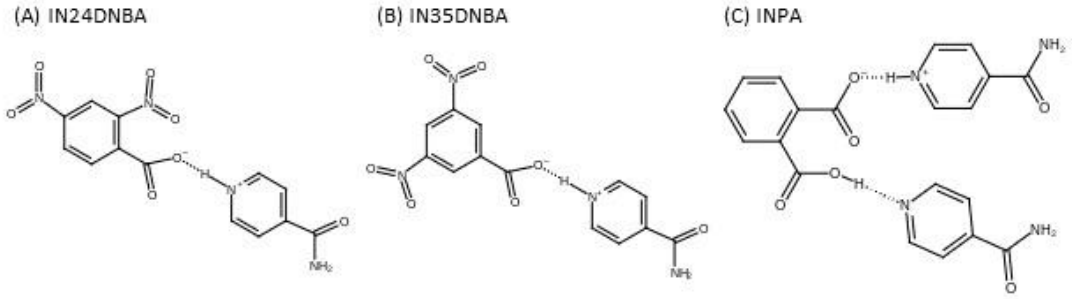



**Figure B1.** Chemical structures for the three systems investigated: (A) isonicotinamide 2,4-dinitrobenzoic acid, (B) isonicotinamide 3,5-dinitrobenzoic acid and (C) isonicotinamide phthalic acid including the relevant donor-proton-acceptor interactions present in the structure.

The three two-component organic crystals studied in this work are: isonicotinamide 2,4-dinitrobenzoic acid ($C_{13}H_{10}N_4O_7$), isonicotinamide 3,5-diniotrobenzoic acid ($C_{13}H_{10}N_4O_7$) and isonicotinamide phthalic acid ($C_{20}H_{18}N_2O_6$). The chemical structures are shown in Figure B1 including all of the relevant intermolecular hydrogen bonding and proton transfer interactions. These interactions have previously been characterized using X-ray diffraction (XRD), and the bond distance

**Table B1.** Hydrogen bond and proton transfer interaction bond distances and angles from XRD for the three systems.

| | N-H / Å | H···O / Å | N···O / Å | N-H···O / ° |
|---|---|---|---|---|
| IN24DNBA<br>N-H···O | 1.046 | 1.572 | 2.600 | 166.2 |
| IN35DNBA<br>N-H···O | 1.084 | 1.484 | 2.535 | 161.4 |
| INPA<br>N-H···O | 1.102 | 1.449 | 2.550 | 176.7 |
| N···H-O | 1.662 | 1.042 | 2.696 | 179.2 |

and angle parameters for each interaction are shown in Table B1.[1]

**NEXAFS normalization procedure for NAP-NEXAFS using the gas phase (with longer path length) as the background I0**

Due to the difficulties of measuring the background signal when using NAP-NEXAFS, measurements of the He gas phase (the chamber with sample removed from the beam path) were taken to provide a background. Significant features are present in this spectrum at the nitrogen K-edge due to the use of a silicon nitride membrane in the beamline.

The absorption coefficient (intensity of a NEXAFS spectrum) is defined by μ, where I is the measured spectrum and $I_0$ is the background spectrum:

$$\mu = \frac{I}{I_0}$$

By using NAP-NEXAFS, the measurement of the $I_0$ background becomes non-trivial, with the effect of the gas phase important in the normalization of the measured spectra. To measure an approximate $I_0$, a gas phase measurement (by removing the sample plate from the beam path) was taken at each of the required absorption edges. Because of the difference in path length through the gas phase compared to the measurements of samples, this approximate $I_0$ is offset by a value n and scaled by a factor m, resulting in $I_0^*$.

$$I_0^* = mI_0 + n$$

Through a series of rearrangements, we can then obtain the true absorption spectrum, divided by a constant m, which is dependent upon a single variable n, which is fit to best model the spectrum baseline.

$$\frac{\mu}{m} = \frac{I}{I_0^*}\left(\frac{1}{1-\frac{n}{I_0^*}}\right)$$

The value of n should be a spectrometer and condition dependent constant, and results in the standard equation where the value is 0 (i.e. no offset). Therefore, to identify the appropriate value of n, we suggest minimizing the value off n while retaining a baseline with physical meaning (i.e. no negative obtrusions). We found a value of n=0.005 appropriate for our experiments.

## Fitted spectra – peak parameters

**Table B2.** Best fit photon energies (eV), Relative intensities and Gaussian peak width for the N 1s → π* transitions. *single peak used to describe transitions from two nitrogen environments

| N1s→π* transition | | <u>N</u>−C=O | C−<u>N</u>=C | C−<u>N</u>H⁺=C | −<u>N</u>O₂ |
|---|---|---|---|---|---|
| IN24DNBA | Energy / eV<br>Relative intensity<br>Width (σ) | 398.9<br>0.348<br>0.35 | n/a | 399.7<br>0.673<br>0.35 | 403.5<br>2.746<br>0.36 |
| IN35DNBA | Energy / eV<br>Relative intensity<br>Width (σ) | 398.8<br>0.111<br>0.27 | n/a | 399.7<br>1.073<br>0.46 | 403.5<br>2.231<br>0.35 |
| INPA | Energy / eV<br>Relative intensity<br>Width (σ) | 398.8<br>0.430*<br>0.22* | 398.8<br>0.430*<br>0.22* | 399.7<br>0.798<br>0.51 | n/a |

Fitting parameters have been determined using the Athena XAS analysis software[2] to fit Gaussian peaks to the NEXAFS spectra. These fits, shown in Table B2 are consistent with our previous XPS study[3] and the relative intensities align well with the calculated DFT spectra.

