Core Level Spectroscopy of Organic Systems – Experimental Studies and Development of Data Analysis Techniques

Laila Hesham Ahmed Al-Madhagi

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

The University of Leeds

School of Chemical and Process Engineering

November 2019

The candidate confirms that the work submitted is her 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 the appropriate credit has been given within the thesis where reference has been made to the work of others.

Some of the work in Chapters 4 and 6 of the thesis has appeared in publication as follows:

X-ray Raman scattering: a new *in situ* probe of molecular structure during nucleation and crystallization from liquid solutions. 2018. Laila H. Al-Madhagi, Sin-Yuen Chang, Mahalingam Balasubramanian, Anna B. Kroner, Elizabeth J. Shotton, Elizabeth A. Willneff, Bhoopesh Mishra and Sven L. M. Schroeder.

I was responsible for planning and carrying out the experiment, interpreting the results, performing the computations and writing the manuscript.

The contribution of other authors was:

Sin-Yuen Chang, Elizabeth A. Willneff and Sven L. M. Schroeder helped with planning and carrying out the experiment. Mahalingam Balasubramanian and Bhoopesh Mishra helped with carrying out the experiment. Bhoopesh Mishra and Sven L. M. Schroeder conceived the original idea and helped with the interpretation of the results. Sven L. M. Schroeder helped writing the manuscript. Anna B. Kroner, Elizabeth J. Shotton, Elizabeth A. Willneff, Bhoopesh Mishra Sven L. M. Schroeder supervised the project. All authors provided feedback and contributed to the final manuscript.

Acknowledgment

First and foremost, I would like to express my sincere gratitude to my supervisor Prof. Sven Schroeder for the continuous support and encouragement during my PhD studies, for his motivation, immense knowledge and patience and for believing in my capabilities to undertake this project. I could not have imagined having a better mentor for my PhD. I would also like to thank Dr. Elizabeth Willneff for her constant support during my PhD from planning experiments to preparing manuscripts. My sincere thanks also go to Dr. Bhoopesh Mishra, Dr. Anna Kroner and Dr. Elizabeth Shotton for their constant support and encouragement throughout the duration of the project. I would like to thank the University of Leeds and Diamond Light Source for funding my PhD studies.

I thank all current and past members of the SLMS lab for the material and moral support and for the sleepless nights when we were working at a synchrotron facility somewhere around the world. I would especially like to thank Dr. Sin-Yuen Chang, Dr. Thokozile Kathyola and Bethan Evans; my synchrotron experiments would not have run as smoothly without your help and my time would not have been as enjoyable if it was not for you.

Most of the experimental work I have done was performed at neutron and synchrotron facilities and without the help and support of the scientists at these facilities my experiments would have been impossible. I would especially like to thank Dr. Mali Balasubramanian, Dr. Sam Callear, Dr. Blanka Detlefs, Dr. Christoph Sahle, Dr. Sofia Diaz-Moreno, Dr. Philip Chater, Dr. Dean Keeble for their help with setting up and running the experiments and for the valuable advice on data analysis. I would also like to thank Dr. Joanna Leng for her guidance on the Python scripts.

My dearest mother and father, I would not where I am today without your love and support. Thank you both for believing in me and I hope that I made you proud. Waddah, Lamia Angela and Adeeb, thanks for supporting me spiritually throughout my studies and my life in general. Nuha, Nora, Muneera, Fatema and Hadeel you are my family here in the UK and I would like to thank you for your love and support and for putting up with my complaints. Lastly, a massive thank you to the rest of my family and to my friends in the UK and back in Yemen for their support.

Abstract

Many natural chemical and physical transformations take place in solution. Tremendous research has thus been invested to understand the nature of intermolecular interactions in the liquid-phase. The research reported in this dissertation deals with the molecular basis of organic crystal nucleation from solution. Our understanding of this process is still incomplete, and any breakthrough in nucleation science will require a fundamental understanding of the molecular evolution taking place prior to nucleation.

The crystallisation of imidazole from water and ethanol was investigated with several analytical and computational techniques probing the local interactions between solute and solvent molecules. FTIR, X-ray Raman scattering (XRS) and total X-ray scattering and the associated X-ray pair distribution function (XPDF) analysis were combined with density functional theory (DFT) and empirical potential structure refinement (EPSR) simulations to obtain a realistic model of the intermolecular interaction between solute and solvent molecules in aqueous and ethanolic solutions of imidazole.

While hydrated imidazole molecules with very few direct imidazole-imidazole interactions are predicted to be the most probable structural motifs in aqueous solutions, direct imidazole-imidazole interactions are anticipated in the ethanolic solutions. The change in the solvation structure between aqueous and ethanolic solutions is expected due to the difference in the number of hydrogen bonds that can be formed by ethanol (2 bonds) and water (4 bonds). This should facilitate direct imidazole – imidazole interactions in the ethanolic solutions.

The local solvation structure in aqueous and ethanolic solutions of imidazole does not seem to be disturbed by changing the concentration of imidazole in the solution, as indicated by FTIR, XRS and XPDF. For the first time, the molecular structure of the solutions was also probed during cooling crystallisation, utilising XRS and XPDF. Negligible changes are observed in the XRS and XPDF data acquired in supersaturated solutions, indicating that the average solvation structure around imidazole molecules does not change significantly while traversing the thermodynamically metastable zone. This indicates that breakage of solute-solvent bonds during desolvation is the key step in nucleation of imidazole from the solution-phase.

Table of Contents

List of Tables

List of Figures

Abbreviations

APS	Advanced photon source
ATR	Attenuated total reflectance
CNT	Classical nucleation theory
DFT	Density functional theory
DLS	Diamond Light source
EELS	Electron energy loss spectroscopy
EPSR	Empirical potential structure refinement
ESRF	European synchrotron radiation facility
EXAFS	Extended X-ray absorption fine structures
FEL	Free electron laser
FTIR	Fourier transform infrared
FWHM	Full-width-at-half-maximum
HF	Hartree-Fock
IP	Ionisation potential
ISEELS	Inner-shell electron energy loss spectroscopy
LCF	Linear combination fitting
LINAC	Linear (electron) accelerator
MD	Molecular dynamics
NEXAFS	Near-edge X-ray absorption fine structures
NIMROD	Near and intermediate range order diffractometer
NMR	Nuclear magnetic resonance
PCA	Principal component analysis
RMC	Reverse Monte Carlo
TD-DFT	Time-dependent density functional theory
TEM	Transmission electron microscopy
XAS	X-ray absorption spectroscopy
XPS	X-ray photoelectron spectroscopy
XRS	X-ray Raman scattering

Chapter 1 Introduction

Non-covalent bonding interactions between molecules (e.g. hydrogen bonding) play a central role in all fields of chemistry and biology. They are responsible for the properties of molecular structures and the energy transfer between molecules and between molecular clusters. The existence of condensed phases is attributed to these intermolecular interactions. The vast majority of chemical reactions take place in solutions; which has driven a tremendous research to elucidate the effect of solute-solvent and solute-solute interactions on the solvation structures and consequently on the chemical reactivity [1-9]. Phase transformation and the creation of crystalline materials from liquid phases is one form of these chemicals reactions that take place in solution. Crystallisation is central to the process engineering of materials [10]; the ability to control molecular assemblies in solutions is important for utilising structural prediction procedures and for process design [11]. Understanding and controlling the intermolecular interactions between solute molecules is capable of defining the final solid-state form of crystals namely polymorphs, solvates and hydrates [12] and it could be used for purifying crystals during intermediate and final stages of synthesis. The rate and mechanism by which crystallisation occurs are determined by numerous thermodynamic, kinetic and molecule recognition factors and many of these factors directly depend on the solution chemistry [12].

1.1 Hydrogen Bonding

Molecular bonds can be categorised into intramolecular and intermolecular bonds (Figure 1.1). Intermolecular bonds are weaker in nature than their counter part however they are very essential as they keep the molecules packed together in the liquid and solid phases. The dispersion forces are quantum mechanical in origin and arise from the uneven distribution of the electron cloud; hence they are transient at any instance and attain a spherical symmetry with time [13]. This means that they are non-directional and lead to organising the molecules in the closest packing. Other intermolecular interactions (e.g. dipole–dipole, ion–dipole), which can be treated by simple electrostatics, are inherently directional [13]. Hence, they impose orientational conditions on the molecular structure which opposes the non-directionality of the dispersion forces.

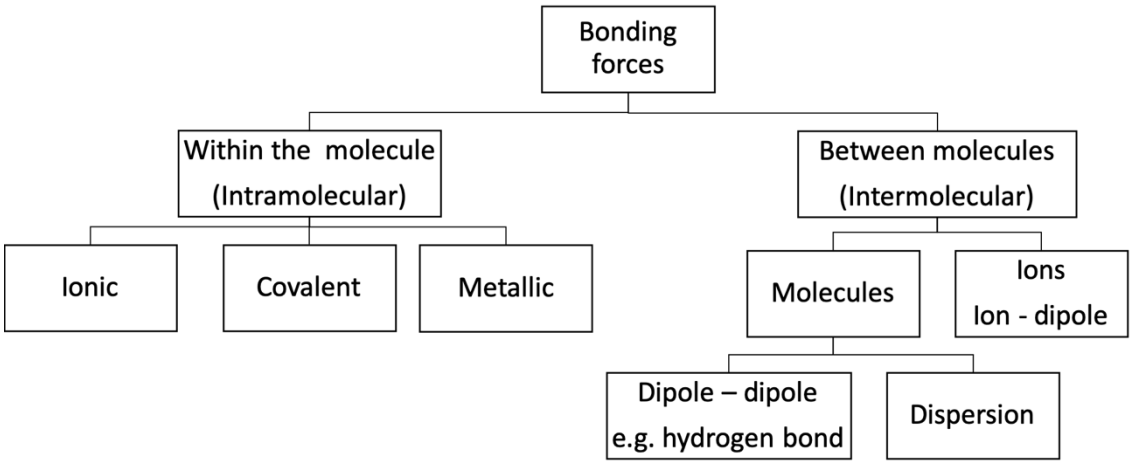

Figure 1.1. Types of molecular bonds.

A hydrogen bond is a dipole-dipole interaction between an X – H group (a proton donor) and an atom or group of atoms Y (proton acceptor) between different molecules or in the same molecule [14]. It is a complex interaction, which geometry is described in terms of three significant structural features: the intermolecular H…Y distance, the linearity of the hydrogen bond and the orientation of the proton acceptor molecule [15, 16]. The discovery of the hydrogen bond cannot be attributed to a single author; elaborate studies on the concept were published from the 1920s with pioneering roles attributed to G.N. Lewis, W.M. Latimer and W.H. Rodebush, M.L. Huggins, and L. Pauling [17-19]. By the end of the 1930s, a "classical" view of the hydrogen bond was established based on water where hydrogen-bonding exists between strongly polar groups X-H on one side (X is O, N or a halogen) and an atom Y on the other (Y is O, N, S or a halide) [16]. The research continued in the following decades and experienced a peak in the 1950s and 1960s followed by an era of stagnation in the 1980s. An intense revival occurred from 1990s onwards and this could be attributed to the prime importance of the hydrogen bond for the structure and dynamics of many chemical systems ranging from biological to inorganic chemistry. A historical narrative of the development of the concept of the hydrogen bond can be found in many books [19-27].

The hydrogen bond energy is split into contributions from electrostatics, charge transfer, polarisation, dispersion and charge repulsion [28]. These constituents differ in their distance and angular characteristics; for example, the electrostatic term is directional and of long distance while the charge transfer and polarisation terms reduce with distance. Since the electrostatic contribution diminishes the slowest with distance,

the hydrogen bond potential is dominated by electrostatics at longer distance. The dispersion term is isotropic (as mentioned previously) and the charge repulsion term increases drastically with reducing distance. These two terms are usually combined as the van der Waals contribution, which is described in terms of the Lennard-Jones potential. The van der Waals contribution is always present and for weaker hydrogen bonds it contributes as much as the electrostatic term.

The energy of the hydrogen bond cannot be measured directly; hence, computational chemistry was utilised to estimate the energies of hydrogen bonds. These results covered a wide range between 0.2 to 40 kcal.mol^{-1} [16]. Based on these energies, hydrogen bonds were classified into strong, moderate and weak; see Table 1.1 for the characteristics of each category. Moderate bonds (also called normal) are associated with energies between 4 and 15 kcal.mol^{-1} and resemble those between water molecules [25]. Hydrogen bond with energies above and below this range are classified as strong and weak respectively.

Table 1.1. Characterisation of strong, moderate and weak hydrogen bonds[16, 25].

	Strong	**Moderate**	**Weak**
Interaction type	Strongly covalent	Electrostatic	Electrostatic and dispersion
Bond length (Å) H…Y	1.2 – 1.5	1.5 – 2.2	> 2.2
Lengthening of X-H	0.08 – 0.25	0.02 – 0.08	< 0.02
X…Y (Å)	2.2 – 2.25	2.5 – 3.2	> 3.2
X-H vs. H…Y	X-H ≈ H…Y	X-H < H…Y	X-H << H…Y
Directionality	Strong	Moderate	Weak
Bond angles (°)	170 – 180	> 130	> 90
Bond energy (kcal.mol^{-1})	15 – 40	4 – 15	< 4

Chemical, crystallographic, spectroscopic and theoretical approaches have been utilised to study the hydrogen bond. The chemical approach involves studying the thermodynamic properties of liquids to derive the energies and entropies of interactions and consequently the type of association [13, 29]. This method has loosely attributed interactions where molecules with X-H bonds are involved to hydrogen bonding which

is not necessarily accurate. Crystallographic studies showed that the X-Y distance is smaller than the know van der Waals' radii which indicates the presence of an attractive force able to overcome the repulsive force between molecules [16, 30-32], however, these geometric criteria are prone to errors [32].

Vibrational (i.e. IR and Raman) and NMR spectroscopy became standard methods to probe hydrogen bonds [25]. Hydrogen bonding effects the vibrational modes of the involved groups is several ways [13, 24]. The X-H stretching frequency is the easiest to identify and in most cases is very sensitive to the formation of hydrogen bond with a noticeable red shift of the band combined with broadening or intensification. A direct effect of the hydrogen bond can also be observed on the acceptor atom where the formation of the H…Y hydrogen bonds weakens the Y-R bond and consequently lower the vibrational frequency of the Y-R bond. The effects mentioned above show many anomalies such as the blue shifting hydrogen bonds where the X-H stretching band shows a blue shift which coincides with a shortening of the X-H bond [33-35]. In NMR spectroscopy, ^{1}H downfield shifts and shifts in the NMR spectra of the isotopes of the X and Y atoms can be correlated with the length of the hydrogen bond [25, 36, 37]. Core-level spectroscopy, which probes the local structure via the excitation of a core-level electron to unoccupied molecular states has also been utilised to probe hydrogen bonding [38-45].

Intermolecular interactions between molecules and clusters of molecules have been successfully modelled with empirical and semiempirical potential energy functions (e.g. Lennard-Jones-Coulomb potentials) [46]. However, the fundamental understanding of these interactions can only be achieved through *ab-initio* calculations [47, 48]. The motivation to apply these calculations to molecular clusters is to determine the structure of the cluster, the stabilisation energy, the vibrational frequencies and the potential energy and free energy surfaces [48]. The prime property of a system is its structure and a main goal for theoretical calculations is to determine the equilibrium structure of the system. For non-rigid systems the equilibrium structure concept is useless since these systems are dominated by large amplitude motions. Hence, the potential energy and free energy surfaces become the prime properties of these systems as they provide the global and local minima for the system through the optimisation of the stabilisation energy. Vibrational frequencies have a key role in understanding

intermolecular interactions and they are easily observable and detected. They are highly sensitive to the quality of the potential energy surface (PES) and hence can be used to validate the calculation procedure of the PES [48].

1.2 Solvation and Solution Structure

Solutions are homogenous liquid phases consisting of more than one substance in variable ratios; where one of these substances is called the solvent and it is treated differently from the other components of the solution called the solutes. Since the majority of chemical reactions take place in condensed phases, understanding the nature of intermolecular interactions in solution phase is crucial in a number of research fields; such as, chemistry, physics, food science, pharmacy and biological sciences. Moreover, understanding the nature of interactions in solution phase enables applying the same approach to a number of apparently disconnected phenomena; such as, chemical reactions, and crystallisations processes.

Understanding solvation dates back to the 1900s when Arrhenius won the 1903 Nobel prize for his work on defining acid-base dissociation in terms of H+ and OH- ions [49]. After 20 years, Brønsted–Lowry defined acids as a hydrogen ion donors and bases as hydrogen ion acceptors [50, 51]. Lewis extended the definition of acid-base dissociation beyond hydrogen and hydroxyl ions and described it in terms of electron lone pairs [52]. Since then tremendous work has been done on proton transportation, hydration shell size and interfacial dielectrics which gave birth to the contemporary liquid-state science [53, 54].

The effect of solvent on the structural and orientational parameters of the solute molecules has been widely observed in a variety of reactions in the different areas of chemistry; ranging from pharmaceuticals [55], biotechnology [56, 57] to catalysis [58, 59]. The solvent has a major impact on the chemical dynamics of reactions in solution phase through modification of the reaction activation energy and through collisional effects [1, 60]. The solvent can provide the energy required to overcome the energy barrier between the reactants and products as well as provided a sink for the products to dispose of their excess energy. The solvent has also a considerable influence on the structural and orientational parameters of the solute molecules [1, 61]. This influence

can be classified into static and dynamic solvation effects [61-64]. The static solvation is associated with stabilising the molecular structure by the solvent to achieve equilibrium. The dynamic solvation, which contributes to the reaction rates, relates to the dynamics induced prior to the structural reorganisation.

1.3 Nucleation Mechanisms and Self-association in Solutions

Understanding the nucleation of organic solutes from solutions is crucial to control the physical and chemical properties of the produced crystals which is essential in many industries. The Classical Nucleation Theory (CNT) which was based on Volmer's concept of nucleation [65] has been historically used to describe the process. This theory states that nucleation occurs as a result of solute molecules aggregating in super-saturated solutions to form nuclei with a critical size, a surface that separates the nuclei from the surrounding environment and with structural properties similar to that of the final crystal.

The thermodynamic description of CNT was introduced by Gibbs [66, 67] who defined the free energy of the system (ΔG) as the sum of the free energy for phase-transformation ΔG_V and the free energy for surface formation ΔG_S. In nucleation from solutions, ΔG_V expresses the spontaneous tendency of the solute molecules to precipitate out of a supersaturated solution and ΔG_S expresses the formation of solid/liquid interface [68]. Since the solid-phase is more stable than the liquid-phase, ΔG_V becomes negative and hence decreases the Gibbs free energy of the system favoring growth (Figure 1.2). Whereas forming a solid/liquid interface increases the free energy of the system (Figure 1.2) favoring dissolution. Small clusters tend to dissolve back in the solution because ΔG_S dominates at small cluster radius (r). When the cluster reaches the critical radius (r_c), the Gibbs free energy goes through a maximum before decreasing continuously as ΔG_V dominates. At this point growth become favorable and crystal nuclei are formed which have the same structure as the mature crystal.

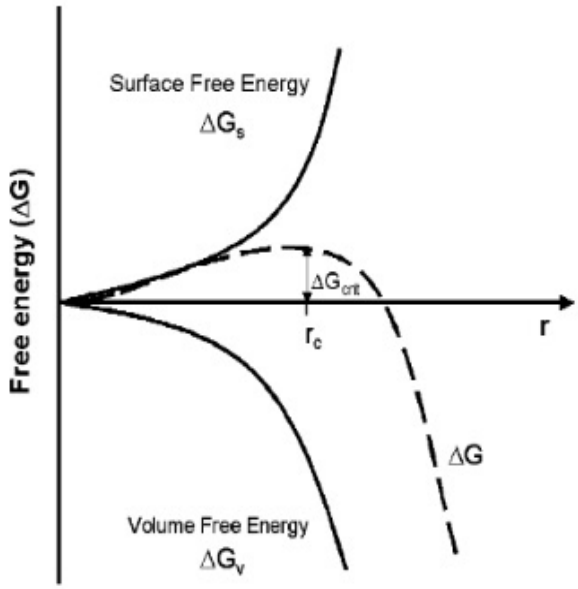

Figure 1.2. Gibbs free energy diagram of nucleation [68].

In many situations, the pre-nuclei differ drastically in terms of structure and composition from the bulk phase which indicates the breakdown of CNT in explaining nucleation [69-72]. A number of "non-classical" nucleation mechanisms have been explored both theoretically and experimentally such as the two-step nucleation mechanism [73] and pre-nucleation cluster mechanism [74]. In the two-step nucleation mechanism, the nucleation occurs inside pre-existing metastable dense liquidlike clusters (Figure 1.3). Numerical simulation of homogenous solutions [75] and DFT calculations[76] showed that this liquid-liquid separation occurs near the metastable critical point. Crystallisation experiments illustrated that this phase separation is not regenerated [77]. All of this indicates that this separation is due a liquid-liquid miscibility gap inside the metastable zone and should not be confused with nucleation mechanisms [77, 78] or with the existence of pre-nucleation clusters suggested by Gebauer and Cölfen [74]. The pre-nucleation clusters are thermodynamically stable associates in the solution with a structure similar or dis-similar to the final product as shown for inorganic system.

As mentioned in the previous section, the self-association of solute molecules in solutions could assist in understanding nucleation since the later could be viewed as the scaled-up version of solution chemistry [78]. The solute associates are either single molecules, self-associated group of molecules or pre-nucleated clusters and they are considered to be the building blocks from which nuclei clusters are formed. The size, composition and intermolecular interactions between solute and solvent molecules are

dependent on solvent, solution concentration and temperature. The correspondence in structure between the solute associates and the final crystals has been illustrated in a number of experimental and computation studies [70, 79-85] which indicates suitability of the CNT. On the other hand, other studies [70-72] showed that there is no one-to-one correspondence between the solution associates and the final crystal indicating a non-classical nucleation mechanism in which rearrangement of the molecular structure takes place prior to crystal formation.

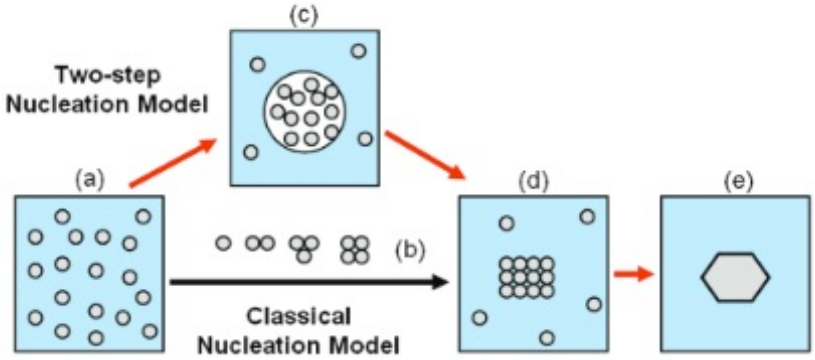

Figure 1.3. Two-step nucleation model and classical nucleation model. (a) supersaturated solution, (b) ordered subcritical clusters of solute molecules. (c) liquid-like clusters, (d) ordered crystalline nuclei, (e) solid eventual crystal [68].

1.4 Overview of the Characterisation Methods for Probing the Dynamic Structure of Solutions

Understanding solution-phase chemistry requires a microscopic description of the electronic structure of the reacting molecules, and of the complex influence of the solvent medium on the reaction energetics and dynamics. This section will list some of the characterisation techniques used to probe the intermolecular interactions in solution phases. The theoretical background to the techniques used in this thesis is discussed further in Chapter 2.

A number of spectroscopic and scattering experimental techniques have been used to probe the solute-solvent interactions and the association of solute molecules during nucleation to get insight on pre-nucleation structures [78]. Raman and FTIR spectroscopy [71], UV/Vis spectroscopy [86] and Nuclear magnetic resonance (NMR) [70, 84] are among the most common spectroscopic techniques used. Infrared and Raman spectroscopy are based on the vibrational modes of the molecules, UV-Vis

probes the transition energy for the excitation of an electron from the highest occupied molecular orbital (HOMO) to lowest unoccupied molecular orbital (LUMO) and NMR measures the resonant frequency that causes a spin flip of the probed proton [87]. On their own, these techniques have limited success in providing molecular level information on solvation structures and structural development prior to nucleation. X-ray absorption spectroscopy (XAS) [39, 88-90] and photoelectron spectroscopy (XPS) [91-93] are highly sensitive to changes in the local environment around the absorbing atom as well as hydrogen-bonding and protonation state. However, they are not suitable to study supersaturated solutions because the current setups, liquid jets [94-98], do not allow precise temperature control which is needed for crystallisation experiments and they are very prone to crystal formation which could block up the jet nozzle. X-ray scattering (SAXS and WAXS) [99-101] and neutron scattering [72, 102-112] have also been utilised to understand the structural properties of solutions. However, they have less atomic sensitivity compared to spectroscopic techniques, X-ray scattering does not accurately determine the proton positions and neutron scattering measurements, which are more accurate, are not time-resolved and hence it is not suitable to study the dynamic structural evolution of solute molecules in solutions.

This raises the need for an alternative approach to unravelling the solute-solute and solute-solvent interactions in the solution-phase and prior to nucleation to obtain experimentally determined structural models of these systems. X-ray Raman scattering (XRS) utilises X-rays in the hard X-ray region which circumvents challenges associated with soft X-rays. The deeper penetration depth for hard X-rays makes XRS a bulk sensitive technique without suffering from self-absorption effects of concentrated samples. This bulk sensitivity makes the technique suitable to study bulk structure of liquids [40, 113] and substances where surface oxidation is an issue [114]. Working in the hard X-ray region also enables conducting measurements at ambient conditions or at elevated temperatures and pressures [115]. XRS has been extensively used to study the structure of water and its condensed phases [45, 116], investigate hydrogen bond network of bulk water and aqueous solutions [40, 113] and explore the effect of isotope substitution on the local structure of water [117]. More recently, XRS has been applied to investigate the effects of organic solutes of the structure of water [118]. Bulk sensitivity and enabling measurements at ambient conditions make XRS the only

available techniques to study the molecular level chemistry of super-saturated solutions. Combining X-ray Raman scattering and total X-ray scattering will give an overview of the structural properties of the probed systems in real time and under real reaction conditions from the molecular level (provided by XRS) to the intermediate and longer radial range (provided by total X-ray scattering and the accompanied X-ray pair distribution function analysis). Total X-ray scattering and X-ray pair distribution function has been previously used to probe the intermolecular interactions in the liquid-phase and the chemistry of nucleation for inorganic systems [119-122]. The application of XRS, total X-ray scattering and XPDF analysis to characterise solvation and nucleation of organic molecules in solutions is a young field and studies applied to different systems are needed.

1.5 Imidazole and Self-Association in Solutions

Imidazole is a five-membered heterocycle that exhibits both aromatic and basic character. It has the following chemical formula $C_3N_2H_4$ (Figure 1.4) and a molecular mass of 68.08 $g.mol^{-1}$. The crystal structure of imidazole is monoclinic with $P2_1/C$ space group and four molecules in each unit cell [123-125]. The molecules in the crystal are connected by hydrogen bonds forming chains along the c-axis (Figure 1.4). The strong hydrogen bonds in imidazole were highlighted by a low frequency Raman spectroscopy study on a single imidazole crystal in 1968 [126].

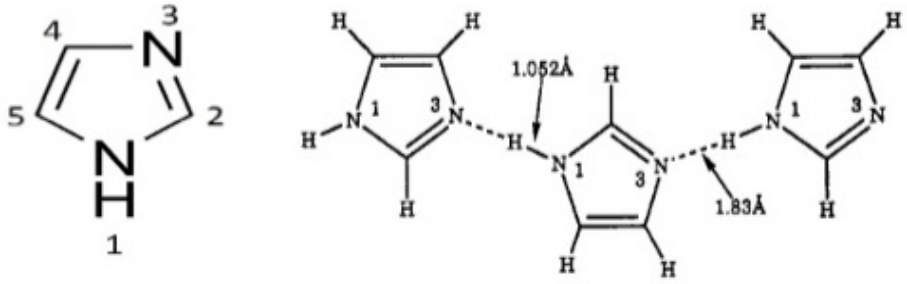

Figure 1.4. (Left) chemical representation of 1-h imidazole. (Right) Hydrogen-bonded imidazole molecules [127].

Understanding the solvation of imidazole in aqueous environments, including solute-solvent and solute-solute interactions, is of special interest due to its biological and pharmaceutical relevance. Imidazole is the side chain of histidine, one of the 20 naturally occurring amino acids, where it plays an important role in enzyme catalysis [128] and as

a metal binding ligand [129-135]. It has been used in the investigation of histidine adsorption on metals and metal oxides [136-138] and in the elemental speciation of amino acids [139]. Imidazole is also one of the two constituents in purine nucleobases, which are the building blocks for RNA and DNA. Derivatives of imidazole are part of antibacterial, antifungal, antiprotozoal and anthelminthic medications [140, 141], and are also potential anticancer agents [142]. Imidazole has been utilised as a spectroscopic marker to elucidate the effect of protonation, methyl substitution and H/D exchange [143] and in the study of natural organic matter in soil and sediments [144, 145]. More recently imidazole containing molecules have been used in making electronic materials with field effect carrier mobility and p-type semiconductor properties [146] and in the production of nanocellulose which is a promising bio-based building material [147].

The solubility of imidazole is very high in water with concentration of 11.02 $mol.L^{-1}$ achievable at 298.5 K [148]. At high concentrations the availability of water molecules becomes limited which leads to the self-association of the solute molecules. The intermolecular interactions in aqueous imidazole solutions have been investigated since the 1940s and evidence for self-association of imidazole molecules at concentrations above 10^{-4} $mol.L^{-1}$ has been established [149]. The first IR study of imidazole solutions showed that the strong absorption band of the non-bonded NH moiety decreased as the concentration of the solution increased, which indicated self-association of imidazole molecules [150]. Other IR studies of Imidazole in non-polar solvents also showed evidence of self-association in solution [151]. Ultraviolet spectroscopy [86] and energy dispersive X-ray diffraction [152] indicated the presence of stacked π-π assemblies of imidazole molecules in aqueous solutions.

The binding preference of water molecules to imidazole has been investigated by supersonic jet FTIR with O18 substitution and it was found that the water molecules prefer to act as hydrogen-bond donors (O-H…N) [153]. Nuclear magnetic resonance (NMR) has been extensively used to elucidate the structural properties of imidazole and imidazole containing molecules as well as the proton transfer mechanism [154-159]. However, the signal from the two pseudoequivalent nitrogen atoms in imidazole is not distinguishable by ^{15}N NMR [154-156] due to the low time resolution of NMR (10-5 seconds), which cannot capture the rapid tautomeric exchange between the nitrogen moieties (picoseconds). X-ray spectroscopic techniques have also been utilised to

examine the local environment around imidazole in aqueous solutions and to probe the changes induced by phase transition [88, 160], solvation [88, 92, 93, 160], and protonation [91, 161]. A particular problem with X-ray techniques is the insensitivity to hydrogen atoms. Neutron scattering is sensitive to scattering from all constituent atoms, C, N, O and H, and by measuring different isotopes one can create contrasts that allow determination of the structure and coordination around different atomic species in the solution. However, due to the limited availability of isotopes, obtaining all site-site correlations for multicomponent systems is usually not feasible. Neutron and X-ray scattering combined with molecular dynamics modelling have been used to study the solvation structure of neutral and protonated imidazole [110], and the related solutes pyrazole [111] and pyridine [112] in aqueous solutions.

Spectroscopic, scattering and diffraction techniques have so far provided limited molecular level information, or they only provide qualitative information on intermolecular interactions. Quantitative structural analysis has therefore been attempted by utilising electronic structure calculations to identify structural models representative of the interactions present in the solution. Theoretical DFT studies using explicit and implicit solvation models gave poor representation of the intermolecular interactions in aqueous imidazole solutions [88, 92, 160]. This is because of the difficulty associated with the construction of an appropriate solvation model, particularly for complex systems or for systems with complex electronic structure. Other simulation approaches such as molecular dynamics (MD) have been utilised to probe the structural and dynamic properties of imidazole in aqueous solutions [162, 163]. In MD simulations, the atomic configurations are generated by integration of the classical Newtonian equations of motion and empirical potential functions (i.e. force fields) are used to provide the interatomic forces required for such calculations. Quantitative and qualitative differences in the predicted local structures by MD simulations arise from differences in the description of the force fields used [162, 163]. Moreover, it is challenging to construct these empirical potential functions in the first place. especially for complex systems.

In this work a number of scattering and spectroscopic techniques combined with theoretical calculations have been used to examine the solvation structure and the

nature of the intermolecular interactions of imidazole in aqueous and ethanolic solutions.

1.6 Aims and Objectives

This project aims to provide further insight on the solvation structure of solute molecules (namely imidazole) in solution and the self-association of these molecules prior to crystallisation. This will consequently provide more understanding of the nucleation mechanism. X-ray Raman spectroscopy (XRS) and total X-ray scattering were the main techniques chosen to study the solvation structure and the self-association of imidazole molecules in solutions. XRS is analogous to X-ray absorption spectroscopy in terms of providing element specific information on the local bonding environment around the probed atoms with the advantage of being bulk sensitive. This feature enables *in-situ* probing of solutions, which is crucial for nucleation studies. On the other hand, total X-ray scattering and the associated X-ray pair distribution functions (XPDF) can provide information on shorter radial distances between molecules and on the average longer-range in the crystalline or amorphous structures. Imidazole was chosen to be studied in this work because of its relevance to biological and pharmaceutical studies. High concentrations of imidazole in solution can be achieved, which increases the signal/noise ratio above the detection limit for XRS and total X-ray scattering. Moreover, there is a wealthy pool of literature on experimental and computation studies on the structural properties and the self-association of imidazole in solutions (see section 1.5), which is crucial for comparison.

The project also aims to develop a code capable of automating the analysis of X-ray absorption spectroscopy (XAS) data. This code should be capable of performing data reduction and peak fitting as well as creating interfaces with other software packages to calculate theoretical spectra to be compared with the experimental spectra. This will enable assigning experimental spectral features to specific electronic transitions.

The two main aims of this project are to be achieved by following a stepwise approach expressed by the following objectives:

1. Probe the intermolecular interactions between solute and solvent molecules in concentrated aqueous solutions of imidazole: this was achieved by:

- Utilising neutron and X-ray diffraction and the accompanied empirical potential structure refinement (EPSR) simulation to understand the solvation structure of the solution and to generate a structural model of the system.
- Studying the effects of the intermolecular interactions on the C and N K-edge spectra of imidazole in gas, aqueous solution and solid phases.
- Combining the previous two studies to verifying the structural model generated by neutron diffraction. This is achieved by comparing the experimental K-edge spectra to the ones calculated for the structural model predicted by neutron diffraction.

2. Elucidate the effects of the variations in supersaturation on the structural properties of aqueous imidazole solutions, which would provide further insight on self-association in solutions. This was achieved by using XRS and total X-ray scattering (combined with X-ray pair distribution function, XPDF, analysis) to probe aqueous solutions of imidazole at different concentrations and during cooling crystallisation. The crystallisation experiments will also give insight on the structural evolution of imidazole in aqueous solutions prior to nucleation.
3. Investigate the effect of the solvent on the solvation structure by repeating the experiments listed in objective 2 for ethanolic solutions of imidazole.
4. Utilise the data analysis experience acquired throughout the project in developing a code for automating the analysis of X-ray absorption (XA) data.

1.7 Outline of Thesis Chapters

Chapter 1 gives an introduction to hydrogen-bonding, solvation, nucleation mechanisms and some characterisation techniques utilised in those fields. It also provides a brief overview on imidazole (the molecule studied in this work) and covers previous developments in understanding intermolecular interactions of imidazole in the solution-phase.

Chapter 2 describes the theoretical background for the experimental techniques used in this thesis including infrared spectroscopy, X-ray Raman scattering (XRS), total X-ray scattering, X-ray pair distribution function (XPDF) and neutron scattering. An introduction to the theoretical molecular simulations utilised to compliment the

experimental results is also described. These simulations include the empirical potential structure refinement (EPSR) simulations and electronic structure calculations using ORCA. Chapter 2 also discusses the design of the flow cells used to conduct the *in-situ* crystallisation experiments.

Chapter 3 concentrates on understanding the solvation structure of concentrated aqueous solutions of imidazole using neutron and X-ray scattering combined with the EPSR simulation. The results of this chapter are the corner stone in understanding the results of the following chapters. Hence, the structural model of the concentrated aqueous solution of imidazole suggested in chapter 3 is verified in chapter 4 by using that model to calculate the C and N K-edge X-ray absorption spectra and compare the calculated spectra with the experimental spectra.

Chapter 5 reports on the variations in the structural properties of aqueous imidazole solution as a function of concentrations using infrared spectroscopy, C, N and O K-edge XRS spectra, total X-ray scattering and XPDF analysis. The results of this chapter paved the way for the following chapter (Chapter 6) which probes the structural evolution of imidazole in aqueous solutions during cooling crystallisation. XRS and XPDF were the experimental techniques of choice to conduct the study in chapter 6 as they enabled conducting *in-situ* time-resolved experiments.

Chapter 7 examines the structural properties of imidazole in ethanolic solutions as a function of concentration and during cooling crystallisations. This was conducted as a comparison to the previous studies in aqueous solutions to understand the effect of the solvent on the intermolecular interactions in the solution.

Chapter 8 covers additional work on the development of a code for automating the analysis of X-ray absorption spectroscopy (XAS) data. Currently, the analysis is done manually and across a number of platforms and software packages, hence this code could potentially encourage the use of this technique.

Finally, chapter 9 concludes the thesis by addressing the impact of this work and outlining future work that could build on it.

Chapter 2 Experimental and Theoretical Techniques

This PhD research investigates the structural speciation of imidazole in aqueous and ethanolic solutions at different concentrations and during cooling crystallisation to shed light on the nature of intermolecular interactions in the solution phase and prior to nucleation. The solvation structure of imidazole solutions was investigated using infrared and X-ray core-level spectroscopy as well as X-ray and neutron scattering. Excited-state density functional theory calculations and reverse Monte Carlo simulations were carried out to generate models for the interpretation of the core-level spectroscopy and scattering experiments, respectively. In order to understand the results correctly, it is important to understand how the techniques works. This chapter presents the background for the experimental techniques and theoretical calculations employed in this PhD research. It also discusses the design of the flow cells used for the X-ray *in-situ* experiments. Details regarding specific approaches are given within the relevant chapters.

2.1 Materials

Imidazole used for the experimental work was obtained from Sigma Aldrich (puriss.p.a grade) with purity ≥ 99.5% without further purification. We noticed that some supplied batches contained small amounts of insoluble matter; such batches were not used for these experiments. D-imidazole (used for the neutron diffraction experiments, chapter 3) was obtained from QMX Laboratories Ltd. with chemical purity of ~98 %. Laboratory grade deionised water and ethanol absolute (≥ 99.5%) were the solvents used. The solutions of imidazole were prepared by dissolving imidazole in either water or ethanol without further modification or purification.

2.2 Characterisation Techniques

To determine the relative positions and motions of atoms in a bulk sample or a liquid, we need to “see” inside the sample. Using conventional microscopic techniques (i.e. light microscopy) to gather interatomic information about the matter is not suitable because the sample has to be transparent and the wavelength of the incident beam has to be on a finer length scale than the interatomic distances. The different regions of the

electromagnetic (EM) spectrum interact differently with the matter as each portion of the EM spectrum have a quantum energy appropriate for the excitation of a certain type of a physical process (Figure 2.1). For example, the energy of infrared radiation is sufficient to increase the vibrational activity of molecules while higher energy ultraviolet and X-ray radiation is capable of exciting valence or core-level electrons respectively (Figure 2.1). All physical processes at the atomic and molecular levels have quantised energy levels, and if these energy levels do not match the quantum energy of the incident radiation, then the material will be transparent to the radiation (i.e. the radiation will pass through).

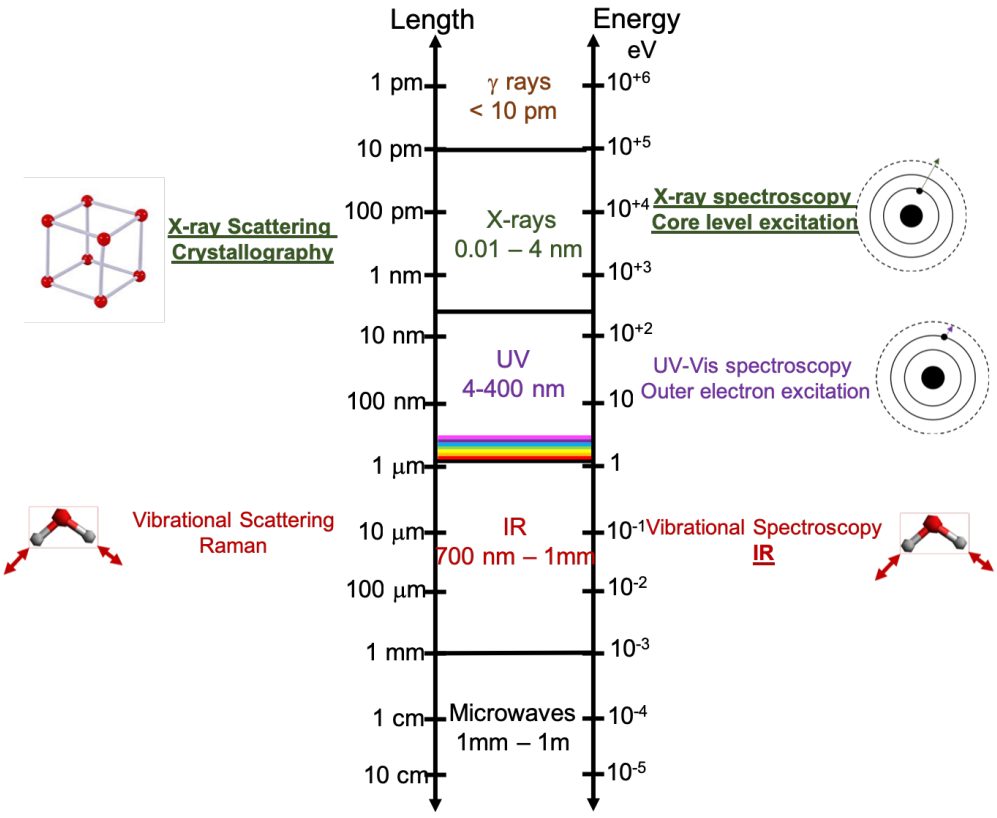

Figure 2.1. Interactions of electromagnetic spectrum with matter. Underlined techniques were used in this thesis.

X-ray radiation has a wavelength shorter than that of visible light; hence it can be used to probe matter at the atomic level. X-rays interact electromagnetically with the electron cloud surrounding the nucleus of atoms through absorption or scattering [164] (Figure 2.2). Absorption occurs when a photon (X-ray) gives all its energy to an atom, enabling the transition of a core electron to a higher energy level. In scattering a photon impinges on the matter interacting with the electrons and a photon is emitted in a

different direction to the direction of the incident photon. The scattering is elastic when the kinetic energies of the emitted and the incident photons are equal and inelastic when the energies are different. Since X-rays interact with electrons, they scatter more efficiently off heavy atoms (i.e. atoms containing many electrons) than light atoms (such as oxygen or even worse hydrogen).

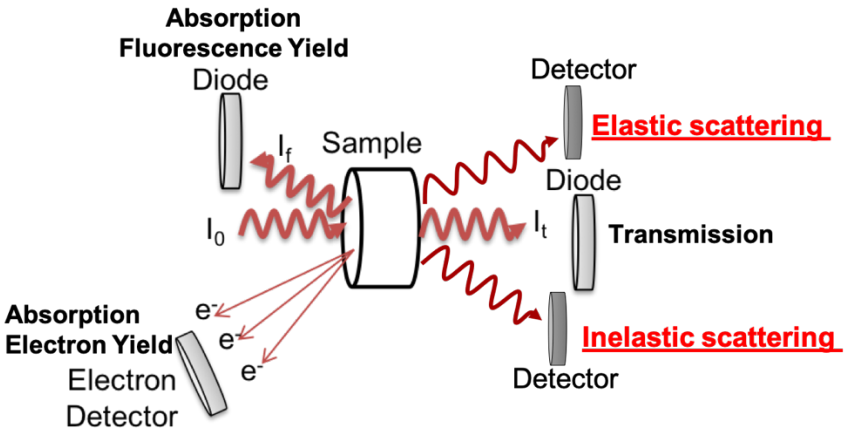

Figure 2.2. Interaction mechanisms of X-rays with the matter.

Neutrons have no charge and their electric dipole is too small; hence they can penetrate the matter better than charged particles. They scatter off the atomic nucleus unless the atom has unpaired electrons. Hence the residual magnetic moment of the atoms will interact with the magnetic moment of the neutrons giving extra scattering in addition to the nuclear scattering. The high penetration power of neutrons is advantageous as it enables neutrons to go deep through the sample even if the sample was contained. A container is necessary for liquid or gaseous samples or if the temperature and pressure is to be controlled. However, the consequence of having to go through a container is that the neutrons are weakly scattered once they pass through since they lose some of their intensity [165]. Another drawback of using neutrons is the challenges associated with detection. Detection of subatomic particles require them to interact with other particles; only few atoms (such as helium-3 and lithium) can strongly interact with neutrons through absorption to produce ionising particles [165]. Inherent low flux of neutron sources is a third drawback of using neutrons for matter characterisation. Neutron sources give 10^{-14} less particles per second per square millimetres than X-ray synchrotron sources [165]. The combination of weak interaction with matter, complex detection and low flux makes neutron characterisation techniques “signal-limited” [165].

The discussion above suggest that using one characterisation technique is not sufficient to fully characterise complex systems because each technique only probes one aspect of the system. Hence, combining different characterisation methods gives complementary information in most cases. Infrared spectroscopy, X-ray absorption spectroscopy and total neutron and X-ray scattering will be discussed further below as they are the techniques employed in this thesis.

2.2.1 Infrared Spectroscopy

The absorption of infrared radiation by samples induces transitions between the vibrational energy levels. Many chemical functional groups have characteristic vibrational frequencies; hence, infrared spectroscopy can be used to identify different species and their concentrations in the measured sample and to probe the bonding nature in the sample; such as hydrogen-bonding. Changing the dipole moment of molecules during vibration and having sufficient energy to induce transition to the next allowed vibrational energy level are the two governing selection rules for infrared excitation [87]. The useful range of infrared spectra from an organic chemistry point of view is between 15.4 – 2.5 µm (650 – 4000 cm^{-1}) [87].

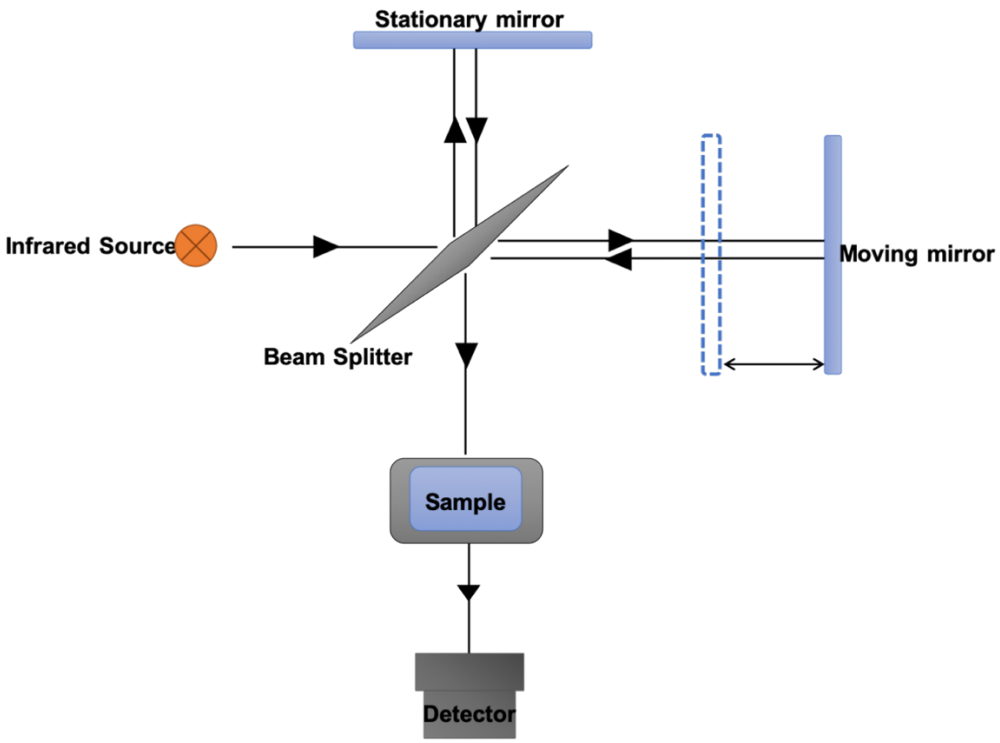

Figure 2.3. Michelson interferometer.

Most IR spectrometers used today employ Fourier transform method (FTIR) because the whole spectrum is measured within seconds eliminating the need to scan each wavenumber successively (like in the older IR spectrometers) [166]. A Fourier transformed infrared spectrometer is based on a Michelson interferometer [166] (Figure 2.3) which depends on the interference of the infrared waves. At the interferometer, an infrared source emits radiation throughout the whole frequency range (typically 4000 – 400 cm^{-1}) [166]. This radiation is divided into two beams of equal intensity by a beam splitter optic. The now split beams are directed along two orthogonal arms where they reflect off mirrors (one moving and one stationary) and are returned to the beam splitter where they are recombined and directed at the sample. Some of the radiation is absorbed by the sample and the remaining passes to the detector. Upon the recombination of the two beams, an interference pattern (interferogram) is generated which is the sum of all the interference patterns created by each wavelength of the incident radiation with respect to time. The interference pattern changes by systematically changing the difference in the two paths (through changing the position of the moving mirror). The detected signal is modified by selective absorbance of some frequencies by the sample. The interferogram is then Fourier transformed from time domain to frequency domain and the conventional FTIR spectrum is generated by taking the ratio of the Fourier transformed sample interferogram over that of the background.

The FTIR spectrum is sometime plotted in absorbance which is calculated from the following equation:

$$A = \log (I_0/I) \quad \text{(equation 2.1) [166]}$$

Where A is absorbance, I_0 is the intensity in the background spectrum and I is the intensity in the sample spectrum. Absorbance is also related to the concentration of molecules in a sample via Beer's Law:

$$A = \varepsilon \mathrm{l} \mathrm{c} \text{ (equation 2.2) [166]}$$

Where ε is absorptivity, l is the path length and c is the concentration.

The FTIR spectrum can also be plotted in percent transmittance (%T), which measures the percentage of light transmitted by the sample and is calculated using the following equation:

$$\%T = 100 \times I/I_0 \text{ (equation 2.3) [166]}$$

The four main FTIR sampling methods are transmission, attenuated total reflectance (ATR), diffuse reflectance (DRIFT) and specular reflectance [166]. The first two will be discussed further as they were the sampling methods used in this thesis.

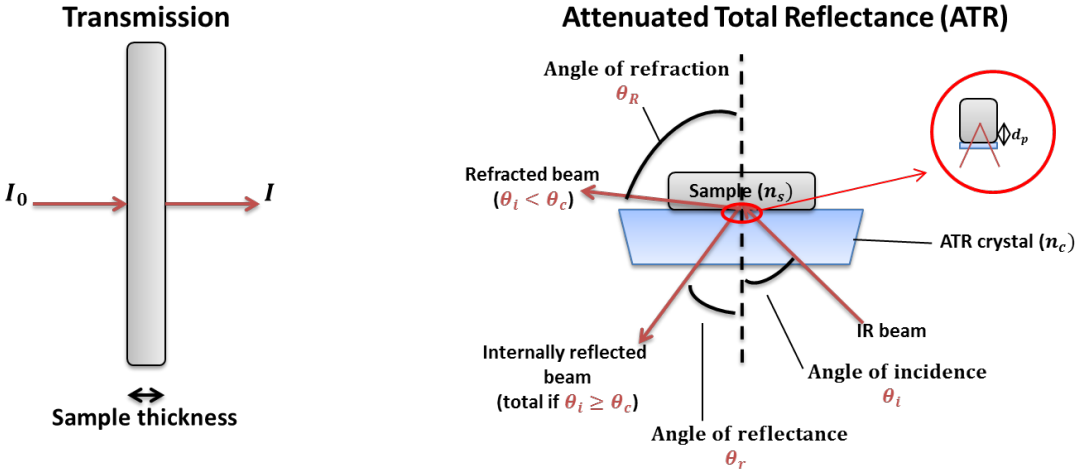

Figure 2.4. Illustration of FTIR transmission (left) and attenuated total reflectance (ATR) (right) sampling modes.

2.2.1.1 Transmission Fourier Transform Infrared (FTIR)

In transmission sampling mode, the infrared light passes through a thin film of the sample and then impinges on the detector (Figure 2.4). Since the bulk contribution is emphasised in transmission mode as the beam passes through the entire bulk of the sample, this mode is preferred for quantitative analysis.

Measuring liquid samples in transmission mode was done using the liquid cell shown in Figure 2.5. A drop of liquid is placed between two IR transparent windows (BaF_2 windows were used in this case) to create a thin film. The thickness of the film is determined by the PTFE spacer placed between the two windows. For the experiments presented here a 6 µm spacer was used. The sample holder is then placed in the FTIR sample compartment for measurement.

Using too much sample is one of the challenges associated with transmission measuring mode as it could lead to saturation of the peaks in the FTIR spectra [166]. Measuring aqueous solutions could also impose challenges because the peaks corresponding to water are broad and strong and can mask features that appear in the same spectral range [166]. Hence, for this work the 6 µm spacer was used to minimise the amount of sample and consequently reduce the saturation of the peaks in the FTIR spectra.

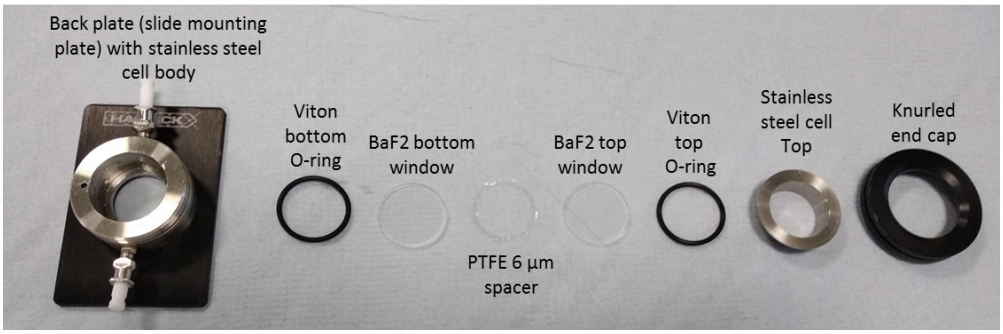

Figure 2.5. The assembly of the demountable cell used for transmission FTIR experiments.

2.2.1.2 Attenuated Total Reflectance - Fourier Transform Infrared (ATR-FTIR)

ATR-FTIR is based on internal reflectance [166] as illustrated in Figure 2.4. The infrared beam travels through a crystal of high refractive index (n_c) encountering a boundary with the sample of low refractive index (n_s). If the angle of incidence (θ_i) is small, some of the beam will reflect off the internal surface of the crystal and some of the beam will refract out of the crystal (Figure 2.4). As the angle of incidence increase, the angle of refraction will increase and at a certain θ_i value, the angle of refraction (θ_R) will be 90° which means that the refracted beam will not leave the crystal and all the incident beam will reflect off the internal surface of the crystal and hence the name "total reflectance". The minimum angle at which total internal reflectance occurs is called the critical angle (θ_c) and it depends on the refractive indices of the crystal and the sample as seen in equation 2.4

$$\theta_c = sin^{-1}(n_s/n_c)$$ (equation 2.4) [166]

At the point of total reflectance, the incoming and outgoing beam occupy the same volume and under the right conditions ($\theta_i \geq \theta_c$) they undergo constructive interference resulting in an infrared amplitude at that point which is greater than the amplitude of either of those beams. Since the point of total reflectance is near the surface of the

crystal, this enhanced amplitude has nowhere to go but up which results in the infrared beam striking the space above the crystal surface called “evanescent wave” or “hot spot” where the sample sits. This beam is absorbed (attenuated) by the sample and then focused on the detector.

Any sample can be brought into contact with the hot spot, a liquid will easily wet the surface and in the case of solids pressure is applied to the sample to obtain good contact between the sample and the hot spot. In the past, ATR crystals were shaped into long thin rods or parallelograms, which resulted in multiple hot spots. The sensitivity of the experiment was subsequently increased due to the increase in the path length (the total path-length is the sum of depth of penetration of the beam into the sample at each of these hot spots) [166]. However, the size of the ATR crystals reduced with the introduction of diamond crystals due to the high cost of diamond; hence most of diamond ATR-FTIR spectrometers contain only one hot spot [166].

For the experiments presented in this thesis, a Nicolet 6700 ThermoScentific FTIR was used; solid samples were measured in ATR mode and liquid samples were measured in transmission and ATR modes. Omnic software was used to collect the background and the sample spectra. The spectra were collected at a resolution of 4 cm^{-1} and 32 scans were accumulated in the 4000 – 500 cm^{-1} range.

2.2.2 X-ray Spectroscopy

2.2.2.1 Theory

As mentioned previously, X-rays interact with the matter through absorption or scattering. Different X-ray spectroscopic techniques are widely used across physics, chemistry, materials, engineering, environmental, geology and biology research to study the structure of materials on the atomic and molecular level. Some of these spectroscopic techniques are illustrated in Figure 2.6 and described in Table (2.1) [167].

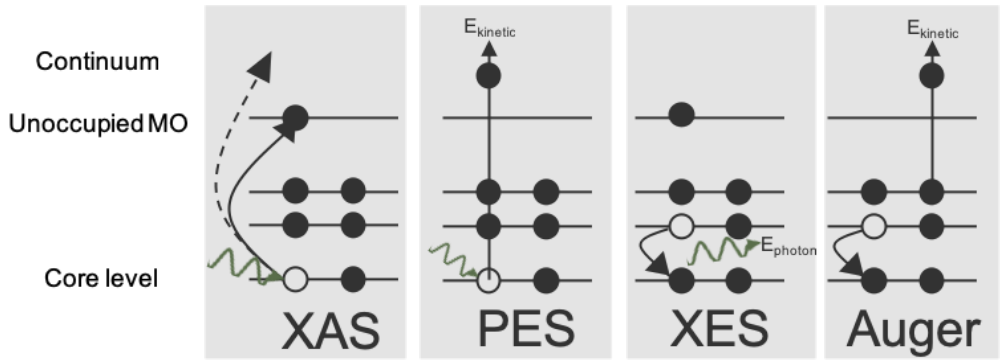

Figure 2.6. Schematic of different X-ray spectroscopic techniques.

Table 2.1. X-ray spectroscopic techniques

Technique	Mechanism
X-ray Absorption Spectroscopy (XAS)	upon X-ray irradiation, a core electron is transferred to one of the unoccupied electronic states or to the continuum creating a core hole. XAS will be discussed further below.
Photoelectron Spectroscopy (PES)	PES occurs when a photon (X-ray) is absorbed by a sample and it ionises and emits an inner-shell electron, which kinetic energy is of interest.
X-ray Emission Spectroscopy (XES)	After the creation of a core hole by X-ray absorption, a higher energy level electron fills the hole and the difference in energy is emitted as a photon or an electron. In XES, the emitted photon energy is of interest as it is characteristic to the transitions revealing information about the excited state electronic structure. The emission spectroscopy has two forms; Non-Resonant X-ray Emission Spectroscopy (NXES) and Resonant Inelastic X-ray Spectroscopy (RIXS). NXES is also called X-ray Raman Scattering which will be described further below.
Auger Electron Spectroscopy	similar to XES; however, the energy difference resulting from core-hole annihilation is emitted as an electron and the kinetic energy of the emitted electron (auger electron) is of interest.

X-ray Absorption Spectroscopy (XAS)

XAS operates by directing a monochromatic X-ray beam on the sample with the wavelength of the beam ramping through the absorption edge of a certain element in the sample and the absorption coefficient is measured. Absorption coefficient measures how strongly X-rays are absorbed as a function of their energy. It decreases smoothly with increasing X-rays energy as higher energy X-rays have higher tendency to penetrate materials [164, 168]. However, sudden increases in the absorption coefficient (X-ray absorption edges) are noticed at certain energies characteristic to the atoms in the material.

Three experimental modes can be used to measure the absorption coefficient: transmission, fluorescence yield (FY) or electron yield (EY) [169, 170] (Figure 2.2). Transmission experimental mode is the most straightforward mode where the transmitted radiation is measured and compared to the incident radiation. Nevertheless for most studies transmission is impractical when soft X-rays are used since very thin samples (~1000 Angstroms) are required to allow operation in transmission experimental mode [171]. The other experimental modes for measuring the absorption coefficient (FY and EY) detect the decayed products (photons, electrons) emitted when a core hole is filled with a high energy level electron. FY can be measured as Total Florescence Yield (TFY) where all or a broad range of photon energies are detected or Partial Florescence Yield (PFY) where an energy sensitive detector detects fluorescence from a selected range of energies [170]. FY is a bulk sensitive experimental mode whereas EY is a surface sensitive mode [171]. Both modes exhibit distortions; however, the distortions associated with EY are less significant than those associated with FY. The distortions associated with FY are caused by self-absorption of emitted photons, which suppresses the absorption peaks [171-173].

The absorption spectrum is typically divided into two regions: Near-edge X-ray absorption fine structures (NEXAFS) (or X-ray absorption near-edge structures (XANES)) and extended X-ray absorption fine structures (EXAFS) as illustrated in Figure 2.7. The term “XANES” is usually used by solids and inorganic compounds community while the term “NEXAFS” are conjugated with the surface analysis community [168]. NEXAFS/XANES acquire narrow energy ranges close to the absorption edge and are used

to determine local structures and oxidation states [174]. On the other hand, EXAFS acquire energy ranges further from the absorption edge, which give information on the local environment surrounding the absorbing atom (atom types, distances between probed atom and neighbouring atoms and disorder) [174]. From this point on, NEXAFS terminology will be used to describe the region near the absorption edge.

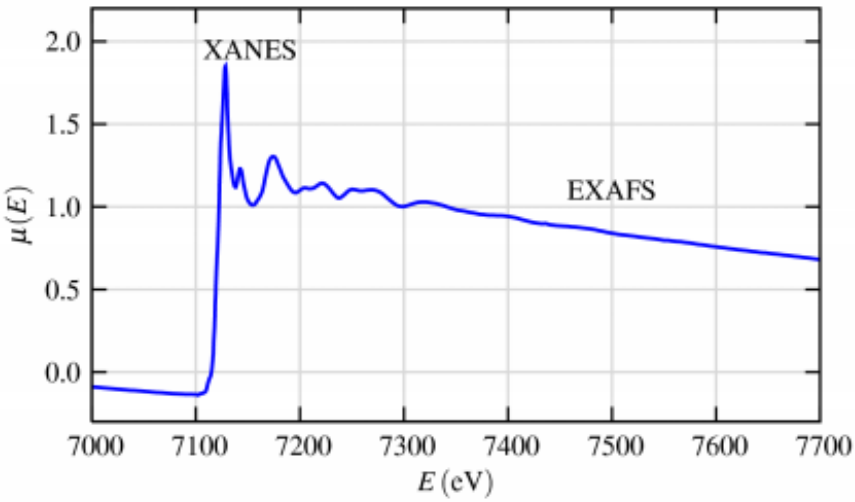

Figure 2.7. XANES (NEXAFS) and EXAFS regions on a measured XAS spectra for FeO [175].

X-ray Raman Scattering (XRS)

XRS is a non-resonant inelastic hard X-ray scattering technique. The cross section for XRS is smaller than that of XAS; hence, beamlines with high X-ray brilliance and large solid angle detection are required [176]. XRS spectrometers are based on spherically bent analyser crystals that act as focusing monochromators with energy resolution in the sub eV range for photon energy of ~ 10 keV. These analyser crystals provide the large solid angle detection required. In XRS, the incident photons with energy (ω_1) and momentum (k1) interact with the sample and the emerging photons have energy (ω_2) and momentum (k2) lower than that of the incident photons (Figure 2.8(a)). The energy ($\omega=\omega_1-\omega_2$) and the momentum (Q = k1- k2) are transferred to the sample. ω is tuned in the soft X-ray region near the absorption edges by varying ω1.

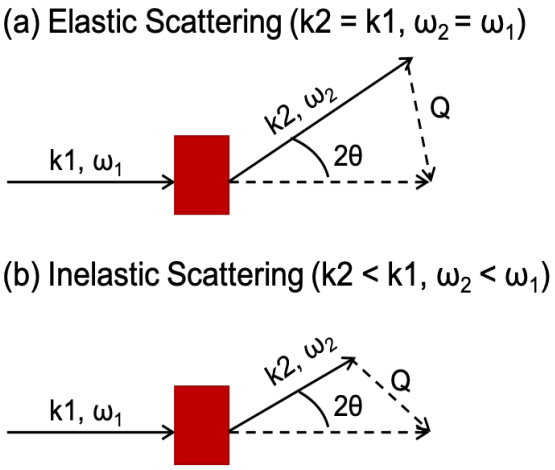

Figure 2.8. Scattering principal (a) elastic (b) inelastic.

XRS probes the double differential cross section [177] (equation (2.5)). The first term in equation (2.5) is Thomson scattering cross section and the second term is the dynamic structure factor

$$\frac{d^2\sigma}{d\Omega d\omega} = \left(\frac{d\sigma}{d\omega}\right)_{Th} S(q,\omega) = \left(\frac{d\sigma}{d\omega}\right)_{Th} \sum_f \left|\langle f\left|e^{iq.r}\right|i\rangle\right|^2 \delta(E_f - E_i - \hbar\omega) \quad \text{(equation 2.5)}$$

q is the momentum transfer, ħω is the energy difference between the incident and scattered photon, r is the position relative to the initial state and i and f label the electron initial and final states. Equation (2.5) gives a general cross section term for all non-resonant inelastic X-ray scattering events, including scattering from phonons, plasmons, valence electrons (Compton scattering) and core-level electrons (X-ray Raman Scattering). The momentum q is also given by equation (2.6) in which 2θ is the scattered angle.

$$\boldsymbol{q^2 = |q|^2 = \frac{1}{c^2}\left(\omega_1{}^2 + \omega_2{}^2 - \omega_1\omega_2 \cos(2\theta)\right)} \qquad \textbf{(equation 2.6 [176])}$$

At low momentum transfer (q), XRS and XAS probe the same transition matrix element for core-level electrons and hence have similar directional sensitivities [177]. The dynamic structure factor $S(q,\omega)$ for XRS in low q range (the dipole limit) is presented in equation (2.7) and is analogous to the X-ray absorption coefficient [177]. The advantage of using XRS over conventional XAS is its use of hard X-rays (~ 10 keV) which makes the technique bulk sensitive and unrestricted by the sample environment, especially the high vacuum requirements associated with soft X-ray techniques.

$$S(q,\omega) \rightarrow q^2 \sum_f |\langle f|\hat{q}.r|i\rangle|^2 \delta(E_f - E_i - \hbar\omega) \quad \text{(equation 2.7)}$$

XRS spectrum (Figure 2.9) consists of three main regions: the elastic Rayleigh scattering peak, the broad Compton spectrum and the XRS spectrum. The elastic Rayleigh scattering peak represents the energy loss measurements conducted by adjusting the spherical analysers so that the radiation is scattered elastically into the corresponding detectors at a defined energy value (elastic energy) ($\omega_1 = \omega_2$). In the Compton spectrum, the energy of the incident radiation is not high enough to excite core-level electrons; however inelastic scattering off valence electrons takes place. In the XRS spectrum region, the energy of the incident radiation is sufficient to excite a core-level electron. XRS spectra are affected by the scattering angle achieved in each experiment; hence the scattering angle from each analyser needs to be considered individually which imposes the main challenge associated with XRS data analysis.

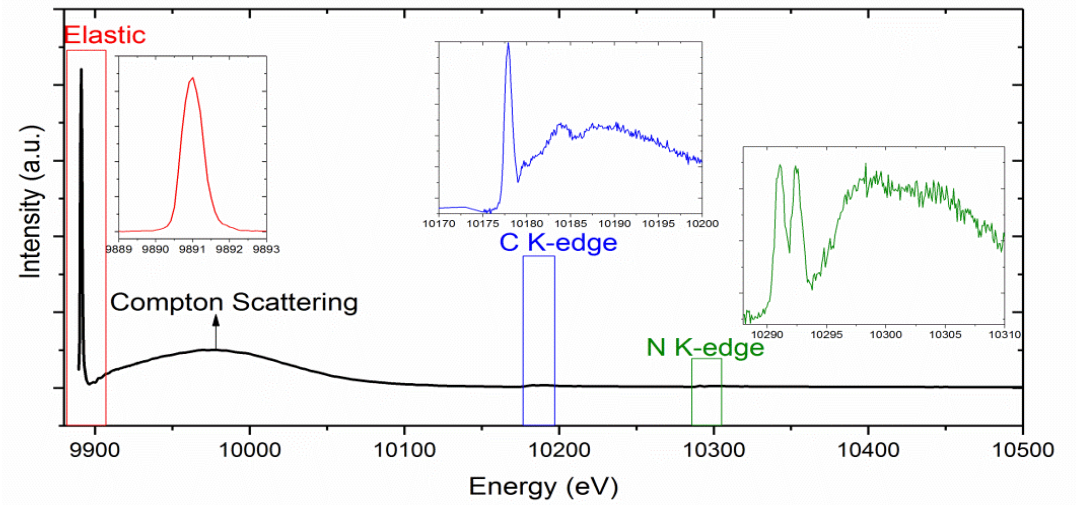

Figure 2.9. Rayleigh, Compton and X-ray Raman scattering profiles for crystalline imidazole.

2.2.2.2 X-ray Sources

X-rays are short wavelength electromagnetic radiation with the wavelengths in the order of angstroms and the energies in the order of keV [164]. X-ray radiation is artificially produced from particle accelerators which, as the name suggests, accelerate charged particles (in this case electrons) to high speed to generate X-ray radiation. Linear accelerators and synchrotrons are examples of particle accelerators used for generating X-rays.

Linear Accelerators (LINAC)

In linear accelerators (LINACs), charged particles are accelerated in a linear beamline to high speeds by subjecting them to an oscillating electric potential supplied by a radio

frequency (RF) generator, which is an AC voltage source. The energy of DC voltage-based particle accelerators is limited to few tens of MeV when compared to radio frequency (RF) linear accelerators [178]. The later can achieve voltage gains higher than the maximum voltage applied through repeated application of electric fields. LINACs consist of a straight hollow pipe maintained under vacuum so that the accelerated particles do not collide with air molecules, lose their energy and consequently decelerate. However, under vacuum electromagnetic waves (injected by the RF generator) cannot be used to accelerate particles because their polarity is continuously alternating causing the particles to gain or lose energy depending on the phase [178, 179]. Hence, the energy is delivered to the particles utilising resonant cavities (or RF cavities) which are cylindrical electrodes capable of maintaining particles' acceleration through providing an electric field with the correct phase and by grouping the accelerated particles into bunches [178, 179]. The field strength per cavity is limited to a few $MeV.m^{-1}$; hence, a number of RF cavities are needed to achieve high energies and this explains why high energy LINACs are very long machines [178]. Linear accelerators could be as small as cathode ray tubes or as large as the Stanford Linear accelerator (SLAC), which at 3.2 km was the longest linear accelerator in the world.

A typical LINAC has a particle source, cylindrical electrodes (RF cavities), electric oscillator and a target sitting inside the hollow pipe (Figure 2.10). The electrons from DC thermionic source (hot cathode) or photocathode are injected into the RF accelerating part of the LINAC. While the first few RF cavities are used to capture slow electrons, accelerate them to velocities close to the speed of light and group them in bunches, the following RF cavities are designed to accelerate relativistic electrons [178]. To generate X-rays, the now accelerated electrons collide with the target (a heavy metal not shown in Figure 2.10) where they excite electrons in the target material to unoccupied energy levels generating holes that are annihilated by higher energy electrons. A fraction of the energy difference is then emitted as X-ray radiation while the rest is dissipated as heat.

LINACs can accelerate particles to speeds not possible by ring-type accelerators (e.g. synchrotrons) and are consequently better to send electrons up to relativistic speeds. Ring-type accelerators are limited by the strength of the magnetic fields needed to keep electrons in the curved path and electrons keep losing energy as they travel in the ring. Nonetheless, constructing linear accelerators require a lot of land and hence is costly.

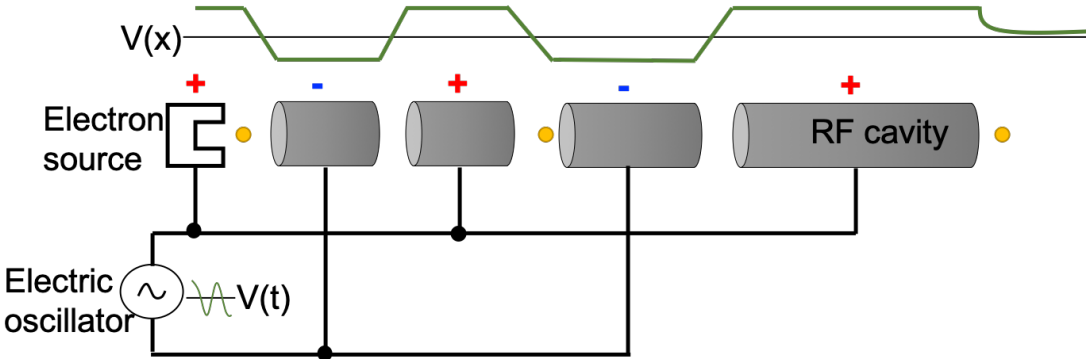

Figure 2.10. Schematic view of LINAC main components and how it operates.

Synchrotrons

Synchrotron radiation is emitted from electrons travelling at the speed of light when they are forced to change the direction of their motion under the influence of a magnetic field [180]. It has high brilliance (i.e. high brightness and angular speed), high energy and short wavelengths which enables probing samples at the molecular and atomic scales [181]. It can also be polarised to enable certain experiments and pulsed at high frequencies enabling time-resolved study of chemical reactions [181]. Therefore, synchrotron radiation is preferred over laboratory radiation to conduct experiments.

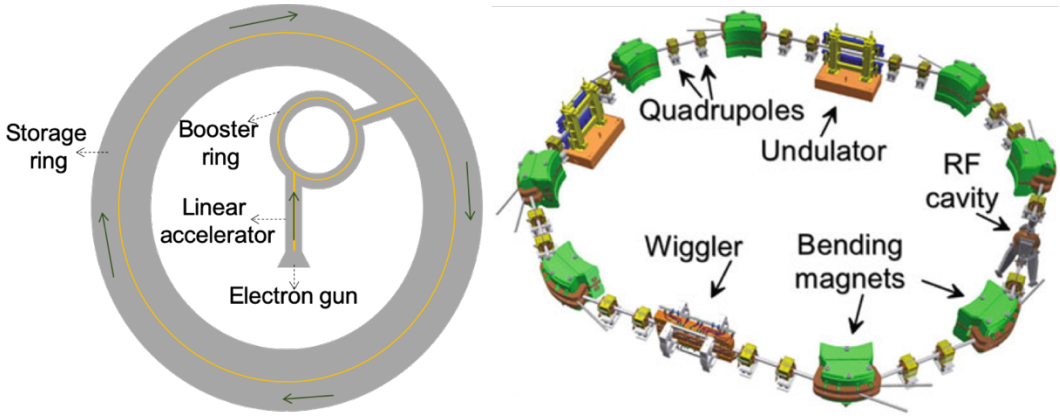

Figure 2.11. (left) Schematic planner view of a synchrotron radiation facility (right) schematic view of a storage ring from reference [180] showing main elements such as bending magnets, focusing quadrupole magnets, insertion devices (wigglers and undulators) and RF cavities.

One of the artificial sources of synchrotron radiation is high energy circular electron accelerators (storage rings). Before entering the storage ring, the electrons are injected from the electron gun and accelerated in a linear accelerator (LINAC) until their energy reaches several MeV (Figure 2.11 (left)). The energy of the accelerated electrons is boosted to billion eV (GeV) in a booster ring before they are transferred to the storage

ring (Figure 2.11 (left)). In the storage ring, the electrons are forced to follow a circular, oval-shaped or polygon-shaped path by utilising magnets placed along the circumference of the storage ring. Storage rings consist of arrays of magnets for focussing (e.g. quadrupole magnets), directing (bending magnets) and generating synchrotron radiation (bending magnets, wigglers and undulators) (Figure 2.11 (right)). Storage rings also contain a number of RF cavities to further accelerate the electrons (if needed) by utilising radio frequency electric fields and to compensate for the energy lost in the form of synchrotron radiation. Joining the storage ring are a number of beamlines and experimental chambers, which accommodate the different types of synchrotron radiation experiments.

Synchrotron radiation (SR) originated from particle accelerators where it was considered an inconvenience as the objective of these accelerators was to increase the energy of particles and the emittance of synchrotron radiation was a result of a loss in energy. First generation synchrotron radiation facilities were created to make use of the huge amount of radiation generated by particle accelerators [182]. The research utilising SR was a parasitic operation of these particle accelerators and hence it was limited as the machines were tuned to meet the particle physics experiments' requirements and not the SR production. As the SR user community grew, there was an increased pressure to build storage rings designed for the production of SR radiation [182]. The use of storage rings was an important development to the second generation SR sources because they deliver a stable SR spectrum due to current and energy stabilities [182]. Adding to that, they provide a beam that is available for more time since the beam is accumulated and left to circulate for many hours (energy losses by SR are compensated for by RF cavities in the storage ring) [182]. The SR source at Daresbury laboratory in the UK was the first second generation SR source [182]. The success of second generation SR sources, developed pressure for construction of sources able to produce more intense SR radiation and capable of covering hard X-ray energies [182]. This meant construction of sources with higher energy electrons and with insertion devices (i.e. wigglers and undulators) to produce more intense X-ray beams [182]. Using more sophisticated magnetic field devices resulted in a noticeable increase in the brightness of the synchrotron sources (Figure 2.12). Examples of third generation SR sources are: Diamond Light Source UK, European Synchrotron Radiation Facility (ESRF) France and

Advanced Photon Source (APS) USA. Fourth generation synchrotron sources are being developed which are based on free electron lasers (FEL) that are capable of producing very short coherent light pulses that have very high intensity and brightness.

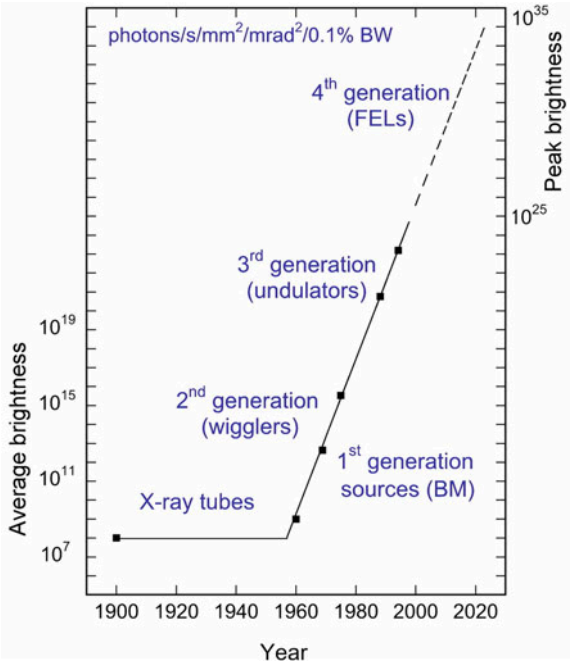

Figure 2.12. Average brightness of SR sources of different generations [180].

2.2.2.3 X-ray Raman Scattering Instruments

European Synchrotron Radiation Facility (ESRF), ID-20

The X-ray Raman scattering (XRS) measurements of imidazole aqueous solutions at different concentrations (presented in chapters 4 and 5) and of ethanolic solutions of imidazole (presented in chapter 7) were collected at the inelastic scattering beamline ID-20 of the ESRF, Grenoble, France [183]. ID-20 is an undulator beamline that uses Si(111) pre-monochromator which can be coupled by a consecutive monochromator to form a fixed exit ensemble. A Si(311) post-monochromator was used to achieve an energy resolution of 0.6 eV needed for this experiment. All data were collected using the inverse scanning mode, where the incident photon energy was scanned and the analyzer energy was fixed at 9690 eV. The scattered radiation was collected using 72 spherically bent Si(600) analyser crystals grouped into 6 vertical and horizontal chambers (Figure 2.13) based on the scattering angle. The data presented in chapters 4, 5 and 7 was from

the VD chamber (containing 12 analyser crystals) at an angle of 42° with q value of 3.6 $Å^{-1}$. Hence, the data presented here were collected at low momentum transfer.

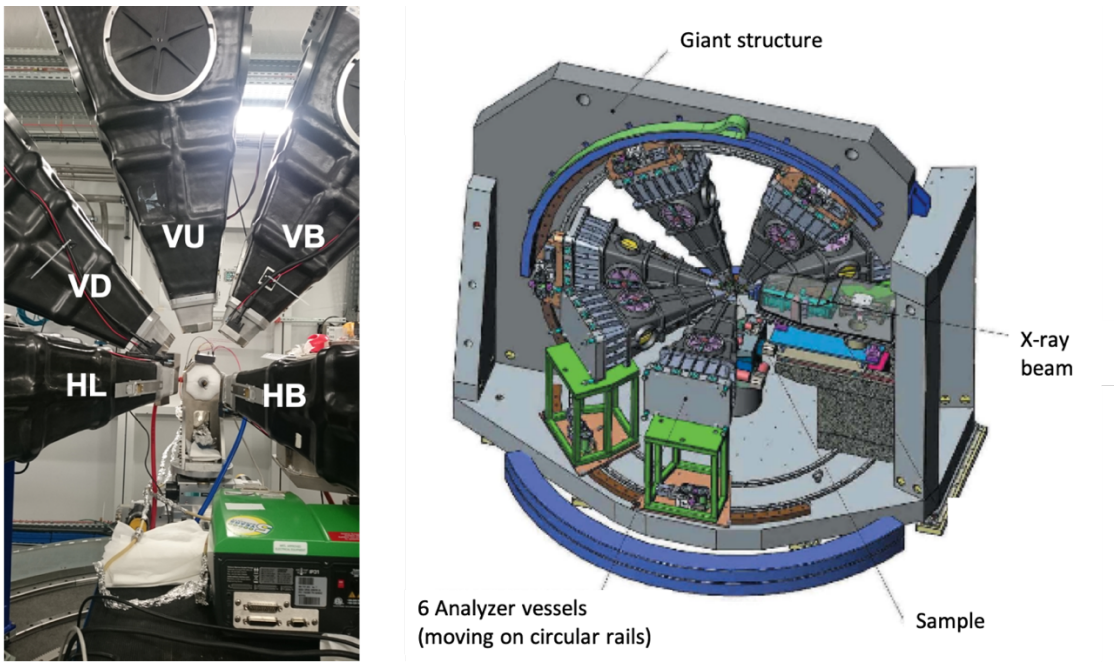

Figure 2.13.(right) ID-20 end station setup (left) Schematic of spectrometer setup from reference [183] showing the six movable crystal analyser chambers, each hosting 12 analyzer crystals. Three chambers are movable in the horizontal (H) plane while the other three move in the vertical (V) plane.

Advanced Photo Source (APS), 20-ID

The XRS measurements of the imidazole aqueous solutions during cooling crystallisation were performed at sector 20-ID of the Advanced Photon Source (APS), Chicago, USA, using the low-energy resolution inelastic X-ray scattering (LERIX) spectrometer [184, 185]. 20-ID is an undulator beamline equipped with liquid nitrogen cooled Si(111) and Si(311) monochromators. The Si(311) monochromator was used as it achieves high spectral resolution which is needed for this study. The combined instrumental resolution of the monochromator and the analyzer crystal was ~0.5 eV.

The monochromatised beam (L in Figure 2.14 (right)) was focused using a toroidal mirror to a spot size of 600 µm (horizontal) and 400 µm (vertical). The scattered radiation (M in Figure 2.14 (right)) was analyzed with a combination of Si diced and undiced spherically bent crystal analyzers (SBCA). The SBCAs were tuned to

use the Si(555) reflection, and the scattered radiation (N in Figure 2.14 (right)) was detected using NaI scintillators coupled to custom-built amplifier and discriminator electronics. Similar to the experiments at the ESRF, all data were collected using the inverse scanning mode where the incident photon energy was scanned and the analyzer energy was fixed at 9891 eV.

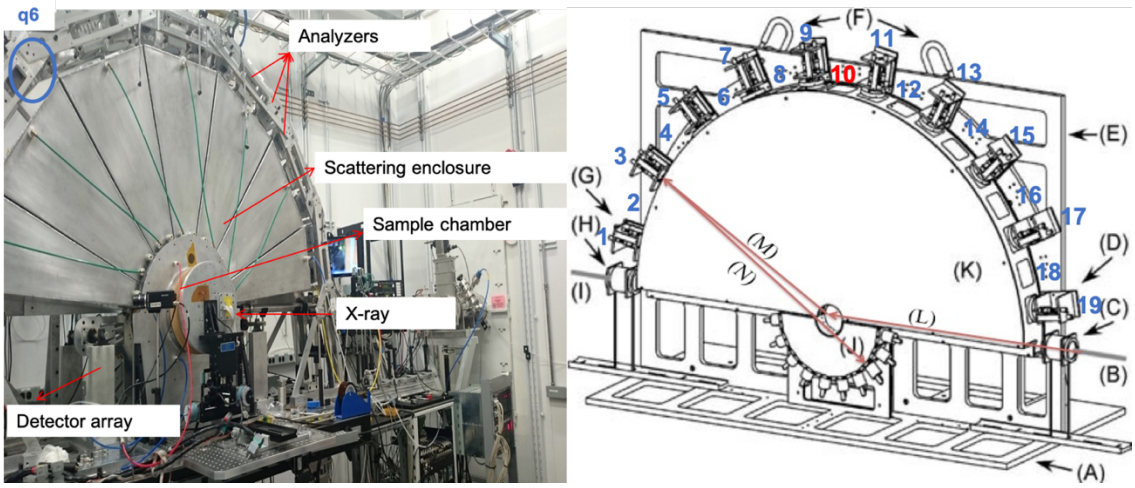

Figure 2.14. (right) LERIX end station setup (left) Schematic of LERIX setup from reference [184] where (A, E) support frame, (B) incoming X-ray, (C) gas ionization chamber to monitor the incident intensity, (D,G) analyzers 19 and 1, (H) gas ionization chamber to monitor the transmitted intensity, (I) transmitted X-ray beam, (J) detector assembly and (K) Helium filled scattering enclosure.

Data were collected from 18 of the 19 crystal spectrometers (identified in blue in Figure 2.14 (right)). The data presented this thesis (chapter 6) were taken from *q6* diced crystal analyzer which is one of the low momentum transfer (*q*) analyzers with the best signal to noise ratio. This way, the data were collected at a scattering angle of 54° with *q* value of ~4.6 $Å^{-1}$. The monochromator drift was small (±0.137 eV) and it was monitored through frequent measurement of the elastic scattering peak. The monochromator was detuned by 15% to suppress the contribution from the harmonic component of the beam.

2.2.3 Total Scattering

2.2.3.1 Theory

Over the past 100 years, scattering has been a fundamental technique for many scientific advances [186, 187]. Each material has a characteristic scattering pattern that

contains information on structure, geometry, bond lengths and atomic positions. X-rays were first produced and recorded by Roentger in 1895 [186, 187]. Diffraction of X-rays by crystals was first observes by Max von Laue in 1912 where photographic plates were used to record the diffraction pattern [186, 187]. In the same year, William H. Bragg and William L. Bragg published Bragg's law the key for using diffraction in interpreting crystal structures [186, 187]. They assumed that X-rays diffract off layers in crystals which work as semi-transparent mirrors (Miller planes). Diffraction is possible because the distance between these layers is of the same order as the wavelength of the incident X-ray radiation. Neutron radiation was discovered in 1930s; 35 years after the discovery of X-rays [187] and the first neutron diffraction experiment was performed in 1946 [186].

During a total scattering experiment, a monochromatic beam (of wavelength λ) impinges on the sample and elastically scatters into scattering angle 2θ (Figure 2.8(a)). The data is measured versus the momentum transfer (Q) which is calculated using equation 2.8. The wavelength of the incident beam is selected by:

- Utilising a monochromator for the incident radiation.
- Measuring the travel time taken by the incident beam to go from source to detector (such as in the case of time of flight neutron source)
- Detecting the energy of the scattered radiation (such as in the case of energy dispersive X-ray scattering).

$$Q = \frac{4\pi \sin(\theta)}{\lambda} \quad \text{(equation 2.8)}$$

In powder diffraction, Bragg peaks provide information on the average structure. Liquids have no long-range order meaning that the Bragg peaks are absent and to understand short-range ordering total scattering (Bragg and diffuse scattering) is essential. The scattering intensity between and underneath the Bragg peaks is called diffuse scattering (Figure 2.15) and it gives information on the local atomic correlation and on any deviations from the average structure.

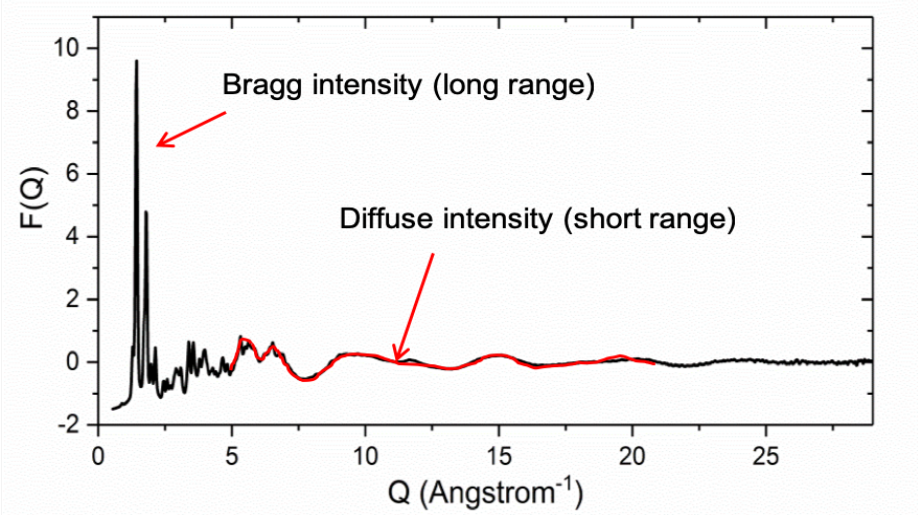

Figure 2.15. Total scattering pattern of solid imidazole showing Bragg and diffuse scattering intensities.

The differential cross section ($\frac{d\sigma}{d\Omega}(Q)$) is the quantity obtained by scattering after the relevant corrections are applied [188]. For neutrons, it is the sum of single atom cross section and interference differential cross section (equation 2.9)[188]. The first term in equation (2.9) arises from interferences between waves from the same atom while the second involves interferences of waves between different atoms.

$$\frac{1}{N}\left[\frac{d\sigma}{d\Omega}(Q)\right] = \sum_{\alpha}^{n} c_\alpha \overline{b_\alpha^2} + F(Q) \qquad \text{(equation 2.9)}$$

N is the number of atoms in the system, Q is the vector between the incident and scattered radiation, c_α is the atomic fraction of species α and b_α^2 is neutron diffraction lengths. F(Q) is the interference differential cross section and is the weighted sum of the partial structure factors $S_{\alpha\beta}(Q)$(equation 2.10) [188]. The partial structure factor ($S_{\alpha\beta}(Q)$) is related to $g_{\alpha\beta}(r)$ (the radial distribution functions RDFs) through Fourier transformations (equation 2.11).[188] The RDFs gives the distances between pairs of atoms in real-space.

$$F(Q) = \sum_{\alpha\beta}^{n} c_\alpha c_\beta \overline{b_\alpha}\,\overline{b_\beta}\left[S_{\alpha\beta}(Q) - 1\right] \qquad \text{(equation 2.10)}$$

$$S_{\alpha\beta}(Q) = 1 + \frac{4\pi\rho}{Q}\int r\left[g_{\alpha\beta}(r) - 1\right]\sin(Qr)\,dr \qquad \text{(equation 2.11)}$$

ρ is the atomic number density of the sample

X-ray diffraction also provides a measure of F(Q) and the partial structure factors ($S_{\alpha\beta}(Q)$) but these are now weighted by both the atomic concentration and the form factor (which is Q dependent) because X-rays do not scatter from point-like atomic nuclei like neutrons but from extended electron density distribution [188]. While the electron density at atoms and hence the X-ray scattering cross sections increase with atomic number, there are considerable variations of neutron scattering cross sections across the periodic table due to their more complex dependence on nuclear properties. Hence, the X-ray and neutron scattering techniques are often complementary and their combination can increase the contrast between the measured F(Q) in the determination of the partial structure factors $S_{\alpha\beta}(Q)$ [188].

An effective way to describe the average local structure of disordered systems is the atomic pair distribution function (PDF) which is the Fourier transformation of the total scattering patterns. PDF is a real-space histogram of interatomic distances where the position of the peaks corresponds to bond lengths and the area under the peaks gives the coordination numbers. The reduced atomic pair distribution function G(r) gives the number of atoms in a spherical shell from a reference atom (equation 2.12).

$$G(r) = 4\pi r(g(r) - \rho_0) \quad \text{(equation 2.12)}$$

g(r) is the pair distribution function related to the local density of atoms at distance r from a central atom and ρ0 is the average atomic number density.

The pair distribution function (G(r)), accessible from diffraction data using X-rays, neutrons or electrons, is generated through Fourier transformation of the structure factor F(Q) (equation 2.13) where F(Q) is related to the coherent elastic part of the X-ray scattered intensity.

$$G(r) = (2/\pi)\int_{Q=Q_{min}}^{Q_{max}} Q(F(Q) - 1)\sin(Qr)\,dQ \text{ (equation 2.13)}$$

The real-space information provided by PDF is similar to that obtained from EXAFS but PDF is not element specific, it can provide information to a longer range and

it is easier to calculate [189, 190]. Transmission electron microscopy (TEM) is also a powerful tool in providing information on atomic correlations but it is a projection along one axis meaning that it provides limited information on 3D atomic arrangement [189].

In total scattering experiments, the diffuse scattering is of interest, but it has weak intensity compared to the intensity of the Bragg peaks. Hence it is important to measure the background of the sample environment to be subtracted from the measured data to remove as much of the background noise as possible. The experiment limitation of finite Q (Qmax) introduces truncation effects in G(r); this majorly affects the data in real-space [189]. Hence, a large Q value minimises the contribution of this artefact. Q is the wave vector of the momentum transfer of incoming and outgoing radiation and its magnitude (equation 2.8) is proportional to θ (half of the angle between the incoming and outgoing radiation) and 1/λ (where λ is the wavelength of the incident radiation) [191]. Therefore, collecting over a wider range of 2θ and using short wavelength radiation will improve the quality of the PDF generated.

2.2.3.2 Neutron Sources

The two primary sources for neutrons will be listed in this section. Man-made sources for X-ray radiation have been discussed in section 2.2.2.2.

Neutron Reactors

Neutron reactors produce neutrons through neutron induced nuclear fission. In the process, a nucleus (uranium or plutonium) captures a neutron and fissions producing two daughter nuclei, at least two high energy (2 MeV) neutrons and a lot of kinetic energy (200 MeV for each fission event). These neutrons are very energetic to sustain a nuclear reaction and hence the nuclear fuel is typically surrounded by a moderator to slow the neutrons and enable them to sustain the chain reaction. The moderator is made of hydrogenous material; such as water or heavy water. Neutron absorbing control rods can be readily inserted into the reactor to shut the chain reaction.

Spallation Sources

In spallation sources, neutrons are generated by accelerating bunches of protons in a linear accelerator (LINAC) or a synchrotron (or both) to sufficiently high energies (~800 MeV at ISIS, the UK neutron spallation source). Once the protons reach the set energy, the collide with a heavy metal "target" nucleus producing highly excited nuclear states. These states decay either immediately or after a delay by emitting nuclear particles such as neutrons, γ-rays and neutrinos. The maximum energy of the neutrons produced correspond to the energy of the proton beam used and hence neutrons from spallation sources are more energetic than those generated from nuclear reactors.

A uranium target can produce up to 30 neutrons per impinging proton but using uranium results in a significant delayed neutron background that can be challenging to account for in the diffraction data and it significantly offsets the increased neutron flux. Hence, non-fissioning targets such as tantalum or tungsten are used in most spallation sources nowadays although they produce half the number of neutrons produced by uranium.

2.2.3.3 Total Scattering Instruments

ISIS Spallation Source, NIMROD

The neutron total scattering experiment was conducted at NIMROD (Near and Intermediate Range Order Diffractometer), which is one of the end stations at target station 2 in ISIS neutron and muons spallation source at the Rutherford Appleton Laboratory in Oxfordshire, UK. The neutrons come through the incident flight line (Figure 2.16). Neutrons should be slowed so that their wavelength is short and comparable to the separation between atoms in the probed molecules. Helium and solid methane are used as a moderator to slow down the neutrons. Some of the slowed down neutrons will scatter off the sample (most of the incident neutrons will penetrate the sample without moderation and will be absorbed by the beam stop wall at the far end of the assembly, see Figure 2.16). The diffracted neutrons are detected by ZnS scintillators, which are placed in the detector blockhouse, which is held in vacuum to prevent neutron from interacting with air. The detectors are arranged in two arrays; wide-angle detector array to cover scattering angles between 3.5° to 40° and low-angle detector bank to cover scattering angles between 0.5° to 2.2°.

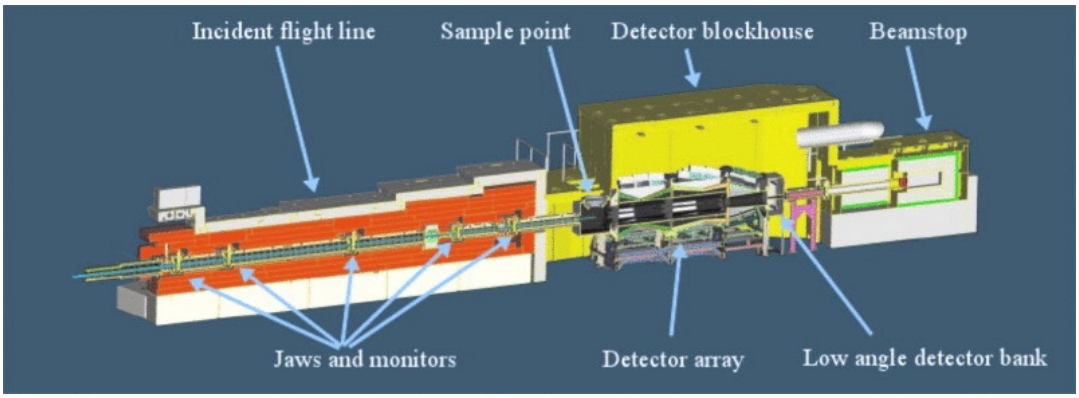

Figure 2.16. NIMROD instrument setup.

The sample holder is made of TiZr alloy because of its high neutron transparency and chemical stability in the presence of hydrated samples. A thin layer of the sample is injected inside the sample holder. A number of sample holders are placed in an assembly which can hold up to 10 sample holders and which temperature is controlled using an oil bath circulator (for the experiment presented in this thesis, the temperature was maintained at 25°C).

Diamond Light Source, I15-1

Total X-ray scattering data (presented in chapters 5, 6 and 7) were collected at the high energy XPDF beamline I15-1 (Figure 2.17) at Diamond Light Source [192]. The energy used for these experiments was 76 keV (λ=0.163 Å). A large area Perkin Elmer 2D detector located close to the sample was used, given access to Qmax= 40 $Å^{-1}$. Instrument calibration was done using silicon wool. The solid samples were measured in capillary mode while the liquid samples were measured using Flow-cell 2 (presented in section 2.4 below).

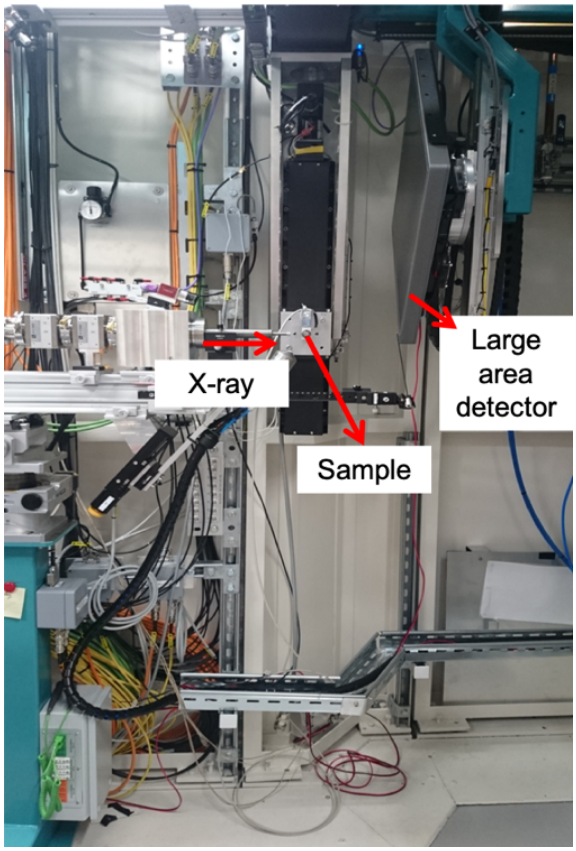

Figure 2.17. I15-1 XPDF beamline end station setup.

2.3 Theoretical Calculations

2.3.1 Empirical Potential Structure Refinement (EPSR)

"Small-box" and "Big-box" approaches are the most common analytical methods for processing total scattering data [193]. As the name suggests, one way to distinguish these methods is by the size of the model; the small box approach consists of a few hundred unit-cells while the big box approach consists of thousands of unit cells. The small box approach is based on a real-space Rietveld refinement of a structural model with symmetry restrictions. The use of a structural model keeps the number of refined parameters to a minimum; however, this approach struggles to deal with situations where the interaction of short-range and long-range order are important. Another disadvantage of using small-box approach is the need to provide a structural model before the refinement which means that this approach cannot be used with disordered materials. PDFGui [194] is one of the most commonly used small box programs which refines the local structure using a least-square method and where periodic boundary conditions allow calculation of the PDF at distances larger than the unit cell used. TOPAS

[195] is another software that applies the small-box approach to refine PDF data. The program fits the data to local and average structure at the same time and it is faster and more stable than other programs [195]. In this thesis, PDFGui was used to refine solid-state PDF data, which was used as a reference for the analysis of solution phase PDF data.

The big-box approach attempts to unify the local and average structure by using a supercell larger than the correlation length of disorder in the material. The most common way to create big-box models is by using Monte Carlo (MC) algorithms. The idea of MC methods is to use random rather than deterministic movement of atoms and deciding whether the move is accepted or rejected based on the minimisation of the configuration energy [196]. Reverse Monte Carlo (RMC) methods were proposed with the aim to minimise the difference between experimental and numerically determined scattering data [197]. RMC methods does not require suggesting a starting model for the disordered state and it attempts to build the structural model of the scattering system with minimal prior information. However, it is not safe to rely solely on the scattering data to supply all the required structural constraints [198]. Moreover, there are usually far fewer X-ray and neutron isotopes available compared to the number of partial structure factors needed to define the structures investigated by total scattering experiments [198]. Hence, imposing harmonic constrains on the structures of molecules available in the studied system by assuming an initial guess for any interatomic forces (repulsive, dispersive, Coulombic) is a way of restricting the structural model [198]. The difference between experimental and simulated data is then used to perturb these forces so that the simulation reproduces the scattering data as close as possible. The empirical potential structure refinement (EPSR) is one of the programs capable of doing that.

EPSR is a reverse Monte Carlo (RMC) simulation code where a random configuration of the system is generated. A Cubic box of a defined side length is constructed with a certain number of solute and solvent molecules (depending on the system studied) and a specific atomic number density [199]. Bond and angle restraints are employed, as well as torsional restraints for the molecules in the system. However, it is worth noting that the molecules are not held rigid during the EPSR simulation in order to reproduce the experimental diffraction data [198].

The initial reference potential used by EPSR for each atom within the 3-dimensional model comprises of a Lennard-Jones potential together with a Coulomb charge to define the intermolecular interactions. The reference potential is required in EPSR to generate the initial configuration of atoms using a standard Monte Carlo simulation. After equilibration, an additional empirical potential is refined iteratively by perturbation of the reference potential and F(Q) is calculated for each new atomic configuration. The atomic configuration is either accepted or discarded according to the goodness of the agreement between the experimental and calculated F(Q). Finally radial, spatial and angular distribution of atomic and molecular pairs were calculated using routines within EPSR and Dlputils [200].

Other RMC software include RMC Modelling [201] and RMCprofile [202]. For this PhD thesis, the EPSR software was used to generate the structural model for imidazole (the molecule studied in this research project) in different solvents (water and ethanol) as EPSR has been extensively used for the study of water [198, 203-205], alcohols [206, 207] and aqueous solutions of organic molecules [102-109]. The reference potential parameters used in the EPSR simulation will be defined in the respective chapters.

2.3.2 ORCA

ORCA is a general purpose quantum chemistry software package that features all modern electronic structure methods including density functional theory (DFT), semi-empirical (based on HF formalism) methods and many-body perturbation and coupled cluster theories [208]. ORCA uses contracted Gaussian basis functions and it is most efficient with segmented contracted bases where the calculation speed is comparable to alternative programs. Closes-shell (RHF), spin-unrestricted (UHF), restricted open-shell (ROHF) self-consistent field (SCF) calculations on the basis of various DFT and HF methods are available. Generalised Gradient Approximation (GGA), meta-GGA, hybrid DFT(HDFT) and double hybrid DFT (DHDFT) functionals are available in ORCA.

In this thesis, ORCA has been used to calculate the theoretical K-edge X-ray absorption (XA) spectra of imidazole in gas, aqueous solution and solid-phases as well as calculating the molecular orbitals energy using DFT method. ORCA has been the quantum chemistry package chosen to be used in this project because of user friendliness, flexibility and

efficiency as well as being available free of charge (Version 3 and older) and for academic users (Version 4 and older).