**Isonicotinamide 2,4-dinitrobenzoic acid & Isonicotinamide 3,5-dinitrobenzoic acid cluster calculations**

In addition to the INPA calculations on a larger cluster of molecules, the equivalent calculations were completed on the other two systems. The overall result was consistent with the INPA, with a similar decrease in energy of the amide groups. However, in these cases, this shift was larger than observed in the experimental spectrum, as seen in Figure B2. This amide nitrogen is highly affected by the longer range interactions, and in these cases the DFT calculation overestimates the relaxation effect of the additional interactions in the system leading to the shift indicated by the arrows in Figure B2.

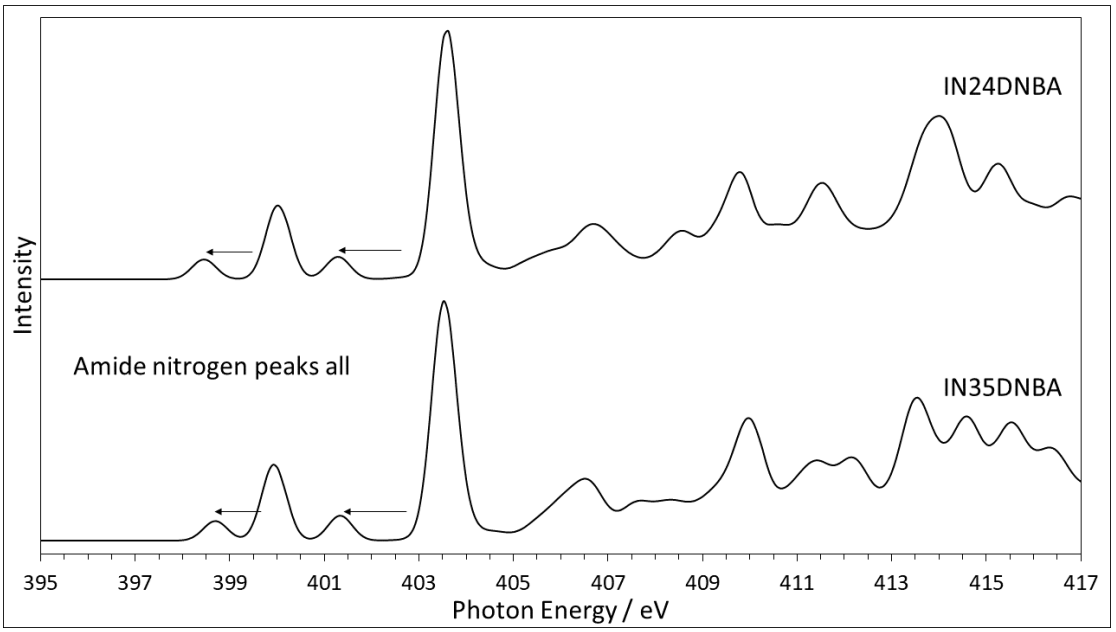



**Figure B2.** DFT Calculated spectra based on a cluster of molecules to observe the effect of longer-range interactions on the spectra. It is clear that the energy of the amide peaks (shown with arrows), rather than the pyridine (399.9 eV) or nitro (403.5 eV), are affected by the longer-range interactions.

**Crystal component spectra**

The Individual components of the three crystals were also measured to determine the difference spectra between the sum of the components and the formed crystal. These are shown in Figure B3. Isonicotinamide has three clear peaks corresponding to the transitions from the core level N 1s orbitals to the 1π*, 2π* and 3π* orbitals respectively. In pure isonicotinamide, the core level of both the pyridine ring and amide nitrogen are at the same energy (as observed in XPS measurement) and therefore both contribute to each peak.

The spectra of 35DNBA and 24DNBA are very similar, with slightly different σ* resonances. The spectrum of 24DNBA also has an additional feature at 400 eV, although this is anticipated to be some form of contamination, as there should not be a feature at this energy. Importantly, this small peak does not appear to affect the difference spectra as seen in the manuscript.

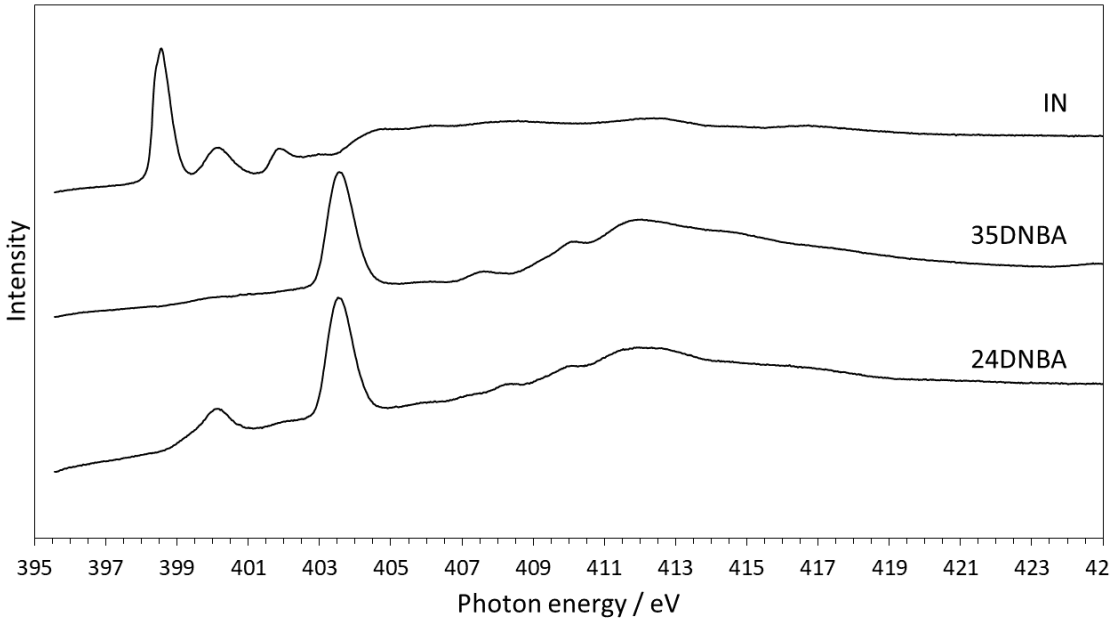



**Figure B3.** Experimental NEXAFS spectra of the pure components of the systems. Used to determine the difference spectra. Note: Phthalic acid does not contain nitrogen so does not contribute directly to the N K-edge spectrum.

**Density Functional Theory Calculation Details**

Example input files and the xyz files for the calculations are included here for reference purposes. *.xyz files were obtained from the CSD based on XRD structure refinement. Most TDDFT calculations were done directly using these structures, but some were also carried out after an optimization calculation of the hydrogens in the structure.

*Example input file for a hydrogen atom only constrained geometry optimization*

```
! B3LYP def2-TZVP def2/J RIJCOSX opt freq

%geom
optimizehydrogens true
end

* xyzfile 0 1 in35dnba.xyz
```

*Example input file for a TDDFT excited state calculation*

```
! B3LYP def2-TZVP def2/J RIJCOSX

%tddft
orbwin[0]=7,10,-1,-1
doquad true
nroots 800
end

* xyzfile 0 1 in35dnba_opt.xyz
```

*.xyz files for the three complexes from XRD studies.[1]

48
Isonicotinamide Phthalic Acid

| | | | |
|---|---|---|---|
| H | 7.615970 | 9.656497 | 7.321674 |
| H | 7.315375 | 8.071444 | 3.343656 |
| H | 7.548208 | 10.003317 | 4.855522 |
| H | 7.508525 | 7.340543 | 8.211975 |
| H | 7.314089 | 3.954033 | 8.466086 |
| O | 6.060540 | 4.877835 | 3.928695 |
| O | 7.146430 | 4.932295 | 8.127826 |
| O | 7.757451 | 4.055500 | 6.152620 |
| O | 8.110379 | 5.229528 | 3.103202 |
| C | 7.407626 | 6.403413 | 6.292070 |
| C | 7.202081 | 5.472445 | 3.909463 |
| C | 7.452692 | 5.012980 | 6.840144 |
| C | 7.495985 | 9.009148 | 5.258077 |
| C | 7.359272 | 6.606489 | 4.899763 |
| C | 7.481025 | 7.507791 | 7.150013 |
| C | 7.533677 | 8.805069 | 6.638318 |
| C | 7.396943 | 7.909930 | 4.398123 |
| H | 7.894837 | 0.298093 | 11.610475 |
| H | 7.256790 | 2.667074 | 10.990738 |
| H | 8.074566 | 2.036492 | 7.020032 |
| H | 9.043512 | -2.275826 | 8.020021 |
| H | 8.656249 | -0.358285 | 7.420393 |
| H | 9.236816 | -3.468199 | 9.314340 |
| O | 8.569074 | -1.954516 | 11.143844 |
| N | 7.647250 | 2.425733 | 8.985276 |
| N | 9.001606 | -2.516451 | 8.993319 |
| C | 8.644537 | -1.645818 | 9.947422 |
| C | 7.935156 | 0.649786 | 10.569719 |
| C | 8.323041 | -0.226866 | 9.555836 |
| C | 7.593482 | 1.957383 | 10.239558 |
| C | 8.378546 | 0.256245 | 8.248904 |
| C | 8.045077 | 1.589352 | 8.015085 |
| H | 6.107399 | 0.478669 | 0.488112 |
| H | 6.336583 | 1.901777 | 5.164477 |
| H | 6.066225 | 2.892077 | 1.188287 |
| H | 6.412641 | -0.561791 | 4.648943 |
| H | 5.735133 | -3.538423 | 2.665418 |
| H | 5.287425 | -2.293024 | 3.806174 |
| H | 6.131878 | 3.542722 | 3.477110 |
| O | 6.669769 | -1.971857 | 0.974066 |
| N | 5.795765 | -2.569047 | 2.984976 |
| N | 6.184137 | 2.439061 | 3.189727 |
| C | 6.245115 | -1.679210 | 2.101439 |
| C | 6.145669 | 0.744660 | 1.531427 |
| C | 6.225676 | -0.229302 | 2.526115 |
| C | 6.304006 | 0.167391 | 3.863943 |
| C | 6.127950 | 2.079486 | 1.898334 |
| C | 6.276818 | 1.523141 | 4.165768 |