The calculations were performed using ORCA version 3.0.3 and 4.1.0 , and they were carried out using the B3LYP functional [209]. 6-31G* basis set was used for the monomeric imidazole molecule, while a minimally augmented def2-SVP and def2-TZVP basis sets [210] were used for simulating the XA spectra of solvated and solid-state imidazole. To speed up the calculations, the D3 dispersion correction [211] was applied to all calculations. RIJCOSX approximation [212] with AutoAux and def2/J auxiliary set was applied to the calculations for solvated and solid-state imidazole.

The structure of the monomeric molecule used to represent imidazole in the gas-phase was constructed in Avogadro version 1.1.1, a molecular builder and visualisation tool with ORCA support [213]. The intramolecular structure was initially optimised in Avogadro using the UFF force field and then further refined using ground-state geometry optimisation calculations in ORCA. Frequency calculations were run on the optimised structure to ensure that the structure corresponded to the global energy minimum. Finally, time-dependent DFT (TD-DFT) calculations were performed on the optimised structure to generate the theoretical near-edge spectra.

To simulate the water solvation shell around the imidazole molecule, both implicit and explicit solvation models were used, which will be discussed further in Chapter 4. To perform excited state simulations (TD-DFT) of solid-phase imidazole, a cluster of imidazole molecules was taken from the crystal structure of imidazole [125]. The geometry optimisation and frequency calculations for the solvated imidazole molecules and for imidazole in the solid-state were performed using the same procedure as described above for the monomeric imidazole.

2.4 Flow-cell Design

Sample environments commonly used in hard X-ray experiments include microfluidic cells,[214] closed-loop flow cells utilising peristaltic pumps [215, 216] and liquid jets [94]. Since cryogenic cooling is not an option for reducing radiation damage of the sample when studying liquids, circulating or flushing samples is required to reduce the contact time between the sample and the X-ray beam. The use of microfluidic cells is

one way of achieving this. However, their stock volume is small, which is convenient for precious or hazardous samples, but may not be sufficient for diluting the influence of radiation-induced chemical changes. Closed-loop flow cells often use higher stock volumes. However, maintaining the equilibrium temperature, and thus the metastable state of the supersaturated system, is challenging across a combination of a stock solution reservoir, flow tubing and the region where X-ray probing takes place, making these cells less attractive for crystallisation studies, which is one of the main focuses of this thesis.

We have therefore designed a different type of crystalliser [160] (Flow-cell 1 shown in Figure 2.18) built around a jacketed vessel of the type commonly used in laboratory-scale crystallisation studies. The design minimises the potential for radiation damage to the sample through circulation of a large stock volume (~250 ml) and provides the required temperature control, i.e. the thermodynamic driving force for crystallisation. It is therefore possible to follow the structural evolution of a solution from undersaturation through supersaturation to nucleation and crystal growth in a classic cooling crystallisation experiment (as shown in Chapter 6).

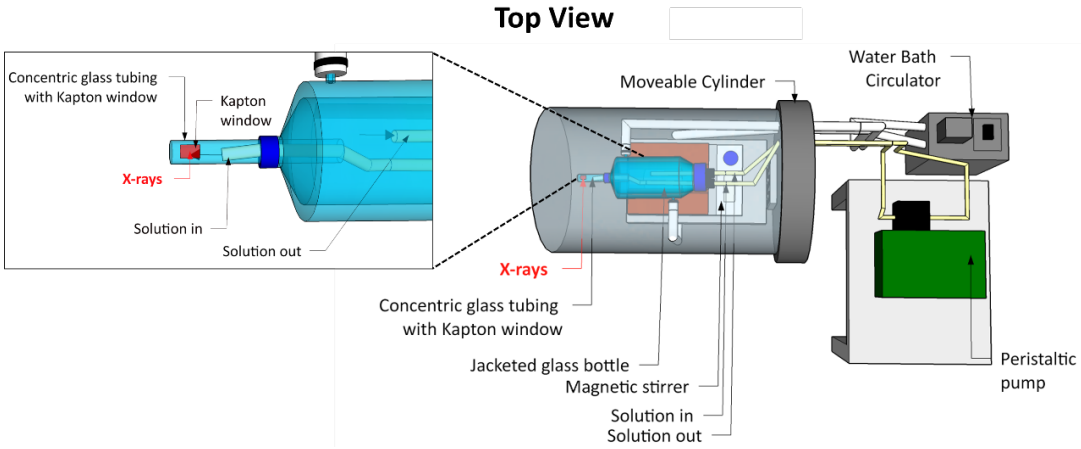

Figure 2.18. Flow-cell 1, the crystalliser cell used for XRS measurements of solution phase samples [160]. The shaded grey area is encapsulated in He gas for experiments at the APS and in air for experiments at the ESRF.

At the heart of the crystalliser is a standard jacketed glass vessel. The temperature of the solution is controlled through the temperature of the fluid in the jacket. We used water as the temperature control fluid, with a water bath circulator to control its temperature. This jacketed vessel was modified by fitting a system of two concentric

tubes to its bottom. The outer cylindrical glass tube has a closed end and was attached to the jacketed vessel using a screw cap connection. It has a Kapton window to allow probing of the sample volume directly underneath using X-rays. The Kapton window was fixed over a 10 mm wide cut-out slot covering a 180° arc. The solution flows continuously past this window and is recycled back into the main bulk of the cell via an inner concentric tube using a peristaltic pump at a flow rate of up to ~450 ml/min to minimise radiation damage. The inner concentric tube diameter is 3.2 mm, which is sufficiently large to minimise turbulence that may induce nucleation. In addition to internal solution circulation, magnetic stirring was used to ensure uniform temperature distribution in the stock solution within the jacketed vessel and to mimic stirring in a real-world crystalliser.

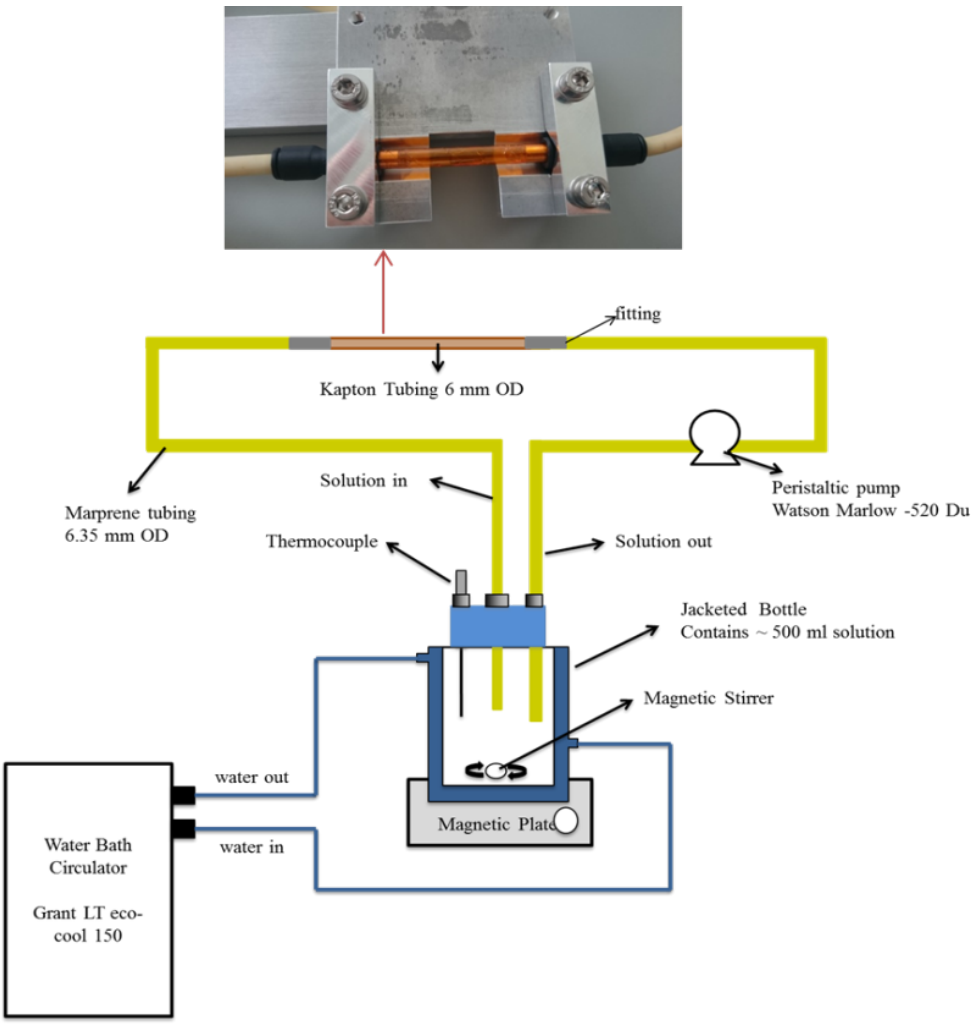

Figure 2.19. Flow-cell 2 which was used for the XPDF experiments.

Flow-cell 1 illustrated in Figure 2.18 was modified for the X-ray scattering experiments conducted at the XPDF beamline, Diamond Light Source (I15-1 beamline) (Figure 2.19).

Similar to Flow-cell 1, Flow-cell 2 consisted of a jacketed vessel which enabled controlling the temperature of the liquid samples through the temperature of the fluid in the jacket (in the case the temperature control fluid was water). The liquid samples were continuously circulated through a Kapton tube (with outer diameter of 6mm) using a peristaltic pump. The X-rays probed the sample volume flowing inside the Kapton tube. Continuous circulation of the sample ensured minimisation of beam induced damage and uniform temperature distribution. Magnetic stirring was also incorporated to ensure uniform temperature distribution inside the jacketed vessel and to mimic real-world crystallisers. Flow-cell 2 operated with similar flow rates (~450 ml/min) and sample volumes (250 ml) as Flow-cell 1.

Chapter 3 Solvation Structure of Imidazole in Aqueous Solutions

3.1 Introduction

The solvation of organic molecules in aqueous systems influences virtually all biological processes, with water being the dominant component of living organisms. Water forms hydration shells around soluble organic molecules and hydrophilic functional groups that may be at the surface of suspended objects. Likewise, when molecules aggregate or, in the case of biological macromolecules, when enzymes or substrates (e.g. drugs) interact with solvated molecules then desolvation takes place, i.e., water molecules are expelled from the hydration shell. These hydration and dehydration processes are mainly governed by the competitive breaking and forming of hydrogen bonds in solvent-solvent, solute-solvent and solute-solute pairs. Unravelling the structure and dynamics of these hydrogen-bonding networks in multi-component and macromolecular systems is not trivial using conventional spectroscopic and diffraction techniques. However, examining the hydrogen-bonding of models for the constituent functional groups, for example in the aqueous solutions of small molecules (such as imidazole), can provide structural information that can then be extrapolated to more complex macromolecular systems.

As mentioned in Chapter 1, the intermolecular interactions in aqueous imidazole solutions have been investigated since the 1940s and evidence for self-association of imidazole molecules at concentrations above 10^{-4} mol.L^{-1} has been established [149]. A wide range of spectroscopic and scattering techniques have been utilised since then to probe the structural properties of solutions of imidazole; such as ultraviolet spectroscopy [86], energy dispersive X-ray diffraction [152], FTIR spectroscopy [153] and core-level X-ray spectroscopy techniques [88, 91-93, 160, 161, 217]. These spectroscopic and scattering techniques are limited in providing molecular level information. Electronic structure calculations [88, 92, 160] and molecular dynamics (MD) simulations [162, 163] were utilised to identify the structural and dynamic properties of aqueous solutions of imidazole. Theoretical

molecular simulation studies are very important in providing information on the thermodynamics of solvation; however, unless bias and directed they are hindered by the length and time scales provided by currently available computational ability [218].

This raises the need for an alternative representation of the solute-solute and solute-solvent interactions in aqueous solutions of imidazole that is more synonymous with the bulk structure of the solution than models generated from theoretical molecular simulations. Neutron diffraction is sensitive to C, N, O and H and by measuring different isotopes neutron diffraction gives the possibility to interpret the structure and coordination around different sites in the solution. However, due to the limited availability of isotopes, obtaining all site-site correlations for multicomponent systems is usually not feasible. Hence, the empirical potential structure refinement (EPSR), a reverse Monte Carlo computational method, can be used to model the measured diffraction data [198, 219]. In EPSR, experimental neutron diffraction data is directly used to refine the structural model. Neutron diffraction augmented by isotopic substitution and combined with EPSR simulations has been extensively used to probe the structure of water [198, 203-205] and the intermolecular interactions in aqueous solutions [102-109, 220]. In this chapter, this technique has been employed to probe the solvation structure of 5 M aqueous imidazole solutions and the nature of the intermolecular interactions between solvent and solute molecules.

3.2 Methodology

3.2.1 Neutron Diffraction with Isotopic Substitution

Aqueous solutions of imidazole with concentration of 5 M were prepared with a series of H/D isotopomeric contrasts (see Table 3.1 for the isotopic composition of measured samples). The concentration was chosen to enable isotopic substitution of the imidazole [221]. D2O (98% purity D) and H-imidazole (puriss > 99.5 %) were obtained from Sigma-Aldrich and D-imidazole was obtained from QMX Laboratories Ltd. Neutron diffraction data were collected for the seven samples (Table 3.1) using NIMROD at the ISIS neutron spallation source,

Rutherford Appleton Laboratory (UK), which is ideally suited to measurement of hydrogen containing liquids owing to the forward scattering arrangement of its detectors. The H/D isotopic substitution technique was employed in order to constrain the solvent diffraction contributions to the EPSR model as much as possible [222, 223]. The solutions were loaded into TiZr cells with a 1 mm path length and data were collected at 25 °C across the range $0.05 < Q < 50$ Å^{-1}. In order to produce F(Q) for each isotopic substitution of the solution, the raw data were normalised to a vanadium niobium plate and corrected for instrumental and TiZr cell backgrounds, absorption, multiple scattering and the effect of the incoherent scattering from hydrogen using GudrunN [199].

Table 3.1. The isotopic compositions of imidazole and water

Sample number	Imidazole	Water
1	D	D
2	50:50 H/D	50:50 H/D
3	H	H
4	50:50 H/D	D
5	H	D
6	D	H
7	D	50:50 H/D

3.2.2 Total X-ray Scattering

Total X-ray scattering data were collected for a 5 M aqueous imidazole solution using the PANalytical X'Pert Pro Diffractometer at ISIS, Harwell Oxford (UK), equipped with an Ag source. Data were collected in capillary mode, with the sample held in 2 mm inner diameter sealed quartz capillary. Data were collected across the range $0.03 < Q < 18$ Å^{-1}. The raw data were corrected for instrument and quartz capillary backgrounds, absorption, multiple scattering and Compton scattering using GudrunX [224].

3.2.3 Empirical Potential Structure Refinement (EPSR)

The EPSR simulations for the results presented in this chapter (only) were carried by Dr. Samantha K. Callear at ISIS neutron spallation source. The interpretation of the simulation results was done by me.

For this study, a cubic box of side length 60.916 Å was constructed with 5530 water molecules and 500 imidazole molecules, with an atomic number density of 0.093 atoms/Å [199]. The water molecules were constructed according to Soper [204] and imidazole molecules were constructed in Jmol and the intramolecular structure was minimised using MOPAC7 (PM3 calculation) [225]. Bond and angle restraints were employed, as well as torsional restraints for the imidazole molecule. However, it is worth noting that the molecules are not held rigid during the EPSR simulation in order to reproduce the experimental diffraction data [198].

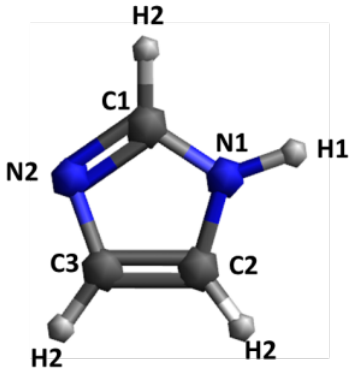

Figure 3.1. Atom labelling used for the Lenard-Jones potentials in the EPSR simulation of imidazole.

Table 3.2. Lennard Jones potentials and charges for water [226] and imidazole [227] The labelling scheme used for imidazole is shown in Figure 3.1.

Atom	ε / kJ mol-1	σ /Å	Charge / e-
Ow	0.650	3.166	-0.848
Hw	0.000	0.000	0.424
N1	0.711	3.250	-0.570
C2	0.293	3.550	-0.015
H1	0.000	0.000	0.420
C1	0.293	3.550	0.015
C3	0.293	3.550	0.295
H2	0.126	2.420	0.115
N2	0.711	3.250	-0.490

The initial reference potential used by EPSR for each atom within the 3-dimensional model comprises of a Lennard-Jones potential together with a Coulomb charge to define the intermolecular interactions (Table 3.2). The parameters for water were obtained from the SPC/E model [226] and those for imidazole were taken from the OPLS [227]. The reference potential is utilised by EPSR to generate the initial configuration of atoms using a standard Monte Carlo simulation. After equilibration, an additional empirical potential is refined iteratively by perturbation of the reference potential and F(Q) is calculated for each new atomic configuration. The atomic configuration is either accepted or discarded according to the goodness of the agreement between the experimental and calculated F(Q). Finally radial, spatial and angular distribution of atomic and molecular pairs were calculated using routines within EPSR and Dlputils [200].

3.3 Results and Discussion

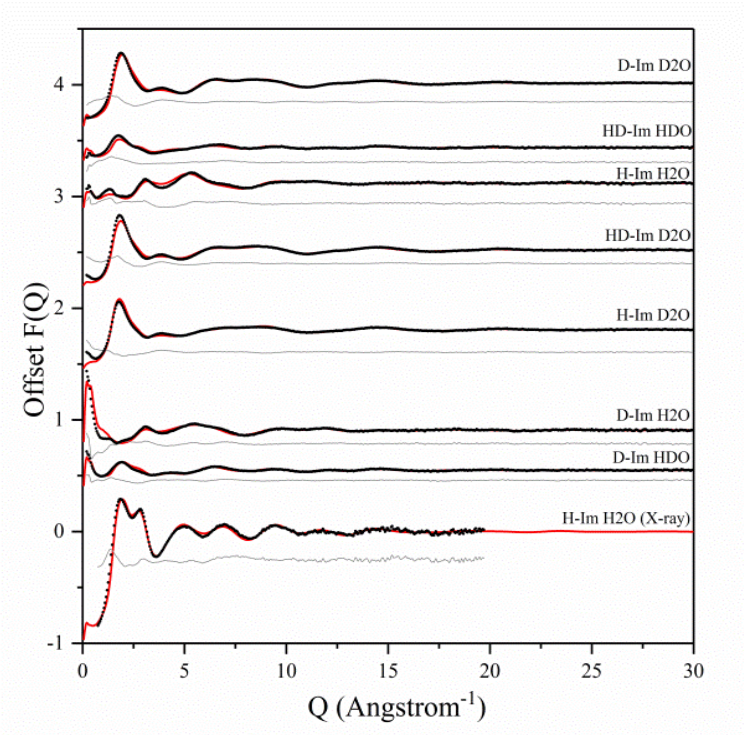

Figure 3.2. F(Q) values from experimental scattering data (black dots) and EPSR model (red line) and the difference between the two patterns (grey line).

The interference differential cross section (F(Q)) values obtained from the neutron and the X-ray scattering data (experimental data) and from the EPSR structural model (calculated) are shown in Figure 3.2 alongside the difference between the experimental

and calculated values. For all the datasets, the calculated F(Q) values from the EPSR fits well with the experimental data, with small discrepancies at the low Q values (< 2.5 Å-1) where the background and the inelasticity corrections are hardest to remove due to the hydrogen presence.

3.3.1 Water – Water Interactions

The radial distribution functions (RDFs) for water-water atomic pairs in the aqueous solutions of imidazole are similar to those of pure water (Figure 3.3)). There is a slight increase in the intensity of the RDF peaks from the imidazole solution when compared to pure water, which could be attributed to the excluded volume effect [228]. In solutions, the presence of the solute increases the local water density, resulting in a higher average density over the whole volume. Comparison of the coordination numbers for the water-water atomic pairs in the imidazole solution (listed in Table (3.3)) reveals that they are lower than those reported for pure water [204]. This deviation from pure water can be attributed to the larger volume of imidazole molecule (compared to water) and to the fact that the solution is fairly concentrated (1:10 imidazole:water).

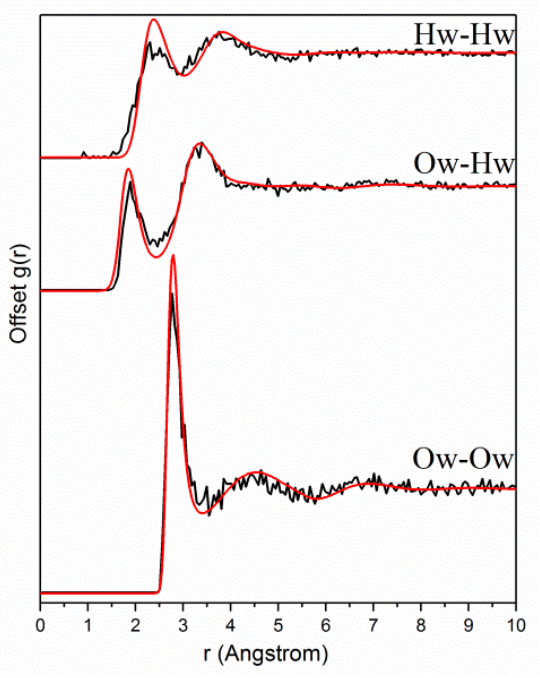

Figure 3.3. The radial distribution functions (RDFs) of the atomic pairs in water in 5 M imidazole aqueous solution (red) and pure water (black) [204].

Table 3.3. Coordination numbers (cog = centre of geometry)

Central atom	Surrounding atom	Coordination number	Cutoff / Å
Ow	Ow	3.60 5.4[204]	3.40 3.54[31]
Ow	Hw	1.40 1.8[204]	2.40 2.4 [31]
H1	Ow	0.75	2.40
N1	Ow	3.20	3.70
N2	Hw	1.30	2.50
N2	Ow	2.70	3.60
N1	N2	0.08	3.30
Imidazole (cog)	Imidazole (cog)	2.70	6.40
Imidazole (cog)	Ow	0.90, 17	3.35, 5.75

The positions of first and second peaks in the RDFs for the imidazole solution are almost identical to those in pure water. This could indicate that upon solvation, imidazole is incorporated into the water network with minimal disruption to the hydrogen-bond network. The absence of disruption to the bulk water network has also been reported for other large organic molecules such as proline [105] and acetylcholine [102]. Note, however, the limitations of the diffraction analysis and the associated modelling in determining the water structure have been highlighted by inconsistensies between RDFs of water atomic pairs determined by EPSR and molecular dynamics for DMSO solutions [104, 229]. Likewise, discrepancies have been observed when comparing experimental X-ray absorption (XA) spectra with calculated spectra that were based on diffraction data and the associated EPSR simulations [230]. Other techniuqes such as FTIR can then be helpful in providing complementary information on the effect of the solute molecules on the hydrogen-bond network of water in the presence of salts, denaturants and osmolytes [231].

3.3.2 Imidazole – Water Interactions

Although there appears to be minimal disruption to the water network by imidazole, significant interaction between imidazole and water is expected, since imidazole has high solubility in water (11 mol.L^{-1} at 25 °C [148]). Imidazole can act as both a hydrogen-bond donor (via the pyrrolic N1-H moiety) and acceptor (via the pyridinic N2). Indeed, the experimentally determined radial distribution functions (RDFs) for hydrogen-bonding pairs between solute and water molecules (Figure 3.4 (a)) and the coordination numbers at the minima after the first RDF peak (Table 3.3) confirm that hydrogen-bonding takes place between the nitrogen moieties (N1 and N2) and water molecules.

The difference of ~1 Å between the N2…Ow and N2…Hw and between N1…Ow and H1…Ow indicates collinear N-H-O hydrogen-bonding.

It is also noted that the minima for the first peak in the RDF of imidazole-water bonding pairs do not go to zero, which indicates significant disorder around the centroid distances of the first hydration shell. The second peak in the N1…Ow RDF (at ~4.70 Å) corresponds to correlation involving water molecules bound to the N2 site of the imidazole ring. Similarly, the second peak in the N2…Ow corresponds to correlation with water molecules at the N1 site. The C…Ow RDFs show a broad peak at ~3.50 Å, which is attributed to the distance between carbon atoms in the imidazole ring and water molecules bonded to the nitrogen moieties.

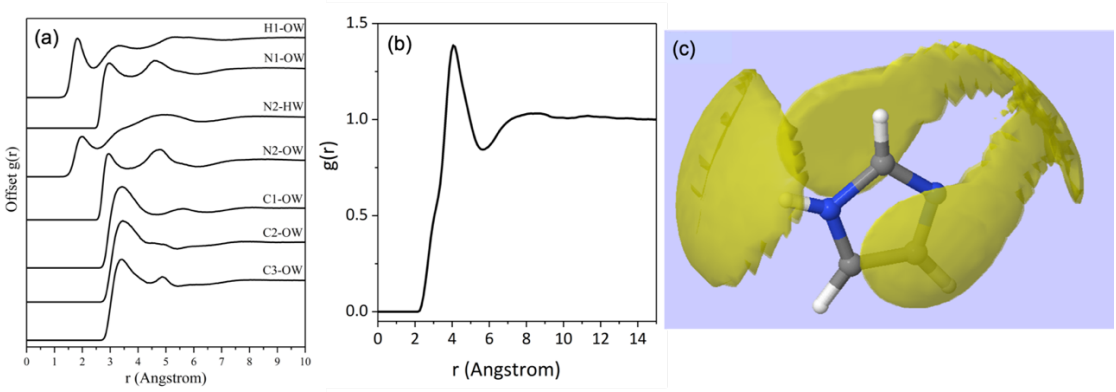

Figure 3.4. The radial distribution functions (RDFs) of (a) hydrogen-bonded imidazole-water pairs and (b) water centre of geometry (cog) around imidazole (cog) (c) Spatial density function (SDF) of water around imidazole for distances up to 5.6 Å in 5 M aqueous imidazole solution.

Applying a cut-off radius of 2.40 Å and 2.50 Å, the coordination numbers of Ow around H1 and Hw around N2 are 0.75 and 1.30 respectively (Table 3.3). This indicates that there is approximately one water molecule per nitrogen atom in the first hydration shell. At a cut-off radius that is ~ 1.3/1.2 Å longer, the presence of three water molecules per nitrogen atom is predicted, as seen in Table 3.3. This could indicate the presence of water molecules in locations that are not collinear to the nitrogen moieties. This is further illustrated by the RDF of the water centre of geometry (cog) around the imidazole cog, which has an intense peak at ~4 Å with a slight shoulder at ~3 Å (Figure 3.4(b)). The water molecules bonded to the nitrogen atoms seem further away (at ~4 Å) than in the atomic pair RDFs (~ 2.4/2.5 Å) because the radial distribution here is

determined from the centre of geometry of the two molecules. The shoulder at 3 Å is attributed to water molecules over and below the plane of the imidazole ring.

The spatial density function (SDF) of water around imidazole (Figure 3.4(c)) confirms the structural characteristics obtained from the RDFs. It shows that the most common location of water molecules for distances up to 5.60 Å (which is the minimum after the first peak in the RDF shown in Figure 3.4(b)) is adjacent to the N1-H1 and N2 moieties of the imidazole ring, as well as above and below the plane of the ring, with no water molecules adjacent to the hydrophobic C-H moieties. Water molecules prefer the formation of a collinear Ow-H1-N1 bond in the plane of the aromatic ring, while the hydrogen bonds to the N2 site have a wider angular distribution with a significant out-of-plane population. The higher mobility of the water molecules around the N2 side of the imidazole ring is indicated by the featureless correlation of the N2…Hw RDF between 3–6 Å, which hints a non-directional hydrogen-bonding. The reason for this is probably that the hydrogen bonds to N2 are formed with its lone pair of electrons, resulting in partial sp^3 character of the nitrogen centre, which imparts out-of-plane positions. Due to the mirror symmetry of the molecule such bonds are formed with equal likelihood above and below the aromatic ring plane, and this is evident through a symmetric angular distribution of water positions relative to N2. These observations about the positions of the hydrated molecules are in line with previous studies combining molecular dynamics modelling with neutron and X-ray diffraction results of neutral and protonated imidazole [110] and the related solutes pyrazole [111] and pyridine [112] in aqueous solutions. A recent MD simulation study of aqueous solutions of imidazole at different concentrations came to similar conclusions [162, 163].

3.3.3 Imidazole – Imidazole Interactions

The peak at ~5 Å in RDF of the imidazole cog around imidazole cog (Figure 3.5(b)) is comparable to the intermolecular distances between the imidazole molecules in the crystal structure [125, 232] associated with hydrogen-bonded chains (HB) and π-π stacks (parallel-displaced (PD) or edge-to-face (EF)) (Figure 3.6). More insight is provided by the RDFs of the intermolecular bonding pairs of imidazole Figure 3.5(a) where the sharp peak in the H1…N2 and N1…N2 pair correlations at about 2 Å and 3 Å, respectively, is associated with intermolecular hydrogen-

bonded imidazole chains. The difference of ~1 Å is in line with the distance in the collinear N1-H…N2 hydrogen bond. However, the fraction of imidazole molecules exhibiting such intermolecular hydrogen bond is low, as indicated by the low coordination number (0.08) for N1 around N2, below a cut-off radius of 3.3 Å (Table 3.3). In line with this, the imidazole-imidazole spatial density function (SDF) (Figure 3.5(c)) reveals that (for distances up to 6.4 Å) the nearest imidazole neighbours are most likely positioned above and below the plane of the imidazole ring. This suggests self-association through hydrophobic interactions, such as π-π stacking, which has previously been evident in ultraviolet spectroscopy and energy dispersive X-ray diffraction data[86, 152].

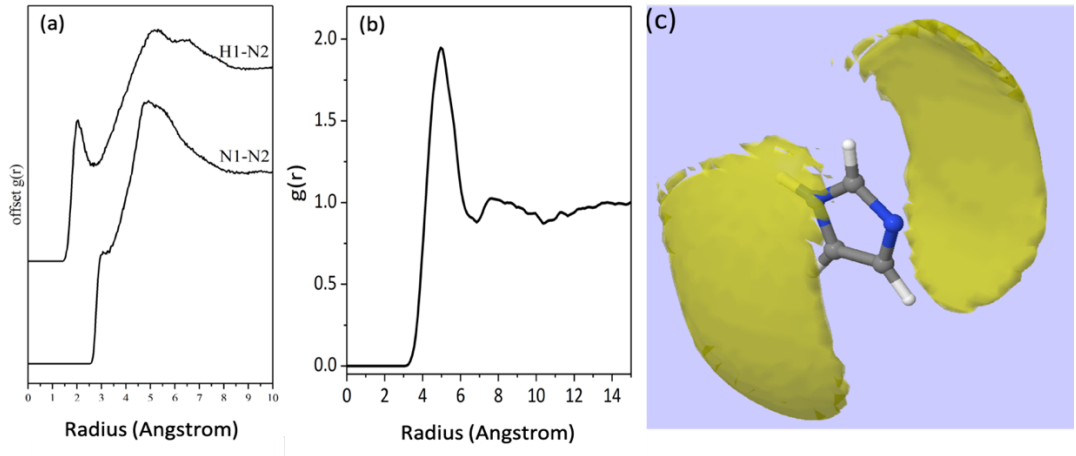

Figure 3.5. The radial distribution function (RDF) of (a) potential imidazole-imidazole bonding pairs and (b) imidazole (centre of geometry) around imidazole (centre of geometry) (c) Spatial density function (SDF) of imidazole around imidazole.

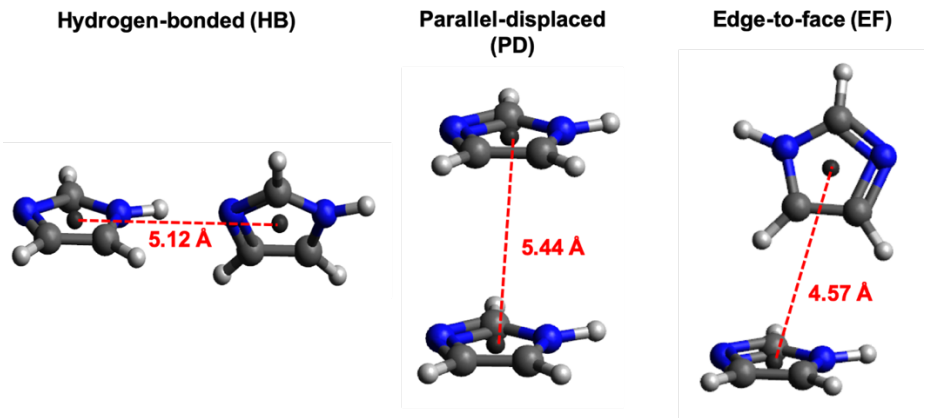

Figure 3.6. Hydrogen-bonded (HB), parallel-displaced (PD) and edge-to-face (EF) assemblies with annotated ring…ring centroid distances from crystal structure [125, 232].

The broad maxima centred at ~5 Å in the intermolecular H1…N2 and N1…N2 RDFs and the peak at ~ 6.5 Å in the H1…N2 RDF (Figure 3.5 (a)) indicate that at these distances the imidazole molecules are assembled such that N2 is either at the same distance from N1 as from H1 or is closer to N1 than H1. These arrangements indicate stacked assemblies; such as parallel-displaced (PD) or edge-to-face (EF) stacks. The orientation of the imidazole molecules within the stack was investigated by determining the angular radial distribution function (ARDF) between the z-axes of imidazole molecules (Figure 3.7). The z-axis is defined as perpendicular to the lines passing through the two nitrogen atoms (x-axis) and the C1 atom (y-axis). The plot shows that at short distances (3.5–3.8 Å) there is a high probability for finding parallel π-π stacked imidazole molecules. At longer distances (up to ~5 Å) the likelihood of observing both parallel and edge-to-face stacked molecules becomes equal, suggesting that the influence of π-π stacking does not extend beyond the first imidazole neighbours (Figure 3.7). It is interesting to note here that previous studies on π-π stacking in aromatic liquids [233] and in amino acids with cyclic side chains (e.g. histidine) [234] likewise suggested parallel stacking at smaller separations and edge-to-face stacking at larger distances.

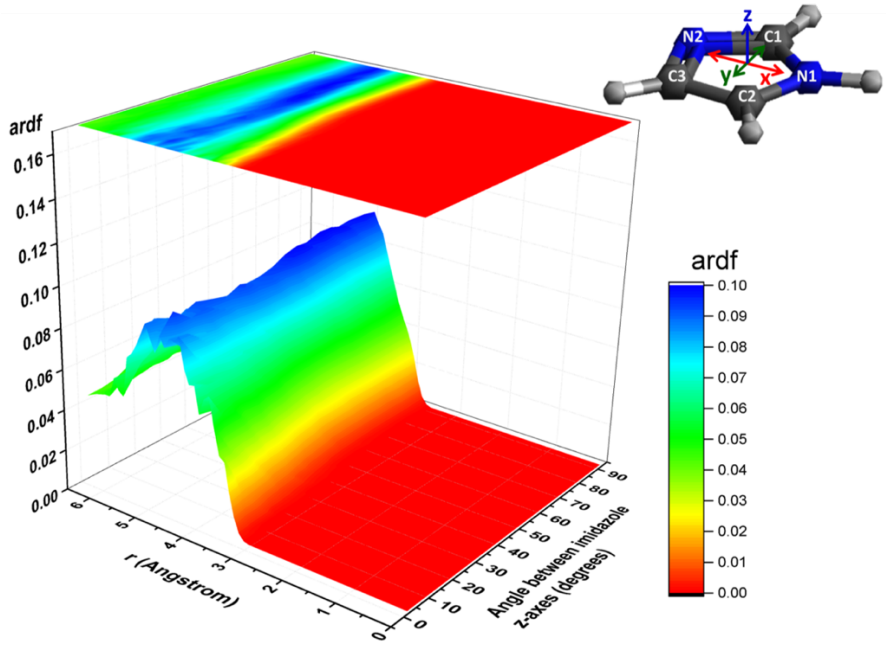

Figure 3.7. The angular radial distribution function (ARDF) between the z-axes of imidazole molecules. The z-axis is defined as perpendicular to the lines passing though the two nitrogen atoms (x-axis) and the C1 atom (y-axis).

A molecular dynamics simulation utilising a multipolar electrostatic potential derived from quantum chemical topology predicted a higher probability for the presence of hydrogen-bonded imidazole 'chain' dimers over stacked ones in aqueous solutions [162]. The discrepancy between the predicted self-assembled structures of imidazole in the higher rank MD simulation and this study might indicate that hydrogen-bonded assemblies are localised and perhaps transient features which are not as significant within the average structure of the solution. A recent neutron diffraction and isotopic substitution study on imidazolium protic ionic liquids (PILs) attributed the absence of NH…N correlations to the fact that only the average ring position can be identified, while the short lived NH…N correlations responsible for the rapid proton exchange are averaged out and not picked up by neutron diffraction [235]. The pre-defined geometrical parameters for the stacked dimers in the above-mentioned MD simulation could be another reason for the discrepancy between their results and the experimental findings presented here. The distance between the stacked imidazole molecules in the MD simulation was limited to 3.8 Å [162], which is shorter than any π-π stack distances in the crystal structure and also shorter than the distances suggested by the experimental RDFs presented here. The presence of edge to face assemblies in aqueous imidazole solutions was not investigated in the high rank MD study [162]. These dimeric assemblies, which are predicted in this work, were proven to be the most abundant structure in another MD simulation study [163].

Finally, the first peak in the RDF of the imidazole…imidazole cog is centred at 5 Å (Figure 3.5(b)), which is at a longer radius than the centre of the first peak in the imidazole…water cog RDF (Figure 3.4(b)). The coordination number of Ow around imidazole (cog) is 17 at a cut-off radius of 5.8 Å, while the coordination number of imidazole…imidazole (cog) is merely ~3 at a cut-off radius of 6.4 Å (Table 3.3). These observations suggest that the imidazole in 5 M aqueous solution is predominantly hydrogen bonded to water, with water molecules forming a

hydration shell around the imidazole molecule. Direct hydrogen-bonding between imidazole molecules is limited.

3.4 Conclusions

The results of a combined study with neutron diffraction augmented with isotopic substitution, total X-ray scattering and EPSR simulations strongly indicates the dominance of hydrogen-bonding between imidazole and water molecules in aqueous solutions, rather than hydrogen-bonding between imidazole molecules. The solvent molecules prefer to bind through a collinear Ow…H1-N1 bond and a more angularly dispersed distribution around pyridine-like N2 moiety. The presence of imidazole in the aqueous solution appears to cause minimal disruption to the water structure, as the peak positions in the water-water RDF for the aqueous imidazole solution were identical to those of pure water. No evidence of hydrogen-bonded imidazole assemblies was found. The imidazole-imidazole spatial density function indicated that the most probable positions of the nearest imidazole neighbours are above and below the plane of the ring, in line with the presence of hydrophobic π-π stacking. The orientation of imidazole molecules in these stacks was investigated through analysing their angular distribution, which indicated the presence of parallel π-π stacking at short distances (up to ~3.5-3.8 Å) and both parallel and edge-to-face stacking at longer distances (~5 Å). The coordination number of water around imidazole (~17 within a distance of ~5.8 Å) was found to be much higher than for imidazole around imidazole (~3 within 6.4 Å distance) which indicates that the most probable structural motif in the aqueous solution is hydrated imidazole molecules with minimal direct solute-solute interactions.

Chapter 4 Theoretical Calculations of the X-ray Absorption Spectra of Imidazole in Aqueous Environments

4.1 Introduction

The results in Chapter 3 suggested that the solvation of imidazole in aqueous solutions proceeds through the formation of hydrated clusters of imidazole with minimal direct interactions between the solute molecules. Similar structural models with dominant solute-solvent interactions have been predicted for protonated imidazole [110], pyrazole [111], and pyridine [112] using neutron scattering. The solute – solute interactions reported in Chapter 3 for aqueous solutions of imidazole are different from those reported by a molecular dynamics (MD) simulation study [162]. The MD simulation predicted direct imidazole – imidazole interactions to proceed through hydrogen-bonding whereas the study in Chapter 3 suggested that these interactions occur via hydrophobic π-π stacking at shorter radial distances and via π-π stacking and edge-to-face assemblies at longer radial distances (up to ~5 Å). The π - π stacking was postulated in aqueous imidazole solutions by energy dispersive X-ray diffraction [152] and ultraviolet spectroscopy [86].

Verification of the structural model of aqueous imidazole solutions reported in Chapter 3 can be achieved by comparing the theoretical XAS spectra generated from this model with the experimental core-level X-ray absorption spectra of imidazole in aqueous solutions. Core-level X-ray absorption spectroscopy (XAS) is an ideal technique to understand the intermolecular interactions in solutions because it probes the local structures around the absorbing atom. This is done via the excitation of core-level electrons (usually 1s) to unoccupied states in the valence regions of molecules. The involvement of valence levels makes the technique highly sensitive to the chemical and physical changes around organic molecules in solutions and in the solid-state [88, 89, 161, 236-238].

Core-level XAS utilising soft X-ray synchrotron radiation (energy range between ~100 eV – 2000 keV [164]) requires end-stations equipped with ultra-high vacuum (UHV) spectrometers. The most commonly used designs are liquid microjets that

expand into a vacuum chamber [95-98]. In a liquid microjet system, evaporative cooling and adiabaticity of the sudden pressure drop prevent control of the temperature, which is required to maintain well-defined supersaturation conditions. Moreover, turbulence and interfaces in the microjet nozzles act as nucleation sites, often resulting in blockage by growth of solid deposits from supersaturated solutions. Hence, core-level XAS is not an ideal technique to probe supersaturated solutions. An alternative which circumvents the experimental challenges associated with the use of soft X-rays is X-ray Raman scattering (XRS). XRS is a non-resonant X-ray scattering technique, which utilises hard X-rays and probes both dipole and higher-order transitions, depending on the momentum transfer involved. The deep penetration depth (~2 mm at 10 keV) of hard X-rays makes XRS a bulk-sensitive technique, which allows the design of ambient-pressure environmental cells with precise temperature and pressure control. It has been shown previously that these characteristics make the technique suitable for studying molecular structure in the bulk of liquid water and aqueous solutions [40, 113].

This chapter will start by comparing the C and N K-edge spectra of imidazole in gas, aqueous solution and solid-phases to highlight the ability of XRS to pick up the differences in the spectral features arising from phase transformation. After that, excited-state DFT calculations are performed on a monomeric imidazole molecule and on a cluster of eight imidazole molecules taken from the crystal structure of imidazole to establish the origin of the features in the experimental spectra. Finally the effect of solvation on the core-level XA spectra is studied by utilising a number of solvation models: (1) implicit solvation, (2) explicit solvation utilising micro-solvation structural models, (3) combination of implicit and micro-solvated (explicit) models, and (4) explicit solvation utilising the structural model generated in Chapter 3 by refining neutron and X-ray scattering data.

4.2 Methodology

4.2.1 X-ray Raman Scattering (XRS)

The C and N K-edge spectra of imidazole in the solid-state and in aqueous solution (concentration of 5 M) were collected at the inelastic scattering beamline ID-20, ESRF, Grenoble, France [183]. The spectra presented in this chapter are from 12 spherically bent Si (660) analyser crystals at an angle of 42° with q value of 3.6 $Å^{-1}$, which means that the data were collected at low momentum transfer (q). The monochromator drift was small, and it was monitored through frequent measurement of the elastic scattering peak.

The aqueous imidazole solution (concentration of 5 M) was measured using a jacketed flow-cell crystalliser (flow-cell 1, Chapter 2), which ensured precise temperature control and continuous circulation of the stock solution to minimise the beam-induced damage to imidazole. The total volume of the solution was ~250 ml and the solution was pumped at a rate of ~450 ml/min. The solid sample was pressed into a 13 mm diameter pellet with a thickness of ~2 mm. For data analysis, the elastic energy of the incident photons was subtracted from the XRS data to obtain the near-edge spectra. The spectra were then background subtracted and normalised to the edge-step height using Athena software [239]. After normalisation, the contribution from the Kapton window to the C and N K-edge spectra of the aqueous imidazole solution was removed by subtracting a reference spectrum of Kapton from those of the solution. The subtraction coefficients were defined as the value required to remove a peak significant to Kapton in the C and N K-edge; these peaks are circled in Figure 4.1.

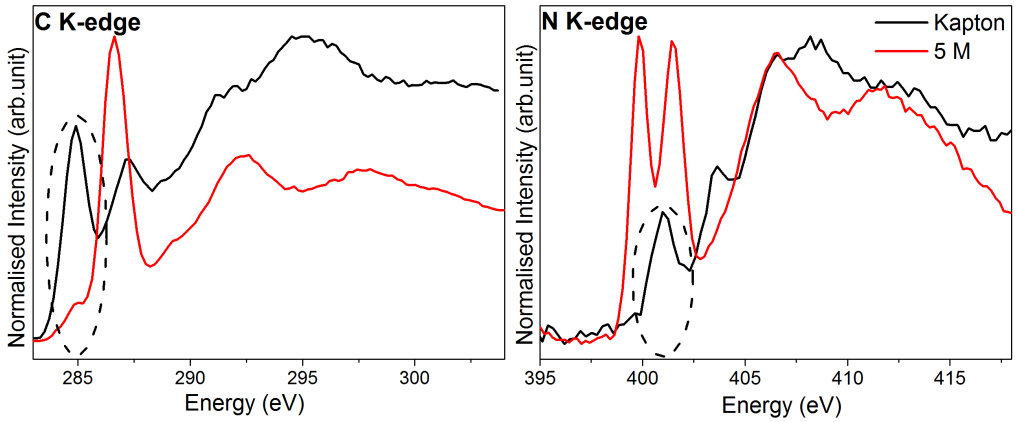

Figure 4.1. The C and N K-edge XRS spectra of Kapton and 5 M aqueous imidazole solutions. The peaks used to define the subtraction coefficient to remove the Kapton contribution to the K-edge signal of the aqueous solution of imidazole are circled in black.

4.2.2 Computational Methods

Time-dependent density functional theory (TD-DFT) (or excited-state DFT) calculations were performed to generate the theoretical C and N K-edge X-ray absorption (XA) spectra of imidazole in the gas, aqueous solution and solid phases. Calculations were performed using ORCA version 4.1.0 (unless stated otherwise) [208]. Calculations were carried out using the B3LYP functional [209] and the 6-31G* basis set for the monomeric imidazole molecule, while minimally augmented def2-SVP and def2-TZVP basis sets [210] were used for simulating the XA spectra of solvated and solid imidazole respectively. D3 dispersion correction [211] was applied to all calculations expect those of the monomeric molecule. RIJCOSX approximation [212] and def2/J auxiliary set were applied to speed up the calculations of the solid-state.

Imidazole in the gas-phase was represented by a monomeric molecule, which was constructed in Avogadro, a molecular builder and visualisation tool [213]. The intramolecular structure was initially optimised in Avogadro using the UFF force field and further refinements were performed by geometry optimisation calculations in ORCA. Frequency calculations were run on the optimised structure to ensure that the structure corresponded to the global energy minimum. Finally, time-dependent density functional theory (TD-DFT) calculations were performed on the optimised structure to generate the theoretical X-ray absorption spectra.

To simulate the water solvation shell around the imidazole molecule, implicit and micro-hydrated explicit models were used. In Models 2-6 (Figure 4.2), water molecules were explicitly added around the imidazole molecule. The positions of the water molecules (except for Model 5) were optimised in ORCA. For comparison, Model 1 used implicit solvation CPCM (conductor-like polarisable continuum model) [240]. Models 7 and 8 used a hybrid explicit/implicit solvation which imposed CPCM on the structures from Models 4 and 6 respectively. The construction of the solvation models (2 – 4 and 6 – 7) as well as the geometry optimisation and frequency calculations were performed using the same procedure as described above for the monomeric imidazole.

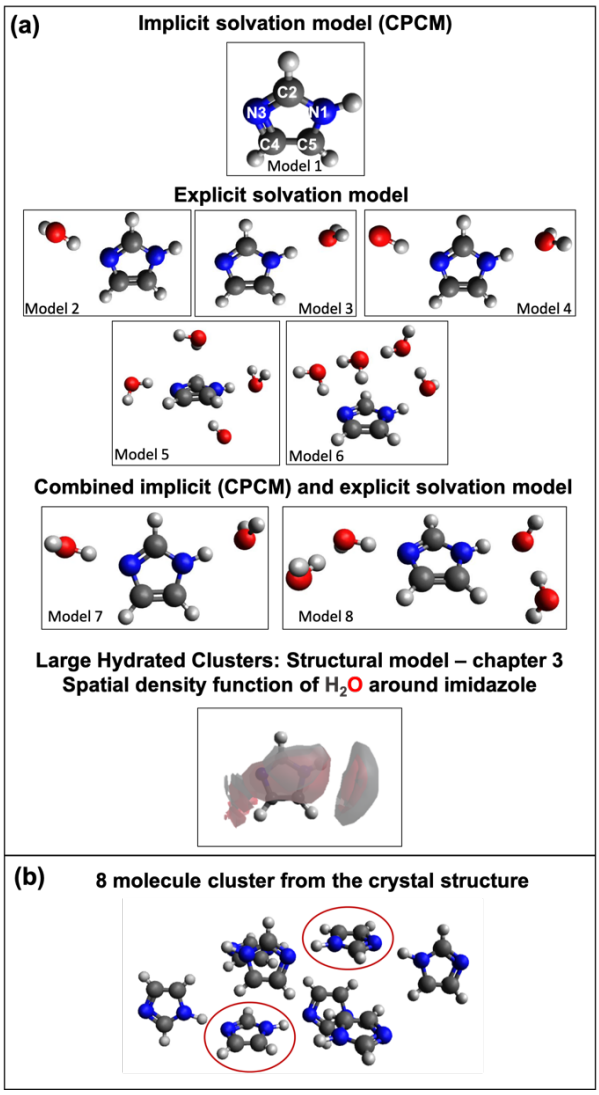

Figure 4.2. The structural models representing imidazole in (a) aqueous solution and (b) solid-phase. The two imidazole molecules used in the solid-phase TD-DFT calculations are circled.

Solvation models based on the structural model of the 5 M aqueous imidazole solution reported in Chapter 3 were also utilised to generate the theoretical C and N K-edge X-ray absorption (XA) spectra. In model 5, four water molecules were explicitly added around the imidazole molecule mimicking the spatial density function of water around imidazole, which was presented in Chapter 3 (Figure 3.4(c)). Larger hydrated clusters from the above-mentioned structural model were also used to calculate the C and N K-edge XA spectra. 500 simulation boxes each centred on one of the 500 imidazole molecules in the structural model from Chapter 3 were randomly taken. Excited-state DFT calculations were performed on 60 and 200 of the 500 simulation boxes to generate the C and N K-edge spectra respectively. The calculations were performed in ORCA utilising the same parameters applied for the solvation models mentioned previously. Calculations for Model 5 utilised version 4.1.0 of ORCA while version 3.0.3 was used the larger cluster models.

To perform excited-state simulations (TD-DFT) on solid imidazole, a cluster of eight imidazole molecules (Figure 4.2) was taken from the crystal structure of imidazole [125]. The calculations were performed for the two circled imidazole molecules in the eight-molecule cluster (Figure 4.2(b)). These two molecules are involved in hydrogen-bonding as well as edge-to-face and π-π stacking. Hence calculations on these molecules should account for the effects of the different types of imidazole-imidazole interactions that exist in the crystal structure.

It should be noted that the excited-state DFT calculations presented here only model the electronic states below the continuum threshold. In addition, relative changes in the core-level potential energies calculated using TD-DFT are more accurate than absolute values, which do not necessarily reflect the experimental values [241, 242]. Therefore, an offset was applied to the calculated 1s core-level energy to compare them to the experimental near-edge spectra. This offset was ~10.4 eV (±0.4) for the C K-edge spectra and ~12.3 eV (±0.5) for the N K-edge spectra.

4.3 Results and Discussion

4.3.1 K-edge X-ray Raman Scattering (XRS) Spectra

The C and N K-edge spectra of solid and aqueous imidazole measured using X-ray Raman scattering (XRS) are compared with the gas-phase imidazole spectrum measured using inner-shell electron energy loss spectroscopy (ISEELS) [127] (Figure 4.3). At low momentum transfer, ISEELS [168] and XRS [177] probe dipole transitions and hence are analogous to X-ray absorption (XA) spectra. The strong resonances in the near-edge region of the C and N K-edge spectra correspond to transitions from the 1s core-level of the absorbing atom to the unoccupied valence orbitals with π* character. These are followed by the ionisation potentials (IP) and broad σ* shape resonances corresponding to multiple scattering at the potential barrier to the bound neighbouring atoms. The energy difference between the centroid of the σ* shape resonances and the IP (usually termed δ) provides a highly sensitive measure of the bond length [168, 243]. The sensitivity of δ values to bond length variations as low as 0.001 Å have been previously illustrated for p-aminobenzoic acid [39] and saccharides [237].