34
Isonicotinamide 2,4 Dinitrobenzoic Acid

| | | | |
|---|---|---|---|
| H | 3.797843 | 3.469588 | 4.716098 |
| H | 5.107453 | 7.042507 | 3.678873 |
| H | 2.176963 | 3.523550 | 2.941065 |
| H | 0.846284 | 6.484525 | -0.290378 |
| H | 2.244801 | 6.985101 | 0.279079 |
| H | 3.522494 | 7.241130 | 1.884632 |
| H | 5.276526 | 5.260640 | 5.003087 |
| O | 0.669273 | 4.583025 | 1.307604 |
| N | 1.604780 | 6.400598 | 0.321675 |
| N | 4.540562 | 5.279354 | 4.259855 |
| C | 2.691414 | 5.384406 | 2.244271 |
| C | 1.572953 | 5.431249 | 1.232355 |
| C | 4.497605 | 6.358463 | 3.471541 |
| C | 3.582994 | 6.439749 | 2.437705 |
| C | 3.709143 | 4.234689 | 4.075346 |
| C | 2.767271 | 4.264884 | 3.068854 |
| H | 8.346000 | 4.903578 | 7.747553 |
| H | 6.943716 | 1.145814 | 10.130460 |
| H | 9.549559 | 4.192898 | 9.658172 |
| O | 6.116094 | 5.050651 | 6.315437 |
| O | 6.000592 | 2.913674 | 5.665306 |
| O | 4.374058 | 2.110341 | 7.859411 |
| O | 10.265007 | 2.615165 | 11.365187 |
| O | 8.641616 | 1.288524 | 11.922102 |
| O | 5.517797 | 0.271758 | 8.083578 |
| N | 5.405267 | 1.497250 | 8.109452 |
| C | 6.296811 | 3.812529 | 6.478816 |
| N | 9.152252 | 2.125152 | 11.187797 |
| C | 8.383075 | 2.573144 | 10.013856 |
| C | 6.964781 | 3.405409 | 7.778399 |
| C | 8.069418 | 4.124585 | 8.231705 |
| C | 6.592426 | 2.273258 | 8.504343 |
| C | 7.269652 | 1.838354 | 9.634105 |
| C | 8.801301 | 3.709658 | 9.338869 |

34
Isonicotinamide 3,5 Dinitrobenzoic Acid

| | | | |
|---|---|---|---|
| O | 3.807306 | 1.827214 | -0.045287 |
| N | 4.765867 | 0.392795 | 1.428422 |
| N | 6.132029 | 4.986923 | 2.971947 |
| C | 4.510099 | 1.612532 | 0.955636 |
| C | 5.307673 | 3.976291 | 0.998407 |
| C | 5.108689 | 2.785911 | 1.692490 |
| C | 5.949276 | 3.852386 | 3.662571 |
| C | 5.836801 | 5.067785 | 1.669217 |
| C | 5.444619 | 2.721047 | 3.048054 |
| H | 5.327050 | 1.906723 | 3.593592 |
| H | 6.002327 | 5.932174 | 1.243291 |
| H | 5.280926 | 0.250876 | 2.196201 |
| H | 5.138745 | 4.060263 | 0.083864 |
| H | 6.167620 | 3.880510 | 4.586338 |
| H | 4.375508 | -0.343477 | 0.992746 |
| H | 6.493154 | 5.783573 | 3.501342 |
| O | 9.097109 | 12.227690 | 8.412026 |
| O | 6.485519 | 8.714984 | 3.933034 |
| O | 9.807369 | 7.770221 | 10.112587 |
| O | 9.086823 | 6.065025 | 8.978006 |
| O | 7.046218 | 6.659588 | 4.645777 |
| O | 7.621909 | 12.525933 | 6.848792 |
| N | 9.247943 | 7.268552 | 9.138397 |
| N | 8.331047 | 11.811668 | 7.556398 |
| C | 8.726815 | 8.175216 | 8.103614 |
| C | 8.131130 | 7.623617 | 6.976161 |
| C | 8.831648 | 9.540269 | 8.321766 |
| C | 8.250354 | 10.355272 | 7.360155 |
| C | 7.614807 | 9.860762 | 6.232283 |
| C | 7.580326 | 8.481048 | 6.024614 |
| C | 6.969659 | 7.920054 | 4.753753 |
| H | 7.225718 | 10.448573 | 5.587470 |
| H | 9.291378 | 9.945743 | 9.087764 |
| H | 8.112045 | 6.693352 | 6.858017 |

**Effect of Basis set and Exchange Correlation functional on the calculated NEXAFS spectra**

As mentioned in the main manuscript, to ensure an appropriate basis set and

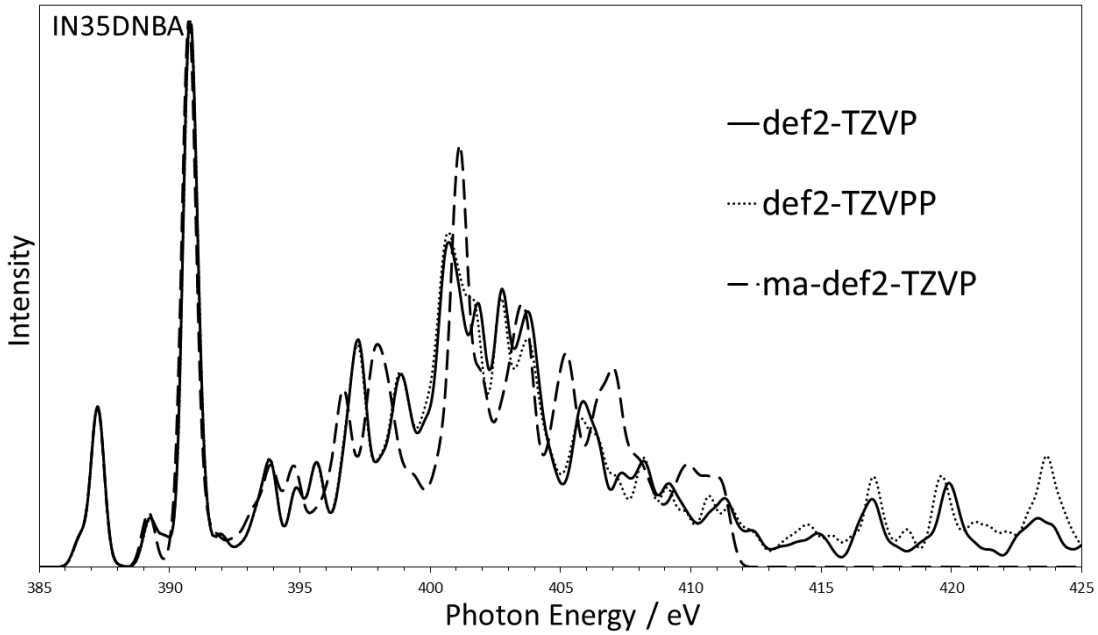



**Figure B4.** DFT calculated NEXAFS spectra demonstrating the effect of basis set on the N 1s → π* transitions and σ* resonances. Photon energies have not been calibrated to allow for direct comparison of the calculated spectra. The primary N 1s → π* peaks are unaffected by the change in basis set, hence the decision to utilize the smaller def2-TZVP basis set for the calculations completed for this work.

exchange correlation functional were utilized in the calculations, a series of tests were completed to ensure the basis set convergence was reasonable (at least in the region of the spectrum of interest) and the exchange correlation functional gave reasonable results. Figure B4 shows the calculated NEXAFS spectrum of IN35DNBA using three different basis sets. The standard triple zeta valence polarized, with additional polarization functions, and augmented with additional basis functions. The important feature to note is that the basis set does not affect the position or intensity of the main N 1s → π* transitions.

The two exchange correlation functionals tested were B3LYP and SRC1 (Utilizing the error function based range separation method mixing BLYP and Hartree Fock exchange components). These calculations were done on the “crystal structure” from

the database, so no geometry optimization was required, and therefore it is the NEXAFS spectrum calculation (TDDFT) that we are probing. Figure B5 shows the

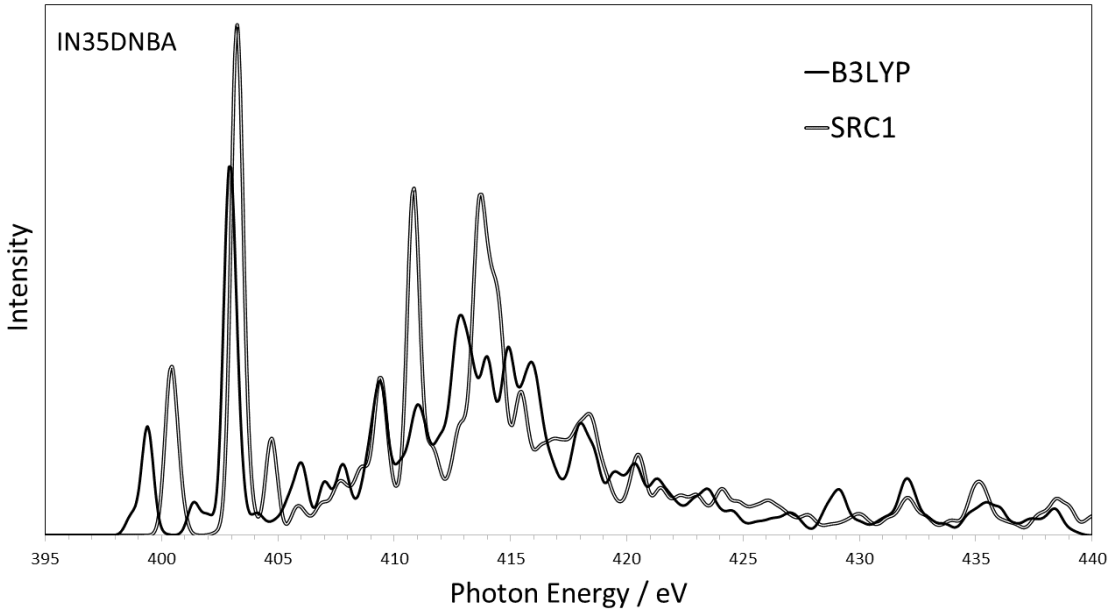



**Figure B5.** DFT calculated NEXAFS spectra demonstrating the effect of exchange correlation functional on the N 1s → π* transitions and σ* resonances. The B3LYP Photon energies have been calibrated to the Nitro peak, the SRC1 functional Photon energies have not been calibrated since the calculated values are close to experiment. Relative peak positions are more consistent with experiment for the B3LYP calculation.

effect of the two functionals, with B3LYP reproducing a spectrum resembling experimental data, while SRC1 does not require any photon energy scale calibration. However, the relative peak positions are off compared to experiment and therefore the B3LYP functional was used.