The spectral features in the C and N K-edge spectra were deconvoluted with non-linear least-squares fitting using Fityk software [244] (Figure 4.3). Gaussian functions were used to model the π* and σ* shape resonances and arctan functions were used to model the ionisation potentials (IPs) [168]. The full-width-at-half-maximum (FWHM) of the Gaussian peaks modelling the π* resonances vary by ~1 eV for the C K-edge and ~0.3 eV for the N K-edge. The peaks in the near-edge region of the C K-edge are broader because several features contribute to each π* resonance; as will be illustrated by the excited-state DFT calculations illustrated in this chapter. The FWHMs of the σ* shape resonances increase with energy due to lifetime broadening [168]. Three and two arctan functions were used to model the C K-edge and N K-edge IPs respectively because imidazole has three different carbon moieties and two different nitrogen moieties. Since the ratio of the different carbon moieties in the imidazole molecule is equal, the width and height of the arctan functions in the C K-edge spectra were constrained to be equal; a similar procedure was followed for the N K-edge. The energy difference between the arctan functions (ΔIP) in the N K-edge spectra was constrained to be equal to the energy

difference between the two 1π* peaks (ΔE (1π*)) because it has been illustrated that the transitions to π* states are dominated by the core-level binding energy shifts [88]. Core-level binding energies determined from C1s X-ray photoelectron spectra [91] were used to constrain the relative positions and centroids of the arctan functions in the C K-edge spectra (ΔIP (C2-C5) and ΔIP (C5-C4)). The resulting centroids of the fitted Gaussian and arctan curves are listed alongside the values predicted by TD-DFT calculations in Tables 4.1 and 4.2.

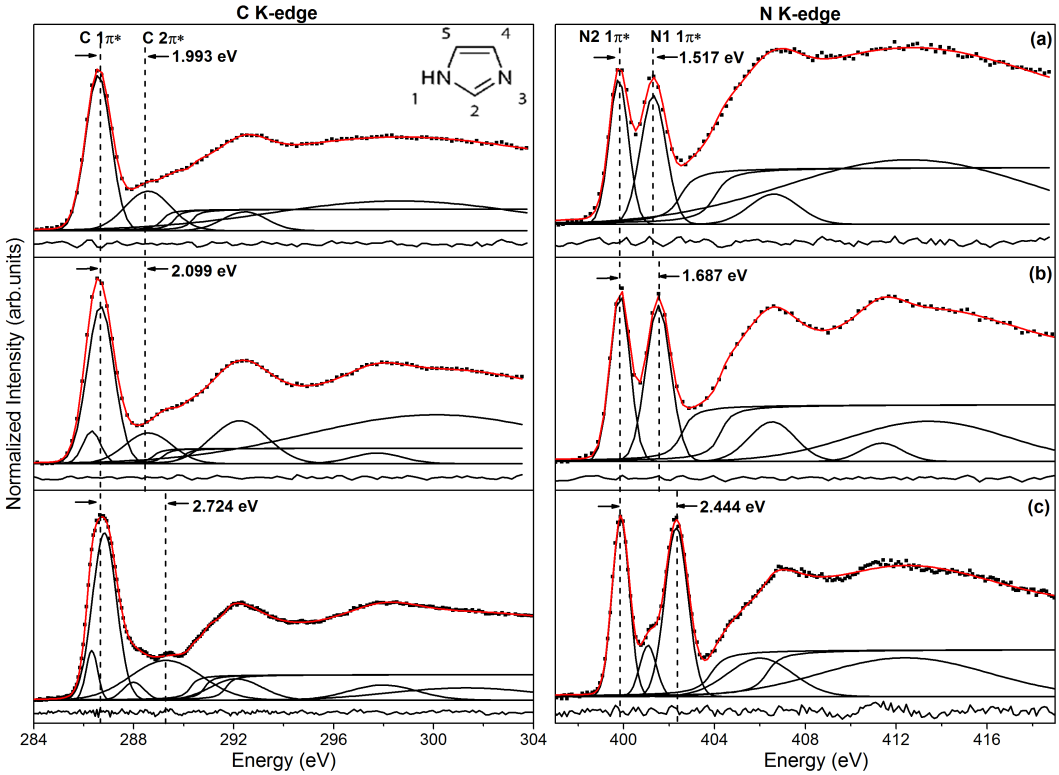

Figure 4.3.Fitted C and N K-edge spectra of the different phases of imidazole. Residual is shown at the bottom of each spectrum. (a) XRS spectra of solid imidazole, (b) XRS spectra of 5M aqueous solution of imidazole and (c) ISEELS spectra of imidazole in the gaseous phase [127].

4.3.1.1 Carbon K-edge XRS Spectra

Excited-state DFT calculations were performed on an isolated imidazole molecule to define the origin of features in the C K-edge spectra. Based on these calculations, a molecular orbital diagram was constructed to summarise the atomic and molecular orbital energies associated with the C K-edge absorption spectra (Figure 4.4). The near-edge region in the C K-edge spectra is expected to be dominated by transitions into the unoccupied π* states. The calculated XA spectrum for the monomer (Figure 4.5) shows

a sharp peak corresponding to C4, C5 and C2 1s → 1π* transitions and C4 1s → 2π* transition and a less dominant feature corresponding to C5 and C2 1s → 2π* transitions. The energy of the centroids for these transitions is listed in Table 4.1. The predicted near-edge features in the C K-edge corresponding to the 1s → 1π* and 1s → 2π* transitions are observed in the experimental spectra of imidazole in gas, aqueous solution and solid-phases. ORCA calculations for the hydrated imidazole molecule and the molecules in the crystal structure reproduced the spectral features in the C K-edge spectra of the solution and solid-phases and enabled the identification of the contributions from the different carbon atoms (Figure 4.5).

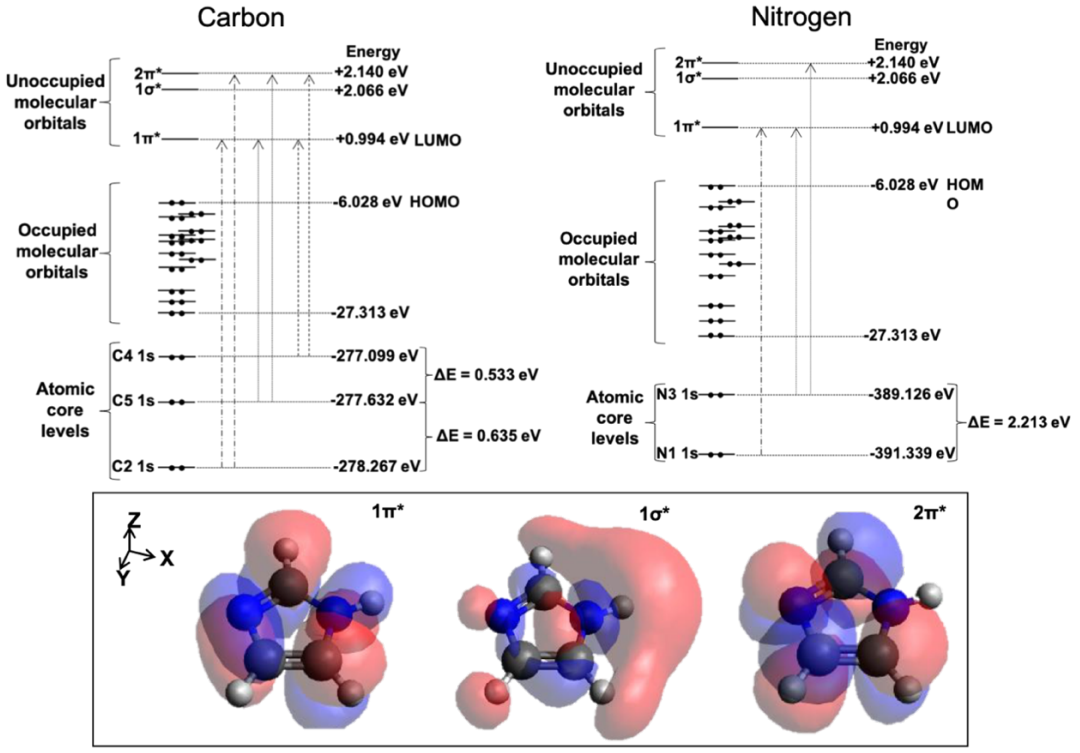

Figure 4.4. Schematic representation of the calculated atomic core-levels as well as the occupied and unoccupied molecular orbitals for a monomeric imidazole molecule. The electron density of the three lowest unoccupied molecular orbitals is shown at the bottom of the figure.

The observed energy difference between the two features in the near-edge region of the solid-state spectrum (1.993 eV) agrees with previously published total electron yield (TEY) data [127]. The energy difference between the two C K-edge peaks in the near-edge region for the solution phase is 2.099 eV, which is significantly lower than that observed in the gas-phase (2.724 eV) and comparable to that of the solid-phase (1.993

eV). This indicates that the electron density at the carbon moieties of imidazole in aqueous solutions which is present in a hydrogen-bond network is similar to that in the crystalline phase where hydrogen-bonding exists between imidazole molecules in the same plane.

Table 4.1. Peak position and assignment of the experimental and calculated C K-edge spectra

	Gas		5M Aqueous Solution		Solid	
	Exp/ eV	Calc/ eV *	Exp/ eV	Calc / eV *	Exp/ eV	Calc/ eV *
E (C 1s–1π*)	286.320 286.827	276.281 (C4) 276.453 (C5) 276.872 (C2)	286.322 286.674	286.538	286.58	275.777 (C4) 275.886 (C5) 276.290 (C2)
E (C 1s–2π*)	286.827 287.989 289.297	277.049 (C4) 278.121 (C5) 278.859 (C2)	288.597	288.335	288.573	275.922 (C4) 276.983 (C5) 277.705 (C2)
ΔE (C2 2π* -C 1π*)	2.724	2.324	2.099	2.031	1.993	1.721
IP (C2)	291.690	278.267	289.998	-	290.684	278.606
IP (C5)	290.841	277.632	289.198	-	289.931	277.920
IP (C4)	290.266	277.099	288.798	-	289.261	277.572
ΔIP (C2-C5)	0.849	0.635	0.800	-	0.753	0.686
ΔIP (C5-C4)	0.575	0.533	0.400	-	0.670	0.348
σ (C-H)	292.103	-	292.232	-	292.442	-
σ (C-C)	297.987	-	297.74	-	298.615	-
σ (C-N)	301.266	-	300.084	-	-	-

*** The energy shift of ~ 10.4 eV (±0.4) has not been applied to the calculated transitions presented here.**

In terms of the C1s binding energies (BE), C2 is expected to have the highest 1s core-level BE as it is bonded to the two electronegative nitrogen moieties in the imidazole ring. The other two carbon atoms are expected to have lower BE values due to bonding to only one nitrogen neighbor, and C5 should have a slightly higher value than C4 as it is bonded to the protonated nitrogen atom (N1). This is reflected in the experimental IPs and the ORCA-calculated core-level orbital energy values and the π* transition energies, which are the highest for C2 followed by C5 and C4 (Table 4.1). Similarly, in solid-state theophylline [245] and histidine [246] the C1s binding energy for the N=C-NH (equivalent

to C2 in imidazole) was higher than the other carbon atoms in the imidazole ring of these molecules.

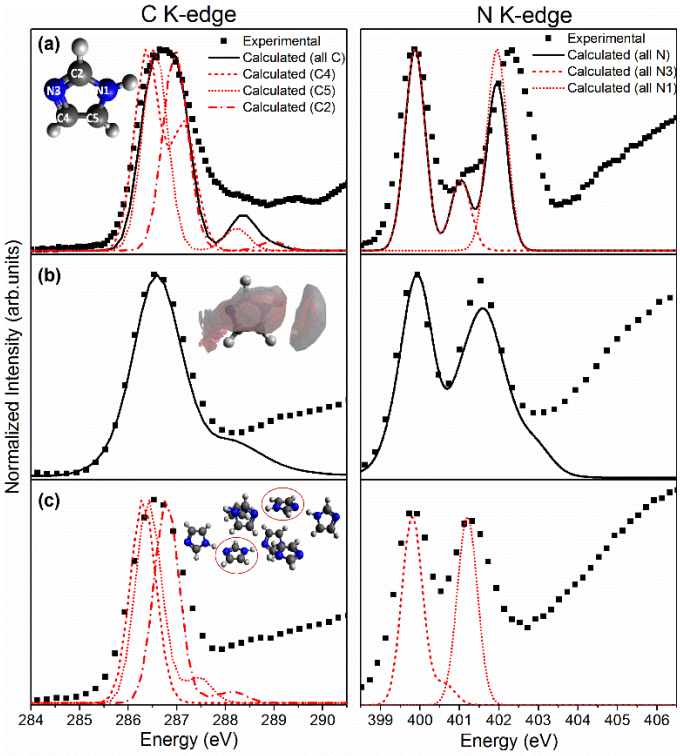

Figure 4.5. Comparison of the experimental and calculated C and N K-edge spectra. (a) ISEEL spectra of imidazole in the gaseous phase reported in reference [127] (b) XRS spectra of the 5M aqueous solution (c) XRS spectra of solid-state imidazole. For the aqueous solution sample, the TD-DFT simulated spectra used the large hydrated-clusters model based on neutron diffraction data reported in Chapter 3.

4.3.1.2 Nitrogen K-edge XRS Spectra

The near-edge region of the imidazole N K-edge spectra is dominated by excitations from 1s to the 1π* orbital of the two nitrogen moieties (N3 and N1-H) [88, 127]. An additional transition into the 2π* valence orbital is observed at N3 for the gas-phase spectrum [127]. Comparison of the N K-edge spectra of the three phases reveals a decrease in the energy difference between the two 1π* peaks, from 2.444 eV in the gas-phase, to 1.687 eV in the aqueous solution and 1.517 eV in the solid-phase (Figure 4.3). The energy difference between the two 1π* transitions and IPs for the solution is closer to that of the hydrogen-bonded solid than the gas-phase monomer, indicating that hydrogen-bonding plays an important role in the aqueous solution phase of imidazole.

The features associated with the 1π* peaks in the N1s spectra of imidazole broadly agree with previously reported values. The ΔE(1π*) in the solid-state (1.517 eV) spectra is in good agreement with previously reported total electron-yield (TEY) NEXAFS spectra for imidazole in the solid-state (1.5 eV) [127]. Similarly, the 1π* energy difference for the solution-phase sample (1.687 eV) agrees with previously reported X-ray excited optical luminescence (XEOL) NEXAFS of aqueous imidazole (1.7 eV) [88]. However, the relative intensity of the two π* peaks in the solution phase XRS spectrum is equivalent which is different than that previously reported in NEXAFS spectra measured using XEOL which had a less intense N1 1π* peak [88]. The relationship between core-level absorption oscillator strengths and optical photon emission is complex, involving multiple excited state decay pathways that may include the transfer of energy from the absorbing atom to nearby atoms [247]. This may explain the differences in relative intensities between the near-edge features measured using XRS and XEOL.

Table 4.2. Peak position and assignment of the experimental and calculated N K-edge spectra

	Gas		5M Aqueous Solution		Solid	
	Exp/ eV	Calc/ eV*	Exp /eV	Calc /eV*	Exp/ eV	Calc/ eV*
E (N1 1s - 1π*)	402.316	390.098	401.527	389.167	401.304	388.996
E (N3 1s - 1π*)	399.872	388.020	399.840	387.513	399.787	387.600
ΔE (1π*)	2.444	2.078	1.687	1.654	1.517	1.396
E (N3 1s - 2π*)	401.072	389.180	-	388.573	-	388.475
ΔE (N3 2π*- 1π*)	1.200	1.160	-	1.060	-	0.875
IP (N1)	406.301	391.339	404.253	390.935	404.023	391.175
IP (N3)	403.857	389.126	402.567	389.175	402.506	389.652
ΔIP	2.444	2.213	1.686	1.760	1.517	1.523
1σ* (C-N)	406.012		406.538		406.619	
2σ* (C-N)	412.400		411.364		412.587	
3σ* (C-N)	-		413.441		-	
*** The energy shift of ~12.3 eV (±0.5) has not been applied to the calculated transitions presented here.**						

The experimental and ORCA-calculated features (Table 4.2) show that the variations in ΔE(1π*) and ΔIP in the N K-edge spectra of the three phases stem from larger changes in the energy of 1s → 1π* transitions and the IP associated with the protonated N1 moiety when compared to N3, an interpretation that is in line with previously published results combining NEXAFS and StoBe calculations [88]. The energy shifts of the near-edge resonances and the IP values associated with the N1 moiety are reflected by more significant changes in the lengths of the bonds formed by N1 when compared to N3 (Table 4.3). As one goes from gas to solution to solid-phases, there is an increase in the N1-H bond length and an associated decrease in the N1-C2 and N1-C5 bond lengths. On

the other hand, the C2-N3 bond length is almost identical and the decrease in the N3-C4 bond length is less severe when compared to that associated with N1-C bond lengths. These observations indicate that the formation of hydrogen-bonds with water molecules (N1-H…O-H/N3…H-O) in the solution sample and with imidazole molecules (N1-H…N3) in the solid sample changes the electron density at the N1 site much more than at the N3 site. A likely explanation is that the hydrogen-bonds formed at the N1 centre occur through proton donation with changes in electron density located in the N-H bond. In contrast, the N3 centre acts as a proton acceptor and participates in the aromatic system, which therefore allows significant delocalisation of the electron density variation across the aromatic ring resulting in a smaller core-level shift due to hydrogen-bonding.

Table 4.3. Internal bond lengths for imidazole monomer, imidazole in solution phase using the five solvation models and in the solid-state imidazole. Intermolecular hydrogen-bond lengths were included as well where relevant

Structural model	N1-H	N1-C2	C2-N3	N3-C4	C4-C5	C5-N1	N3…HO	N1H…O	N1H…N3
Imidazole Monomer (orca optimised)	1.009	1.368	1.316	1.380	1.373	1.382			
Model 1 (Implicit)	1.016	1.360	1.324	1.381	1.379	1.377			
Model 2 (Explicit)	1.013	1.364	1.317	1.376	1.377	1.381	1.874		
Model 3 (Explicit)	1.022	1.366	1.318	1.375	1.380	1.378		1.916	
Model 4 (Explicit)	1.023	1.361	1.320	1.375	1.379	1.378	1.855	1.891	
Model 5 (Explicit)	0.936	1.445	1.380	1.371	1.338	1.413	2.015	1.833	
Model 6 (Explicit)	1.035	1.355	1.324	1.378	1.381	1.376	1.801	1.782	
Model 7 (Combined)	1.033	1.355	1.327	1.380	1.379	1.377	1.753	1.818	
Model 8 (Combined)	1.029	1.355	1.326	1.380	1.378	1.377	1.703	1.859	
Solid	1.038	1.337	1.316	1.368	1.358	1.362			1.828

4.3.2 The Effects of Solvation

4.3.2.1 Implicit and Micro-Hydrated Explicit Solvation Models

The sensitivity of the C and N near-edge spectra to intermolecular interactions between imidazole and water has been explored through the calculated spectra (Figure 4.6) of the 8 solvation models illustrated in Figure 4.2. These models applied (1) implicit solvation (Model 1), (2) explicit solvation where water molecules were added around a single imidazole molecule forming micro-hydrated clusters (Models 2 – 6) or (3) a combination of implicit and explicit solvation (Models 7 and 8).

The implicit solvation model accounts for the long-range effects of the solvent, which were previously found necessary to determine the vertical and adiabatic ionisation potentials of imidazole in aqueous solutions [92]. The implicit solvation model (Model 1) gave some agreement with the N K-edge experimental data; however, it shows an extra feature at 400.826 eV, which is attributed to the N3 1s→2π* transition. This extra feature is only visible in the spectrum of the gas-phase species (Figure 4.3). The calculated energy differences between the N3 2π* and 1π* peaks are 1.160 eV for the monomeric molecule (analogous to gas-phase system), 0.988 - 2.148 eV for the geometry optimised explicit solvation models (Models 2 – 5, Table 4.4) and ~ 0.816 eV for the implicit model (Model 1). The transition from the N3 1s core to the 2π* valence orbital is shifted to lower energy in the implicit solvation model and is associated with the longer N3-C bond lengths when compared to the monomeric molecule and the geometry optimised explicit solvation models (Models 2 – 4 and 6) (Table 4.3). These changes in the internal bond lengths seem to affect the C K-edge as well; the second feature in the C K-edge spectra of the implicit solvation model is shifted to lower energy when compared to the experimental spectrum as seen in Figure 4.6 and Table 4.4.

The explicit solvation model accounts for local effects of the water molecules on the electronic structure of imidazole. The best agreement between theoretical and experimental N K-edge spectra was found using Model 4 where explicit solvation was defined by adding one water molecule at each side of the imidazole ring (Figure 4.6 and Table 4.4). Models 2 and 3, where water was added at one side of the imidazole ring only, resulted in theoretical N K-edge spectra with greater separation between the two 1π* peaks than in the experimental spectrum. This is associated with longer hydrogen-bond lengths between the imidazole and water molecules (Table 4.3) when compared to Model 4. The variations between the calculated C K-edge spectra for the geometry-optimised explicit solvation models (Models 2-4) were minimal, indicating that the interaction of water with the N moieties results in similar effects on the C members of the aromatic ring (Figure 4.6).

Model 5, an explicit solvation model where 4 water molecules were added to form a cluster around the imidazole molecule, is based on the solvation model of aqueous solutions of imidazole predicted in Chapter 3. Model 5 gave good agreement with the

experimental N K-edge spectrum (Figure 4.6 and Table 4.4). However, this is not the case for the C K-edge spectrum calculated using Model 5, where the energy of the C 1s - 2π* transition is overestimated when compared to the experimental data or the other explicit solvation models (Figure 4.6 and Table 4.4).

Model 6 is an explicit solvation model generated by geometrically optimising the structure of Model 5. Upon optimisation, the water molecules asymmetrically assemble near the nitrogen moieties in the imidazole ring as seen in Figure 4.2. The hydrogen-bond lengths between imidazole and water are shorter than those reported for Model 5 (Table 4.3) and the ΔE(1π*) in the N K-edge spectrum is smaller than the experimental ΔE(1π*) (Figure 4.6 and Table 4.4).

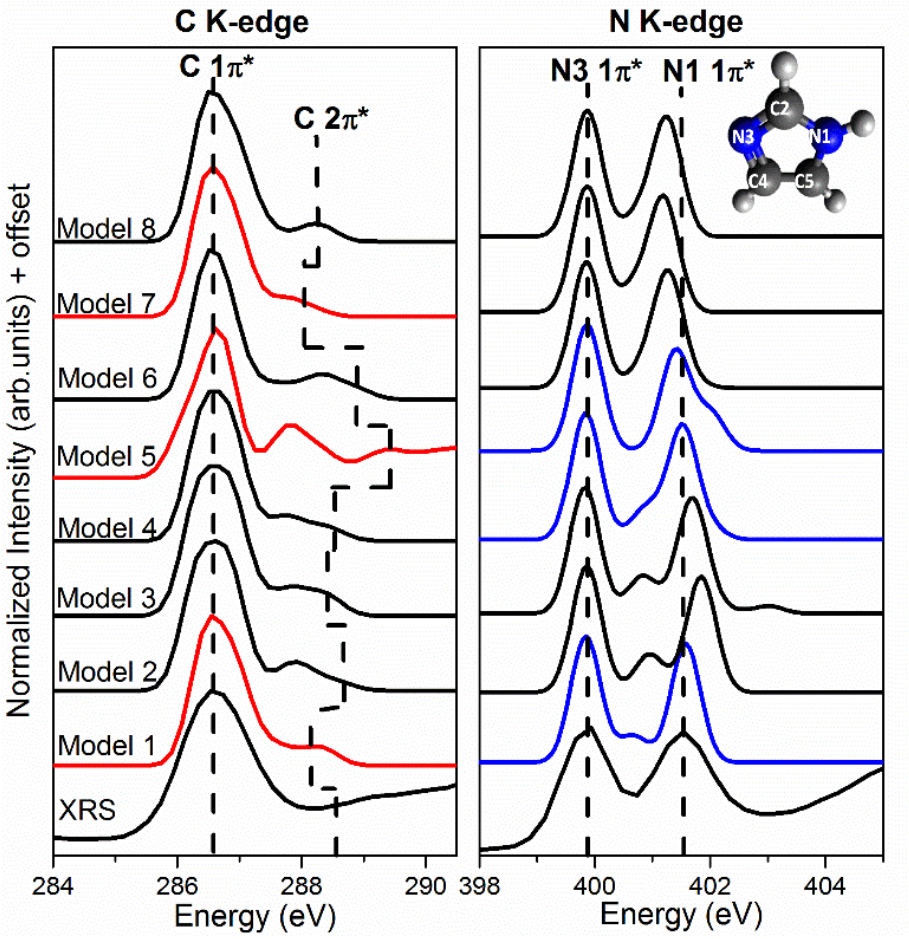

Figure 4.6. Calculated C K-edge and N K-edge X-ray absorption spectra using the structural models shown in Figure 4.2. The C K-edge spectra in red have the least agreement with the experimental XRS spectrum while the blue N K-edge spectra show the highest agreement with the experimental spectrum.

In the Combined models, implicit solvation was used to account for the long-range effects of the solvent and explicit water molecules were added to account for local effects of the solvent (Models 7 and 8). The relative energies for the two 1π* peaks in N K-edge spectra were underestimated for these models when compared with the experimental spectrum (Figure 4.6 and Table 4.4) and, similarly to the implicit solvation model (Model 1), the calculated transition from N3 1s→2π* is shifted to lower energy (Table 4.4). The N3 1s→2π* transition was not resolved as an individual peak in the N K-edge spectra generated by Models 7 and 8 because the energy of the N1 1s→1π* transition is underestimated, and the two peaks overlap. The second peak in the C K-edge spectra calculated using Models 7 and 8 is shifted to lower energy similar to that of the implicit solvation model (Model 1).

Table 4.4. Energy difference (ΔE) between 1s→π* spectral features in the near-edge region of the C and N K-edge spectra for the monomeric imidazole molecule and the eight solvation structural models shown in Figure 4.2

Structural models	ΔE (C2 2π* - C 1π*) eV	ΔE (N1 1π* - N3 1π*) eV	ΔE (N3 2π* - N3 1π*) eV
XRS experimental spectrum	**2.099**	**1.687**	-
Monomeric molecule (model for gas-phase)	2.324	2.078	1.160
Model 1: Implicit solvation model (CPCM)	**1.651**	**1.733**	0.816
Model 2: Explicit solvation model (1 water molecules at the N3 side)	2.169	1.981	1.094
Model 3: Explicit solvation model (1 water molecules at the N1 side)	2.058	1.850	0.999
Model 4: Explicit Solvation model (1 water molecules at the N3 side + 1 water molecule at the N1side)	2.057	**1.678**	1.045
Model 5: Explicit Solvation model (4 water molecules, based on solvation model in chapter 3)	**3.086**	**1.529**	2.148
Model 6: Explicit Solvation model (Model 5 + geometry optimisation)	2.376	1.426	0.988
Model 7: Combined Solvation model (Model 4 + CPCM)	**1.510**	1.327	0.847
Model 8: Combined Solvation model (Model 5 + CPCM)	**1.581**	1.393	0.883

Although the C and N K-edge spectra generated from Model 4 broadly agree with the experimental spectra, the solvation of imidazole by two water molecules is not realistic. Increasing the number of water molecules is expected to decrease the hydrogen-bond length between water and imidazole molecules and consequently the separation between the two 1π* peaks in the N K-edge spectra

as seen in Model 6. Micro-hydrated clusters of imidazole have been previously used to interpret the photoelectron spectra of imidazole [92] and to calculate the N K-edge X-ray absorption (XA) spectra of aqueous solutions of imidazole [88, 160]. The greatest agreement between the calculated and the experimental XA spectra was for the model with three water molecules [88] although solvation of imidazole by three water molecules is not plausible. When more water molecules were added to the model, the separation of the two 1s→1π* peaks in the pre-edge region of the calculated N K-edge spectra reduces [88] similar to what has been reported in this chapter. The water molecules in these clusters were oriented in a solvation ring at both sides (adjacent to the nitrogen moieties) and above the imidazole molecule [88, 92]. These structural models are not in agreement with the model predicted by neutron and X-ray diffraction and the associated EPSR simulation (presented in Chapter 3) where the water molecules form an isotropic solvation shell around the imidazole molecule. Model 5, which is based on the solvation model from Chapter 3, gave a good agreement with experimental N K-edge spectrum. However, it overestimated the transition to the 2π* valence orbitals in the C and N K-edge spectra.

Moreover, explicit solvation models do not account for the long-range effects of the bulk solvent nor, naturally, for solute-solute interactions. Hence, structural models utilising polarisable continuum solvent model (implicit solvation) were used as well as combinations of implicit and explicit solvation models, which probe the local and the bulk effects of the solvent. These models have been previously utilised to calculate the vibrational ionisation potential [92]. However, the vibrational ionisation potential in the photoelectron spectra and the transitions to the 2π* valence orbitals in the X-ray absorption spectra for these models were underestimated**.**

4.3.2.2 Large Hydrated-Clusters: Simulation Boxes from the Structural Model Generated by Neutron and X-ray scattering

It has been previously illustrated that the best agreement between the theoretical and experimental N K-edge XA spectra was achieved with a sufficiently large

solvation cluster containing one imidazole molecule and 15 water molecules [217]. This suggests that self-association of imidazole in aqueous solutions occur through secondary interactions between individually hydrated molecules. This structural image of solvation agrees with the structure predicted by neutron and X-ray scattering and the accompanied empirical potential structure refinement (EPSR) simulation presented in Chapter 3. Therefore, the solvation structure predicted from experimental results presented in Chapter 3 was used to generate simulation boxes for calculating theoretical C and N K-edge XA spectra. This method will also enable accounting for secondary imidazole-imidazole interactions; especially that the models presented in section 4.3.2.1. based on explicit, implicit and a combination of explicit/implicit solvation did not account for solute-solute interactions and therefore their ability to adequately reflect the experimental spectra is questionable.

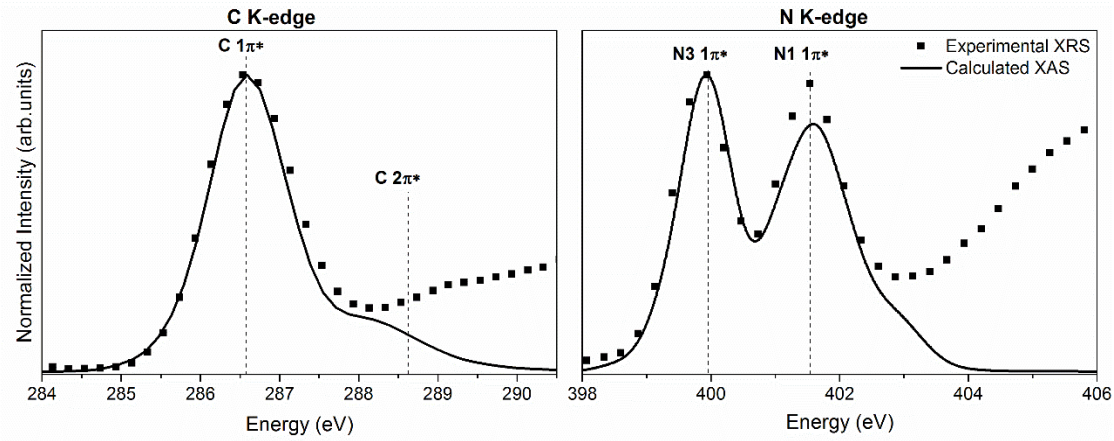

Figure 4.7. Calculated C K-edge and N K-edge X-ray absorption (XA) spectra (solid lines) of 5M aqueous solution of imidazole using 60 and 200 simulation boxes respectively from the structural model predicted in Chapter 3. The calculated spectra are compared with experimental XRS spectra (scatter).

The number of simulation boxes utilised to calculate the C and N K-edge spectra were 60 and 200 boxes respectively. The calculated and experimental spectral features in the C and N K-edge spectra are in good quantitative agreement in terms of the relative energies and the intensity and width of the peaks (Figure 4.7). When more than one molecule is considered (as done here) the width of the simulated peaks increases and the intensity of the spectral line between the two

1π* peaks does not drop massively as in the calculated XA spectra presented in section 4.3.2.1 (Figure 4.6). This suggests that the local environment around the imidazole molecules varies within the solution.

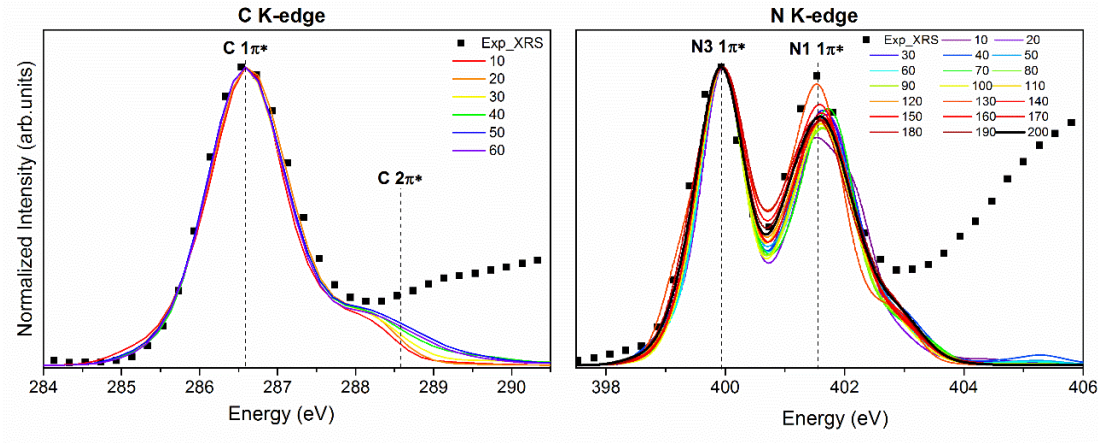

Figure 4.8. Comparison of the C and N K-edge X-ray absorption spectra averaged from different number of simulation boxes (solid coloured lines) with experimental XRS spectra (scatter).

The variations between the calculated C and N X-ray absorption K-edge spectra with respect to the number of simulation boxes used is minimal as illustrated in Figure 4.8. The most significant dissimilarities in the N K-edge spectra were in the intensity of the second peak attributed to N1 1s→1π* transition and in the intensity of the spectral line between the two peaks. This is expected because N1 interacts through its localised σ bond and the variations induced by the differences in the local environment around the imidazole molecule in each simulation box will result in dissimilarities in the spectral features associated with N1.

The effect of the variations in the local environment around the imidazole molecules on the C and N K-edge spectra can be studied by averaging the K-edge spectra calculated from simulation boxes containing the same number of imidazole molecules and comparing these averaged spectra (Figure 4.9). It was observed that the second peak attributed to the N1 1s→1π* transition in the N K-edge spectra is more prominently affected by the variation in the number of imidazole molecules in the simulation boxes than the first peak associated with the N3 1s→1π* transition. This is expected because the changes in the N1 electronic density as a result of solvation is localised while the changes in N3 is

delocalised through the aromatic system. Variations in the C K-edge spectra are noticeable for simulation boxes with less than 3 imidazole molecules which indicates that the C K-edge is affected by intermolecular interactions between the imidazole molecules. This is expected because of the hydrophobic nature of the carbon moieties.

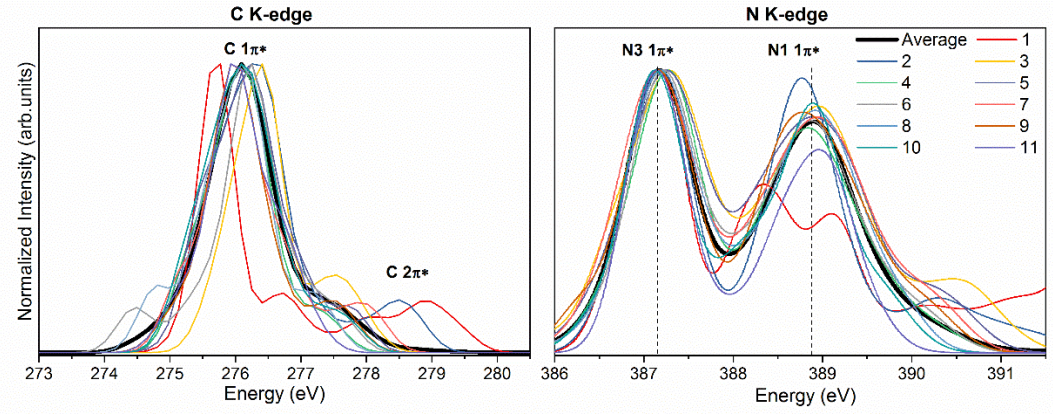

Figure 4.9. Calculated C and N K-edge X-ray absorption spectra based on the number of imidazole molecules in the simulation boxes (coloured lines) compared with the averaged spectra from all simulation boxes (black line).

The correlation between the peak separation of the two 1π* peaks in the N K-edge spectra (ΔE 1π*) and the hydrogen-bond length of the imidazole – water pairs in the 200 simulation boxes used to calculate the N K-edge spectra has been explored (Figure 4.10 (top)). It can be seen that the N3…H-O hydrogen-bond length has a mean value of ~2.5 Å and it does not differ much between the different simulation boxes unlike the N1-H…O-H bond length, with a bond length range between 1.6 Å – 3 Å. Hence the variation ΔE 1π* in the N K-edge spectra could be attributed to variations in the N1-H…O-H bond length. This is in line with the previous observation where the dissimilarities in N K-edge spectra of the different simulation boxes stems from changes in the spectral features associated with the N1 moiety in the imidazole ring (Figure 4.8 and 4.9).

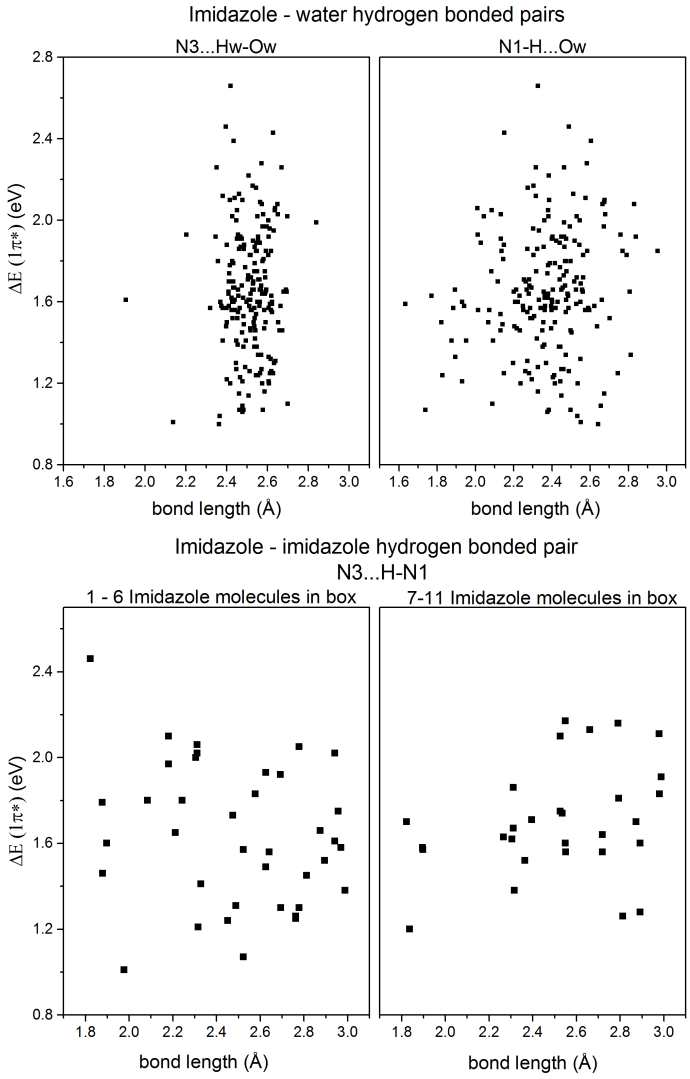

Figure 4.10. Correlation between (top) imidazole – water and (bottom) imidazole – imidazole hydrogen-bonds' length and the energy difference between the two 1π* peaks in the calculated N K-edge spectra.

To explore the effect of direct hydrogen-bonding between imidazole molecules on the N K-edge spectra, the correlation between ΔE 1π* in the N K-edge spectra and the N3…H-N1 hydrogen-bond length was investigated (Figure 4.10 (bottom)). For simulation boxes containing less than 7 molecules, it was difficult to find a link between the N3…HN1 bond length and ΔE 1π*. However, for simulation boxes containing more than 7 molecules, it was found that ΔE 1π* in the N K-edge spectra decreases as the N3…HN1 hydrogen-bond length decreases. The smallest ΔE 1π* value is associated with N3…HN1 bond length of ~ 1.8 Å. This is in agreement with the observations for imidazole in the solid-state where the N3…HN1 bond length of 1.828 Å (Table 4.3) correlated with a peak separation (ΔE

1π*) in the N K-edge spectrum of 1.517 eV (less than that observed for the aqueous solution spectrum)(Table 4.2).

4.4 Conclusions

Comparison of the C and N K-edge X-ray Raman scattering (XRS) spectra of imidazole in aqueous solution and in the solid-phase illustrated the sensitivity of the near-edge spectra to changes in the local bonding environment of imidazole. The energy differences between the features attributed to π* transitions in the C and N K-edge spectra of the 5 M aqueous imidazole solution are closer to those of the hydrogen-bonded solid than the gas-phase monomer. This indicates that hydrogen-bonding plays an important role in the aqueous solution phase of imidazole. The nature of the solute – solvent bonding environment was examined by excited-state DFT calculations performed on a number of solvation models (1) implicit solvation, (2) explicit solvation utilising micro-solvation structural models, and (3) combination of implicit and explicit solvation models. These calculations confirmed the sensitivity of the N K-edge spectra to hydrogen-bonding between imidazole and water. Nonetheless, these models did not reflect the nature of intermolecular interactions in aqueous solutions of imidazole adequately. The explicit solvation models did not account for the long-range effects of the bulk solvent nor, naturally, for solute-solute interactions and the implicit solvation, which accounted for the effects of the bulk solvent, underestimated the transitions to the 2π* orbitals.

For these reasons, excited-state DFT calculations were performed on large hydrated-clusters (simulation boxes) taken from the structural model of aqueous imidazole solutions reported in Chapter 3, which was generated from neutron and X-ray scattering data. The calculated C and N K-edge spectra were in good quantitative agreement with the experimental spectra. This provided an independent validation of the structural model reported by neutron and X-ray scattering. It also illustrated that scattering and spectroscopic techniques are highly complementary, giving information on the structural properties and intermolecular interactions of imidazole in aqueous solutions as well as shedding light on aspects of the local electronic structure which plays a central role in the chemical reactivity of the system.

Chapter 5 Concentration Effects on the Solvation Structure of Aqueous Imidazole Solutions

5.1 Introduction

Understanding the solvation structure of imidazole in aqueous solutions is not trivial because both imidazole and water can donate and accept hydrogen bonds. Therefore, a number of spectroscopic and scattering techniques have been employed to unravel the nature of the hydrogen-bond network in the system. The binding preference of water to imidazole has been investigated by FTIR spectroscopy, which showed that water prefers to act as a hydrogen bond donor (Ow-Hw…N3) [153]. Since imidazole can also donate protons through the N1-H moiety, it will compete with water to form hydrogen bonds with other imidazole molecules in the aqueous solution. Hydrogen-bonded dimers of imidazole (N1-H…N3) were predicted by a recent MD simulation study of aqueous imidazole solutions [162]. The presence of these dimers indicates that imidazole molecules self-associate in aqueous solutions through direct solute-solute interactions. π-π stacking is another form of direct interactions, which can exist in solutions due to the aromatic nature of imidazole. The presence of π-π stacks of imidazole molecules in aqueous solutions has been predicted by energy dispersive X-ray diffraction [152] and ultra-violet spectroscopy [86]. However, the solvation of imidazole in water through dominant solute-solvent interactions with minimal direct interactions between the imidazole molecules is also possible. This was predicted by the study presented in Chapter 3 where neutron diffraction and total X-ray scattering were combined with the empirical potential structure refinement (EPSR) to probe the solvation structure of aqueous imidazole solutions. The suggested structure was independently validated by generating theoretical X-ray absorption (XA) spectra and comparing them with the experimental spectra (in Chapter 4).

It is worthwhile to study the structural evolution of imidazole in aqueous solutions as a result of increasing the concentration of imidazole especially that imidazole has a very high solubility in water (~11 mol.L^{-1} at 25°C [148]). This will provide insight into changes in the structural configuration of the solution prior to phase separation during crystallisation processes, which will be discussed further in Chapter 6 of this thesis.

Changes in the local coordination in aqueous imidazole solutions as a function of varying the imidazole concentration were probed by Raman spectroscopy and found to be minor [217]. This finding is in line with the outcome of a core-level X-ray absorption spectroscopy (XAS) study of aqueous imidazole solutions at the N K-edge in the concentration range 0.5 – 8 M [88, 217]. While Raman spectroscopy examines the average bonding environment in the system, core-level (XAS) probes the local structure around the absorbing atom via the excitation of core-level electrons (usually 1s) to unoccupied states. Hence, core-level XAS is highly sensitive to chemical and physical changes around organic molecules in solutions [88, 92, 160, 161, 217].

In an extension to the previous work by Raman spectroscopy and N K-edge X-ray absorption spectroscopy (XAS), this chapter will conduct a systematic study on the effects of increasing the concentration of imidazole on the solvation structure of aqueous imidazole solutions. FTIR spectroscopy, X-ray Raman scattering (XRS) and total X-ray scattering will be combined since they provide highly complementary information on the nature of hydrogen-bonding in the system. FTIR spectroscopy provides information on the average hydrogen-bond network in the probed solution [231]. It has been recently used to probe the dynamic structure of water and the rigidity of the hydrogen-bond network in the presence of salts, denaturants and osmolytes [231]. X-ray Raman scattering (XRS), which is analogous to core-level XAS as discussed in Chapter 2, probes the local changes in the hydrogen-bond environment around the absorbing moieties (in this case: C, N and O). Finally, total X-ray scattering offers information on short-range order between molecules as well as the average long-range structure. These are provided from (1) the pair distribution functions (PDFs), which gives real-space information on correlations between atoms in the probed systems, and (2) the structural models from the empirical potential structure refinement (EPSR) simulation where the experimental total scattering data are utilised to refine the generated structures [198, 219].

5.2 Methodology

Aqueous imidazole solutions with concentration between 0.5 M and 11 M were characterised with FTIR spectroscopy, X-ray Raman scattering (XRS) and total X-ray scattering.

5.2.1 Fourier Transform Infrared (FTIR) Spectroscopy

Solid imidazole was measured in attenuated total reflectance (ATR) mode and liquid samples were measured in transmission and ATR modes on a Thermo Scientific Nicolet 6700 FTIR spectrometer. Omnic software was used to collect the background and sample spectra. The spectra were collected at a resolution of 4 cm^{-1} in 32 scans over 4000-500 cm^{-1}.

5.2.2 X-ray Raman Scattering (XRS)

The C, N and O K-edge spectra of aqueous imidazole solutions were collected at the inelastic scattering beamline ID-20 of the ESRF, Grenoble, France [183]. ID-20 is an undulator beamline that uses Si(111) pre-monochromator which can be coupled by a consecutive monochromator to form a fixed exit ensemble. Si(311) post-monochromator was utilised to achieve an energy resolution of 0.6 eV which is needed for this experiment. All data were collected using the inverse scanning mode, where the incident photon energy was scanned and the analyser energy was fixed at 9690 eV. The scattered radiation was collected employing a 12 spherically bent Si(660) analyser crystals at an angle of 42° with q value of 3.6 $Å^{-1}$. Hence, the data presented here were collected at low momentum transfer (q). The monochromator drift was small, and it was monitored through frequent measurement of the elastic scattering peak.

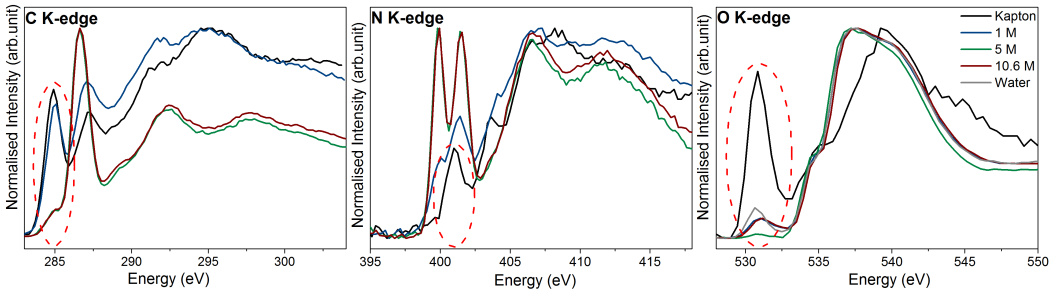

Figure 5.1. The C, N and O K-edge XRS spectra of Kapton, the aqueous imidazole solutions (1 M, 5 M and 10.6 M) and water. The peaks used to define the subtraction coefficient to remove the Kapton contribution to the K-edge signal of the aqueous solutions of imidazole and of water are circled in red.

The aqueous imidazole solutions (concentrations of 1 M, 5 M and 10.6 M) were measured using a jacketed flow-cell crystalliser (flow-cell 1, Chapter 2) which ensures precise temperature control and continuous circulation of the stock solution to minimise

radiation induced damage to imidazole [160]. The total volume of the solution was ~250 ml and the solution was pumped at a rate of ~450 ml/min. For data analysis, the elastic energy was subtracted from the XRS data to obtain the near-edge spectra. The spectra were then background subtracted and normalised to the edge-step height using Athena software [239]. After normalisation, the contribution from the Kapton window to the K-edge spectra of the aqueous imidazole solutions was removed by subtracting a reference spectrum of Kapton from those of the solutions. The subtraction coefficients were defined as the value required to remove a peak significant to the K-edge spectra of Kapton. These peaks for the C, N and O K-edges are circled in Figure 5.1.

5.2.3 Total X-Ray Scattering and X-ray Pair Distribution Function (XPDF)

Total X-ray scattering data of imidazole in solid and aqueous solutions states were collected at the X-ray pair distribution function (XPDF) beamline I15-1 of Dimond Light Source (DLS), UK [192]. The solid sample was measured in a borosilicate glass capillary while the liquid samples (water and aqueous imidazole solutions with concentrations of 1 M, 5 M and 10.7 M) were measured using flow-cell 2 (illustrated in chapter 2). As mentioned above, the flow-cell ensured continuous circulation of the solution to minimise beam damage and to ensure uniform temperature distribution.

The acquisition time per sample was 10 minutes. The raw data was automatically corrected for the internal dark current during data collection using DLS generic data acquisition (GDA) software. Further corrections for instrument and sample holder backgrounds, absorption, multiple scattering and Compton scattering were performed in GudrunX.[224] Moreover, GudrunX was utilised to convert total scattering data to pair distribution functions (PDFs) using Fourier transformation routines [224, 248].

5.2.4 Empirical Potential Structure Refinement (EPSR)

The total X-ray scattering data were used to refine structural models generated by the EPSR software. The EPSR simulation starts by utilising reference potentials (based on the Lennard-Jones potentials and Coulomb charges) to generate the initial configuration of atoms using a standard Monte Carlo simulation. After equilibration, an additional empirical potential is refined iteratively by perturbation of the reference potential and the structure factor (F(Q)) is calculated for each new atomic configuration. The atomic

configuration is either accepted or discarded according to the goodness of the agreement between the experimental and calculated F(Q). Finally radial and spatial distribution of atomic and molecular pairs were calculated using routines within EPSR and Dlputils [200].

The side length of the cubic box for the generated models, the number of imidazole and water molecules and the number density of water and aqueous imidazole solutions (concentrations of 1 M, 5 M and 10.7 M) are shown in Table 5.1. The water molecules were constructed according to Soper [204] and the imidazole molecules were constructed in Jmol. Bond and angle constraints were employed, as well as torsional constraints for the imidazole molecules. However, it is worth noting that the molecules were not held rigid during the EPSR simulation in order to reproduce the experimental diffraction data [198]. The Lennard-Jones potential and Coulomb charge parameters for water were obtained from the SPC/E model [226] and those for imidazole were taken from the OPLS force field [227] and they are listed in Table 5.2.

Table 5.1. The side length of the cubic box, the number of imidazole and water molecules and the number density of the water, 1 M, 5 M and 10.7 M aqueous imidazole solutions

	Water	1 M	5 M	10.7 M
Side length of cubic box	31.05	32.05	34.94	35.39
Number of imidazole molecules	-	20	128	286
Number of water molecules	1000	1040	1000	500
Number density (atoms.Å^{-3})	0.1002	0.1002	0.0974	0.0919

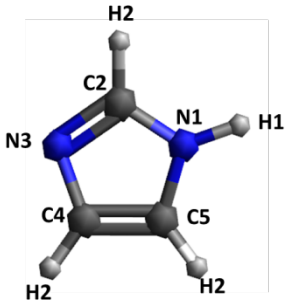

Figure 5.2. Atom labelling used for the Lennard-Jones potentials in the EPSR simulation of imidazole.

Table 5.2. Lennard-Jones potentials and Coulomb charges for water [226] and imidazole [227] The labelling scheme used for imidazole is shown in Figure 5.2

Atom	ε / kJ mol-1	σ /Å	Charge / e-
Ow	0.650	3.166	-0.848
Hw	0.000	0.000	0.424
N1	0.711	3.250	-0.570
C5	0.293	3.550	-0.015
H1	0.000	0.000	0.420
C2	0.293	3.550	0.015
C4	0.293	3.550	0.295
H2	0.126	2.420	0.115
N3	0.711	3.250	-0.490

5.3 Results and Discussion

5.3.1 Fourier Transform Infrared (FTIR) Spectroscopy

Extensive vibrational data of imidazole in solid [249-253] and aqueous solution phases [252, 253] has been reported in the literature. In the FTIR spectrum of solid imidazole, the N-H stretch band (3100 – 2200 cm^{-1}) is broad with a number of sub-bands (Figure 5.3). These have been attributed to Fermi interactions of the stretch band with overtones and combinations of the lower frequency vibrations [254-256]. These sub-bands are detected in the 5M aqueous imidazole solution FTIR spectrum alongside other bands in the range between 1600 – 1000 cm^{-1}, which are attributed to stretching and bending vibrational modes of the imidazole ring (Figure 5.3, Table 5.3). The vibrational frequencies of the N-H stretch sub-bands shift to lower values (red shift) while the N-H bending frequency (1533 cm^{-1}) shifts to a higher value (blue shift) in the solid FTIR spectrum when compared to the spectrum of the aqueous solution (Table 5.3). This indicates that the intramolecular N-H bond in imidazole is weaker in the solid-state than in aqueous solution, which could reflect stronger hydrogen-bonding in the solid.

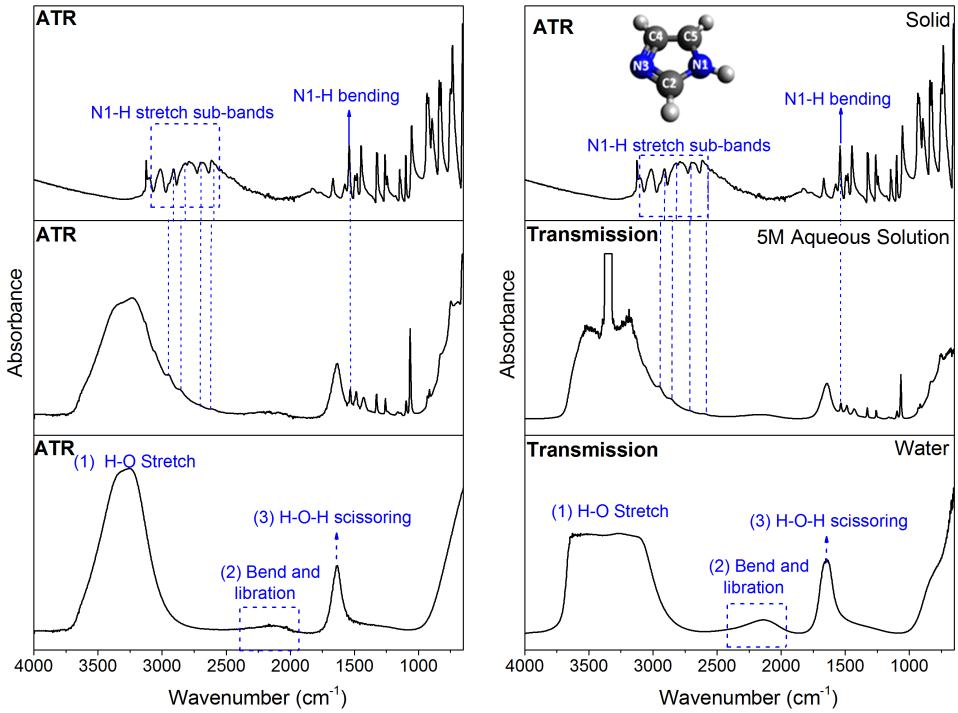

Figure 5.3. FTIR spectra of solid imidazole, 5 M aqueous solution of imidazole and water in the range 4000 – 650 cm^{-1} collected using ATR and transmission.

The FTIR spectrum of water has four significant vibrational bands; (1) Ow-Hw stretch (3700 – 2900 cm^{-1}), (2) Ow-Hw bend and libration (2130 cm^{-1}), (3) Hw-Ow-Hw in-plane bending (scissoring) (1650 cm^{-1}) and (4) Ow-Hw libration (less than 1000 cm^{-1}). The Ow-Hw stretch band is a broad strong band, which could mask other vibrational bands in the same frequency range especially if the data is collected in transmission mode. Adding to that, the signal from the Ow-Hw stretch band can saturate the FTIR spectra as seen in Figure 5.3. Hence the FTIR spectra of water and aqueous imidazole solutions were collected in ATR mode to overcome these challenges. The frequency of the Ow-Hw bend and libration band depends on frequency of both the Hw-Ow-Hw bending mode and the intermolecular libration motion. Hence, this band reports on the structure and dynamic of the hydrogen-bond network in water [231]. The four vibrational bands of water were less intense in the FTIR spectrum of the 5M aqueous imidazole solution than in the spectrum of pure water (Figure 5.3).

Table 5.3. Summary of FTIR vibrational bands frequencies and assignments for imidazole solid, 5 M aqueous imidazole solution and water

Imidazole Solid	5M aqueous imidazole solution	Water	Assignment
1054.9	1066.5		ν(ring), δ(CH) out of plane
1145.5	1160.9		ν(ring) δ(CH) out of plane
1261.2	1259.3		ν(ring)
1321.0	1328.7		ν(ring), ν (CN)
1446.4	1430.9		ν(ring), δ(CH)
1540.9	**1533.2**		**ν(ring), δ(N1H)**
1575.6			C-C stretch
	1646.9	1646.9	H-O-H scissoring
	2134.8	2134.8	H-O-H Bend and libration
2510.9	2516.7		N-H stretching sub-bands
2582.3	**2599.6**		
2699.9	**2701.8**		
2817.5	**2856.1**		
2908.2	**2958.3**		
2946.7			
3012.3			
3043.2			

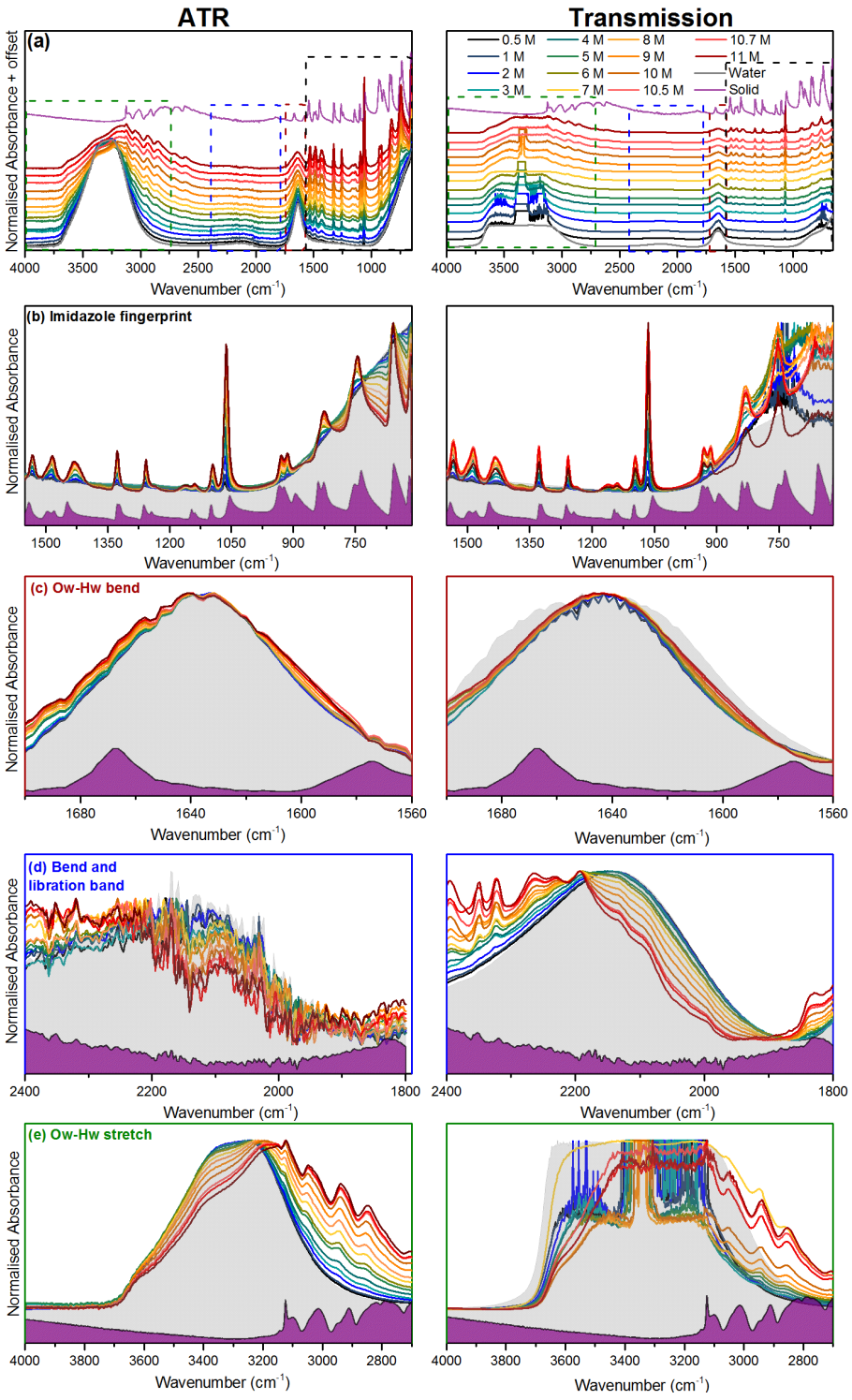

Figure 5.4. FTIR spectra of aqueous imidazole solutions in concentration range (0.5 M – 11 M) overlaid with those of water (grey) and solid imidazole (purple) collected using ATR (left) and transmission (right) (a) FTIR spectra in range between 4000 – 650 cm^{-1} (b) Fingerprint region (1550 – 650 cm^{-1}) (c) Ow-Hw bend band (1700 – 1560 cm^{-1}) (d) water bend and libration band (2400 – 1800 cm^{-1}) (e) Ow-Hw stretch band (4000 – 2700 cm^{-1}).

The intensity of the vibrational bands corresponding to imidazole are expected to increase as a result of increasing the concentration of imidazole in aqueous solutions, which is evident in the fingerprint region (Figure 5.4(b)), the water bend and libration region (Figure 5.4(d)) and the Ow-Hw stretch region (Figure 5.4(e)). These vibrational bands (associated with imidazole) are well represented in the concentration series (0.5 M–11 M) shown in Figure 5.4 and no significant shifts in the frequency of these features are observed as a function of increasing the concentration of imidazole. This is in line with previously reported Raman spectra of aqueous imidazole solutions [217]. On the other hand, the vibrational bands associated with water show a decrease in intensity and shifts in frequency values as a consequence of increasing the concentration of imidazole. The bend and libration band of water shifts to higher wavenumber (blue shift) (Figure 5.4(d)-transmission) whereas the Ow-Hw stretch band shifts to a lower wavenumber (red shift) (Figure 5.4(e) ATR) as a result of increasing the concentration of the aqueous imidazole solutions from 0.5 M to 11 M. These shifts indicate that the intramolecular Ow-Hw bond becomes weaker as a result of increasing the concentration of imidazole, which could reflect stronger intermolecular hydrogen-bonding and more rigid hydrogen-bonded network of water. Since the number of water molecules available to maintain the hydrogen-bond network in water is not sufficient upon increasing the concentration of imidazole in solution, imidazole is expected to play a role in strengthening the water network. Especially that hydrogen-bonding between imidazole and water pairs becomes dominant at higher concentrations.

5.3.2 X-ray Raman Scattering (XRS)

C and N K-edge

The near-edge region of the C and N K-edge spectra of solid imidazole and aqueous imidazole solutions are dominated by transitions from the 1s orbitals to unoccupied π* states (Figure 5.5). These are followed by the ionisation potentials (IP) and broad σ* shape resonances corresponding to multiple scattering at the potential barrier to the bound neighbouring atoms. The features in the K-edge spectra were deconvoluted with non-linear least squares fitting using Fityk software [244]. Gaussian functions were used to model π*and σ* resonances and arctan functions were used to model the ionisation potentials [168]. It is worth noting here that the C and N K-edge spectra of the 1 M

aqueous imidazole solution were dominated by features from the Kapton window (Figure 5.1) and separating the weak signal from imidazole from the signal of Kapton was difficult; hence the Kapton contribution to the 1 M aqueous solution spectra was not subtracted.

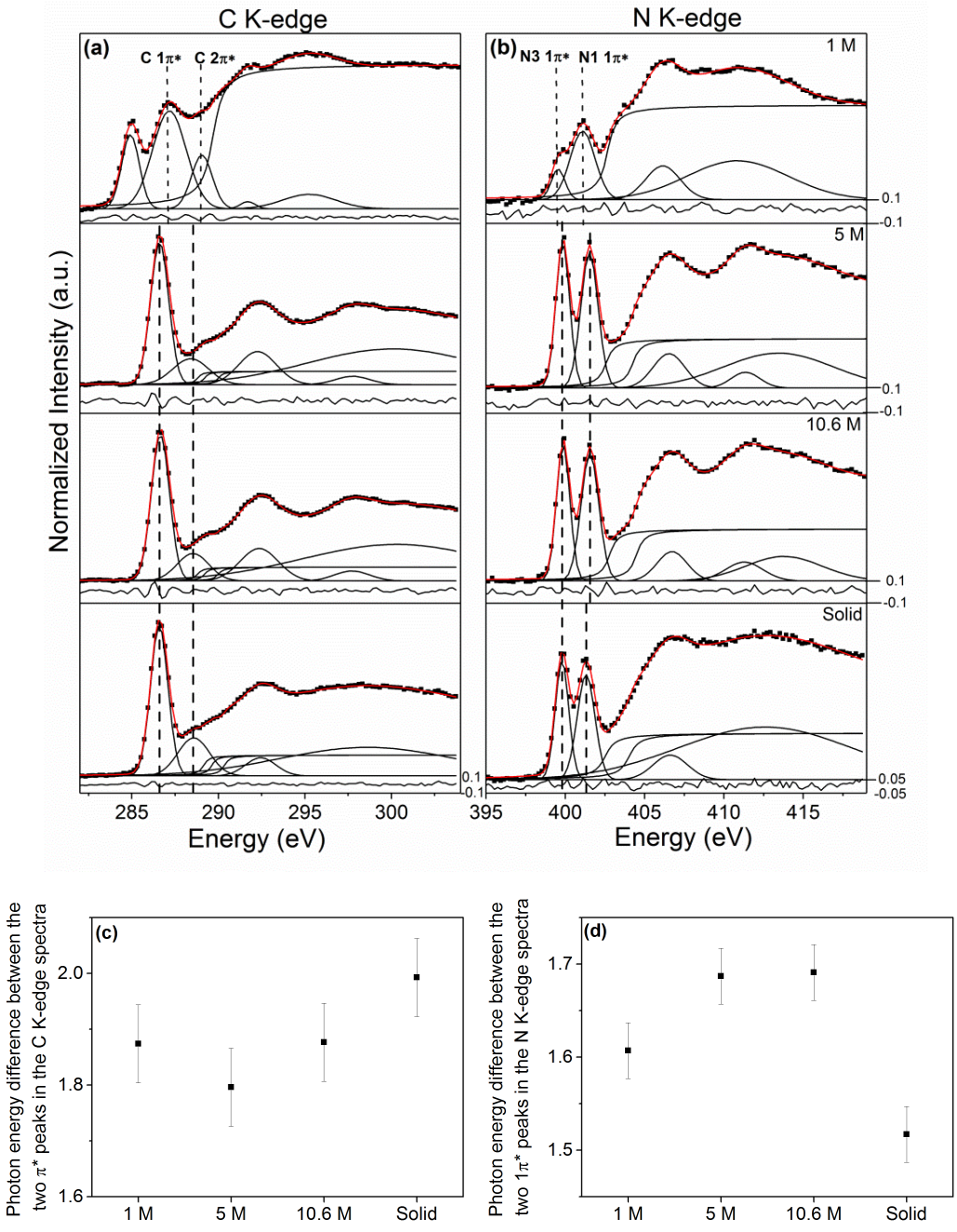

Figure 5.5. Fitted (a) C K-edge (b) N K-edge spectra of 1 M, 5 M and 10.7 M aqueous imidazole solution and solid imidazole. The photon energy difference between the two π* peaks in the (c) C K-edge (d) N K-edge of the previously mentioned spectra.

The near-edge region of the C K-edge spectra is dominated by two features corresponding to transitions from the 1s orbital of the carbon atoms to the unoccupied 1π* and 2π* orbitals. The energy difference between these two features is 1.874 eV, 1.796 eV, 1.877 eV and 1.993 eV for the 1 M, 5 M, 10.6 M aqueous imidazole solutions and solid imidazole respectively. Similarly, the near-edge region for the N K-edge spectra is dominated by two intense peaks corresponding to 1s → 1π* transitions on N3 and N1 moieties. These peaks are separated by 1.607 eV, 1.687 eV, 1.691 eV and 1.517 eV for the 1 M, 5 M, 10.6 M and solid correspondingly. Their separations are comparable to values reported for aqueous solutions of imidazole [161, 217] and solid imidazole [127].

The energy differences between the peaks in the near-edge region of the C and N K-edges are affected by the local bonding environment as shown for imidazole in the gas, aqueous solution and solid-phases [88, 160, 217]. The energy difference between the two features (ΔE (π*)) in the near-edge region of the C K-edge spectra of the 5 M and 10.6 M aqueous solutions is comparable to that of the solid-phase (Figure 5.5(c)). This indicates that the electron density at the carbon moieties in the solution phase is comparable to that of the crystalline phase. On the other hand, the electron density at the nitrogen centres in the solution phase is lower than that of the solid-phase as indicated by the higher ΔE (1π*) values in the N K-edge spectra of the solution phase when compared to the solid-phase (Figure 5.5(d)).

It is challenging to directly compare the K-edge spectra of the 1 M aqueous solution to those of the higher concentration solutions and the solid because of the significant contribution from the Kapton to the C and N K-edge signals. The variation in the energy difference between the π* peaks in the near-edge region of the C and N K-edge spectra for the 5 M and 10.6 M aqueous imidazole solutions is minimal (Figure 5.5(c) and (d)). This suggests that the local structure around the imidazole molecules in aqueous environments does not change as a result of varying the concentration of imidazole. Hence, the solvated clusters of monomeric imidazole molecules, which were predicted in chapter 3 by the neutron and X-ray diffraction combined with the empirical potential structure refinement (EPSR) simulation, are not disturbed as more imidazole molecules are added to the solution.

O K-edge

The O K-edge spectra of aqueous imidazole solutions only report on changes in the local environment around the water molecules since there is no direct contribution from imidazole to the O K-edge signal. The spectral features in the O K-edge spectrum of water and their origin will be discussed and compared to those of ice before embarking on explaining the changes in the O K-edge spectra of aqueous imidazole solutions. The O K-edge of water has three main regions: the pre-edge (534 eV), the main-edge (536.5 – 539.5 eV) and the post-edge (541 eV) and the variations in these spectral features are mostly correlated [41, 42, 45, 257]. For example, the O K-edge spectrum of bulk ice has a prominent post-edge feature and less intense pre- and main-edge features (Figure 5.6(a)). On the other hand, the spectrum of ice surface shows a reduction in the intensity of the post-edge region correlated with enhanced intensity in the pre- and main-edge regions (Figure 5.6(b)).[41] While the water molecules in the bulk ice are fully coordinated by four molecules in a tetrahedral structural motif, the molecules at the ice surface have one free dangling OH group. Hence, the variations in the spectral features of the O K-edge could be related to the bonding coordination in the system. In other words, a dominant post-edge feature in the O K-edge spectra is attributed to fully coordinated water molecules whereas intense pre and main-edge peaks are associated with broken or weak hydrogen bonds.

Treatment of the ice surface with ammonia (NH_3) further illustrates the impact of hydrogen-bonding on the O K-edge. Similar to the water molecules in the bulk ice, the water molecules at the ice surface become fully coordinated after treating the surface with ammonia because the free OH groups at the ice surface are now bonded to the NH_3 molecules. These changes in the coordination environment are reflected on the spectral features of the O K-edge (Figure 5.6(c)), which are now comparable to those of the bulk ice (Figure 5.6(a)). Another example that demonstrates the effect of local coordination on the O K-edge spectra is the reduction in the intensity of the pre- and main-edge features in the O K-edge when comparing the spectrum of ice phase III to those of phases II and IX (Figure 5.7(b)) [258]. In ice phase III, only 25% of the protons (i.e hydrogen atoms) in the crystal lattice are ordered compared to complete proton ordering in phases II and IX. This indicates that some of the hydrogen bonds in phase III are not coordinated which could result in the observed enhanced intensity in the pre- and main-

edge features of the O K-edge. However, this enhanced intensity does not always relate to a change in the coordination environment around the water molecules. An increase in density can increase the intensity of the pre- and main-edge features as seen in the case of ice phases Ih, III, VI and VII (Figure 5.7(c)), which have similar hydrogen-bond networks [259].

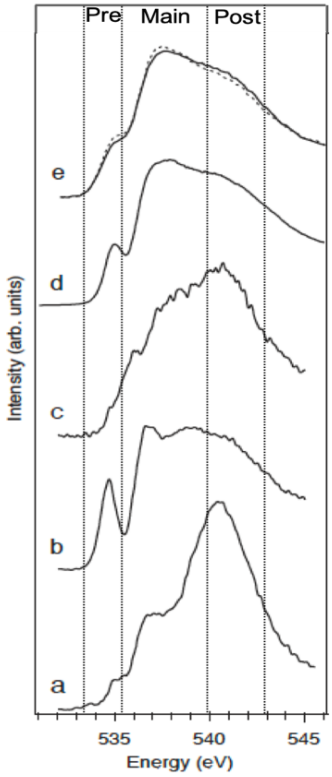

Figure 5.6. Ice and water O K-edge spectra from reference [41] (a) Bulk ice (b) Ice surface (c) Ice surface with first layer treated with NH_3 (d) Liquid water at ambient condition (e) bulk liquid water at 25°C (solid line) and at 90°C (dashed line).

The O K-edge spectra of water have intense pre- and main-edge features and a less intense post-edge peak (Figure 5.6(d) and (e)). The similarities in the O K-edge spectral features of water with those of the ice surface could indicate that the water molecules are not fully coordinated in the liquid state. This agrees with the electronic structure calculations of the O K-edge X-ray absorption (XA) spectra where the structure of water with broken hydrogen-bonds gave the most comparable representation of the experimental XA spectrum [260, 261]. The effect of coordinated hydrogen bonds on the O K-edge of water is further illustrated by probing the changes in the spectral features as a result of increasing temperature [257, 262] and pressure [257]. The hydrogen-bond network of water is distorted as a result of increasing temperature and/or pressure as

illustrated from molecular dynamics simulations [257]. These distortions were associated with enhancing the intensity of the pre- and main-edge features and reducing that of the post-edge peak in the O K-edge spectra of water [257, 262].

The changes in the hydrogen-bond network of water induced by the addition of ions (Li^+) [215], osmolytes (TMAO) [118], and denaturants (urea) [118] were investigated by the O K-edge spectra of aqueous solutions. The O K-edge spectra of aqueous LiCl solutions in a range of concentration showed an increase in the intensity of the pre- and main-edge regions correlated with a decrease in the intensity of the post-edge region as a function of increasing the solute concentration [215]. These spectral changes in the O K-edge have been attributed to the disturbance of the water network by Li^+ ions. On the other hand, the reduction in the intensity of the main-edge feature in the O K-edge of aqueous TMAO solutions was taken as evidence for strengthening the hydrogen-bond network of water by TMAO [118]. The O K-edge spectra of aqueous urea solutions showed minimal variation when compared to that of water even at high concentrations [118]. Consequently, the effect of urea on the hydrogen-bond network of water was concluded to be minor.

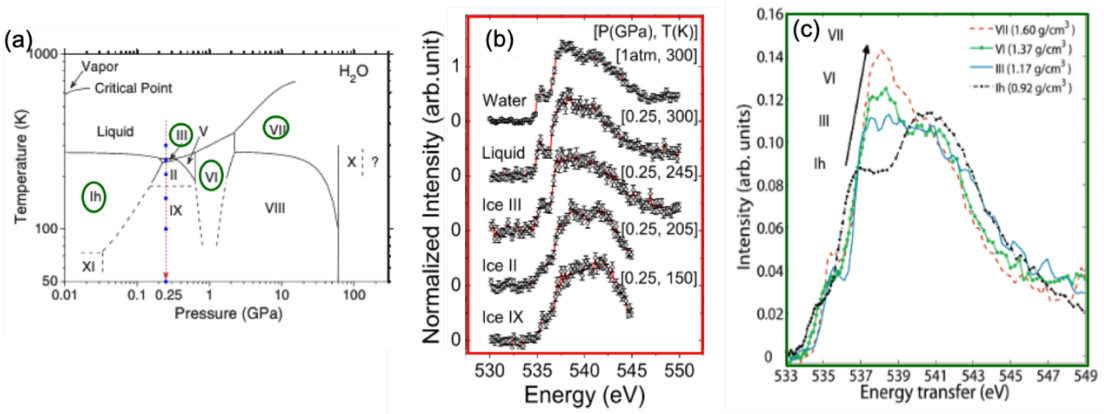

Figure 5.7.(a) Phase diagram of water from reference [258] the dashed red arrow and solid blue dots indicate water and ice phases presented in (b) and the green circles indicate the ice phases presented in (c) (b) O K-edge spectra of water at ambient pressure and 0.25 GPa alongside with ice phases III, II and IX from reference [258] (c) O K-edge spectra of ice phases Ih, III, VI and VII (from reference [259]) which illustrate the evolution in the spectral features with increasing density.

In this chapter the effect of imidazole on the coordination environment of water has been examined by comparing the O K-edge spectra of three solutions of aqueous

imidazole (concentrations of 1 M, 5 M and 10.6 M) with the spectrum of water (Figure 5.8). Increasing the concentration of imidazole reduces the number of available hydrogen bonds formed by water from 208 accepting and donating bonds in the 1 M aqueous imidazole solution (containing 52 water molecules per solute molecule) to 31 in the 5 M solution (with 7.8 water molecules per imidazole molecule) to 6 hydrogen bonds in the 10.6 M solution (which contains 1.6 water molecules per imidazole molecule). This would indicate that increasing the concentration of imidazole in the solution will heavily disturb the water network resulting in adverse changes in the O K-edge spectra. However, the variations in the O K-edge spectra as a result of increasing the imidazole concentration are minor. Adding to that, the changes in the intensity of the pre-, main- and post-edge features reported here do not correlate in the same manner as the patterns discussed above. A gradual increase in the pre-edge and a systematic decrease in the post-edge are observed as a result of increasing the concentration of imidazole. However, these are not associated with an increase in the intensity of the main-edge feature. All of this indicates that the presence of imidazole does not disturb the hydrogen-bond network of water. On the contrary, imidazole plays an important role in maintaining that network especially at higher imidazole concentrations.

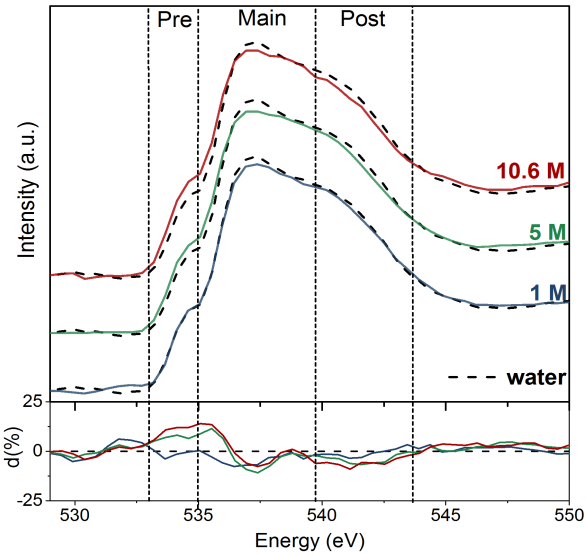

Figure 5.8. O K-edge spectra of 1 M, 5 M and 10.6 M aqueous imidazole solution and water. The difference spectra at the bottom of the figure illustrate the increase in the intensity of the pre-edge feature and the decrease in the intensity of the post-edge peak.

5.3.3 Total X-ray scattering

The X-ray pair distribution function (XPDF) patterns of 1 M, 5 M and 10.7 M aqueous imidazole solutions and water were generated through Fourier transformation of the Total X-ray scattering patterns. At 1 M imidazole concentration, there are ~52 water molecules per imidazole molecule resulting in minimal disruption to the water network. This is evident from the resemblance of the XPDF pattern of the 1 M solution to that of water (Figure 5.9) as well as the high similarity index of 94% between the two patterns (Figure 5.10). As the concentration increases from 1 M to 5 M, the XPDF pattern develops features attributed to imidazole (Figure 5.9). For example, the PDF peaks associated with the intramolecular bond lengths in imidazole are more prominent in the 5 M solution. The PDF peaks corresponding to water-water correlations in the second and third hydration shells (centred at 4.5 Å and 6.8 Å respectively) are slightly distorted in the 5 M XPDF pattern when compared to the 1 M and the water patterns. Nonetheless, the features in the range longer than 12 Å resemble those of water. The XPDF pattern of the 10.7 M solution has strong features attributed to imidazole (Figure 5.9) and a similarity index of 74% with the PDF pattern of solid imidazole (Figure 5.10). The reduction in the intensity of the peak corresponding to the water in first hydration shell (centred at 2.9 Å) and the distortion of the peaks attributed to the second and third hydration shells (centred at 4.5 Å and 6.8 Å respectively) indicate that the imidazole molecules heavily disturb the intermolecular interactions of water - water in the short and the intermediate ranges. Features between 3.7 Å and 22 Å are not similar to those of water nor solid imidazole, indicating strong imidazole – water correlation.

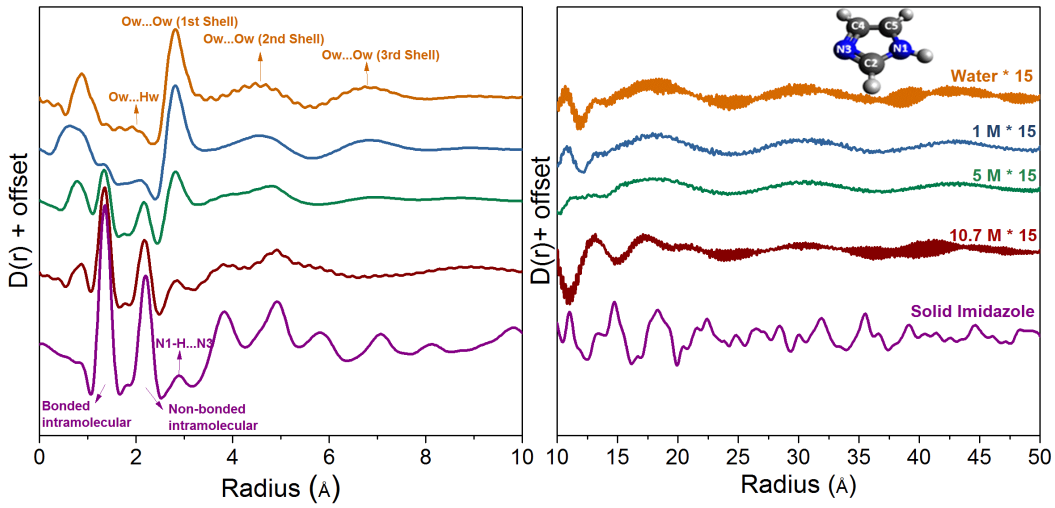

Figure 5.9. XPDF patterns of solid imidazole, aqueous imidazole solutions (1 M, 5 M and 10.7 M) and water in range 0 -10 Å (left) and 10 – 50 Å (right).

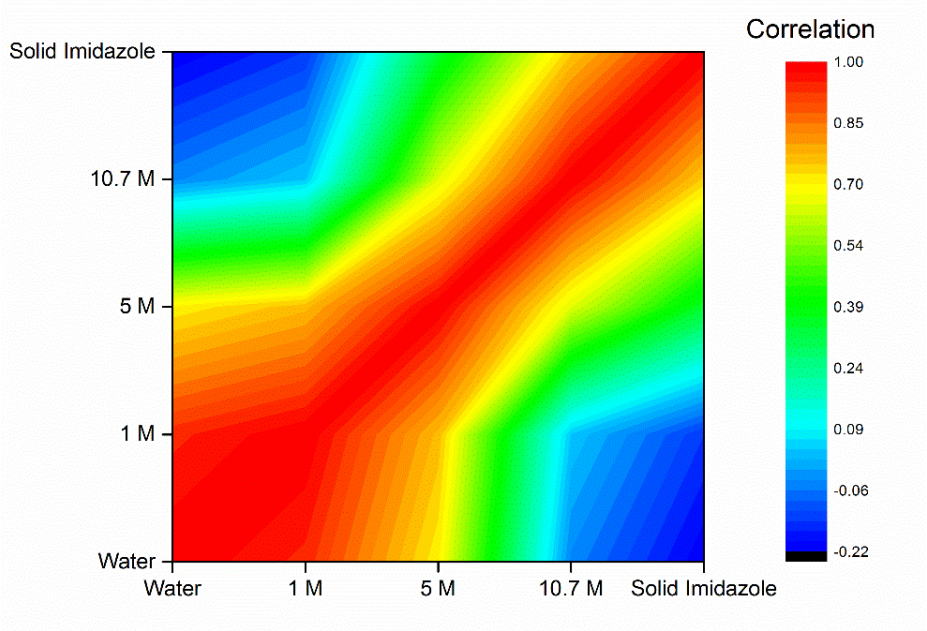

Figure 5.10. Correlation matrix of solid imidazole, water, 1 M, 5 M a 10.7 M aqueous imidazole solution XPDF patterns.

Quantitative analysis of the atomic correlations in aqueous imidazole solutions from the XPDF patterns shown in Figure 5.9 is not feasible. The reported XPDF patterns represent the sum of 78 partial PDFs, which are attributed to all the atomic pair-pair correlations present in aqueous imidazole solutions. Therefore, the empirical potential structure refinement (EPSR) simulations were utilised to generate structural models for the probed aqueous imidazole solutions and for water. These models were refined utilising the same total X-ray scattering data, which were used to generate the XPDF patterns in

Figure (5.9). The structure factor F(Q) patterns calculated from the EPSR simulation compare well with those obtained from the experimental scattering data (Figure 5.11).

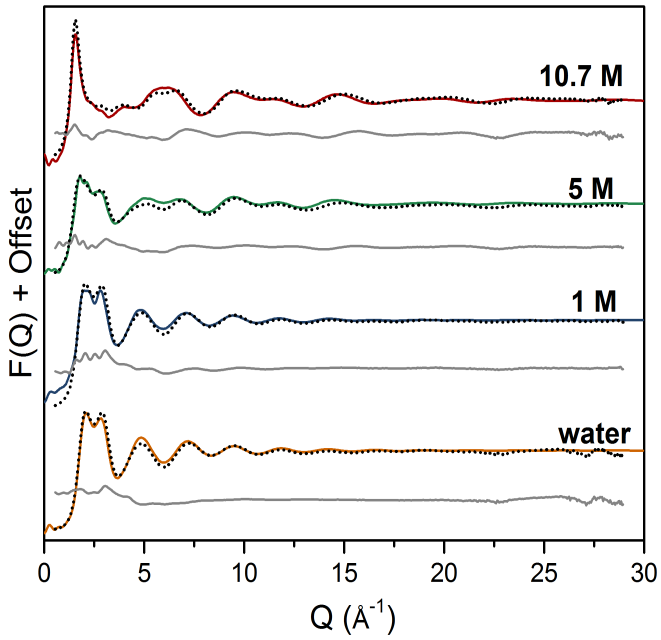

Figure 5.11. F(Q) values from experimental data (black dots) and EPSR models (coloured lines). The difference patterns (grey lines) illustrate the high comparability between the experimental and calculated F(Q) patterns.

Strong imidazole – water interactions are anticipated in aqueous imidazole solutions because imidazole and water are strong hydrogen-bond donors and acceptors. The difference of 1 Å between the first peak in the N1…Ow radial distribution function (RDF) and H1…Ow RDF and between N3…Ow RDF and N3…Hw RDF (shown in Figure 5.12 (a)) indicates collinear hydrogen-bonding. This is also inferred from the intense peak at ~ 4 Å in the radial distribution function of the centre of geometry (cog) of water around the cog of imidazole (cog) (Figure 5.12 (b)). The shoulder at ~ 3 Å in the RDF of the cog of water around imidazole is attributed to water molecules that are not collinearly bonded to the imidazole molecule. The location of water molecules around imidazole is clearly illustrated in the spatial density functions (SDFs) shown in Figure 5.13, which clearly show the presence of water directly adjacent to the N1 site and angularly distributed around the N3 site in the imidazole molecule. The intermolecular interactions between imidazole and water predicted here are in line with those predicted in chapter 3 for 5 M aqueous imidazole solution.

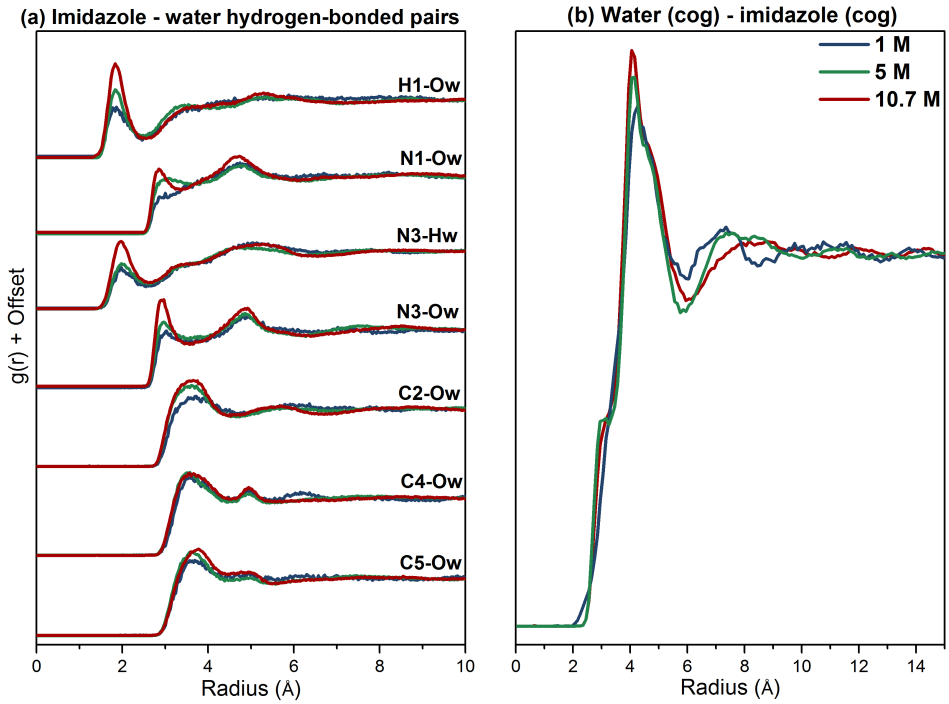

Figure 5.12. (a) The radial distribution functions (RDFs) of imidazole-water hydrogen-bonded pairs (b) The radial distribution function (RDF) of the centre of geometry (cog) of water around the cog of imidazole.

Increasing the concentration of imidazole in aqueous solutions lowers the width of and the area under the peaks in the RDFs of hydrogen-bonded pairs (Figure 5.12 (a)). This indicates a reduction in the number of imidazole – water bonded pairs as a result of increasing the imidazole concentration. The coordination number of Ow and Hw around the nitrogen atoms in the imidazole ring reduces by 12 % as the concentration increases from 1 M to 5 M (Table 5.4). A further 50% reduction is observed as the concentration increases to 10.7 M (Table 5.4). On the other hand, the effect of increasing the imidazole concentration on the positions of the peaks in the RDFs shown in Figure 5.12 is minimal; suggesting that the nature of the bonding environment around the imidazole molecules does not change. This is clearly pictured in the spatial density function of water around imidazole (Figure 5.13) and could explain the minor changes in the C and N K-edge spectra of aqueous imidazole solutions as a result of increasing the concentration of imidazole were (reported in section 5.3.2).

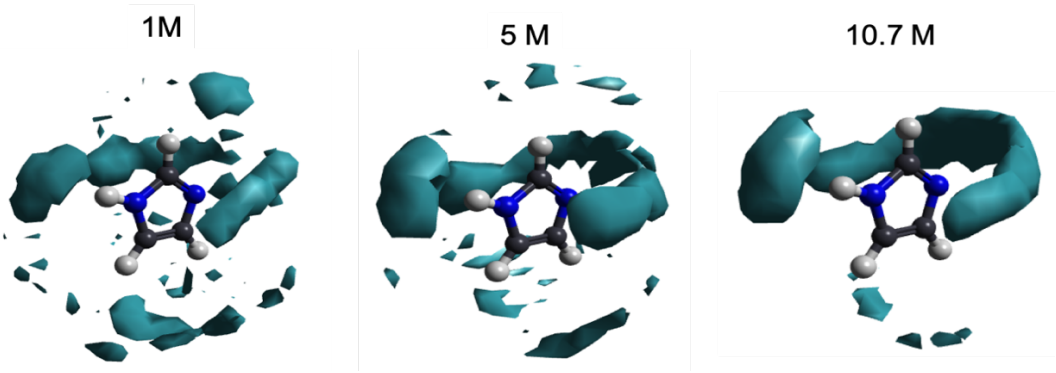

Figure 5.13. Spatial density function (SDF) of water around imidazole for distances up to 6 Å (10 % of water molecules shown).

Table 5.4. Coordination numbers of hydrogen-bonded imidazole – water pairs

Central atom-surrounding atom	Sample	Coordination number	Cut-off / Å
H1-Ow	1 M	0.7	2.4
	5 M	0.7	
	10.7 M	0.4	
N1-Ow	1 M	3.2	3.7
	5 M	2.8	
	10.7 M	1.3	
N3-Hw	1 M	1.4	2.5
	5 M	1.3	
	10.7 M	0.9	
N3-Ow	1 M	2.8	3.6
	5 M	2.5	
	10.7 M	1.3	
Water (cog) – imidazole (cog)	1 M	1.2, 27	3.4, 6
	5 M	1.1, 19.4	
	10.7 M	0.5, 9.7	

The water-water intermolecular interactions in aqueous imidazole solutions are examined through the radial distribution functions (RDFs) of the atomic pair-pair correlations in water (Figure 5.14 (a)) and the RDFs of the cog of water around the cog of water (Figure 5.14 (b)). The intensities and positions of the RDFs (both pair-pair and cog-cog) for the 1 M aqueous imidazole solutions are similar to those of pure water (Figure 5.14). The intensity of the RDF peaks for the 5 M and 10.7 M solutions are higher than those of water and the 1 M solution. This is attributed to the excluded volume effect where the presence of solute molecules increases the local density of water resulting in a higher average intensity [228]. On the other hand, the width of the peaks in the RDFs for the 5 M and 10.7 M solutions is smaller than that of water and the 1 M solution (Figure 5.14). The reduction in the width of the peaks correlates with a

reduction in the area under the peaks which indicates a reduction in the number of water molecules coordinating other water molecules in the named solutions. The reduction in the number of water molecules in the first hydration shell of water (up to 3.6 Å) is clearly illustrated in Table 5.5.

Moreover, the positions of the second and third peaks in the cog RDFs for the 5 M and 10.7 M solutions deviate from those of water and the 1 M solution (Figure 5.14 (b)). This suggest disruption to the intermolecular interactions between the water molecules in the 5 M and 10.7 M solutions, which is clearly evident in the spatial density functions (SDFs) of water around water shown in top panel of Figure 5.15. Nonetheless, the coordination environment around water seems intact as imidazole replaces water in the first and second hydration shells (at distances of ~3.6 Å and ~6 Å respectively) as seen from the bottom panel in Figure 5.14. Preserving the hydrogen-bond network of water by imidazole explains the broad similarity in the spectral features of the O K-edge spectra of aqueous imidazole solutions with different imidazole concentrations (which were discussed in section 5.3.2).

Table 5.5. Coordination numbers of atomic pairs in water

Central atom-surrounding atom	Sample	Coordination number	Cut-off / Å
Ow - Ow	Water	5.8	3.4
	1 M	4.4	
	5 M	3.8	
	10.7 M	2.5	
Ow – Hw	Water	1.8	2.4
	1 M	1.6	
	5 M	1.5	
	10.7 M	1.1	
Water (cog) – water (cog)	Water	6.3	3.65
	1 M	6.1	
	5 M	4.7	
	10.7 M	3	

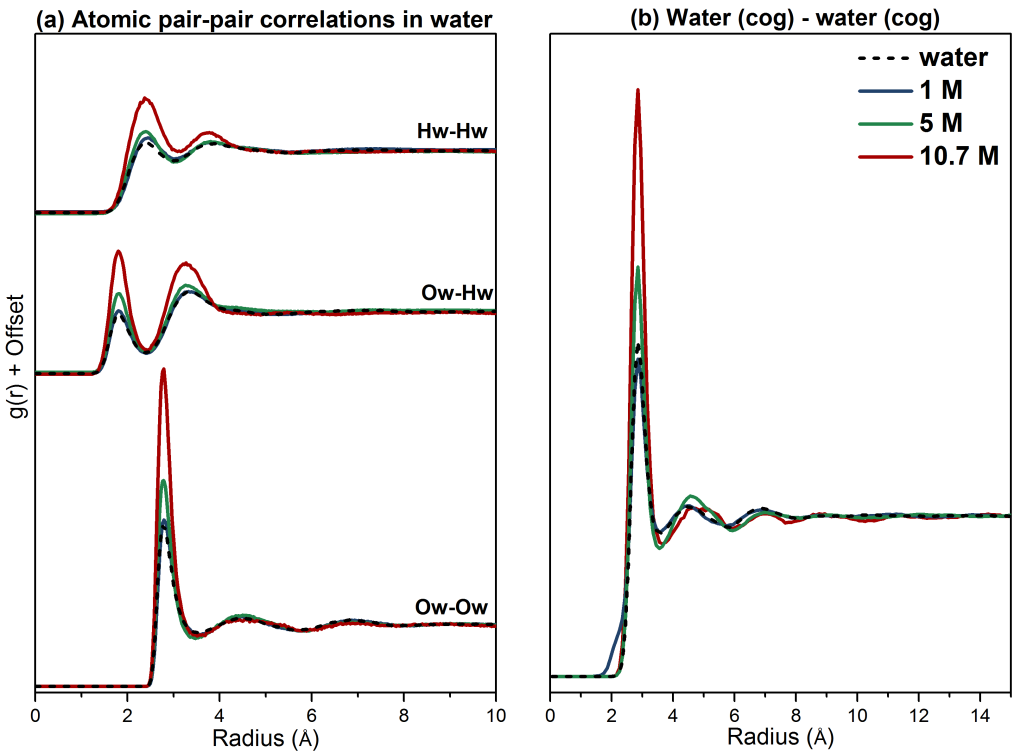

Figure 5.14. (a) The radial distribution functions (RDFs) of the atomic pair-pair correlations in water (b) The radial distribution function (RDF) of the centre of geometry (cog) of water around cog of water.

Direct intermolecular interactions between imidazole molecules are insignificant in the 1 M solution as illustrated from the radial distribution functions (RDFs) of the hydrogen-bonded pairs of imidazole (Figure 5.16(a)) and the RDF of the cog of imidazole around the cog of imidazole (Figure 5.16(b)). Linear N1-H1…N3 bonds in the 5 M and 10.7 M solutions are inferred from the 1 Å difference between the first peak in the N1…N3 RDF and the N3…H1 RDF (Figure 5.16(a)). However, the probability of finding such species is less than 1 as seen from the coordination numbers of H1 around N3 and N3 around N1 at distances of 2.9 Å and 3.5 Å respectively (Table 5.6).

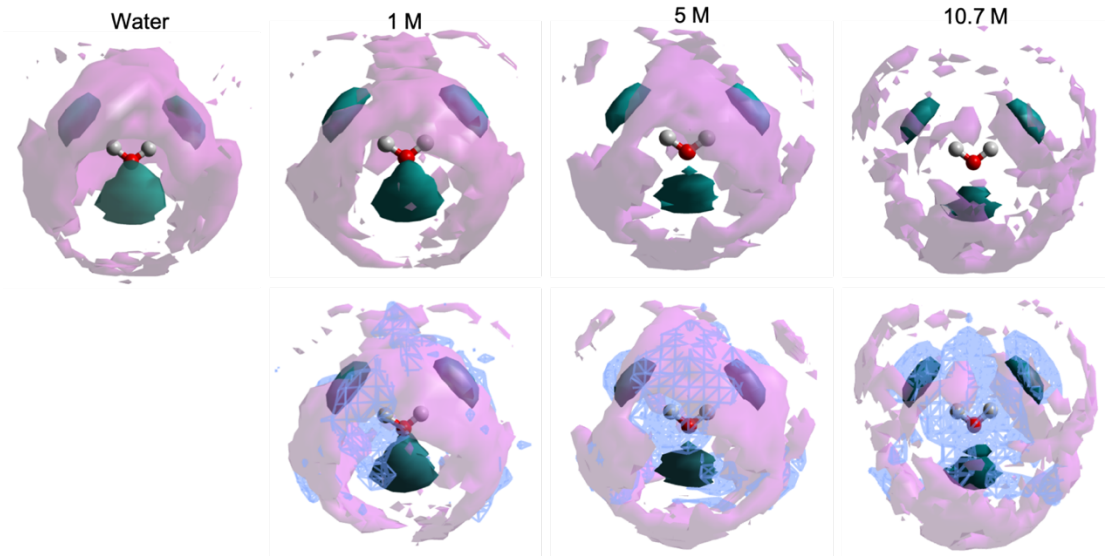

Figure 5.15. (Top) Spatial density function (SDF) of water around water in the first hydration shell (solid green surfaces) for distance up to ~ 3.6 Å (30 % of water molecules shown) and second hydration shell (transparent pink surfaces) for distance up to ~ 5.8 Å (16 % of water molecules shown). (Bottom) SDF of imidazole around water is added (mesh blue surface) for distance up to 6 Å (16 % of imidazole molecules shown).

Table 5.6. Coordination numbers of hydrogen-bonded imidazole pairs

Central atom-surrounding atom	Sample	Coordination number	Cut-off / Å
N1 – N3	1 M	0.03	3.5
	5 M	0.16	
	10.7 M	0.42	
N3 – H1	1 M	0.03	2.9
	5 M	0.15	
	10.7 M	0.43	
Imidazole (cog) – Imidazole (cog)	1 M	0.7	7
	5 M	4.1	
	10.7 M	8.6	

The second peak in N1…N3 RDF centred at ~ 5 Å (Figure 5.16 (a)) and the second peak in the cog RDF centred at 9 Å (Figure 5.16 (b)) are more defined for the 10.7 M solution than for the 5 M solution, which could indicate directional imidazole interactions or clustering of imidazole molecules. This is further illustrated by the increase in the number of imidazole molecules around a central molecule from ~4 molecules in the 5 M solution to ~8 molecules in the 10.7 M solution at a cut-off radius of 7 Å (Table 5.6). Moreover, the spatial density functions (SDFs) show that the probability of finding imidazole molecules coordinating a central molecule increase with the imidazole concentration (Figure 5.17). However, the predicted imidazole clustering is always

mediated by water molecules as illustrated in Figure 5.17. This indicates that the variations in the local environment around the imidazole molecules are trivial, which explains the similarities in the C and N K-edge spectra of aqueous imidazole solutions at 1 M, 5 M and 10.6 M presented in section 5.3.2.

Finally, the solvation of imidazole through the formation of hydrated imidazole clusters was predicted by the study presented in chapter 3 using neutron diffraction with isotopic substitution combined with total X-ray scattering and the EPSR simulation. This means that total X-ray scattering combined with the EPSR simulation can yield similar quality results to the more sophisticated ones produced by the study in chapter 3. The revelation that total X-ray scattering provides complementary information to neutron diffraction is important especially that experiments utilising neutrons are more complex and are currently not applicable to time-resolved studies, which are the theme of the next chapter of this thesis.