**Effect of Exchange Correlation Functional on Geometry Optimization**

For completeness, we also include here the test of the two exchange correlation functionals for geometry optimization. Having optimized the geometry using the two separate functionals, Figure B6 shows the Calculated NEXAFS spectra, using the B3LYP functional for the TDDFT calculation as decided previously. Both exchange correlation functionals result in almost indistinguishable NEXAFS spectra, with only a few very minor differences in peak intensity between the two optimized structures. Therefore, since the choice of functional for geometry optimization had an insignificant effect on the calculated NEXAFS spectra, we used the same B3LYP

functional for all optimizations and TDDFT calculations for consistency and reduced computational cost.

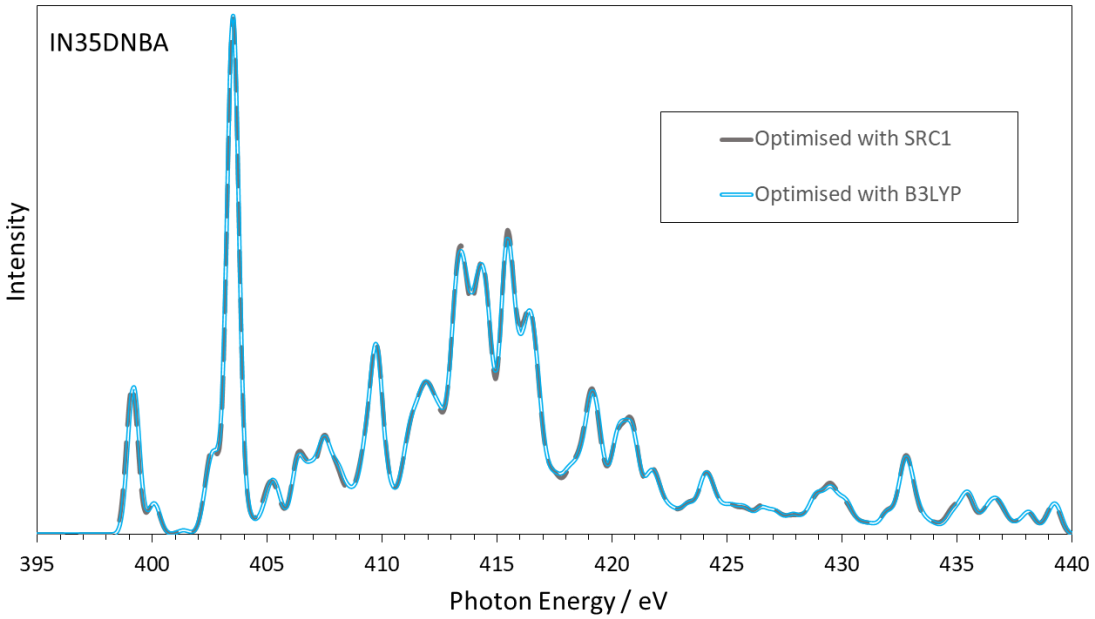



**Figure B6.** DFT calculated NEXAFS spectra demonstrating the effect of exchange correlation functional on the geometry optimization process. Both spectra are almost indistinguishable with only minor changes in peak intensity present. TDDFT calculated using B3LYP functional in both cases, as this is shown to be more consistent with experiment, see Figure B5.

## References

1 Saunders, L. K.; Nowell, H.; Hatcher, L. E.; Shepherd, H. J.; Teat, S. J.; Allan, D. R.; Raithby, P. R.; Wilson, C. C. Exploring Short Strong Hydrogen Bonds Engineered in Organic Acid Molecular Crystals for Temperature Dependent Proton Migration Behaviour using Single Crystal Synchrotron X-ray Diffraction (SCSXRD), *CrystEngComm* **2019**, 21, 5249–5260.

2 Ravel, B.; Newville, M. Athena, Artemis, Hephaestus: Data Analysis for X-ray Absorption Spectroscopy using IFEFFIT, *J. Synchrotron Radiat.* **2005**, 12, 537–541.

3 Edwards, P. T.; Saunders, L. K.; Pallipurath, A. R.; Britton, A. J.; Willneff, E. A.; Shotton, E. J.; Schroeder, S. L. M. Proton Transfer on the Edge of the Salt/Cocrystal Continuum: X-Ray Photoelectron Spectroscopy of Three Isonicotinamide Salts, *Cryst. Growth Des.* **2021**, 21, 6334-6340.