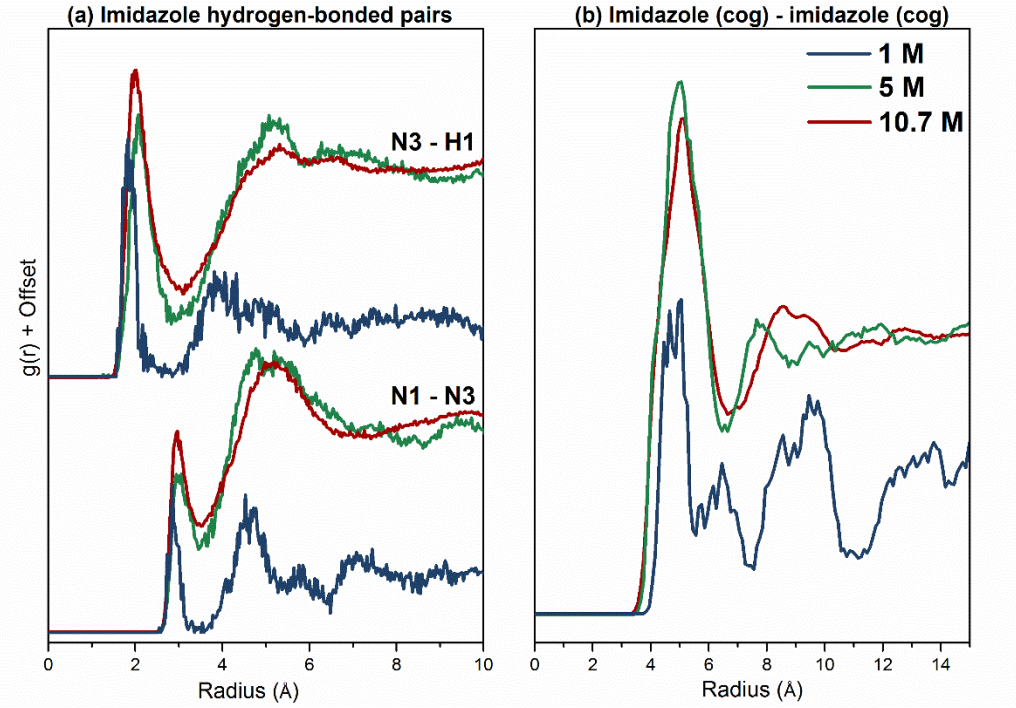

Figure 5.16. (a) The radial distribution functions (RDFs) of hydrogen-bonded pairs in imidazole (b) The radial distribution function (RDF) of the centre of geometry (cog) of imidazole around the cog of imidazole.

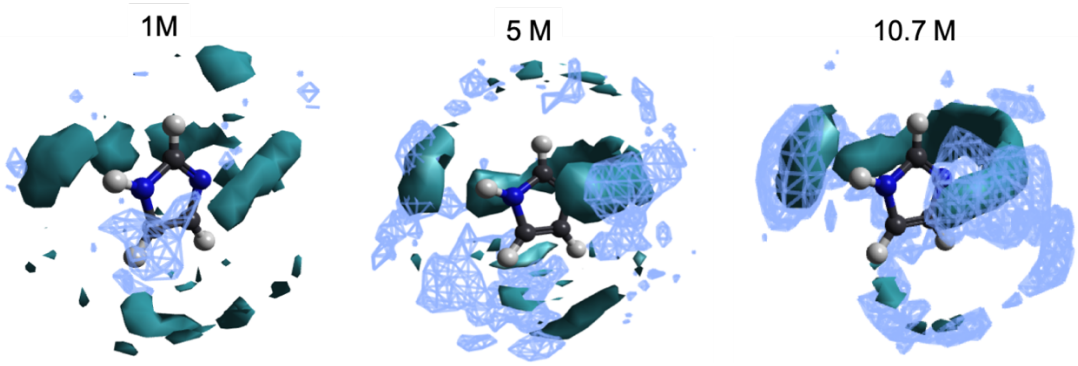

Figure 5.17. Spatial density function (SDF) of water (solid teal surfaces) and imidazole (blue mesh) around imidazole for distances up to 6 Å (10 % of water molecules shown) and 7 Å (10 % of imidazole molecules shown).

5.4 Conclusions

In this chapter, FTIR spectroscopy, X-ray Raman scattering (XRS) and total X-ray scattering were utilised to study the effect of increasing the solute concentration on the solvation structure of aqueous imidazole solutions. The features in C and N K-edge XRS spectra of aqueous imidazole solutions were broadly similar across a range of concentrations, which indicates that the local solvation structure does not change much in response to the concentration changes of this magnitude. The similarities in the local bonding environment around the imidazole molecules have been confirmed by the radial and spatial distribution functions (RDFs and SDFs) of water around imidazole.

The presence of imidazole molecules (even at high concentrations) did not disturb the bonding coordination around water as indicated by the trivial variations in the O K-edge XRS spectra of aqueous imidazole solutions at different concentrations. This is attributed to the role of imidazole in preserving the hydrogen-bond network of water as illustrated by the SDFs of imidazole around water. Imidazole does not only maintain the water network, but it also increases the rigidity of the hydrogen-bond network of water as indicated by the shifts in the frequencies of the bend and libration and the O-H stretch bands of water.

Chapter 6 Structural Evolution of Aqueous Imidazole Solutions during Cooling Crystallisation

6.1 Introduction

Nucleation of crystals from solutions must be associated with the formation of solute-solute bonds at the expense of solute-solvent interactions. Classical nucleation theory (CNT) describes this process by postulating a prenucleation equilibrium between solvated solute molecules and small clusters of solute molecules that are held together by solute-solute interactions. CNT further assumes that the balance between the cohesive energy gain in the nascent crystal volume and destabilisation by the formation of the crystal/solution interface determines a critical cluster size, above which further crystal growth becomes the thermodynamically spontaneous process. Clusters with subcritical size form and re-dissolve again in the prenucleation equilibrium. One would therefore expect that at high supersaturation an unknown mole fraction of the solute molecules exists in prenucleation clusters [78] in which solute-solvent bonds are replaced by solute-solute interactions. This process is possibly mediated by the solvent molecules in the solvation shells surrounding the solute molecules.

The ideal characterisation tool to probe the structural evolution of solutions during crystallisation should (1) provide quantitative and temporal information about the bulk composition and how the structure evolves prior to crystallisation, (2) be operable through a range of temperatures, (3) measure the sample without inducing changes on the state or composition. Near-edge X-ray absorption fine structure (NEXAFS) spectroscopy probes the local structure via excitation of 1s core-level electrons to unoccupied states in the valence region of molecules. The involvement of valence levels makes NEXAFS highly sensitive to the chemical and physical changes around the absorbing atom. The technique has recently been shown to be able to detect the formation of prenucleation equilibrium during the reactive interfacial nucleation of metal nanoparticles [263-265]. However, NEXAFS of organic molecules is a soft X-ray synchrotron radiation technique, soft X-rays cover energy range ~100 eV – 2000 keV, which requires spectrometers housed in ultra-high vacuum (UHV) chambers. Hence, NEXAFS measurements on liquids would involve sample handling in a vacuum chamber. This imposes

constraints on experimental control that, while probably addressable through novel instrument designs, have so far prevented studies of solution structure under controlled supersaturation conditions.

X-ray Raman scattering (XRS) is a non-resonant X-ray scattering technique that utilises hard X-rays; hence it is effectively an alternative technique for measuring X-ray absorption spectroscopy (XAS) that avoids the experimental challenges associated with the use of soft X-rays. The deep penetration depth (~2 mm at 10 keV) of hard X-rays makes XRS a bulk-sensitive technique, with the added benefit of significantly inhibited self-absorption distortions of the spectra [113]. The use of hard X-rays also allows the design of ambient-pressure environmental cells with precise temperature control. On its own, X-ray Raman scattering, and analogous NEXAFS spectroscopy, provides limited information on the extended structural properties of solutions. Therefore, for crystallisation experiments combining X-ray Raman scattering with total X-ray scattering can provide comprehensive information on the structural evolution of the solution prior to crystallisation in the short and long radial ranges. Total scattering combines Bragg and diffuse scattering which enable probing amorphous structures as well as short-range and long-range ordered materials [189, 190]. The interatomic distances in the sample can be obtained from the pair distribution functions (PDFs) which are generated through Fourier transformation of the total X-ray scattering data.

This chapter presents the results of *in-situ* XRS and *in-situ* XPDF experiments, which were used to characterise the speciation of imidazole molecules in aqueous solutions during cooling crystallisation. These techniques were used to follow the structural evolution of the named solution from the undersaturated state through the metastable zone with high supersaturation until the onset of crystallisation. The study presented in this chapter builds on the studies in chapters 3 and 5, which illustrated that the solvation structure of concentrated aqueous solutions of imidazole was dominated by solute-solvent interactions and that this solvation structure is not perturbed by increasing the concentration of imidazole.

6.2 Experimental

6.2.1 X-ray Raman Scattering (XRS)

XRS measurements were performed at sector 20-ID of the Advanced Photon Source (APS), Chicago, USA, using the low-energy resolution inelastic X-ray scattering (LERIX) spectrometer [184, 185]. 20-ID is an undulator beamline equipped with liquid nitrogen cooled Si(111) and Si(311) monochromators. The Si(311) monochromator was used as it achieves high spectral resolution which is needed for this study. The combined instrumental resolution of the monochromator and the analyser crystal was ~0.5 eV. The monochromatised beam was focused using a toroidal mirror to a spot size of 600 µm (horizontal) and 400 µm (vertical).

The scattered radiation was analysed with a combination of Si diced and undiced spherically bent crystal analysers (SBCA). The SBCAs were tuned to use the Si(555) reflection, and the scattered radiation was detected using NaI scintillators coupled to custom-built amplifier and discriminator electronics. All data were collected using the inverse scanning mode, where the incident photon energy was scanned and the analyser energy was fixed at 9891 eV. Data were collected from 18 of the 19 crystal spectrometers. The data presented in this chapter were taken from *q6* diced crystal analyser which is one of the low momentum transfer (q) analysers with the best signal to noise ratio. This way, the data were collected at a scattering angle of 54° with q value of ~4.6 $Å^{-1}$. The monochromator drift was small (±0.137 eV) and it was monitored through frequent measurement of the elastic scattering peak. The monochromator was detuned by 15% to suppress the contribution from the harmonic component of the beam.

The C and N K-edge spectra of a 10.1 M aqueous solution of imidazole were measured at 35, 20 and 13°C utilising Flow-cell 1 (section 2.4 – chapter 2). The solution was also probed after precipitation of crystals had taken place. Deionised water was used for the preparation of the 10.1 M imidazole solution. The solution had a pH value of 11 at room temperature. The typical measurement times at each temperature were 2 – 3 hours. Over this period, several spectra were collected and checked for consistency before averaging. It is worth noting that no evidence

of radiation damage was picked up in the collected spectra (individual spectra are provided in Appendix I). The LERIX sample chamber was flushed with He gas to reduce background scattering from air.

Crystalline imidazole powder was also measured as a reference. The powder sample was pressed into a 13 mm diameter pellet with a thickness of ~2 mm. The pellet was mounted on a spinning platform and measured with an off-centred beam to minimise radiation damage

For data analysis, the elastic energy was subtracted from the XRS data to obtain near-edge spectra which were then subsequently background subtracted and normalised to the edge-step height using Athena software [239].

6.2.2 Total X-ray Scattering and X-ray Pair Distribution Function (XPDF)

Total X-ray scattering data of a 10.7 M aqueous solution of imidazole was collected at the X-ray pair distribution function (XPDF) beamline I15-1 of Dimond Light Source (DLS), UK [192] during cooling crystallisation. The solution sample was measured using Flow-cell 2 (section 2.4 – chapter 2). The flow-cell ensured continuous circulation of the solution, which minimised beam induced damage on the sample and ensured uniform temperature distribution. The total X-ray scattering experiment was repeated three times with acquisition times varying between 10 – 30 seconds per scan. During the third measurement cycle, total X-ray scattering data was collected during crystallisation and dissolution to capture any structural changes as crystal form and as they dissolve back into the solution. Crystalline imidazole was also measured using a borosilicate glass capillary. The acquisition time for the solid sample was 10 minutes.

Raw X-ray scattering data was automatically corrected for the internal dark current during data collection using the DLS generic data acquisition (GDA) software. Further corrections for the instrument and the sample holder backgrounds, absorption, multiple scattering and Compton scattering were performed in GudrunX to generate the interference differential cross sections F(Q)s [224]. GudrunX was also used to generate the pair distribution functions (PDFs) by Fourier transformation of the F(Q)s [224, 248].

6.3 Results and Discussion

6.3.1 X-ray Raman Scattering (XRS)

The strong resonances in the near-edge region of the C and N K-edge XRS spectra correspond to transitions from the 1s core-level of the absorbing atom to the unoccupied valence orbitals with π* character. These are followed by the ionisation potentials (IP) and broad σ* shape resonances corresponding to multiple scattering at the potential barrier to the bound neighbouring atoms. The spectral features in the C and N K-edge XRS spectra were deconvoluted with non-linear least-squares fitting using Fityk software [244] (Figure 6.1) similar to the fitting procedure followed in chapters 4 and 5. Gaussian functions were used to model the π* and σ* shape resonances and arctan functions were used to model the ionisation potentials (IPs) [168]. The full-width-at-half-maximum (FWHM) of the Gaussian peaks modelling the π* resonances vary by ~2 eV for the C K-edge and ~0.3 eV for the N K-edge. Three and two arctan functions were used to model the C K-edge and N K-edge IPs respectively because imidazole has three different carbon moieties and two different nitrogen moieties. Since the ratio of the different carbon moieties in the imidazole molecule is equal, the width and height of the arctan functions in the C K-edge spectra were constrained to be equal; a similar procedure was followed for the N K-edge. The energy difference between the arctan functions (ΔIP) in the N K-edge spectra was constrained to be equal to the energy difference between the two 1π* peaks (ΔE (1π*)) because it has been illustrated that the transitions to π* states are dominated by the core-level binding energy shifts [88]. Core-level binding energies determined from C1s X-ray photoelectron spectra [91] were used to constrain the relative positions and centroids of the arctan functions in the C K-edge spectra (ΔIP (C2-C5) and ΔIP (C5-C4)). The resulting centroids of the fitted Gaussian and arctan curves in the C and N K-edge XRS spectra are listed in Tables 6.1 and 6.2.

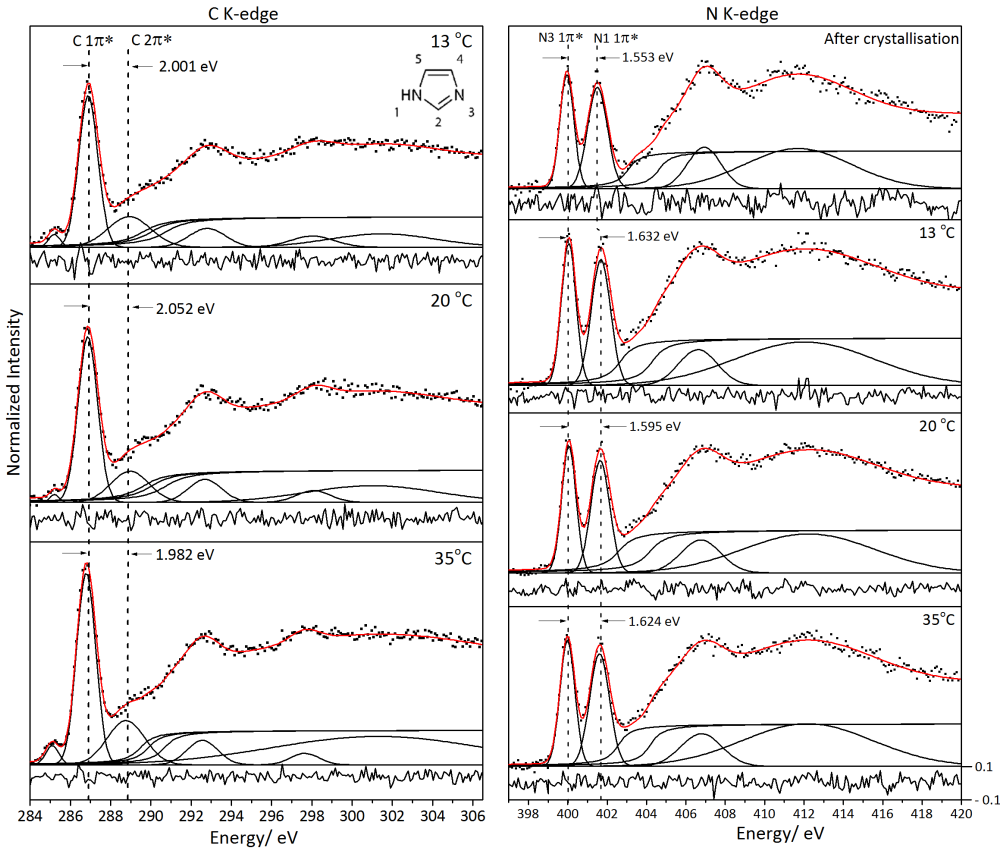

Figure 6.1. Fitted C K-edge and N K-edge XRS spectra of the 10.1 M aqueous solution of imidazole during cooling crystallisation. The residual is shown at the bottom of each spectrum. Peak positions and assignment for the C and N K-edge spectra are presented in Tables 6.1 and 6.2 respectively. Note that the peak at ~ 285.1 eV in the solution C K-edge is from the Kapton window.

Table 6.1. Peak position and assignment of the experimental C K-edge XRS spectra of the 10.1 M aqueous imidazole solution

	Solution		
	Exp @ 35 °C /eV	**Exp @ 20 °C /eV**	**Exp @ 13 °C /eV**
E (From Kapton)	285.106	285.185	285.170
E (C 1s–1π*)	286.787	286.808	286.826
E (C 1s–2π*)	288.769	288.860	288.827
ΔE (C2 2π* -C 1π*)	1.982	2.052	2.001
IP (C2)	290.697	290.697	290.697
IP (C5)	289.897	289.897	289.897
IP (C4)	289.497	289.497	289.497
ΔIP (C2-C5)	0.800	0.800	0.800
ΔIP (C5-C4)	0.400	0.400	0.400
σ (C-H)	292.566	292.497	292.612
σ (C-C)	297.655	297.768	297.729
σ (C-N)	301.302	300.680	301.191

Table 6.2. Peak position and assignment of the experimental N K-edge XRS spectra of the 10.1 M aqueous imidazole solution

	Solution			
	Exp @ 35 °C /eV	**Exp @ 20 °C /eV**	**Exp @ 13 °C /eV**	**Exp after crystallisation /eV**
E (N1 1s - 1π*)	401.610	401.652	401.680	401.509
E (N3 1s - 1π*)	399.986	400.057	400.049	399.956
ΔE (1π*)	1.624	1.595	1.632	1.553
E (N3 1s - 2π*)	-	-	-	-
ΔE (N3 2π*-1π*)	-	-	-	-
IP (N1)	404.192	404.186	404.293	404.625
IP (N3)	402.568	402.591	402.661	403.072
ΔIP	1.624	1.595	1.632	1.553
1σ* (C-N)	406.787	406.759	406.616	406.933
2σ* (C-N)	412.181	412.239	412.029	411.694

To capture the structural evolution of a supersaturated aqueous solution of imidazole, C and N K-edge XRS spectra were collected during cooling crystallisation (Figure 6.1). The spectra for the two edges were collected outside the metastable zone of a 10.1 M aqueous solution of imidazole (at 35°C) and then in the metastable zone (at 20°C and 13°C). A subsequent N K-edge spectrum was significantly different from the spectra in the metastable zone, prompting visual inspection of the crystallisation vessel which indicated the presence of crystals. The photon energy difference between the two peaks in the C K-edge is ~2.002 eV (± 0.051) for the solution at different temperatures (35°C, 20°C and 13°C). Similarly, the photon energy difference between the two 1π* peaks in the N K-edge spectra is ~ 1.624 eV (± 0.019) for the solution at different temperatures and ~ 1.553 eV after crystal formation (Figure 6.1). The variation in the spectral features of the C and N K-edge XRS spectra during cooling is minor and a change is observed in the N K-edge spectra once crystals have formed.

The minimal change observed in the C and N K-edge spectra during cooling crystallisation indicates that the average local coordination around the imidazole molecules does not change significantly during the cooling crystallisation process. The outcome of the *in-situ* measurements in the metastable zone is in agreement with the observations in chapter 5 on the C and N K-edge spectra of aqueous solutions of imidazole, where minimal variation was detected as a function of

concentration. They are also in line with the structural models generated by the EPSR (chapter 5) that suggested the solvation of the imidazole nitrogen moieties remained almost identical for different concentrations. This suggests that the concentration-dependent structure of aqueous imidazole can be described in terms of interactions between hydrated imidazole molecules rather than direct imidazole-imidazole interactions. By implication, desolvation of hydrated imidazole monomers appears to be the key step in nucleation of imidazole crystals from aqueous solutions. A recent simulation of NaCl nucleation from supersaturated solutions similarly suggested that the desolvation of the ions determines the attachment rate of monomers at interface between nuclei and solution, rather than the diffusive flux of molecules as assumed by CNT [266].

Moreover, CNT assumes that the internal structure of critical nuclei is similar to that of the final crystals. It is interesting to note in this context, that the C and N K-edge spectra varied minimally during cooling crystallisation and were distinct from those of the crystalline phase (spectra shown in chapter 4). It therefore appears that any structuring due to aggregation of monomers in the metastable zone prior to nucleation, if it exists, does not reflect the structure of the final crystalline product. This is in line with previous suggestions that the pre-nuclei differ drastically in terms of structure and composition from the bulk phase [69-72]. Other “non-classical” nucleation mechanisms have been explored both theoretically and experimentally such as the two-step nucleation mechanism [73] and pre-nucleation cluster mechanism [74]. In two-step nucleation, formation of metastable dense liquid-like clusters is predicted to precede the nucleation event. Numerical simulation of homogenous solutions [75] and DFT calculations [76] showed that such liquid-liquid separation may occur near the metastable critical point. However, crystallisation experiments did not provide evidence that this phase separation takes place [77]. This indicates that the predicted phase separation is due to a liquid-liquid miscibility gap inside the metastable zone and should not be confused with nucleation mechanisms [77, 78]. Pre-nucleation clusters are thermodynamically stable associates in the solution, with a structure that may be similar or dissimilar to the final crystallised product [74]. However, their presence has only been illustrated for inorganic systems [74].

It is worthwhile to examine how realistic it is that pre-nucleation clusters can be detected by experimental techniques probing the overall composition of the nucleating solution. Arguably, the concentration of solute molecules in prenucleation clusters should be small in comparison with the mole fraction in the remaining bulk saturated solution. An estimate can be made by considering that XRS was able to capture the changes in the N K-edge spectral features induced by nucleation. Based on the assumption that the solution after crystal formation consists of imidazole crystals in a saturated imidazole solution (noting that this might be a simplification requiring further consideration, see below). The balance between solid-state and solution state imidazole molecules is reflected by the photon energy difference between the two 1π* peaks in the N K-edge spectra. For the post-nucleation solution, an intermediate value between those for the crystalline phase and the undersaturated solution (at 35°C) was determined (Figure 6.2). The observed decrease to 1.553 eV (Table 6.2, Figure 6.2) is approximately 1/3 (~36%) of the decrease expected if all imidazole transitioned into the crystalline form. This suggests that the imidazole molecules in the precipitated crystals were approximately 1/3 of all the imidazole molecules previously present in solution, with the remainder staying as solute molecules in a saturated solution. If nucleation of imidazole from aqueous solutions would pass through prenucleation clusters stage it may involve up to 1/3 of the solute molecules. That no significant changes in the spectra were observed prior to the detection of crystals may therefore suggest that the fraction of solute molecules involved in cluster formation before nucleation is much lower than the fraction of molecules assembled in the crystallised product, or it may signify that imidazole-water binding remains the dominant structural motif upon cluster formation, with similar local solvation of imidazole in the undersaturated and in the supersaturated state. If prenucleation clusters are formed then imidazole-imidazole interactions do not appear to change the electronic structure of imidazole to the same extent as the hydrogen bonds to water molecules. Any self-association or cluster formation would thus involve interactions between hydrated imidazole molecules, which have not yet stripped off their inner solvation shell. In this case the N and C K-edges would not appear to be sensitive enough to detect evidence for cluster formation.

The alternative explanation for these observations would be that only a very small fraction of imidazole molecules are involved in self-association and cluster formation, or perhaps that the average lifetime of clusters is so short as to result in an undetectable low concentration over time. It should be noted that these observations were made with solutions that contained imidazole concentrations around 10.1 mol.L^{-1} prior to nucleation. The number of water molecules available per imidazole molecule in these solutions is significantly below four, which suggests that the solubility of imidazole is limited by the availability of solvent molecules. The fact that we do not observe any difference to undersaturated solutions prior to nucleation is remarkable, but also underlines the importance of the local hydrogen-bonding in this system.

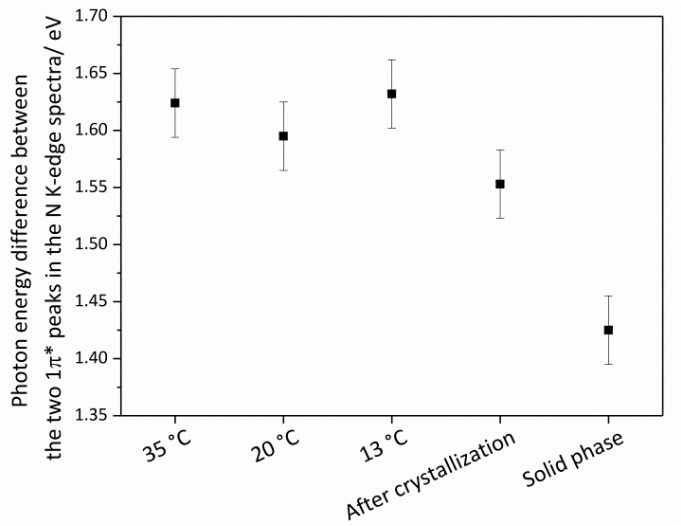

Figure 6.2. The photon energy difference between the two 1π* peaks in the N K-edge spectra for solution during cooling crystallisation (35 °C, 20 °C, 13 °C), after crystallisation and for solid-phase.

However, returning to the point made above that assuming imidazole crystals in saturated imidazole solution may be an unrealistic simplification, the precipitated crystals that were detected after crystallisation event may have a different structure and composition, for example representing unknown imidazole hydrates. This could also explain why the energy difference between the two 1π* peaks in the N K-edge spectra is at an intermediate value between that of the solid-state and solution state. In this case the fraction of precipitated solid might be higher than the estimate given above, because hydrogen bonds to water would be retained, and a lesser effect on the N K-edge would be expected than for non-hydrated imidazole. Unfortunately, examining

other fitted peak parameters such as FWHMs does not allow us to shed further light on this issue. For example, from the N K-edge spectra of the solid and the solution in Figure 6.1 one may expect slight broadening of the peak arising from the N1 nitrogen, due to an energy difference of 0.2 eV between the solid and the solution. However, the expected broadening effect would be within less than 0.1 eV, which we cannot reliably resolve with the quality of the available spectra. Hence, complementary work with other techniques sensitive to short-range structure, and with much faster data acquisition times (e.g. X-ray pair distribution function, XPDF, measurements) was performed and is presented in the following section.

6.3.2 X-ray Pair Distribution Function (XPDF)

Total X-ray scattering data was collected for a 10.7 M aqueous imidazole solution during cooling until the onset of crystal formation to probe the structural evolution of the solution in a time-resolved manner. The experiment was repeated three times with acquisition times varying between 10 – 30 seconds per scan. For the third cooling crystallisation experiment, total X-ray scattering data was collected during crystallisation and dissolution to capture any structural changes as crystals form and as they dissolve back into the solution. The X-ray pair distribution function (XPDF) patterns of the named solution was generated through Fourier transformation of the total X-ray scattering patterns. These are presented in Figures (6.3, 6.7 and 6.11).

6.3.2.1 First Cooling Crystallisation Experiment

Total X-ray scattering data was collected for a 10.7 M aqueous imidazole solution as it was cooled from 15.5 °C to 10.5 °C. The acquisition time per scan for this cooling crystallisation experiment (Figure 6.3) was 30 seconds. The XPDFs patterns (Figures 6.3 (a) and (b)) were generated by Fourier transformation of the total X-ray scattering data.

Upon crystallisation, the XPDF patterns developed features in the longer radial range (10 Å – 50 Å) (scans 7–27, blue XPDF patterns in Figure 6.3(b)). This observation coincided with an increase in temperature of the solution (Figure 6.3 (c)), which is expected since crystallisation is an exothermic process. The developed XPDF features in the 10 Å – 50 Å radial range correspond to peaks in the XPDF pattern of solid imidazole as indicated by the high similarity index (~0.8) shown in Figure 6.4 (d). This indicates that

the structure of the crystals precipitating from the solution is similar to that of solid imidazole.

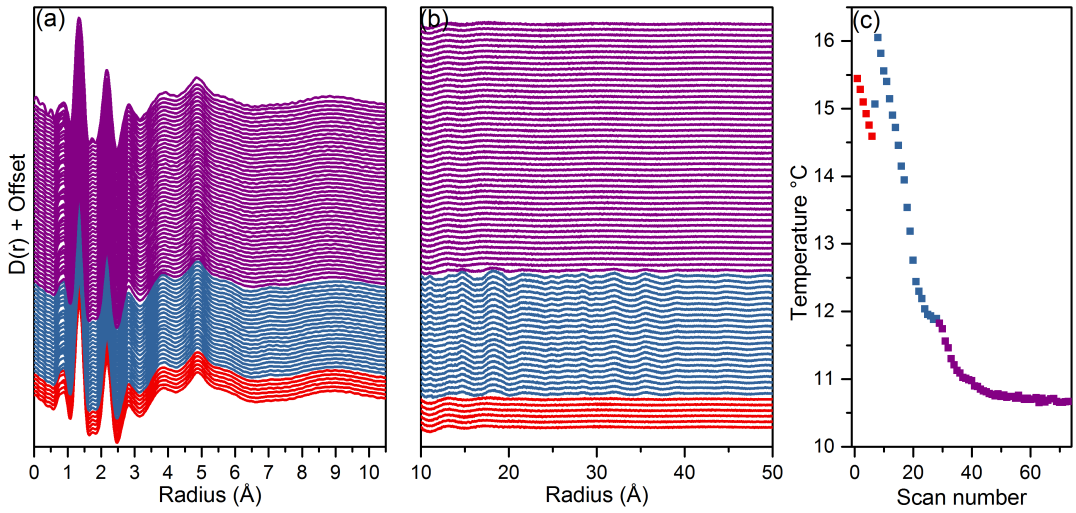

Figure 6.3. XPDF patterns of the 10.7 M aqueous imidazole solution during the first cooling crystallisation experiment in the range: (a) 0 – 10 Å and (b) 10 – 50 Å. (c) The temperature of the solution during cooling crystallisation.

The contour maps of the XPDF patterns of the solution shown in (Figure 6.5 (c) and (d)) illustrates the resemblance between the formed crystals and solid imidazole further. The pink dotted lines highlight regions where the intensity of features in the XPDF patterns of the solution after crystallisation (scans 7 – 27) increase. The regions are centred at 3.9 Å, 4.9 Å, 7.1 Å, 11 Å, 14.7 Å and 18.2 Å and they correspond to peaks in the XPDF pattern of solid imidazole Figure 6.5 (a) and (b). The red dotted lines in the contour maps (Figure 6.5 (c) and (d)) highlight regions of decreased intensity. These features are centred at 12.5 Å, 16.3 Å, 20.2 Å, 24.2 Å, 34 Å and 37.6 Å and they correspond to troughs in the XPDF pattern of the solid Figure 6.5 (a) and (b).

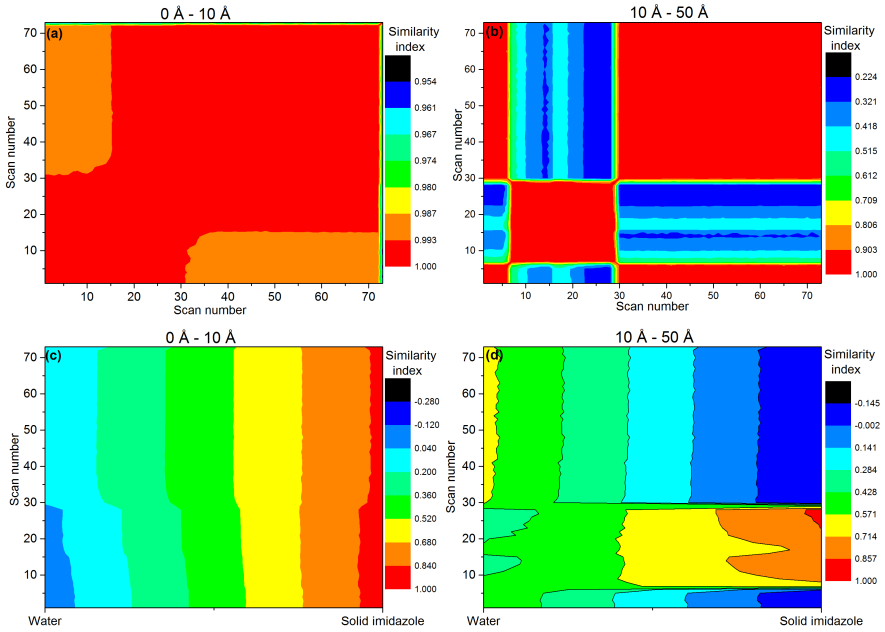

Figure 6.4. Correlation matrix of the XPDF patterns of the 10.7 M aqueous imidazole solution collected during the first cooling crystallisation experiment in the (a) 0 Å – 10 Å range and (b) 10 Å – 50 Å range. Correlation between the XPDF patterns of the aqueous imidazole solution with water and solid imidazole in the (c) 0 Å – 10 Å range and (d) 10 Å – 50 Å range.

It is worth noting here that the purple XPDF patterns in Figure 6.3 lack the defined XPDF features observed in the blue XPDF patterns, which represented the solution after crystallisation. These defined features were attributed to the formation of imidazole crystals and their absence from the purple XPDF patterns encouraged visual inspection of the flow-cell during the experiment. It was found that imidazole crystals were stuck to the walls of the crystalliser and not pumped to the Kapton window where the X-ray measurements took place. Hence, the purple XPDF patterns were actually a measurement of the aqueous imidazole solution without crystals. The similarity index between the purple XPDF patterns and the XPDF patterns of the solution before crystallisation (red XPDFs in Figure 6.3) is high (~1) in the 10 Å – 50 Å radial range (Figure 6.4 (b)), which indicates that both the red and purple XPDFs represented measurements of the solution. The similarity index is slightly lower than 1 in the 0 Å – 10 Å range (Figure 6.4 (b)), which is dominated by features attributed to imidazole-imidazole intra- and intermolecular interactions. This is expected as the number of imidazole molecules dissolved in the solution after crystallisation is reduced and hence the contribution of imidazole to the XPDF features in the purple patterns will be reduced.

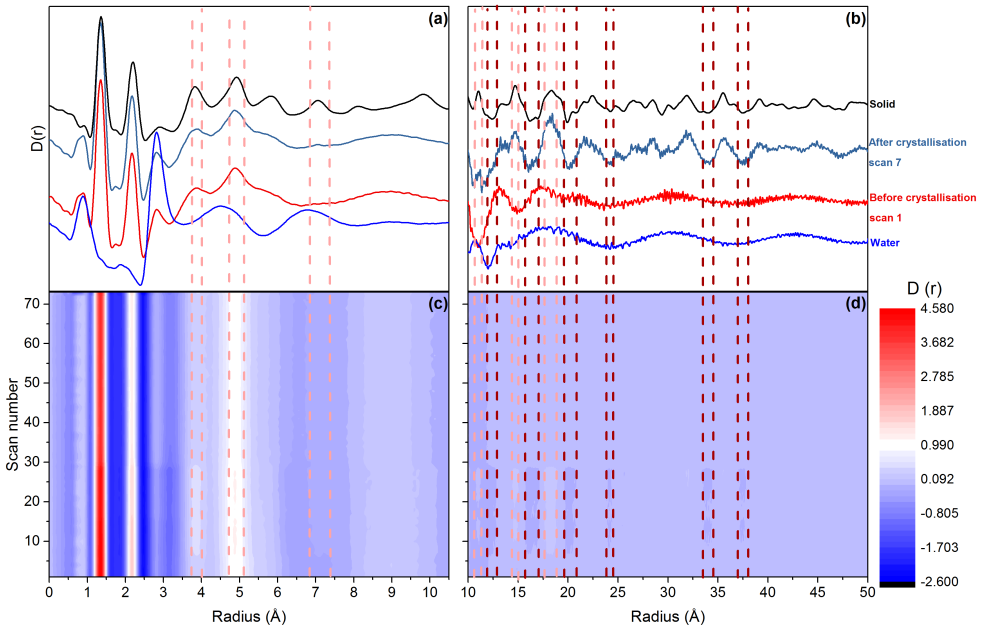

Figure 6.5. XPDF patterns of the 10.7 M aqueous imidazole solution before and after crystallisation, solid imidazole and water cycle in the range: (a) 0 – 10 Å and (b) 10 – 50 Å. Contour plot of the XPDF patterns of the 10.7 M aqueous imidazole solution collected during the first cooling crystallisation experiment in the range: (c) 0 – 10 Å and (d) 10 – 50 Å.

The correlation between the XPDF patterns of the solid and the purple XPDF patterns of the solution in the 0 Å – 10 Å radial range is lower than the correlation between the XPDF patterns of the solid and the XPDF patterns of the solution before crystallisation (red XPDFs) (Figure 6.4 (c)). This indicates the presence of less imidazole molecules in the solution representing the purple XPDF patterns than in the solution representing the red patterns (i.e. the solution before crystallisation). The observed reduction in the above-mentioned correlations suggests that ~¼ of the imidazole molecules crystallised during the cooling process; indicating that ¼ of the imidazole molecule would participate in pre-nucleation clusters (if they exist). The XPDF features for the solution prior to crystallisation do not exhibit variations during cooling and before crystal formation which indicate that (1) imidazole retains its solvation structure before crystallisations both in the short and long radial ranges or (2) the signal from the structuring of ¼ of imidazole molecules in the solutions prior to crystallisation is dispersed by the unstructured solution.

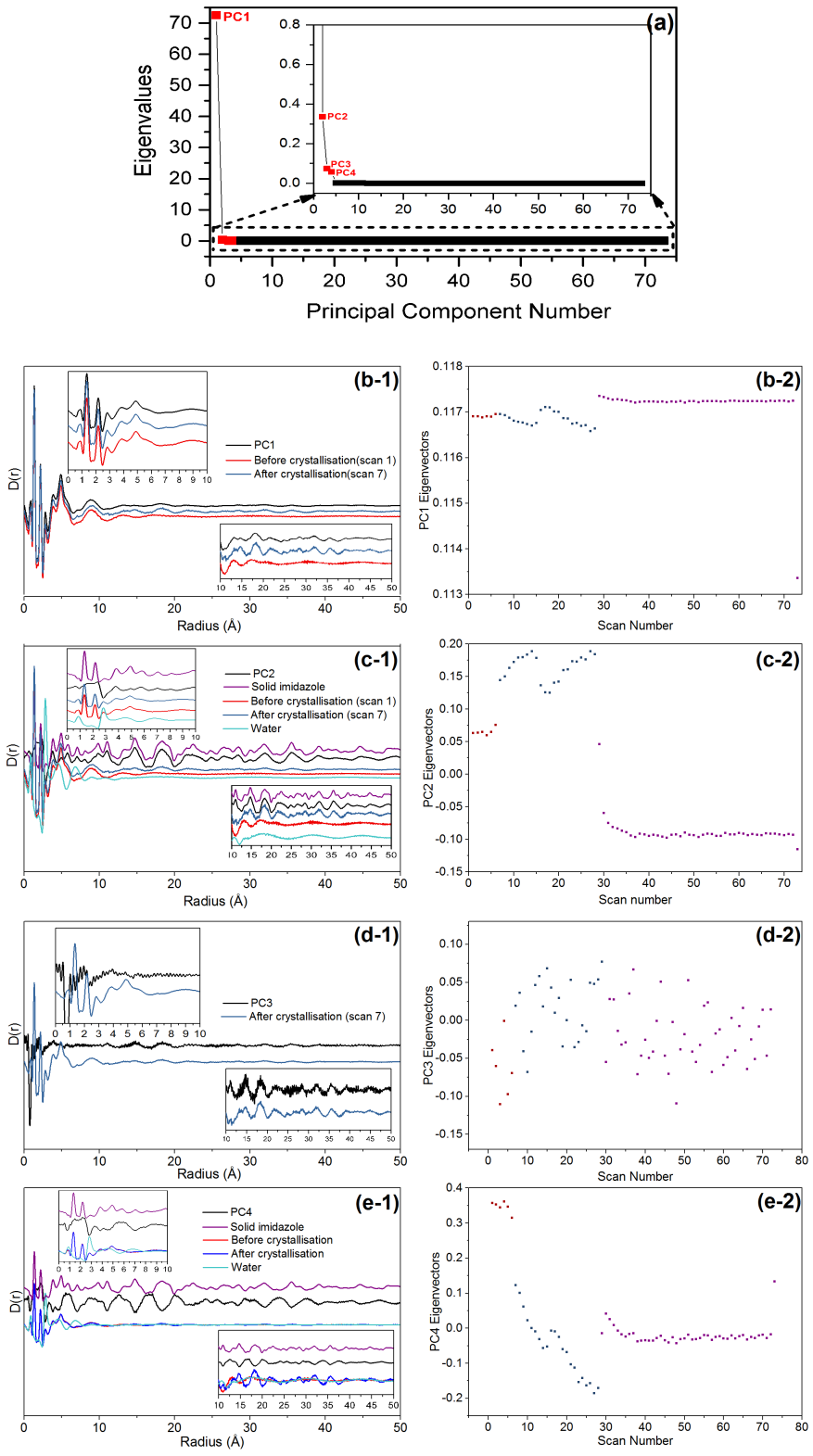

Figure 6.6. (a) Eigenvalues of the principal components from the PCA of the XPDF patterns of the 10.7 M aqueous imidazole solution from the first cooling crystallisation experiment. (b-1) (c-1) (d-1) (e-1) comparison between the patterns of PC1 – PC4 with the XPDF patterns of the solid imidazole, water and the aqueous imidazole solution before and after crystallisation. (b-2) (c-2) (d-2) (e-2) eigenvectors of PC1 – PC4.

Principal component analysis (PCA) was performed on the XPDF patterns of the 10.7 M aqueous imidazole solution from the first cooling crystallisation experiment (Figure 6.6). Four principal components were identified by the PCA analysis (PC1 – PC4) (Figure 6.6 (a)).

- PC1 has the highest eigenvalue and it represents the XPDF pattern of the solution after crystallisation (Figure 6.6 (b-1)). The contributions from the different XPDF patterns to this principal component are almost identical (Figure 6.6 (b-2)).

- PC2 represents the XPDF pattern of the solid in the 3.5 Å – 50 Å radial range (Figure 6.6 (c-1)). The contributions to PC2 from the XPDF patterns of the solution after crystallisation (Figure 6.6 (c-2), blue squares) are three times higher than the contributions from the XPDF patterns of the solution prior to crystal formation (Figure 6.6 (c-2), red squares). This indicates that the crystal structure of the formed crystals is similar to that of solid imidazole.

- PC3 and PC4 have low eigenvalues (Figure 6.6 (a)). While PC3 represents the XPDF pattern of the solution after crystallisation in the 10 – 50 Å radial range (Figure 6.6 (d-1)), PC4 represents features opposing those present in the solution after crystallisation (Figure 6.6 (e-1)).

6.3.2.2 Second Cooling Crystallisation Experiment

The onset of crystallisation was picked up by the XPDF patterns presented in the previous section. Minor variations were noticed in the XPDF patterns of the solution prior to crystallisations, which indicate that (1) the imidazole retains its hydration cluster until the onset of crystallisation or (2) the time resolution of the experiment was not high enough to capture transient structural variations before crystallisation took place. To verify this, total X-ray scattering data was collected for a 10.7 M aqueous imidazole solution during cooling from 15 °C to 11 °C with an acquisition time of 10 seconds; which lower than the acquisition time of the previous measurement (30 seconds). The XPDFs patterns (Figures 6.7 (a) and (b)) were generated by Fourier transformation of the total X-ray scattering data.

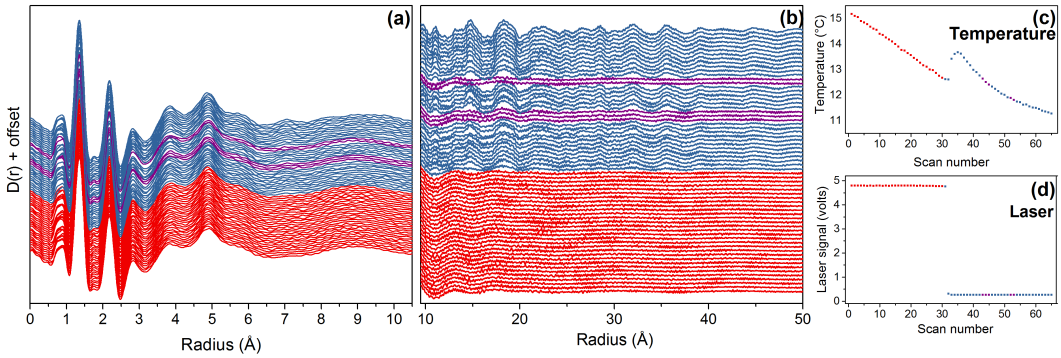

Figure 6.7. XPDF patterns of the 10.7 M aqueous imidazole solution from the second cooling crystallisation experiment in the range: (a) 0 – 10 Å and (b) 10 – 50 Å. (c) The temperature of the solution and (d) The laser signal during the cooling crystallisation experiment.

Similar to the observations from the first cooling crystallisation experiment (section 6.3.2.1), features in the longer radial range (10 Å – 50 Å) are observed as a result of crystal formation (blue patterns (Figure 6.7(b)). These features are similar for the consecutive XPDF patterns (scans 31 – 65) in Figure 6.7 (b)). They are also analogous to the features in the XPDF pattern of solid imidazole as indicated by the high similarity indices depicted in Figure 6.8 (b) and (d). The precipitated crystals increase in number and size as the cooling process continues and this is evident from the increase in the intensity and the definition of the XPDF features in the longer radial range (10 Å – 50 Å) (Figure 6.7 (b)). The five purple XPDF patterns in Figures (6.7) (a) and (b) represent those of the solution without crystals as they lack the defined features in the 10 Å – 50 Å range.

For the 0 Å – 10 Å range, the different XPDF patterns are almost identical as indicated by the high similarity index of ~1 shown in Figure 6.8 (a). This is attributed to the fact that most of the features in the named region are associated with intra-/inter-molecular bonds of imidazole. The correlation between the XPDF patterns of the solution and solid imidazole in the same region (0 Å – 10 Å) are significantly higher than the correlation between the XPDF patterns of the solution and water (Figure 6.8 (c)).

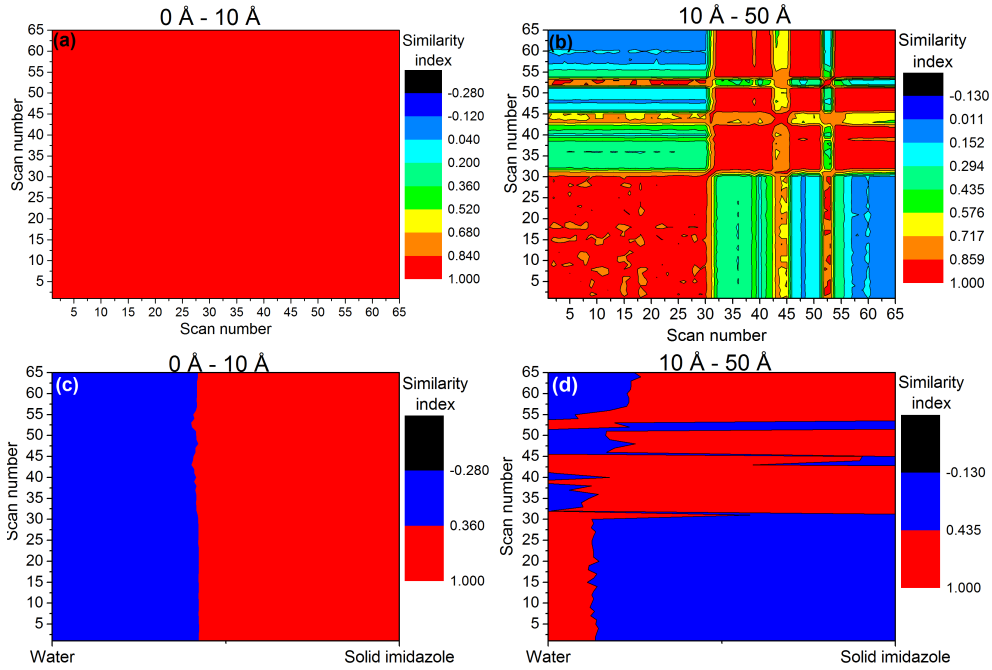

Figure 6.8. Correlation matrix of the XPDF patterns of the 10.7 M aqueous imidazole solution from the second cooling crystallisation experiment in the (a) 0 Å – 10 Å range and (b) 10 Å – 50 Å range. Correlation between the XPDF patterns of the aqueous imidazole solution with water and solid imidazole in the (c) 0 Å – 10 Å range and (d) 10 Å – 50 Å range.

The development of new features in the XPDF patterns of the aqueous imidazole solution after crystal formation in the second cooling experiment is clearly illustrated in the contour maps shown in Figure 6.9 (c) and (d). The pink-dotted highlight regions where the intensity of the features in the XPDF patterns of the solution after crystallisation increase. These features are centred 3.9 Å, 7.1 Å, 11 Å, 14.7 Å, 18.2 Å, 22.1 Å, 26.7 Å, 35.6 Å, 39.2 Å and 44.8 Å and they correspond to peaks in the XPDF pattern of the solid (Figure 6.9 (a) and (b)). The red dotted lines highlight regions where the intensity of the XPDF patterns decrease. These features are centred at 16.8 Å, 20.3 Å and 30.1 Å and they coincide with troughs in the XPDF pattern of the solid (Figure 6.9 (a) and (b)).

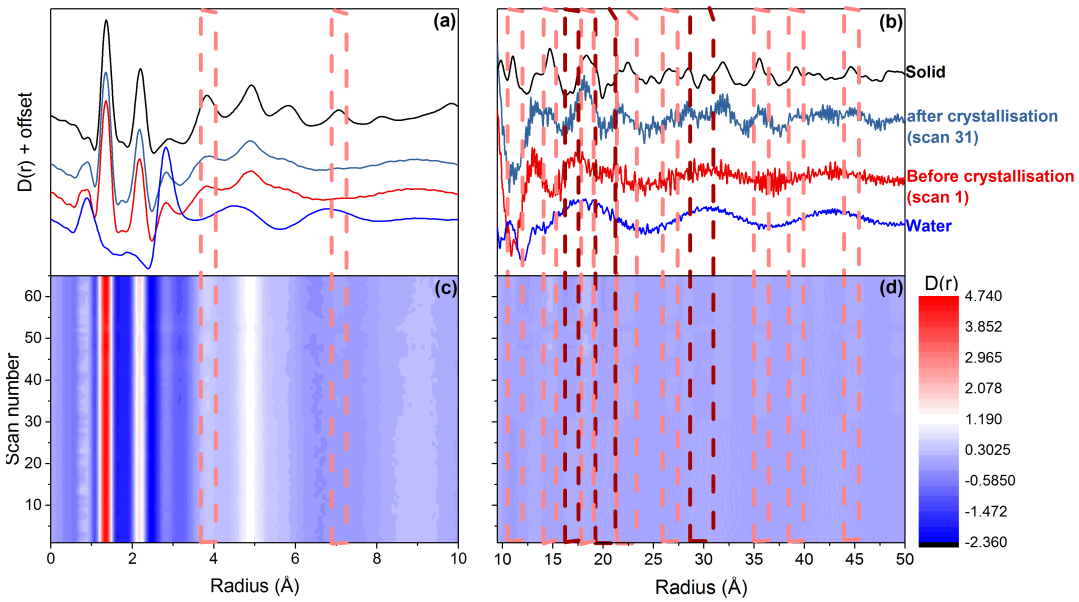

Figure 6.9. XPDF patterns of the 10.7 M aqueous imidazole solution before and after crystallisation from the second cooling cycle, solid imidazole and water cycle in the range: (a) 0 – 10 Å and (b) 10 – 50 Å. Contour plot of the XPDF patterns of the 10.7 M aqueou imidazole solution collected during the second cooling crystallisation experiment in the range: (c) 0 – 10 Å and (d) 10 – 50 Å.

It is worth noting here that the crystallisation event did not coincide with an increase in temperature (Figure 6.7 (c)) similar to the first cooling crystallisation experiment. Rather, two total X-ray scattering scans were collected before the temperature increased. Adding to that, the laser signal did not immediately drop to 0 as a result of crystal formation as seen from (Figure 6.7 (d)). These observations indicate that the number of imidazole crystal that initially precipitated is not enough to block the laser signal or to significantly increase the temperature of the solution. This is further illustrated by the first XPDF pattern collected after crystallisation where some features have not even developed (such as the peak centred at 14.8 Å) and where the intensity of the other features in the 10 Å – 50 Å range is low compared to the consecutive XPDF patterns (Figure 6.7 (b)).

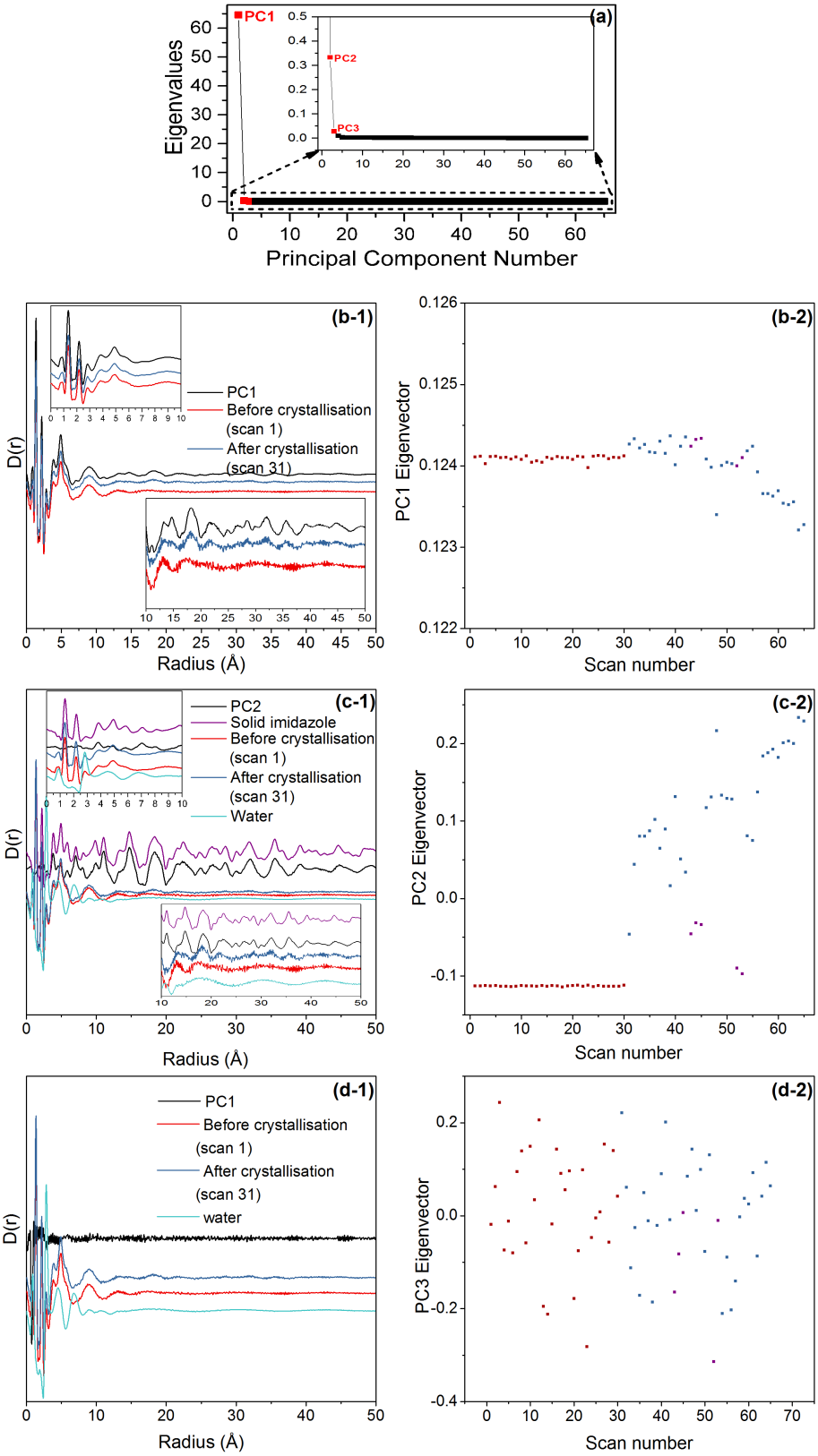

Figure 6.10. (a) Eigenvalues of the principal components from the PCA of the XPDF patterns of the 10.7 M aqueous imidazole solution from the second cooling crystallisation experiment. (b-1) (c-1) (d-1) comparison between the patterns of PC1 – PC3 and the XPDF patterns of the solid imidazole, water and the aqueous imidazole solution before and after crystallisation. (b-2) (c-2) (d-2) eigenvectors of PC1 – PC3.

Principal component analysis (PCA) was also performed for the XPDF patterns collected during the second cooling crystallisation experiment. Although, three principal components were identified by the analysis (PC1 – PC3) (Figure 6.10 (a)), the XPDF pattern of the third principal component shows no features (Figure 6.10(d-1)) and hence is not discussed further.

- PC1 has the highest eigenvalue and it represents the XPDF pattern of the solution after crystallisation (Figure 6.10 (b-1)). The contributions from the different XPDF patterns to this principal component are almost equal (Figure 6.10 (b-2)).

- PC2 represents the XPDF pattern of the solid in the 10 Å – 50 Å radial range with some features corresponding to the solid in the 0 Å – 10 Å range (Figure 6.10 (c-1)). The contributions to PC2 from the XPDF patterns of the solution after crystallisation (scans 31 – 65) increase from 0.05 to 0.2. On the other hand, the contributions from the XPDF patterns of the solution prior to crystal formation (scans 1 – 30) and of scans (43,44,45,52,53) (the purple XPDF patterns in Figure 6.7) are less than 0 (Figure 6.10 (c-2)). This indicates that the crystal structure of the formed crystals highly resembles that of the crystalline imidazole and that the size of the crystals increases as the cooling experiments proceeds. It also indicates that the structure of the pre-nucleation cluster, if present, is not similar to the structure of crystalline solid.

6.3.2.3 Third Cooling Crystallisation Experiment

The variation in the XPDF patterns of aqueous imidazole solutions prior to crystallisation was minor for the second cooling crystallisation experiment where the acquisition time per scan was 10 seconds. This observation suggests that self-association of imidazole molecules in aqueous solutions occurs between micro-hydrated clusters of imidazole where the hydrogen-bonding with water molecules is retained until crystallisation takes place. The cooling crystallisation experiment was repeated for the third time. Total X-ray scattering data was collected for a 10.7 M aqueous imidazole solution in the temperature range 29 °C – 14.3 °C to ensure that the structural evolution of the solution is monitored outside the metastable zone. The solution was also probed during dissolution to monitor any structural changes that manifest as the crystals dissolve back into the solution. The acquisition time for this experiment was 30 seconds to increase

the signal/noise ratio and consequently generate better quality total X-ray scattering data. The XPDFs patterns (Figures 6.11) were generated by Fourier transformation of the total X-ray scattering data.

Similar to the previous two measurements, features in the longer radial range (10 Å – 50 Å) are observed upon the onset of crystal formation (blue patterns (scans 151 – 170) in Figure 6.11(c-2)). These features are similar for the XPDF patterns (151 – 170) and they are analogous to the features in the XPDF pattern of solid imidazole as indicated by the high correlations depicted in Figure 6.12 (b) and (d).

Similar to the second cooling crystallisation experiment, the crystallisation event did not coincide with (1) an increase in temperature (Figure 6.11 (f)) and (2) a loss in the laser signal (Figure 6.11 (g)). These observations indicate that the number of imidazole crystal that initially precipitate is not enough to block the laser signal or to significantly increase the temperature of the solution. This is further illustrated by the first XPDF pattern collected after crystallisation where some XPDF features are missing (such as the peak centred at 14.8 Å) and where the intensities of features in the 10 Å – 50 Å range are lower than consecutive XPDF patterns (Figure 6.11 (c-2)).

During dissolution, the temperature of the solution is raised however there is a delay of ~1 minute, that is time needed to collect 6 total scattering scans, before this increase is observed in the temperature profile (Figure 6.11 (f)). This is expected because dissolution is an endothermic process and the solution needs to absorb the energy (in this case heat) needed to dissolve the formed crystals. The increase in temperature coincides with the disappearance of the features in the 10 Å – 50 Å radial range in the XPDF patterns of the solution during dissolution (scans 177 – 184, Figure 6.11 (c-2)). Although XPDF features in the longer radial range disappeared, the laser signal did not recover. This is attributed to condensation on the crystalliser vessel which blocks the laser signal. The correlation between the XPDF patterns during dissolution and the XPDF pattern before crystallisation ranges between 0.6 – 1 (Figure 6.12 (b)). This highlights the progressive dissolution of the crystals and is further illustrated by the increase in the similarity between the XPDF patterns of the solution during dissolution and that of water (Figure 6.12 (d)). The XPDF patterns of the solution after complete dissolution of the crystals (scans 185 – 207, Figure 6.11) resembles those of the solution prior to

crystallisation and water as indicated by the high similarity indices depicted in Figures 6.12 (b) and (d).

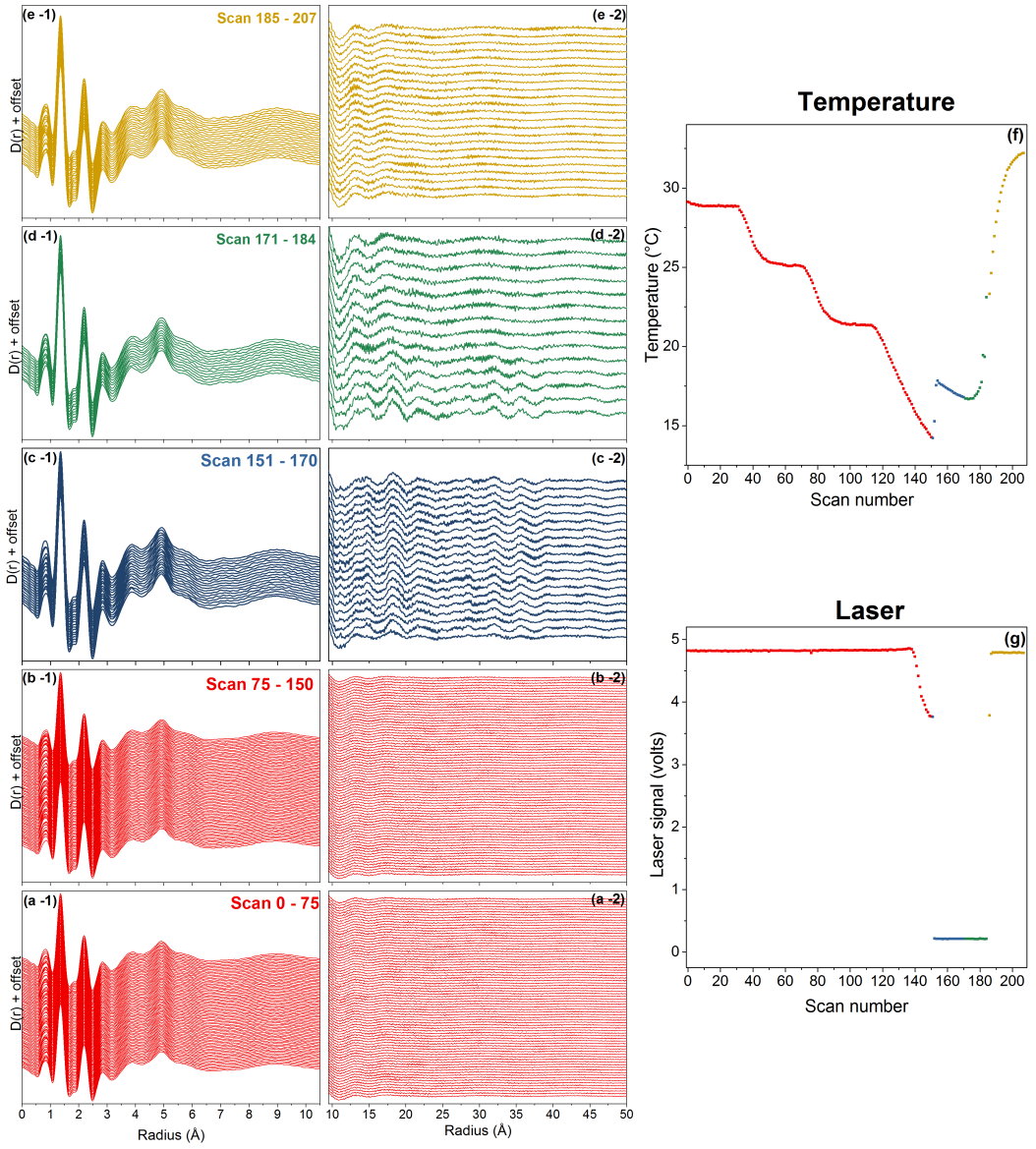

Figure 6.11. XPDF patterns of the 10.7 M aqueous imidazole solution from the third cooling crystallisation experiment (a) and (b) before crystallisation, (c) after crystallisation, (d) during dissolution, (e) after dissolution in the range: (-1) 0 – 10 Å and (-2) 10 – 50 Å. (f) The temperature of the solution and (g) The laser signal during the cooling crystallisation experiment.

The features in 0 Å – 10 Å radial range of the XPDF patterns of the solution before and after crystallisation and during and after dissolution are highly similar as indicated by the similarity index of ~1 (Figure 6.12 (a)). This is attributed to the fact that most of the features in the named region are associated with intra-/inter-molecular bonds of

imidazole. The correlation between the XPDF patterns of the solution and solid imidazole in the same region (0 Å – 10 Å) are significantly higher than the correlation between the XPDF patterns of the solution and water (Figure 6.12 (c)).

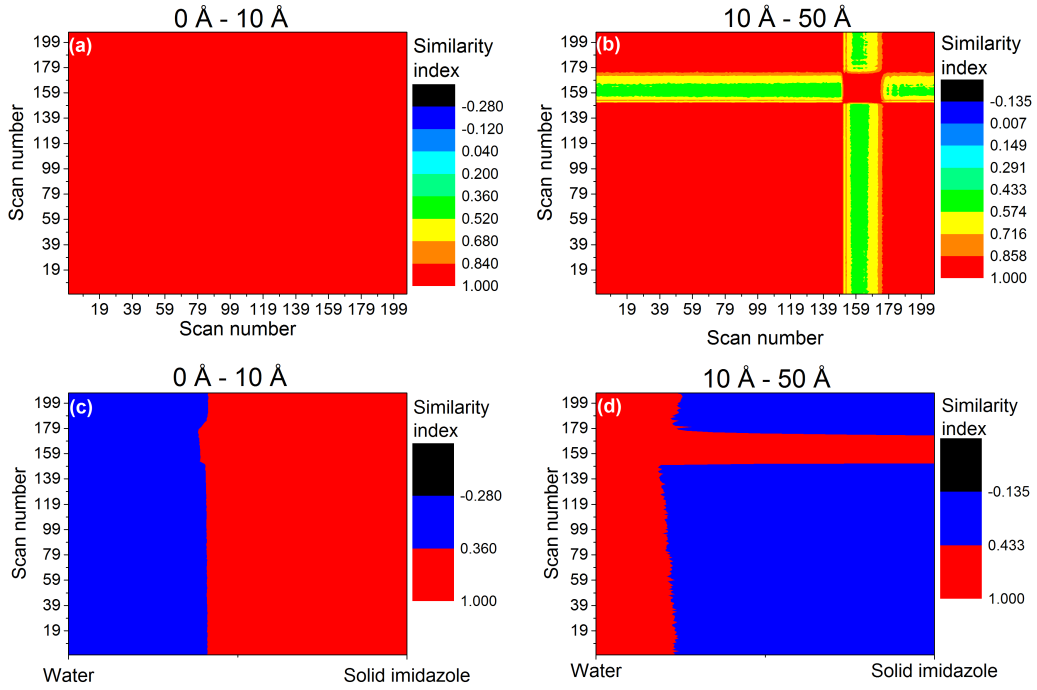

Figure 6.12. Correlation matrix of the XPDF patterns of the 10.7 M aqueous imidazole solution from the third cooling crystallisation experiment in the (a) 0 Å – 10 Å range and (b) 10 Å – 50 Å range. Correlation between the XPDF patterns of the aqueous imidazole solution and water and solid imidazole in the (c) 0 Å – 10 Å range and (d) 10 Å – 50 Å range.

The development of new features in the XPDF patterns of the 10.7 M aqueous imidazole solution after crystal formation in the third cooling cycle is clearly illustrated in the contour maps shown in Figure 6.13. The pink dotted lines highlight regions where the intensity of the XPDF features increase. These features are centred at 6.5 Å, 11 Å, 18.2 Å, and 31.8 Å and they correspond to peaks in the XPDF pattern of the solid. The red dotted lines highlight the region centred at 9 Å, which resembles a trough in the XPDF pattern of the solid. This indicates that the formed crystals have similar structural properties to those of crystalline imidazole.

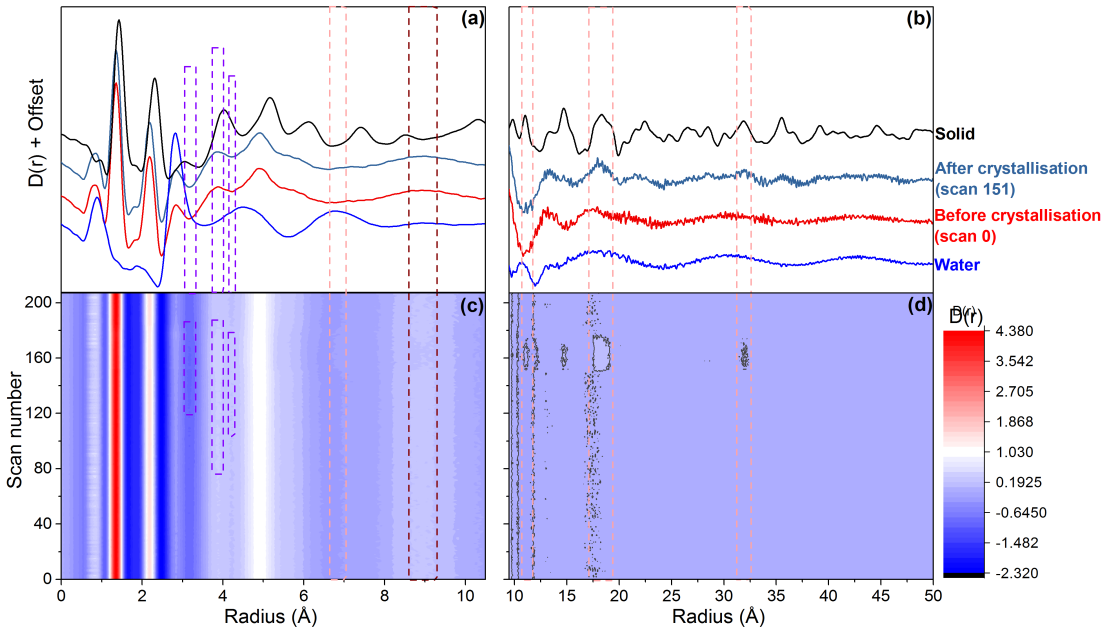

Figure 6.13. XPDF patterns of the 10.7 M aqueous imidazole solution before and after crystallisation from the third cooling cycle, solid imidazole and water cycle in the range: (a) 0 – 10 Å and (b) 10 – 50 Å. Contour plot of the XPDF patterns of the 10.7 M aqueous imidazole solution collected during the third cooling crystallisation experiment in the range: (c) 0 – 10 Å and (d) 10 – 50 Å.

It is worth noting that the purple dotted lines highlight regions where the intensity of the features in the XPDF patterns changes prior to crystallisation. The first highlighted region is centred at ~3.2 Å and its intensity for scans 125 – 183 is lower than the other scans (Figure 6.13 (c)). The intensity of this region in the solid XPDF pattern (black line, Figure 6.13 (a)) is lower than that of the solution before crystallisation (red line, Figure 6.13 (a)). The second and third highlighted regions are centred at 3.9 Å and 4.2 Å respectively and their intensities increase for scans 77 – 181 and scans 105 – 179 respectively. The intensities in these regions is higher in the XPDF pattern of the solid than in the XPDF pattern of the solution before crystallisation (Figure 6.13 (a)). These observations could indicate the presence of short-range direct solute – solute interactions prior to crystallisation.

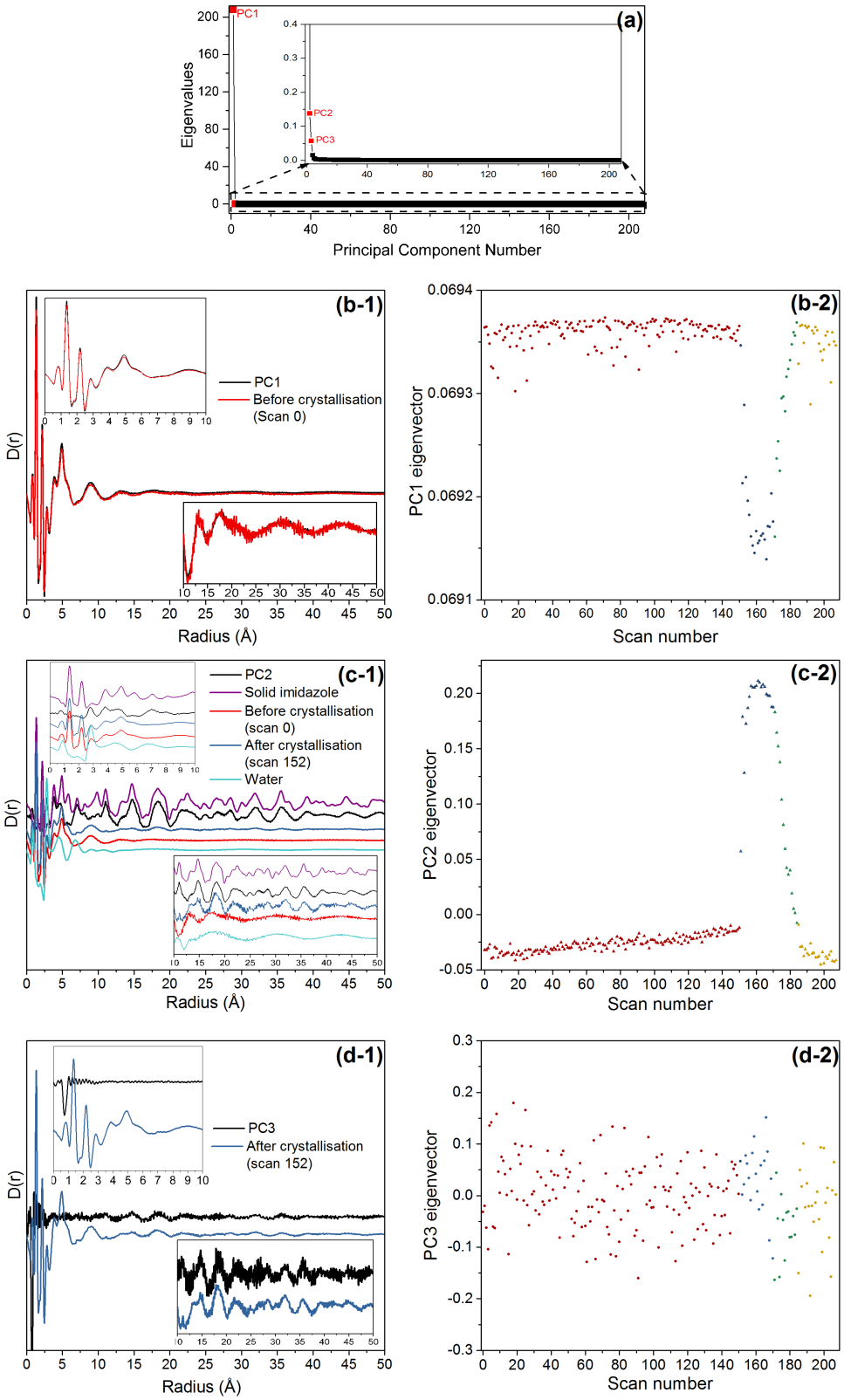

Figure 6.14. (a) Eigenvalues of the principal components from the PCA of the XPDF patterns of the 10.7 M aqueous imidazole solution from the third cooling crystallisation experiment. (b-1) (c-1) (d-1) comparison between the patterns of PC1 – PC3 and the XPDF pattergns of the solid imidazole, water and the aqueous imidazole solution before and after crystallisation. (b-2) (c-2) (d-2) eigenvectors of PC1 – PC3.

Three principal components were identified from the principal component analysis (PCA) for the XPDF patterns collected during the third cooling crystallisation experiment (PC1 – PC3) (Figure 6.14 (a)). PC3 a low eigenvalue and it is difficult to find a pattern for the contributions from the different XPDF patterns to this principal component; hence it will not be discussed further.

- PC1 has the highest eigenvalue and it represents the XPDF pattern of the solution before crystallisation (Figure 6.14 (b-1)) unlike PC1 from the first and the second cooling crystallisation experiments (Figures 6.6 (b-1) and 6.10 (b-1), which represented the XPDF pattern of the solution after crystallisation. The number of XPDF patterns collected for the solution before crystallisation is higher for the third cooling crystallisation experiment (150 scans out of 207) when compared to the first and second measurements. This explains why PC1 resembles the XPDF pattern of the solution before crystallisation. The contributions from the different XPDF patterns to PC1 are almost equal with a slight reduction in contributions as a result of crystal formation (Figure 6.14 (b-2)). The contributions to PC1 increases during dissolution

- PC2 represents the XPDF pattern of the solid in the 10 Å – 50 Å radial range with some features in the 0 Å – 10 Å range corresponding to the solid. The contributions to PC2 from the XPDF patterns of the solution after crystallisation increase. This indicates that the crystal structure of the formed crystals highly resembles that of solid imidazole. During dissolution, the contributions to PC2 reduce until they reach values below 0 which signifies complete dissolution of crystals.

6.4 Conclusions

X-ray Raman scattering (XRS) of the C and N K-edge spectra was used to follow the structural evolution of aqueous imidazole *in-situ* during cooling crystallisation. An abrupt change in the N K-edge XRS spectra indicated the onset of desolvation and crystallisation from the supersaturated solution. The observed relative shift of the π^* resonances in the N-Kedge spectra indicated that approximately 1/3 of dissolved imidazole molecules had formed crystalline product. This also determines that the upper limit for molecules involved in prenucleation clustering in aqueous imidazole is

approximately 1/3 of all solute molecules. There was no evidence for changes due to prenucleation clustering in the C and N K-edge spectra, indicating that molecules with a significantly changed inner solvation shell are at most a small mole fraction.

The XRS measurements were complemented with total X-ray scattering measurements. Total X-ray scattering and the associated X-ray pair distribution function (XPDF) analysis enable probing amorphous structures as well as short-range and long-range ordering. Features appearing in the longer radial range of the XPDF patterns of aqueous imidazole solutions indicated crystal formation. These features coincided with features in the XPDF pattern of solid imidazole, which highlighted the similarity in the structural properties between the formed crystals and crystalline imidazole. Similar to the XRS experiments, very little evidence was found on variations attributed to pre-nucleation clustering in the XPDF patterns.

The observations from the *in-situ* XRS and XPDF experiments suggest that the nucleation of imidazole takes place from highly concentrated aqueous solutions which are dominated by local hydrogen-bonding to water in the inner solvation shell around the imidazole molecules. It follows that the breakage of these solute-solvent hydrogen bonds during desolvation is likely to be the key step in the nucleation of imidazole crystals from aqueous phase.

Chapter 7 Solvent Effects on the Solvation Structure of Ethanolic Solutions of Imidazole

7.1 Introduction

The investigation of aqueous solutions of imidazole at different concentrations (Chapter 5) indicated that their solvation structure consists of individually hydrated imidazole molecules, where solute-solute association occurs in the form of secondary interactions between the hydrated clusters rather than direct imidazole-imidazole interactions. Probing the structural evolution of concentrated aqueous solutions of imidazole during cooling crystallisation (Chapter 6) revealed that the average solvation structure around imidazole molecules does not change significantly while traversing the metastable supersaturated zone. These observations are at odds with previously reported structural models of dilute imidazole solutions, which predicted self-association of imidazole molecules in stacked assemblies [86, 152]. Self-association of imidazole molecules in non-aqueous solutions was also reported in the literature. Infrared studies of imidazole in carbon tetrachloride [151, 267] and benzene [151] solutions predicted self-association of imidazole in the form of hydrogen-bonded oligomers. Similarly, ^{1}H NMR predicted that imidazole molecules exist as trimers in chloroform solutions [268].

This chapter aims to study the intermolecular interactions between solute and solvent molecules in ethanol solutions to establish whether the solvation of imidazole in ethanol differs than that in water. Ethanol was the chosen solvent for these studies for several reasons. Firstly, it has a unique structure in the liquid-phase. The number of hydrogen bonds formed by ethanol is half of those formed by water; hence the structure of ethanol mainly consists of winding chains unlike the tetrahedral structure of water. Secondly, the solubility of imidazole in ethanol is relatively high; a concentration of ~ 7 $M.L^{-1}$ can be achieved at 25 °C. This results in a high signal/noise ratio and enables collecting high quality data using Fourier transform Infrared (FTIR) spectroscopy, X-ray Raman scattering (XRS) and total X-ray scattering. Thirdly, ethanol is widely used within the chemical industry as a primary solvent and reagent with diverse applications in biological, chemical, materials and environmental processes. Lastly, the solvent is chemically different than imidazole. It doesn't contain N making it easier to separate

signatures of the solute, imidazole, from the solvent, ethanol. The structural evolution of imidazole in ethanolic solutions as a function of concentration and during cooling crystallisation is also investigated in this chapter.

7.2 Experimental

Ethanolic solutions of imidazole with concentration between 1 M and 8 M were characterised with FTIR spectroscopy, X-ray Raman scattering (XRS) and Total X-ray scattering.

7.2.1 Fourier Transform Infrared (FTIR) Spectroscopy

The FTIR measurements were collected using a Thermo Scientific Nicolet 6700 FTIR spectrometer. Solid imidazole was measured in attenuated total reflectance (ATR) mode. Ethanol and ethanolic solutions of imidazole (with concentration between 1 M and 8 M) were measured in transmission mode using a liquid flow-cell (Figure 2.5, Chapter 2). Omnic software was used to collect the background and sample spectra. The spectra were collected at a resolution of 4 cm^{-1} in 32 scans over 4000-500 cm^{-1}.

7.2.2 X-ray Raman Scattering (XRS)

The C, N and O K-edge spectra of a 5 M ethanolic solution of imidazole were collected at the inelastic scattering beamline ID-20 of the ESRF, Grenoble, France [183]. ID-20 is an undulator beamline that uses Si(111) pre-monochromator which can be coupled by a consecutive monochromator to form a fixed exit ensemble. Si(311) post-monochromator was utilised to achieve an energy resolution of 0.6 eV which is needed for this experiment. All data were collected using the inverse scanning mode, where the incident photon energy was scanned, and the analyser energy was fixed at 9690 eV. The scattered radiation was collected employing a 12 spherically bent Si(660) analyser crystals at an angle of 42° with q value of 3.6 $Å^{-1}$. Hence, the data presented here were collected at low momentum transfer (q). The monochromator drift was small, and it was monitored through frequent measurement of the elastic scattering peak.

A 5 M ethanolic solution of imidazole was measured using a jacketed flow-cell crystalliser (Flow-cell 1, section 2.4 - chapter 2) which ensured precise temperature control and continuous circulation of the stock solution to minimise radiation induced

damage to imidazole. The total volume of the solution was ~250 ml and the solution was pumped at a rate of ~450 ml/min.

For data analysis, the elastic energy of the incident photons was subtracted from the XRS data to obtain the near-edge spectra. The spectra were then background subtracted and normalised to the edge-step height using Athena software [239]. After normalisation, the contribution from the Kapton window to the K-edge spectra of the ethanolic solutions of imidazole was removed by subtracting a reference spectrum of Kapton from those of the solutions. The subtraction coefficients were defined as the value required to remove a peak significant to the K-edge spectra of Kapton. These peaks for the C, N and O K-edges are circled in Figure 7.1.

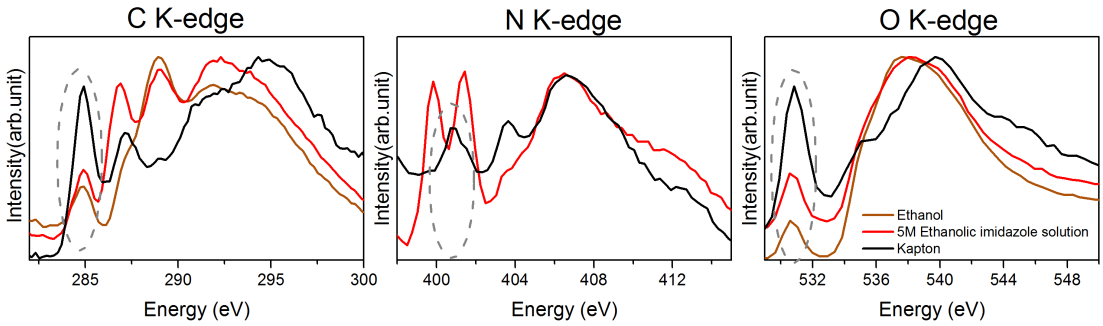

Figure 7.1. The C, N and O K-edge XRS spectra of Kapton, the 5 M ethanolic imidazole solutions and ethanol. The peaks used to define the subtraction coefficient to remove the Kapton contribution to the K-edge signal of the ethanolic solutions of imidazole and of ethanol are circled in grey.

7.2.3 Total X-ray Scattering and X-ray Pair Distribution Function (XPDF)

The total X-ray scattering experiments were performed at the X-ray pair distribution function (XPDF) beamline I15-1 of Dimond Light Source (DLS), UK [192]. Total X-ray scattering data were collected for:

- Crystalline imidazole which was measured in a borosilicate glass capillary. The acquisition time for the solid sample was 10 minutes per scan.
- Ethanol and ethanolic solutions of imidazole (with concentration between 2 M – 8 M) which were measured using Flow-cell 2 (section 2.4 - chapter 2). The acquisition time for these samples was 10 minutes per scan.

- 7 M ethanolic solution of imidazole during cooling crystallisation until the onset of crystal formation. Flow-cell 2 was also used for these measurements. The measurements were repeated twice with an acquisition time of 30 seconds per scan.

A flow-cell was utilised to measure the liquid samples to ensure continuous circulation of the solution, which would minimise the beam induced damage to the samples and guarantee a uniform temperature distribution.

Raw X-ray scattering data was automatically corrected for the internal dark current during data collection using the DLS generic data acquisition (GDA) software. Further corrections for the instrument and the sample holder backgrounds, absorption, multiple scattering and Compton scattering were performed in GudrunX to generate the interference differential cross sections F(Q)s [224]. GudrunX was also used to generate the pair distribution functions (PDFs) by Fourier transformation of the F(Q)s [224, 248].

7.2.4 Empirical Potential Structure Refinement (EPSR)

The empirical potential structure refinement (EPSR) simulation was used to refine structural models for ethanolic solutions of imidazole (with concentration between 2 M – 8 M) utilising the experimental total X-ray scattering data. The EPSR simulation starts by generating an initial configuration of atoms based on reference potentials (Lennard-Jones potentials and Coulomb charges). This is done using a standard Monte Carlo simulation. After equilibration, an additional empirical potential is refined iteratively by perturbation of the reference potential and the structure factor (F(Q)) is calculated for each new atomic configuration. The atomic configuration is either accepted or discarded according to the goodness-of-fit between the experimental and calculated F(Q). Finally radial and spatial distribution of atomic and molecular pairs were calculated using routines within EPSR and Dlputils [200].

The side length of the cubic simulation boxes used in EPSR, the number of imidazole and ethanol molecules and the number density of ethanol and ethanolic solutions of imidazole (with concentration between 2 M – 8M) are shown in Table 7.1. The ethanol molecules were constructed and optimised in ORCA and the imidazole molecules were constructed in Jmol. Bond and angle restraints were employed, as well as torsional

restraints for the imidazole and ethanol molecules. However, it is worth noting that the molecules were not held rigid during the EPSR simulation in order to reproduce the experimental scattering data [198]. The Lennard-Jones potential and Coulomb charge parameters for ethanol and imidazole were taken from the OPLS [227] and they are listed in Table 7.2.

Table 7.1. The side length of the cubic simulation boxes, the number of imidazole and ethanol molecules and the number density of ethanol and ethanolic solutions of imidazole

	Side length of cubic box	Number of imidazole molecules	Number of ethanol molecules	Number density (atoms.Å^{-3})
Ethanol	45.94	-	1000	0.0928
2 M	46.73	123	1000	0.0990
3 M	48.27	203	1000	0.0963
4 M	49.88	299	1000	0.0942
5 M	51.71	403	1000	0.0942
6 M	53.09	541	1000	0.0927
7 M	58.98	852	1000	0.0812
8 M	71.73	1778	1000	0.0677

Table 7.2. Lennard-Jones potentials and Coulomb charges for ethanol and imidazole [227]. The labelling schemes used for ethanol and imidazole are shown in Figure 7.2.

Atom	ε / kJ.mol^{-1}	σ /Å	Charge / e-
OE1	0.711	3.120	-0.683
HE3	0.000	0.000	0.418
HE2	0.126	2.500	0.046
CE2	0.276	3.500	0.145
CE1	0.276	3.500	-0.180
HE1	0.126	2.500	0.069
N1	0.711	3.250	-0.570
C5	0.293	3.550	-0.015
H1	0.000	0.000	0.420
C2	0.293	3.550	0.015
C4	0.293	3.550	0.295
H2	0.126	2.420	0.115
N3	0.711	3.250	-0.490

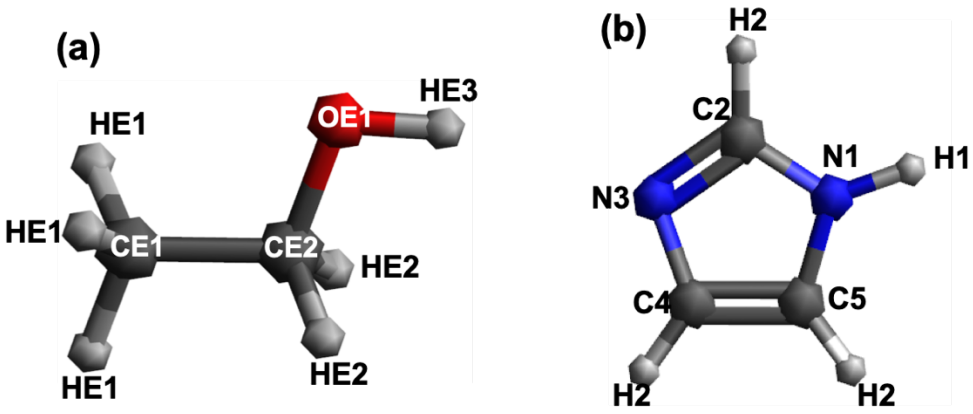

Figure 7.2. Atom labelling used for the Lennard-Jones potentials in the EPSR simulation of (a) ethanol and (b) imidazole.

7.3 Results and Discussion

7.3.1 Effect of Solvent on the Solvation Structure of Imidazole

In this section a 5 M ethanolic solution of imidazole is compared to a 5 M aqueous solution of imidazole to shed light on the effects of the solvent on the intermolecular interactions within the solution.

7.3.1.1 Fourier Transform Infrared (FTIR)

Assigning the vibrational bands attributed to imidazole in the FTIR spectrum of ethanolic solutions of imidazole is not trivial due to the presence of vibrational bands associated with ethanol in the same spectral range. Nonetheless, the sub-bands attributed to the N-H stretching vibrational mode of imidazole (3100 – 2200 cm^{-1}) are detected in the FTIR spectrum of the 5 M ethanolic solution of imidazole (Figure 7.3 (a)). These sub-bands have been attributed to the Fermi resonances of the stretching band with overtones and combinations of the lower frequency vibrations [254-256]. Other bands in the 1600 – 1000 cm^{-1} range, which are attributed to stretching and bending vibrational modes of the imidazole ring are also detected in the FTIR spectrum of the 5 M ethanolic solution of imidazole (Figure 7.3 (a)).

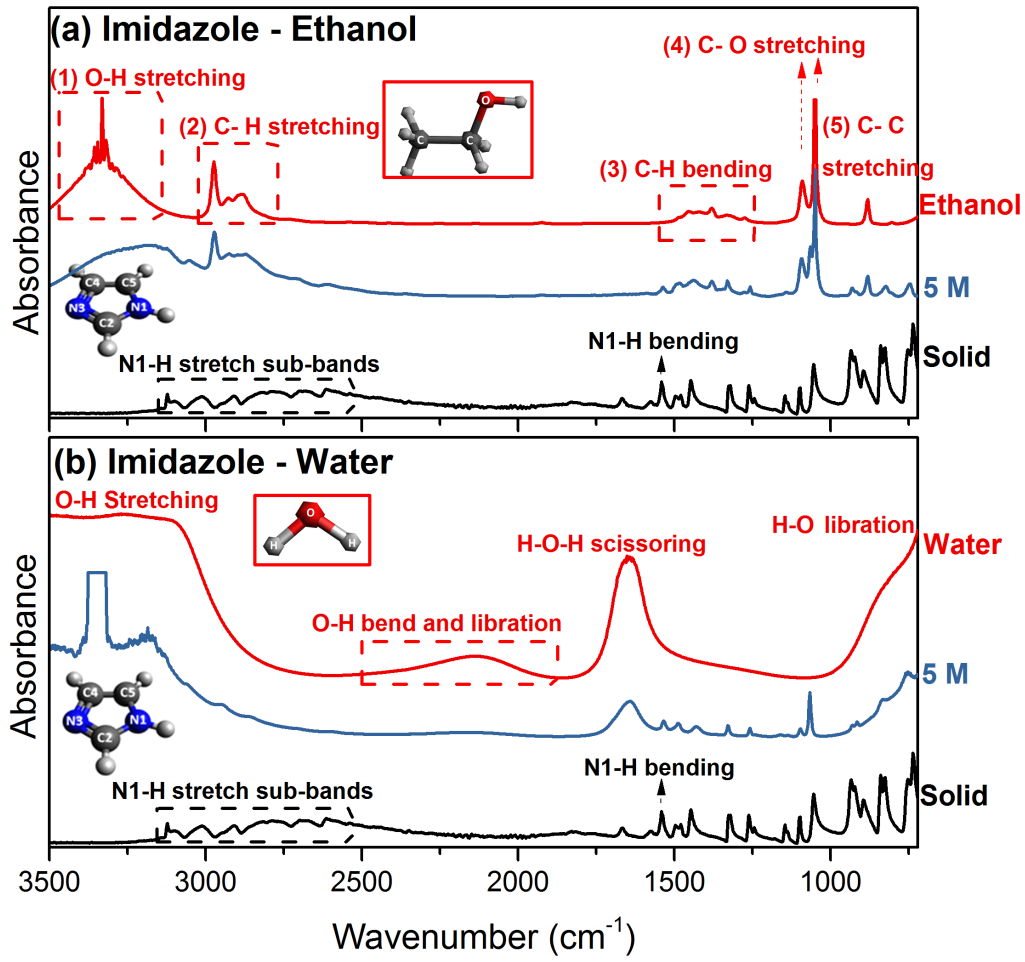

Figure 7.3. (a) FTIR spectra of solid imidazole, 5 M ethanolic solution of imidazole and ethanol. (b) FTIR spectra of solid imidazole, 5 M aqueous imidazole solution and water (presented in Chapter 5). All FTIR spectra are collected in the range between 4000 – 650 cm^{-1} using ATR mode for the solid sample and transmission mode for the liquid samples.

When comparing the FTIR spectrum of solid imidazole to the 5 M ethanolic solution, it is noticed that the vibrational frequencies of the N-H stretching sub-bands shift to lower values (red shifts) (Table 7.3) while the N-H bending frequency (1540.9 cm^{-1}) shifts to higher value (blue shift). The trends in the shifting of the N-H stretching and bending bands are associated with weaker N-H bond. Similar trends were observed in the 5 M aqueous imidazole solution (chapter 5, Figure 7.3(b) and Table 7.3). These indicate that the intramolecular N-H bond is weaker in the solid-state than in the ethanolic or aqueous solutions, which reflects stronger hydrogen-bonding in the solid-state. It is worth noting here that the vibrational frequencies of the N-H stretching sub-bands and the N-H bending band for the ethanolic solution of imidazole have intermediate values between those of solid imidazole and aqueous imidazole solution (Table 7.3). Based on the observations above, it could be established that the intramolecular N-H bond is

weaker in the ethanol solution when compared to the aqueous solution indicating stronger intermolecular hydrogen-bonding involving the N-H bond in the ethanol solutions.

Table 7.3. Summary of FTIR vibrational bands frequencies attributed to imidazole for solid imidazole, 5 M ethanolic solution of imidazole and 5 M aqueous imidazole solution.

Imidazole Solid	5M imidazole ethanol solution	5M aqueous imidazole solution	Assignment
1054.9	1064.5	1066.5	ν(ring), δ(CH) in plane
1145.5	1141.7	1160.9	ν(ring) δ(CH) out of plane
1261.2	1257.4	1259.3	ν(ring)
1321.0	1329.7	1328.7	ν(ring), ν (CN)
1446.4	1438.6	1430.9	ν(ring), δ(CH)
1540.9	**1537.0**	**1533.2**	ν(ring), **δ(N1H)**
1575.6			C-C stretching
2510.9	**2512.8**	**2516.7**	
2582.3	**2593.8**	**2599.6**	
2699.9	**2701.8**	**2701.8**	
2817.5		2856.1	**N-H stretching sub-bands**
2908.2		2958.3	
2946.7			
3012.3			
3043.2	**3050.8**	**3058.6**	

The FTIR spectrum of ethanol has five significant vibrational bands; (1) O-H stretching (3500 – 3056.7 cm^{-1}), (2) C-H stretching (3000 – 2840 cm^{-1}), (3) C-H bending (1520 – 1200 cm^{-1}), (4) C-O stretching (1091.5 cm^{-1} and 1051 cm^{-1}) and (5) C-C stretching (881.3 cm^{-1}). The O-H stretching band is a broad strong band, which can saturate the FTIR spectra as seen in Figure 7.3 (a) and could mask other vibrational bands in the same frequency range especially if the data is collected in transmission mode. To overcome these challenges, the FTIR spectra could be collected in ATR mode. However, this was not feasible with ethanol or ethanolic solutions because ethanol evaporated too quickly to do the FTIR measurements in ATR mode.

The five vibrational bands of ethanol in the FTIR spectrum of the 5M ethanolic solution of imidazole are less intense than those in the spectrum of ethanol (Figure 7.3 (a)). Adding to that, the vibrational frequency of the O-H stretching band shifts to lower value (red shift) in the FTIR spectrum of the ethanolic solution of imidazole when compared

to ethanol. This could indicate that the intramolecular O-H bond of ethanol is weaker in the imidazole - ethanol solutions, which could reflect stronger hydrogen-bonding in the solution phase.

Table 7.4. Summary of FTIR vibrational frequencies attributed to ethanol (red) and water (blue) for 5 M ethanolic solution of imidazole, ethanol, 5 M aqueous imidazole solution and water

5M imidazole ethanol solution	Ethanol	5M aqueous imidazole solution	Water	Assignment
		< 997	< 997	O-H liberation
881.3	881.3			C-C stretching
1049.1 1093.5	1051.0 1091.5			C-O stretching
	1454			C-H bending
	1421.3			
1378.9	1380.8			
	1330.6			O-H in-plane bending
1274.7	1274.7			C-H bending
		1646.9	1646.9	H-O-H scissoring
		2134.8	2134.8	H-O-H Bend and liberation
2871.5 2896.6	2887			C-H stretching
2925.5	2927.5			
2971.8	2971.8			
3056 – 3500	3076 – 3500	3500 – 2825	3500 – 2825	O-H stretching

7.3.1.2 X-ray Raman Scattering (XRS)

N K-edge

The near-edge region of the N K-edge spectra of ethanolic solution of imidazole is dominated by two intense peaks attributed to transitions from the 1s orbital of the N3 and N1 moieties in imidazole to the unoccupied 1π* states (Figure 7.4). These two features were also assigned in the near-edge region of the N K-edge spectra of aqueous imidazole solution and solid imidazole (Figure 7.4). The peaks are separated by 1.687 eV, 1.547 eV and 1.517 eV for the 5M aqueous imidazole solution, 5M ethanolic solution of imidazole and solid imidazole respectively. The separation between the two peaks in the N K-edge spectra of the ethanolic solution of imidazole is more comparable to that

of solid imidazole than imidazole in aqueous environments. This indicates that the electron density at the nitrogen moieties of imidazole in ethanol solutions is similar to that of imidazole in the crystalline phase, which is in agreement with the findings of the FTIR study above (section 7.3.1.1).

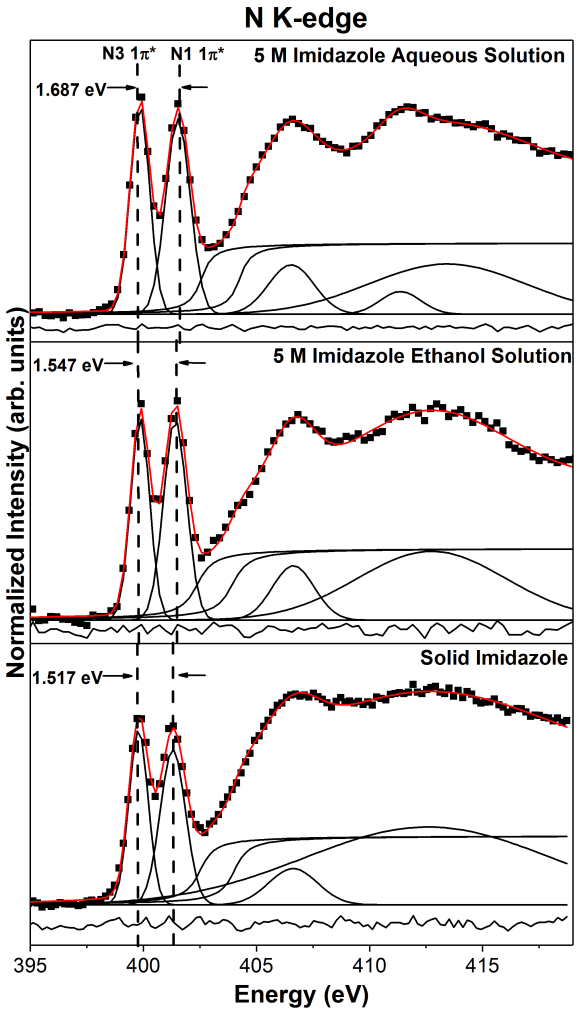

Figure 7.4. Fitted N K-edge XRS spectra of 5 M aqueous imidazole solution, 5 M ethanolic solution of imidazole and solid imidazole.

C K-edge

Unlike the C K-edge spectrum of aqueous imidazole solutions which only reflects the changes in the bonding environment around imidazole, the C K-edge spectrum of ethanolic solutions of imidazole reports on changes in the local environment around both the imidazole and the ethanol molecules. To elucidate the origin of the features in the C K-edge spectra of the imidazole in water and in ethanol solutions, they were compared with the C K-edge spectrum of solid imidazole. The later has two features in the near-edge region corresponding to transitions from the 1s orbitals of the carbon

atoms to the unoccupied 1π* and 2π* orbitals. These are followed by the ionisation potentials (IP) and the broad σ* shape resonances corresponding to multiple scattering at the potential barrier to the bound neighbouring atoms. The spectral features of the C K-edge spectrum of the 5M aqueous imidazole solution are in good qualitative agreement with those of solid imidazole as seen from Figure 7.5 (a). On the other hand, the C K-edge spectrum of the 5M ethanolic solution of imidazole displays features from both the solid imidazole and ethanol as seen in Figure 7.5 (b).

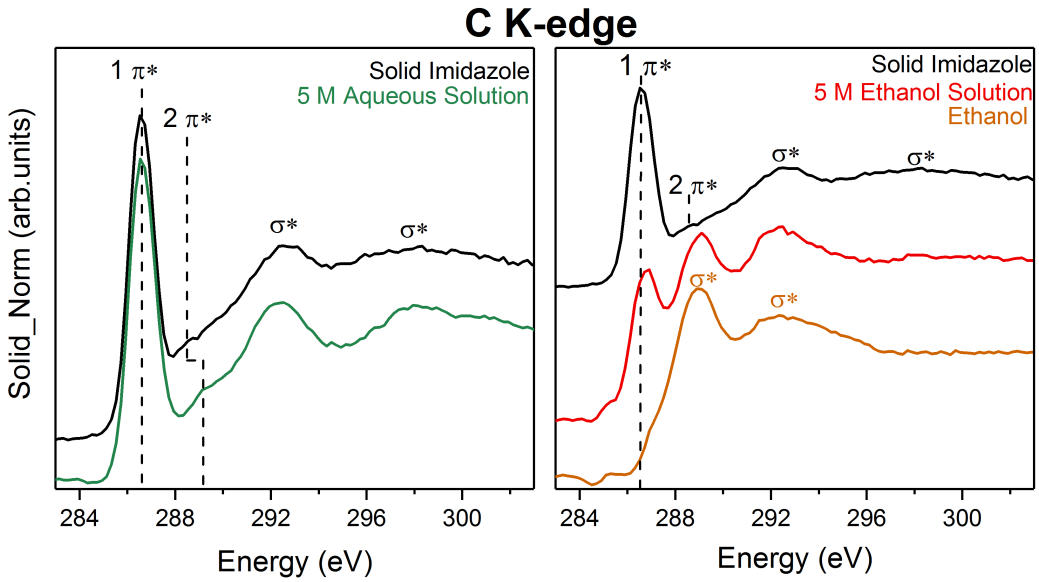

Figure 7.5. C K-edge XRS spectra of (left) solid imidazole and 5 M aqueous imidazole solution (right) solid imidazole, 5 M ethanolic solution of imidazole and ethanol.