# Appendix C: Hydrogen Bonds in Molecular Crystals: Strong Correlation between Proton Positions and 1s Core Level Binding Energies – Supporting Information

*Paul T. Edwards,[1,2] Elizabeth J. Shotton,[2] Sven L. M. Schroeder[1,2,3]**

[1]School of Chemical and Process Engineering, University of Leeds, Leeds, LS2 9JT, UK
[2]Diamond Light Source, Harwell Science & Innovation Campus, Didcot, OX11 0DE, UK
[3]Future Continuous Manufacturing and Advanced Crystallisation Hub, Research Complex at Harwell (RCaH), Rutherford Appleton Laboratory, Didcot, OX11 0FA, UK

* Sven L. M. Schroeder

Email: s.l.m.schroeder@leeds.ac.uk

## Density Functional Theory (DFT) Calculation Details

Example input files and the xyz files for the DFT calculations are included here for reference purposes. *.xyz files were obtained from the CSD based on XRD structure refinement. Calculations were carried out using the B3LYP exchange correlation functional and the def2-TZVP basis set. Single point energy calculations were completed as no optimization of the structure was desired to purely model the effect of the proton position.

**Example input file for a single point energy calculation of isonicotinamide 3,5-dinitrobenzoic acid**

```
! B3LYP def2-TZVP def2/J RIJCOSX

* xyzfile 0 1 in35dnba.xyz
```

*.xyz files for the three complexes from XRD studies. [1]

Calculations were completed on isonicotinamide 3,5-dinitrobenzoic acid over a range of N-H distances from 0.5 Å to 2.0 Å by changing the H atom coordinates with no other changes to the xyz file. From the output file, the orbital energy of the hydrogen bonded nitrogen atom was determined at each N-H distance as this is equivalent to the N 1s Binding Energy and comparable to the electrostatic model proposed.

```
34
Isonicotinamide 3,5 Dinitrobenzoic Acid
O     3.807306    1.827214   -0.045287
N     4.765867    0.392795    1.428422
N     6.132029    4.986923    2.971947
C     4.510099    1.612532    0.955636
C     5.307673    3.976291    0.998407
C     5.108689    2.785911    1.692490
C     5.949276    3.852386    3.662571
C     5.836801    5.067785    1.669217
C     5.444619    2.721047    3.048054
H     5.327050    1.906723    3.593592
H     6.002327    5.932174    1.243291
H     5.280926    0.250876    2.196201
H     5.138745    4.060263    0.083864
H     6.167620    3.880510    4.586338
H     4.375508   -0.343477    0.992746
H     6.493154    5.783573    3.501342
O     9.097109   12.227690    8.412026
O     6.485519    8.714984    3.933034
O     9.807369    7.770221   10.112587
O     9.086823    6.065025    8.978006
O     7.046218    6.659588    4.645777
O     7.621909   12.525933    6.848792
N     9.247943    7.268552    9.138397
N     8.331047   11.811668    7.556398
C     8.726815    8.175216    8.103614
C     8.131130    7.623617    6.976161
C     8.831648    9.540269    8.321766
C     8.250354   10.355272    7.360155
C     7.614807    9.860762    6.232283
C     7.580326    8.481048    6.024614
C     6.969659    7.920054    4.753753
H     7.225718   10.448573    5.587470
H     9.291378    9.945743    9.087764
H     8.112045    6.693352    6.858017
```

**Table of Core Level Binding Energies Calculated from the Electrostatics Model using Crystallographic Data and Compared to Experimental Values**

Experimental XPS data for a variety of structures and over multiple previous publications were combined to form a test set of structures on which to evaluate the electrostatic model. Binding energy (BE) scale calibration was checked for each experiment and BE's reported here have been recalibrated to 285 eV at the excess adventitious carbon peak.

**Table C1**. Experimental and model binding energies. Distances in Å and energies in eV.

| Compound | N-H distance | N-O distance | Experimental BE | Model BE | Error |
|---|---|---|---|---|---|
| Isonicotinamide 2,4-dinitrobenzoic acid | 1.046 | 2.60 | 401.2 | 401.14 | -0.06 |
| Isonicotinamide 3,5-dinitrobenzoic acid | 1.084 | 2.535 | 401.1 | 400.98 | -0.12 |
| Isonicotinamide Phthalic acid (1) | 1.102 | 2.5499 | 401.6 | 400.95 | -0.65 |
| Isonicotinamide Phthalic acid (2) | 1.662 | 2.6955 | 399.9 | 400.00 | 0.10 |
| 3,5-pyridine-dicarboxylic acid | 1.308 | 2.525 | 400.35 | 400.48 | 0.13 |
| 2,3-pyridine-dicarboxylic acid | 1.036 | 2.725 | 401.5 | 401.27 | -0.23 |
| α para-aminobenzoic acid | 1.817 | 2.735 | 399.8 | 399.81 | 0.01 |
| β para-aminobenzoic acid | 1.73 | 2.754 | 400.0 | 399.95 | -0.05 |
| 2-bipyridine BTA | 0.86 | 2.636 | 401.74 | 401.67 | -0.07 |
| 4-bipyridine 5-sulphosalicylic acid | 0.962 | 2.66 | 401.53 | 401.41 | -0.12 |
| 4-bipyridine BUTA | 1.73 | 2.57 | 399.66 | 399.78 | 0.12 |
| 4-bipyridine Phthalic acid | 1.449 | 2.637 | 399.61 | 400.31 | 0.70 |
| 4-bipyridine adipic acid | 1.627 | 2.666 | 399.5 | 400.03 | 0.53 |
| 4-bipyridine thiodipropionic acid | 1.673 | 2.638 | 399.46 | 399.94 | 0.48 |
| 4-bipyridine malonic acid | 1.798 | 2.612 | 399.4 | 399.72 | 0.32 |
| Theophylline 5-sulphosalicylic acid | 0.84 | 2.576 | 402.1 | 401.67 | -0.43 |
| Theophylline oxalic acid | 1.758 | 2.665 | 399.8 | 399.83 | 0.03 |
| Theophylline maleic acid | 1.588 | 2.634 | 399.9 | 400.07 | 0.17 |
| Theophylline malonic acid | 1.719 | 2.65 | 399.9 | 399.88 | -0.02 |
| Theophylline citric acid | 1.882 | 2.746 | 399.7 | 399.72 | 0.02 |
| Theophylline glutaric acid | 1.826 | 2.683 | 399.7 | 399.75 | 0.05 |
| 4-aminobenzoic acid 4-hydroxy-3-nitrobenzoic acid | 1.68 | 2.66 | 399.8 | 399.94 | 0.14 |
| 3,5-diaminobenzoic acid-3,5-dinitrobenzoic acid | 1.03 | 2.807 | 402.1 | 401.36 | -0.74 |
| Isonicotinamide 2,4,6-trihydroxybenzoic acid | 0.835 | 2.93 | 401.8 | 401.99 | 0.19 |
| Glycine | 0.91 | 2.796 | 401.4 | 401.67 | 0.27 |
| Aspartic acid | 0.879 | 2.862 | 401.6 | 401.81 | 0.21 |
| Glutamic acid | 0.933 | 2.819 | 401.5 | 401.62 | 0.12 |

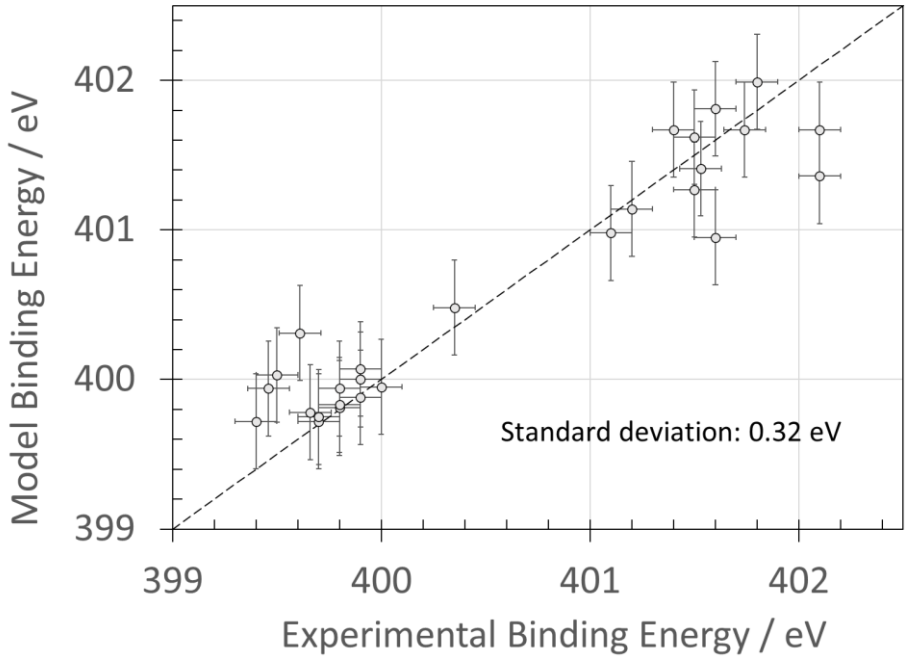



Figure C1. Correlation between experimental and model BEs reported in the accompanying table including error bars for the experimental errors (±0.1 eV) and the standard deviation on model BEs (0.32 eV).

## References

[1] L. K. Saunders, H. Nowell, L. E. Hatcher, H. J. Shespherd, S. J. Teat, D. R. Allan, P. R. Raithby and C. C. Wilson, *Cryst. Eng. Comm.,* **21**, 35, 5249-5260, 2019.

[2] P. T. Edwards, L. K. Saunders, A. R. Pallipurath, A. J. Britton, E. A. Willneff, E. J. Shotton and S. L. M. Schroeder, "Proton Transfer on the Edge of the Salt/Cocrystal Continuum: X-Ray Photoelectron Spectroscopy of Three Isonicotinamide Salts," **21**, 11, 6332-6340, 2021.

[3] J. S. Stevens, S. J. Byard, C. C. Seaton, G. Sadiq, R. J. Davey and S. L. M. Schroeder, *Angew. Chem. Int. Ed.,* **50**, 42, 9916-9918, 2011.

[4] J. S. Stevens, C. R. Seabourne, C. Jaye, D. A. Fischer, A. J. Scott and S. L. M. Schroeder, *J. Phys. Chem.,* **118**, 42, 12121-12129, 2014.

[5] J. S. Stevens, L. K. Newton, C. Jaye, C. A. Muryn, D. A. Fischer and S. L. M. Schroeder, *Cryst. Growth Des.,* **15**, 4, 1776-1783, 2015.