O K-edge

The O K-edge spectra of aqueous and ethanolic solutions of imidazole report on variations in the local environment around the solvent molecules (i.e. water and ethanol) since there is no direct contribution from imidazole to the O K-edge signal. The spectral features in the O K-edge spectrum of ethanol will be discussed and compared to those of water before explaining the origin of the spectral features in the O K-edge spectrum of 5 M ethanolic solution of imidazole.

As mentioned in chapter 5, the O K-edge spectrum of water has three main regions: the pre-edge (534 eV), the main-edge (536.5 – 539.5 eV) and the post-edge (541 eV) and the variations in the spectral features of the O K-edge could be related to the bonding coordination in the system. A dominant post-edge feature is attributed to fully

coordinated water molecules whereas intense pre- and main-edge peaks are associated with broken or weak hydrogen bonds [41, 42, 45, 257].

The X-ray absorption spectrum of ethanol exhibits features similar to those of water [269]. The pre-edge feature in the O K-edge of ethanol, which is associated with broken hydrogen bonds, is manifested as a shoulder when compared to pre-edge feature in the O K-edge of water and exhibits a red shift (by ~ 1eV) (Figure 7.6 (a)). Addition of a small amount of water significantly reduces the intensity of the ethanol pre-edge feature while increasing that of water (Figure 7.6 (b)).[269] The depletion of the ethanol pre-edge feature in the O K-edge spectra as a result of adding a small amount of water suggests that water reduces the number of dangling O-H bonds originating from the ethanol molecule.

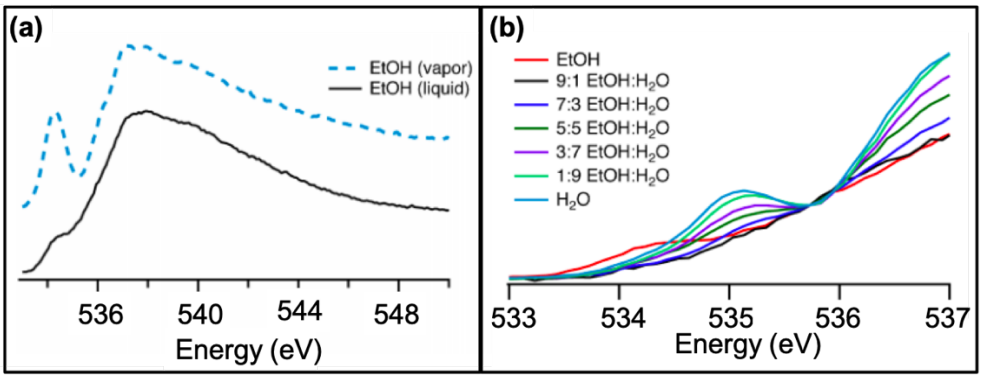

Figure 7.6. (a) O K-edge XA spectra for vapor (dashed-blue) and liquid (solid-black) ethanol (b) O K-edge spectra in the pre-edge region for ethanol: water mixtures [269].

The O K-edge XRS spectrum of ethanol in Figure 7.7 does not show the pre-edge feature associated with broken hydrogen-bonds. The ethanol grade used for the experiment was 99.8% pure with less than 0.2% water, which means that the ethanol: water ratio is higher than the highest ratio in Figure 7.6(b) which shows the absence of the pre-edge feature. This could be attributed to (1) the fact that XRS is a bulk-sensitive technique and the intensity of the pre-edge feature will be even lower than that reported previously using soft X-ray absorption techniques or (2) the presence of water, even at such a low concentration, is sufficient to eliminate the dangling O-H bonds in ethanol.

The effect of imidazole on the coordination environment of ethanol has been examined by comparing the O K-edge spectrum of ethanol to that of a 5 M ethanolic solution of imidazole (Figure 7.7). The post-edge region of the O K-edge spectrum of the ethanolic

solution of imidazole decreased slightly in intensity when compared to that of ethanol. The pre- and main-edge regions in the O K-edge spectrum of the ethanolic solution of imidazole do not seem to be different than those of ethanol. Hence, it seems that the presence of imidazole in ethanolic solutions does not disturb the hydrogen-bond network of ethanol. This is similar to the effect of imidazole on water (Chapter 5) where it was concluded that the presence of imidazole does not disturb the hydrogen-bond network of water.

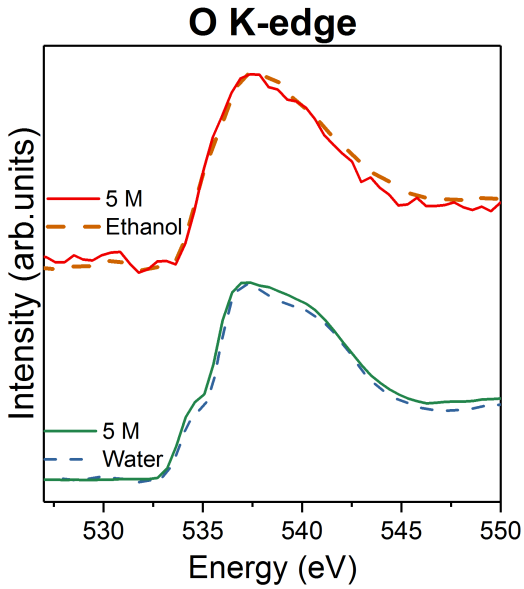

Figure 7.7. O K-edge XRS spectra of ethanol (dashed-orange), 5 M ethanolic solution of imidazole (solid-red), water (dashed-blue) and 5 M aqueous imidazole solution (solid-green).

7.3.1.3 X-ray Pair Distribution Function (XPDF)

The X-ray pair distribution function (XPDF) patterns of ethanol and a 5 M ethanolic solution of imidazole were generated through Fourier transformation of the total X-ray scattering patterns. They are compared with the XPDF patterns of water, 5 M aqueous imidazole solution and solid imidazole in Figure 7.8. The similarity index between the XPDF patterns of the 5 M ethanolic solution of imidazole and solid imidazole is 86% in the 0 – 10 Å range (Figure 7.8 (c)). The features attributed to imidazole in the XPDF pattern of the 5 M ethanolic solution of imidazole are more intense than those in the XPDF pattern of the 5 M aqueous imidazole solution in the 0 – 10 Å radial range (Figure 7.8 (a)). This could be attributed to the reduction in the number of solvent molecules per imidazole molecule from 7.8 in the 5 M aqueous solution to 2.5 in the 5 M ethanol solution, which means direct imidazole – imidazole interactions might take place in the

ethanol solutions. However, evidence for such direct imidazole – imidazole interactions does not appear to extend into the longer radial range from 10 – 50 Å. The features in the XPDF pattern of the 5 M ethanolic solution of imidazole highly resemble those of ethanol in the 10 – 50 Å range as seen from Figure 7.8 (b) and from the similarity index of 97% between the two patterns (Figure 7.8 (d)). The high similarity index either suggests (1) ethanol – ethanol interactions dominate in the longer-range or (2) the nature of imidazole – ethanol interactions is very similar to the ethanol – ethanol interactions.

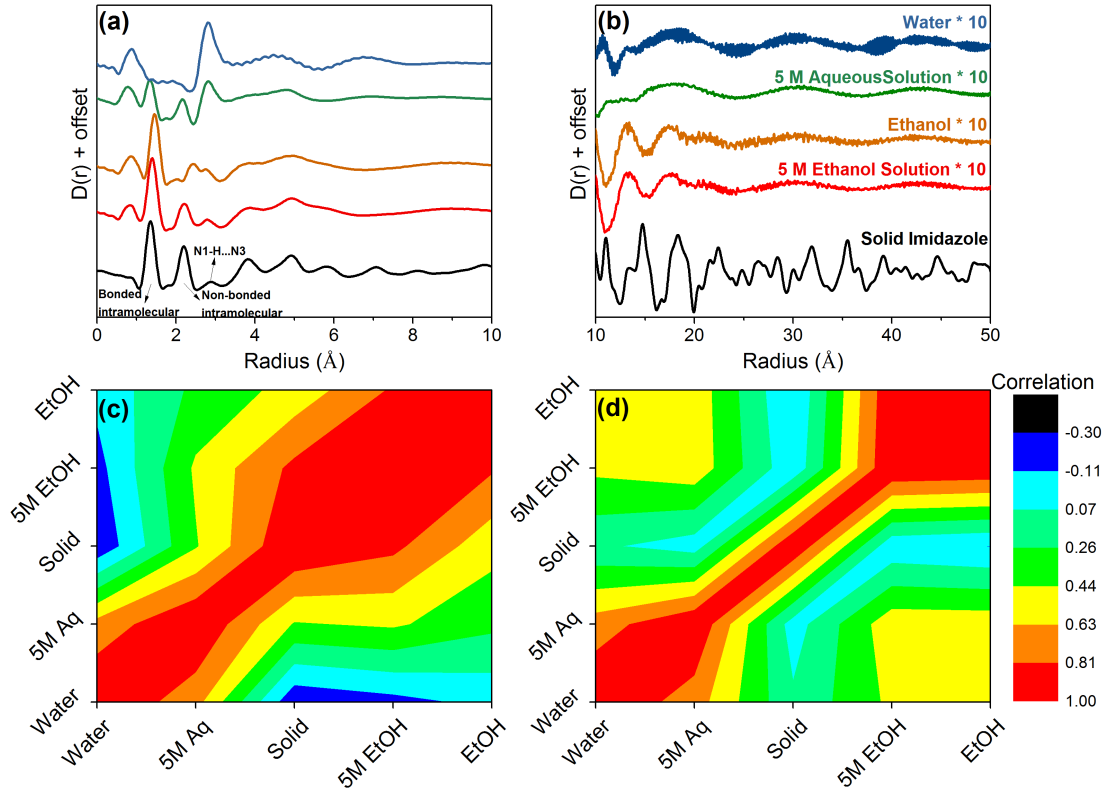

Figure 7.8. (Top) XPDF patterns of solid imidazole, 5 M ethanolic solution of imidazole, ethanol, 5 M aqueous imidazole solutions and water in the range (a) 0 -10 Å and (b) 10 – 50 Å. (Bottom) Correlation matrix of XPDF patterns of solid imidazole, 5 M ethanolic solution of imidazole, ethanol, 5 M aqueous imidazole solutions and water in the range (c) 0 -10 Å and (d) 10 – 50 Å.

Quantitative analysis of the atomic correlations in the ethanolic solution of imidazole from the XPDF patterns shown in Figure 7.8 is not feasible. The reported XPDF patterns represent the sum of 171 partial PDFs, which are attributed to all the atomic pair-pair correlations present in the ethanolic solution of imidazole. Therefore, the empirical potential structure refinement (EPSR) simulation was utilised to generate a structural

model for the 5 M ethanolic solution of imidazole and for ethanol similar to what has been done for the aqueous imidazole solutions and water in chapters 3 and 5. These models were refined utilising the same total X-ray scattering data, which were used to generate the XPDF patterns shown in Figure 7.8 (a) and (b). The structure factor (F(Q)) patterns calculated from the EPSR simulation compare well with those obtained from the experimental scattering data (Figure 7.9).

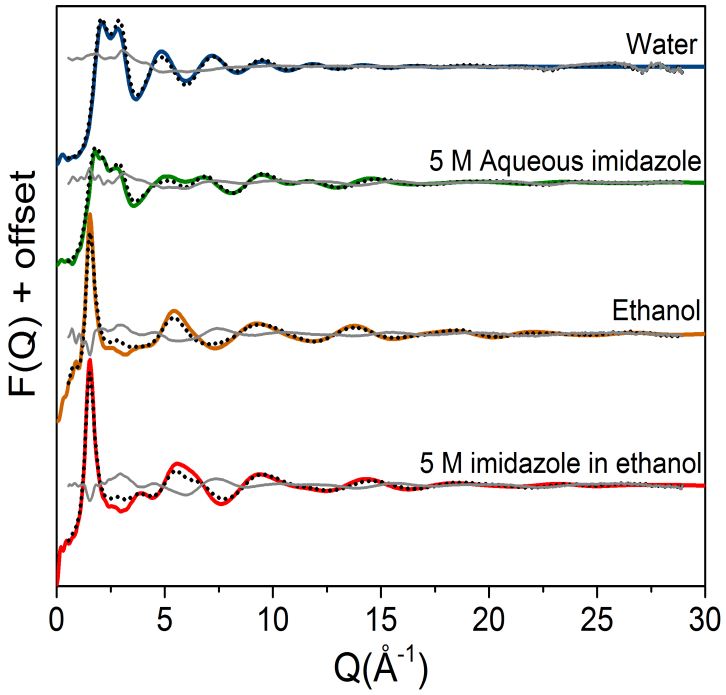

Figure 7.9. F(Q) patterns from experimental total X-ray scattering data (black dots) and EPSR models (coloured lines). The difference patterns (grey lines) illustrate the high comparability between the experimental and calculated F(Q) patterns.

Imidazole – solvent interactions

Imidazole – ethanol interactions are anticipated in ethanolic solutions of imidazole because both imidazole and ethanol can accept and donate hydrogen-bonds. The difference of 1 Å between the first peak in the N1…O_S radial distribution function (RDF) and H1…O_S RDF and between N3…O_S RDF and N3…H_S RDF (shown in Figure 7.10 (a)) indicates collinear hydrogen-bonding between imidazole and ethanol. The RDFs of the hydrogen-bonded pairs of imidazole and ethanol resembles those of imidazole and water pairs in terms of peak position (Figure 7.10 (a)). This indicates that the hydrogen-bond lengths of solute-solvent pairs are similar in aqueous and ethanolic solutions of imidazole.

The solvation environment around the imidazole molecules is different as indicated by the coordination number (CN) of O_s and H_s around the nitrogen moieties in the imidazole ring (Table 7.5). The CN of ethanol molecules around the N1 and N3 sites of the imidazole ring indicate the presence of one ethanol molecules per nitrogen moiety at a distance of 3.7 Å (Table 7.5). Whereas the coordination number of water molecules around the nitrogen moieties in imidazole for the same distance indicate the presence of three water molecules per nitrogen atom (Table 7.5). This is further illustrated by the RDF of the centre of geometry (cog) of the solvent molecules around the centre of geometry (cog) of the solute molecules (Figure 7.10(b)). The shoulder at 3 Å in the water (cog) – imidazole (cog) RDF was assigned to water molecule above and below the plane of the imidazole ring. On the other hand, the RDF of the ethanol (cog) around imidazole (cog) lacks this shoulder, which could be attributed to ethanol molecules in locations not adjacent to the nitrogen atoms in the imidazole ring.

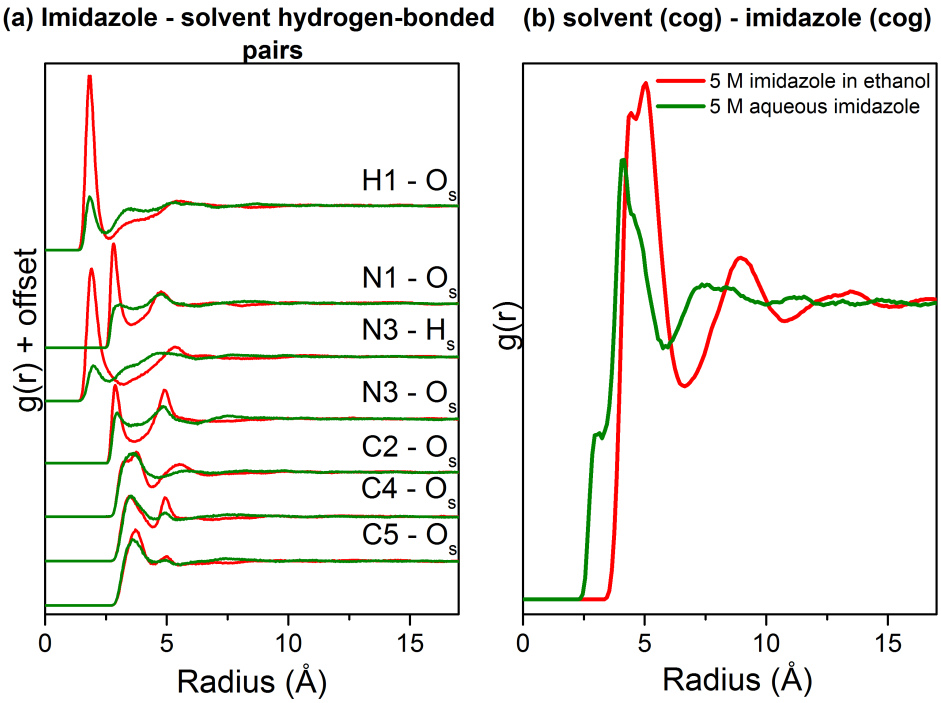

Figure 7.10. (a) The radial distribution functions (RDFs) of imidazole-solvent hydrogen-bonded pairs. O_s is Ow in water and OE1 in ethanol. H_s is Hw is water and HE3 in ethanol (b) The radial distribution function (RDF) of the centre of geometry (cog) of solvent around the cog of imidazole.

Table 7.5. Coordination numbers of hydrogen-bonded imidazole – solvent pairs

Central atom-surrounding atom	Sample	Coordination number	Cut-off / Å
H1-O_s	5 M aqueous solution	0.7	2.4
	5 M ethanol solution	0.7	2.7
N1-O_s	5 M aqueous solution	2.8	3.7
	5 M ethanol solution	1	3.6
N3-H_s	5 M aqueous solution	1.3	2.5
	5 M ethanol solution	0.9	3.3
N3-O_s	5 M aqueous solution	2.5	3.6
	5 M ethanol solution	0.9	3.7
Solvent (cog) – imidazole (cog)	5 M aqueous solution	19.4	6
	5 M ethanol solution	9	6.7

The spatial density function (SPD) of the solvent molecules around imidazole (Figure 7.11) illustrates the presence of the solvent molecules adjacent to the N1 site and angularly distributed around the N3 site. The SDF also shows that the position of the ethanol molecules around imidazole are further away from the imidazole molecule than the water molecules as also seen from the cog-cog RDF (Figure 7.9(b)). This could be attributed to the larger size of the ethanol molecules when compared to water molecules.

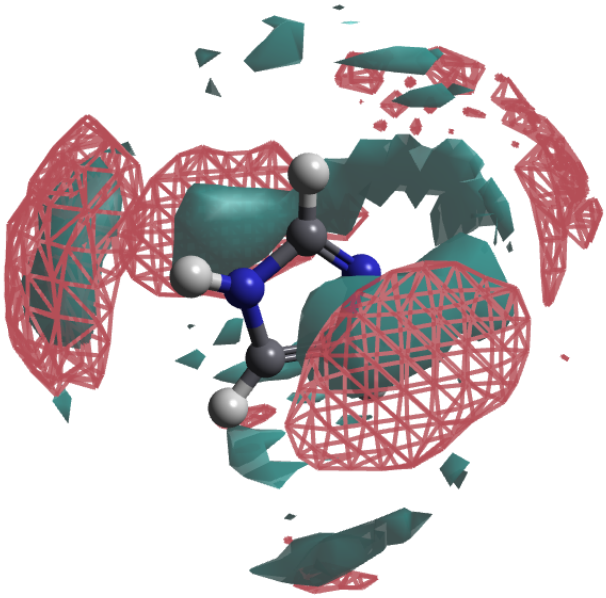

Figure 7.11. Spatial density function (SDF) of (a) water around imidazole for distances up to 6 Å (10 % of water molecules shown) (green surface) and (b) ethanol around imidazole for distances up to 6.7 Å (10 % of ethanol molecules shown) (red mesh).

Solvent – solvent interactions

The intermolecular interactions between the ethanol molecules in ethanol and in the 5 M ethanolic solution of imidazole are investigated through (1) the pair-pair correlations of the hydroxyl group in the ethanol molecule (Figure 7.12 (b)) and (2) the RDFs of the cog of ethanol around the cog of ethanol (Figure 7.12 (c)). The positions of the RDFs (both pair-pair and cog-cog) for the 5 M ethanolic solution of imidazole are similar to those of pure ethanol (Figure 7.12). The height of the peaks in the RDFs of the ethanolic solution of imidazole is smaller than that of ethanol (Figure 7.12 (b)). This correlates with a reduction in the area under the peaks which indicates a decrease in the number of ethanol molecules coordinating other ethanol molecules in the named solution (Table 7.6). The reduction in the number of ethanol molecules suggests a disruption to the ethanol-ethanol interactions in the 5 M solutions when compared to that in ethanol. This disruption is clearly illustrated in the spatial density functions (SDFs) of ethanol around ethanol (top panel of Figure 7.13). Nonetheless, the coordination environment around the ethanol molecules seems intact as imidazole replaces the ethanol molecules in the first and second hydration shells (bottom panel in Figure 7.13). Preservation of the hydrogen-bond network of ethanol by imidazole explains the broad similarity in the spectral features of the O K-edge XRS spectra of the ethanolic solutions of imidazole and ethanol (which were discussed in section 7.3.1.2).

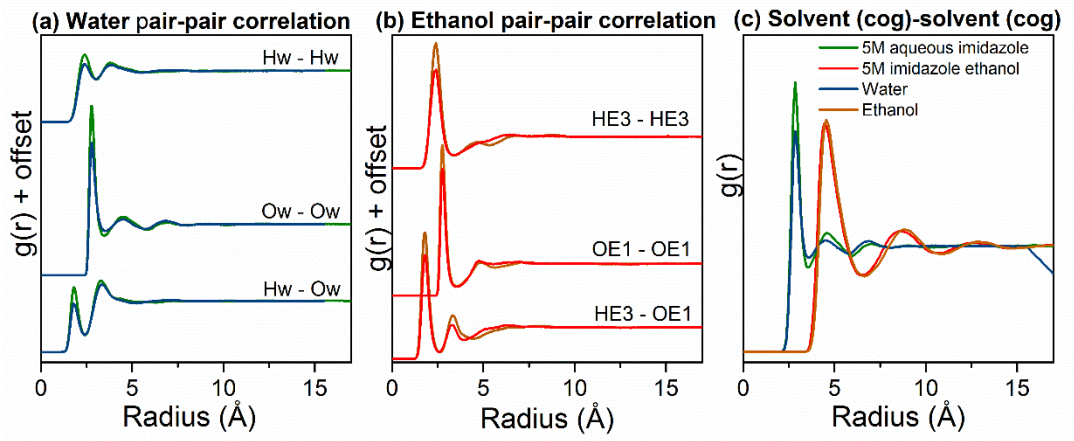

Figure 7.12. (a) The radial distribution functions (RDFs) of the atomic pair-pair correlations in water (b) The radial distribution functions (RDFs) of the atomic pair-pair correlations in ethanol (c) The radial distribution function (RDF) of the centre of geometry (cog) of the solvent (i.e. water or ethanol) around cog of the solvent.

Comparison of the intermolecular interactions between solvent molecules in ethanol/ethanol solutions and water/water solutions illustrates that the positions of the first peaks in the pair-pair RDFs are similar (Figure 7.12 (a) and (b)). This indicates that the lengths of hydrogen bonds between solvent molecules in the named liquids are similar. The first peak in the RDF of the ethanol (cog) around ethanol (cog) occurs at a longer radial distance than that of water. This is attributed to the larger size of the ethanol molecule when compared to the water molecule. Although the hydrogen bond length is similar for water and ethanol, the solvation environment is different as illustrated from the second and third peaks in the pair-pair and cog-cog RDFs (Figure 7.12) and from the spatial density functions (SDFs) (Figure 7.13).

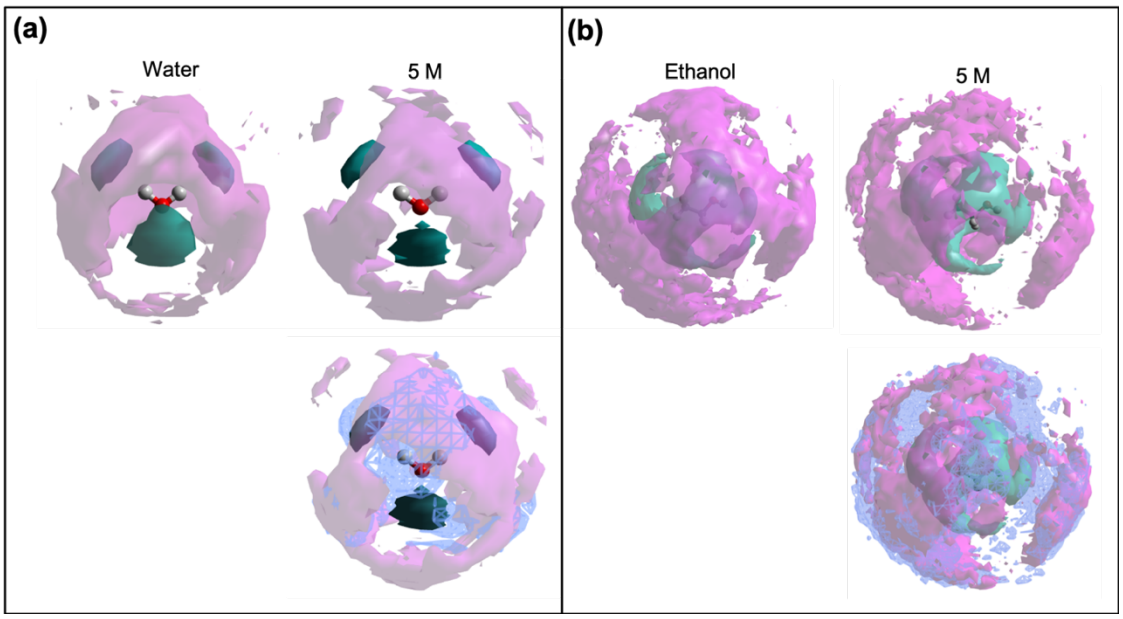

Figure 7.13. (Top) Spatial density function (SDF) of (a) water around water in the first hydration shell (solid green surfaces) for distance up to ~ 3.6 Å (30 % of water molecules shown) and second hydration shell (transparent pink surfaces) for distance up to ~ 5.8 Å (16 % of water molecules shown) (b) ethanol around ethanol in the first hydration shell (solid green surfaces) for distance up to ~ 6.7 Å (30 % of ethanol molecules shown) and second hydration shell (transparent pink surfaces) for distance up to ~ 10.8 Å (15 % of ethanol molecules shown). (Bottom) SDF of imidazole (mesh blue surface) around (a) water for distance up to 6 Å (16 % of imidazole molecules shown) (b) ethanol for distance up to 10.9 Å (15 % of imidazole molecules shown).

Table 7.6. Coordination numbers of atomic pairs in the solvent (i.e. water or ethanol)

	Central atom-surrounding atom	Sample	Coordination number	Cut-off / Å
Water	Ow - Ow	Water	5.8	3.4
		5 M	3.8	
	Ow – Hw	Water	1.8	2.4
		5 M	1.5	
	Water (cog) – water (cog)	Water	6.3, 26	3.6, 5,8
		5 M	4.7, 19	
Ethanol	OE1 – OE1	Ethanol	2.0	3.7
		5 M	1.2	
	OE1 – HE3	Ethanol	1.0	2.7
		5 M	0.6	
	Ethanol (cog) – ethanol (cog)	Ethanol	12.8, 54	6.8, 10.8
		5 M	9.3, 39	

Imidazole – imidazole interactions

Colinear N1-H1…N3 bonds in the 5 M ethanolic solution of imidazole are inferred from the 1 Å difference between the first peak in the N1…N3 RDF and the N3…H1 RDF (Figure 7.14(a)). However, similar to the 5 M aqueous imidazole solution, the probability of finding such species is less than 1 as seen from the coordination numbers of H1 around N3 and N3 around N1 at distances of 3.1 Å and 3.7 Å respectively (Table 7.7). The second peak in N1…N3 RDF centred at ~ 5.1 Å (Figure 7.14 (a)) and the second peak in the cog RDF centred at 9 Å (Figure 7.14 (b)) are more defined for the ethanolic solution of imidazole than for the aqueous solution, which could indicate directional imidazole interactions or clustering of imidazole molecules. This is further illustrated by the spatial density functions (SDFs) (Figure 7.15), which shows that the probability of finding imidazole molecules coordinating a central molecule is higher for the ethanol solutions than for aqueous solutions of imidazole (Figure 7.15). Adding to that, the predicted imidazole clustering in the ethanolic solution of imidazole seems to contain both imidazole and ethanol molecules unlike the clustering in aqueous solutions which is always mediated by water molecules (Figure 7.15). This direct imidazole-imidazole interaction could explain the resemblance between the N K-edge XRS spectra of the ethanolic solution of imidazole and solid imidazole presented in section 7.3.1.2.

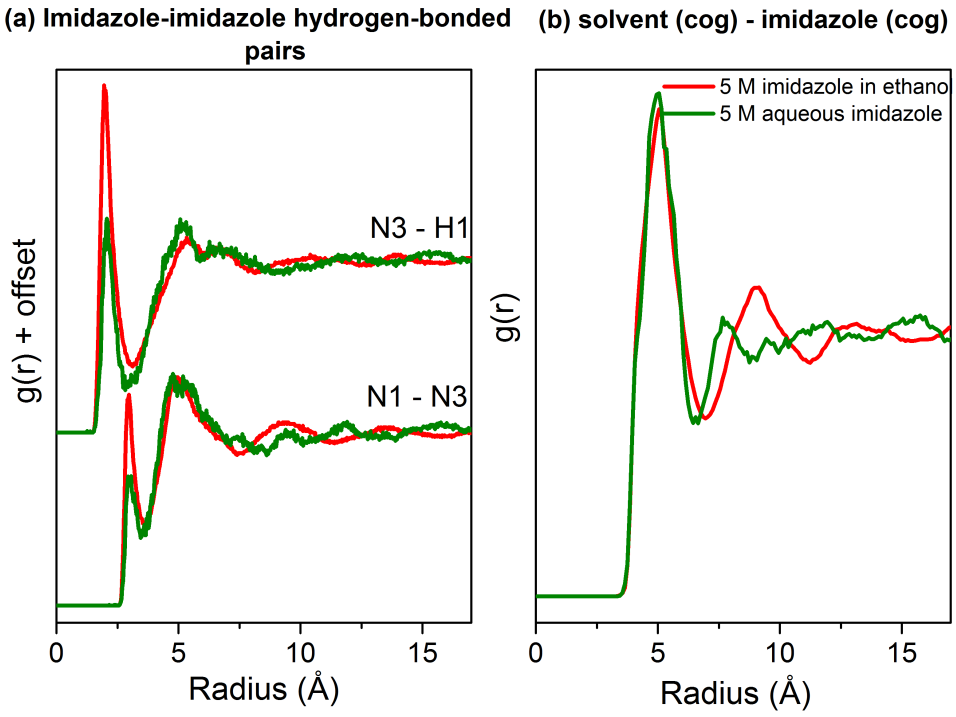

Figure 7.14. (a) The radial distribution functions (RDFs) of hydrogen-bonded pairs in imidazole (b) The radial distribution function (RDF) of the centre of geometry (cog) of imidazole around the cog of imidazole.

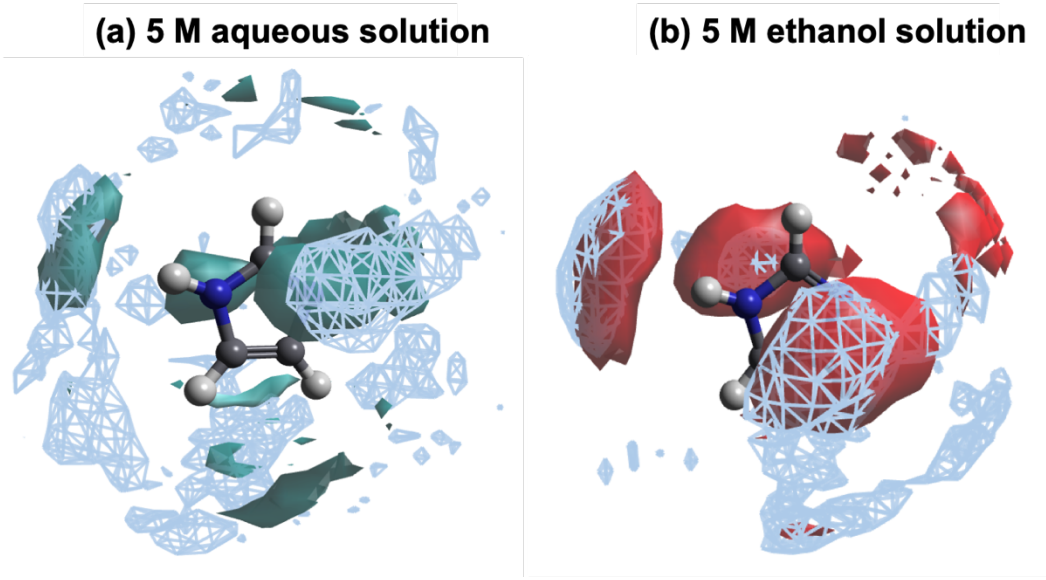

Figure 7.15. Spatial density function (SDF) of (a) water (solid teal surfaces) for distances up to 6 Å (b) ethanol (solid red surfaces) for distances up to 6.65 Å and imidazole (blue mesh) around imidazole for distances up to 7 Å. 10 % of water, ethanol and imidazole molecules are shown.

Table 7.7. Coordination numbers of hydrogen-bonded imidazole pairs

Central atom-surrounding atom	Sample	Coordination number	Cut-off / Å
N1 – N3	5 M aqueous solution	0.16	3.5
	5 M ethanol solution	0.27	3.7
N3 – H1	5 M aqueous solution	0.15	2.9
	5 M ethanol solution	0.27	3.1
Imidazole (cog) – Imidazole (cog)	5 M aqueous solution	4.1	7
	5 M ethanol solution	4.1	7

7.3.2 The Solvation of Ethanolic Solutions of Imidazole as a Function of Solution Concentration

In this section a series of ethanolic solutions of imidazole with concentration between 1 M and 8 M are compared to understand the effect of concentration (i.e. the supersaturation of the solution) on the intermolecular interactions.

7.3.2.1 Fourier Transform Infrared (FTIR)

The intensity of the vibrational bands corresponding to imidazole are expected to increase as a result of increasing the concentration of ethanolic solutions of imidazole, which is evident from Figures 7.16 (b), (c) and (d). The vibrational bands associated with imidazole are well represented in the concentration series (1 M – 8 M) shown in Figure 7.16. No significant shifts are observed in the frequencies of these features as a result of increasing the concentration. This is in line with the results of the FTIR study of aqueous imidazole solutions with concentration between 0.5 M – 11 M (presented in Chapter 5).

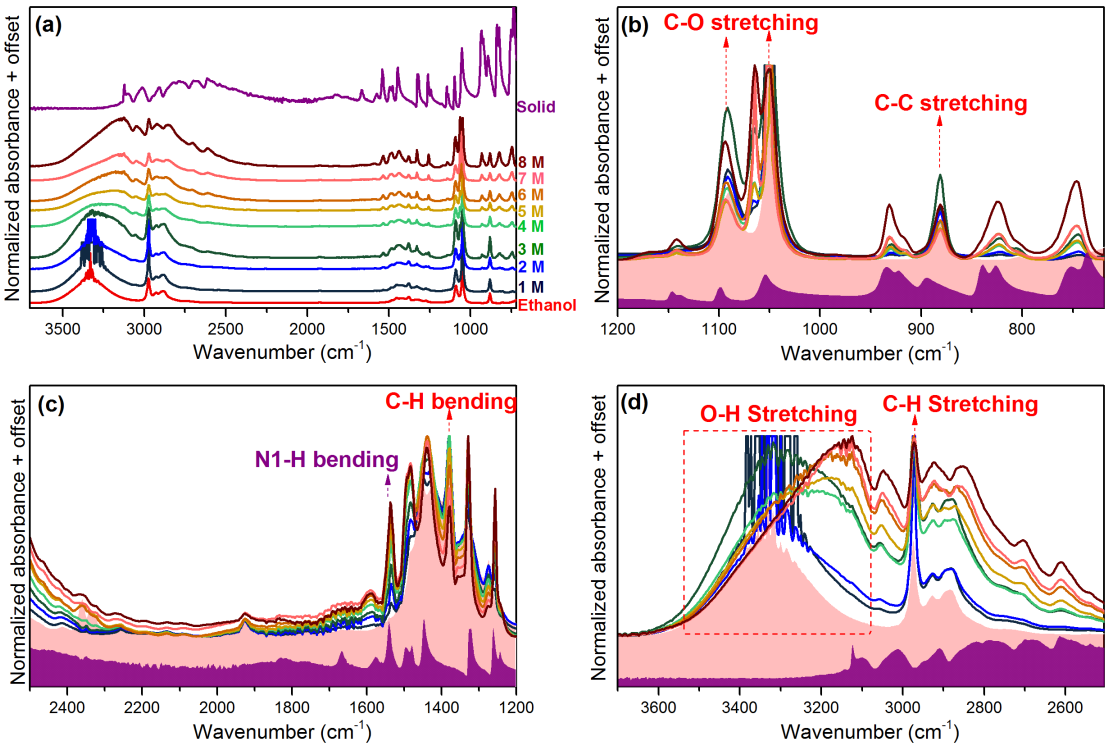

Figure 7.16. FTIR spectra of ethanolic solutions of imidazole with concentration between 1 M and 8 M overlaid with those of ethanol (pink) and solid imidazole (purple). The data was collected using ATR mode for the solid in transmission mode for the liquid samples. (a) FTIR spectra in range between 3700 –719 cm^{-1} (b) Imidazole Fingerprint region and the C-C and C-O stretching bands for ethanol (1200 – 719 cm^{-1}) (c) The C-H bend band for ethanol and N1-H bend band for imidazole (2500 – 1200 cm^{-1}) (d) The C-H and O-H stretching bands for ethanol (4000 – 2500 cm^{-1}).

The vibrational bands associated with ethanol show a decrease in intensity and shifts in vibrational frequencies as a consequence for increasing the concentration of imidazole in the ethanolic solutions. The vibrational frequency of the O-H stretching band shifts to a lower wavenumber (red shift) (Figure 7.15(d)). These shifts indicate that the intramolecular O-H bond in ethanol becomes weaker as a result of increasing the concentration of imidazole in the ethanolic solutions and it could be attributed to stronger hydrogen-bonding involving the O-H bond. Since the number of ethanol molecules available to form hydrogen-bonds with other ethanol molecules is not sufficient upon increasing the concentration of imidazole in solution, imidazole is expected to replace ethanol as hydrogen-bonding between imidazole and ethanol pairs becomes dominant at higher concentrations.

7.3.2.2 X-ray Pair Distribution Function (XPDF)

The X-ray pair distribution function (XPDF) patterns of 2 M – 8 M ethanolic solutions of imidazole and ethanol were generated through Fourier transformation of the total X-ray scattering patterns (Figure 7.17 (a) and (b)). The XPDF features attributed to imidazole are evident in the probed concentration series of ethanolic solutions of imidazole and their intensity increases as the concentration of imidazole is increased in the solution. For example, the first peak in the XPDF patterns, which is associated with the intramolecular bond lengths between bonded atoms in imidazole shifts to lower radial value to coincide with the first XPDF peak of solid imidazole. Furthermore, the second peak, which is attributed to intramolecular interactions between non-bonded atoms in imidazole, is evident even in the lowest concentration solution (i.e. the 2 M solution) and its intensity increases as the concentration of imidazole is increased in the solution. Other Features centred at 5 Å and 5.9 Å, which could be attributed to intermolecular imidazole-imidazole interactions, are also observed in the XPDF patterns of the ethanolic solutions of imidazole with concentrations ranging between 2 M and 8 M.

The similarity index between the XPDF patterns of the ethanolic solutions of imidazole and the XPDF pattern of solid imidazole increases as a function of increasing imidazole concentration in the 0 Å – 10 Å radial range (Figure 7.17 (c)). The correlation between the XPDF pattern of the 2 M, 3 M and 4 M solutions and the XPDF pattern of solid imidazole is 0.81 ± 0.06 in the 0 Å – 10 Å radial range. The correlation value increases to 0.9 for the 5 M and 6 M solution and to 0.93 for the 7 M and 8 M solutions. On the other hand, the XPDF patterns of the ethanolic solutions of imidazole are highly similar to the XPDF pattern of ethanol in the 10 Å – 50 Å radial range as indicated by correlation matrix shown in Figure 7.16 (d). This indicates that (1) the ethanol – ethanol interactions are dominant even at high imidazole concentration or (2) the imidazole – ethanol interactions are similar to those of ethanol – ethanol interactions.

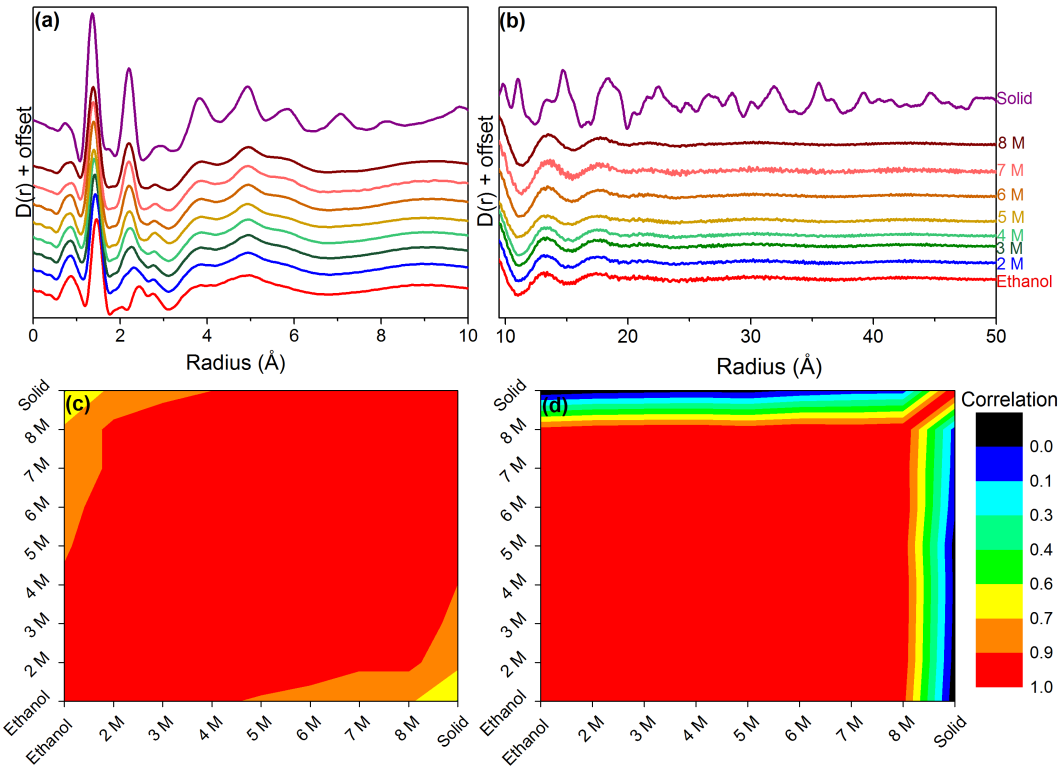

Figure 7.17. (Top) XPDF patterns of solid imidazole, ethanolic solutions of imidazole(concentration range 2 M – 8M) and ethanol in the range (a) 0 -10 Å and (b) 10 – 50 Å. (Bottom) Correlation matrix of XPDF patterns of solid imidazole, ethanolic solutions of imidazole(concentration range 2 M – 8M) and ethanol in the range (c) 0 -10 Å and (d) 10 – 50 Å.

Quantitative analysis of the atomic correlations in the ethanolic solutions of imidazole in the concentration range between 2 M – 8 M from the XPDF patterns shown in Figure 7.17 is not feasible. The reported XPDF patterns represent the sum of 91 partial PDFs, which are attributed to all the atomic pair-pair correlations present in the ethanolic imidazole solutions. Therefore, empirical potential structure refinement (EPSR) simulations were utilised to generate structural models for the probed ethanolic imidazole solutions and for ethanol. These models were refined utilising the same total X-ray scattering data, which were used to generate the XPDF patterns shown in Figure 7.17. The structure factor F(Q) patterns calculated from the EPSR simulation compare well with those obtained from the experimental scattering data (Figure 7.18).

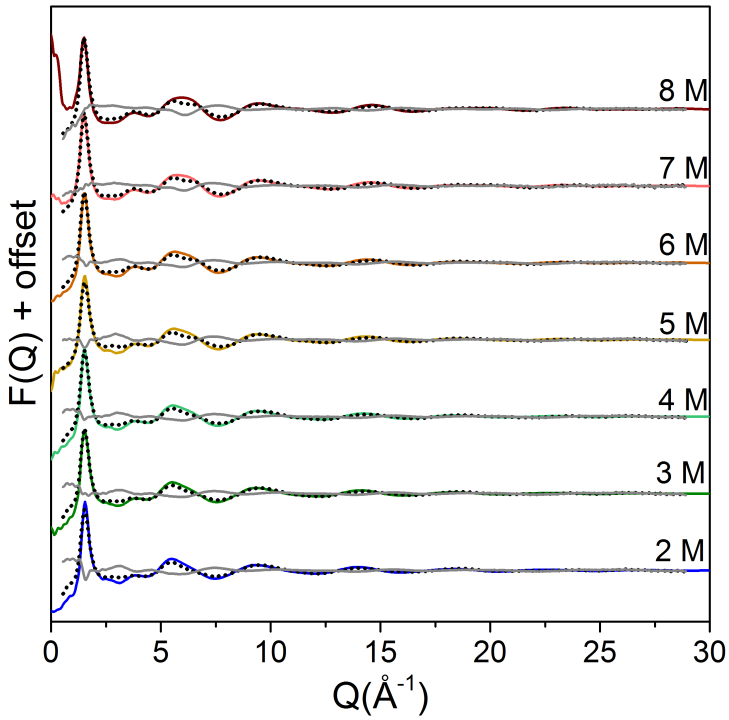

Figure 7.18. F(Q) patterns from experimental total X-ray scattering data (black dots) and EPSR models (coloured lines) of ethanolic solutions of imidazole (concentration range between 2 M – 8 M). The difference patterns (grey lines) illustrate the high comparability between the experimental and calculated F(Q) patterns.

Imidazole – ethanol interactions

Increasing the concentration of imidazole in the ethanolic solutions lowers the width of and the area under the peaks in the RDFs of hydrogen-bonded imidazole – ethanol pairs (Figure 7.19 (a)). This correlates with a reduction in the number of imidazole – ethanol bonded pairs as illustrated by the coordination number of OE1 and HE3 around the nitrogen atoms in the imidazole ring (Table 7.8). On the other hand, the effect of increasing the imidazole concentration on the positions of the RDFs peaks (Figure 7.19) is minimal; suggesting that the nature of the bonding environment between solute and solvent molecules in the ethanolic solutions does not change. This is clearly pictured in the spatial density function of ethanol around imidazole (Figure 7.20).

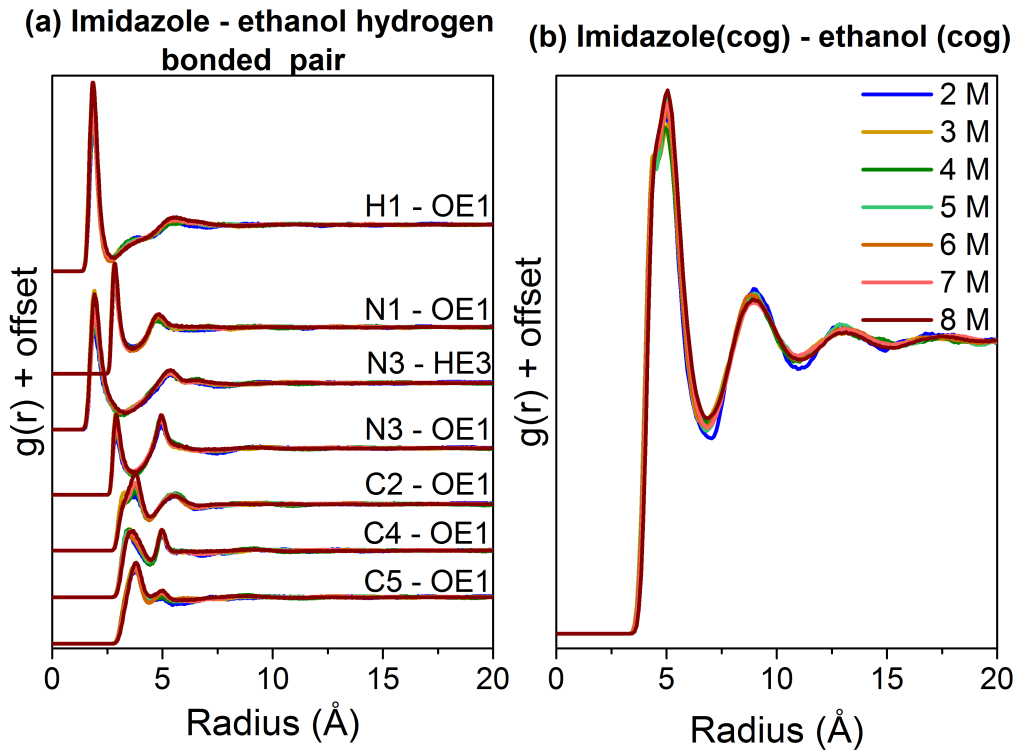

Figure 7.19. (a) The radial distribution functions (RDFs) of imidazole-ethanol hydrogen-bonded pairs (b) The radial distribution function (RDF) of the centre of geometry (cog) of ethanol around the cog of imidazole.

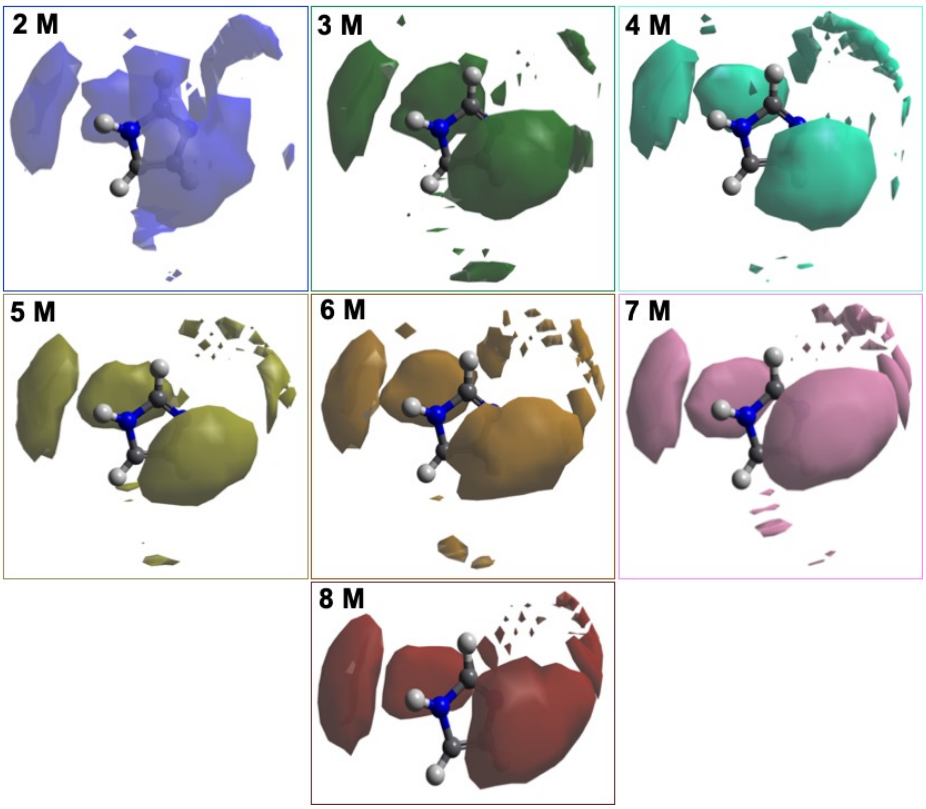

Figure 7.20. Spatial density function (SDF) of ethanol around imidazole for distances up to ~ 7 Å (10 % of ethanol molecules shown).

Table 7.8. Coordination numbers of hydrogen-bonded imidazole – solvent pairs

Central atom-surrounding atom	cut-off /Å	2 M	3 M	4 M	5 M	6 M	7 M	8 M
H1-OE1	2.7	0.8	0.7	0.7	0.7	0.6	0.4	0.3
N1-OE1	3.6	1.2	1.1	1.0	1	0.8	0.6	0.4
N3-HE3	3.3	1.0	0.9	0.9	0.9	0.8	0.6	0.3
N3-OE1	3.7	0.9	0.9	0.8	0.9	0.7	0.5	0.3
Ethanol (cog) – imidazole (cog)	6.7	11.6	10.8	9.9	9	8.9	6	3.4

Ethanol – ethanol interactions

The intensities and positions of peaks in the radial distribution functions (RDFs) (both pair-pair and cog-cog) of ethanol are similar for the series of ethanolic solutions of imidazole with concentrations ranging between 2 M and 8 M (Figure 7.21). The width of the peaks in the above-mentioned RDFs reduces with increasing imidazole concentration (Figure 7.21). This reduction correlates with a reduction in the area under the peaks which indicates a reduction in the number of ethanol molecules coordinating other ethanol molecules in the named solutions (Table 7.9). The reduction in the number of ethanol molecules around a central ethanol molecule as a function of increasing the concentration of imidazole in ethanolic solutions is illustrated in the SDFs (Figure 7.22).

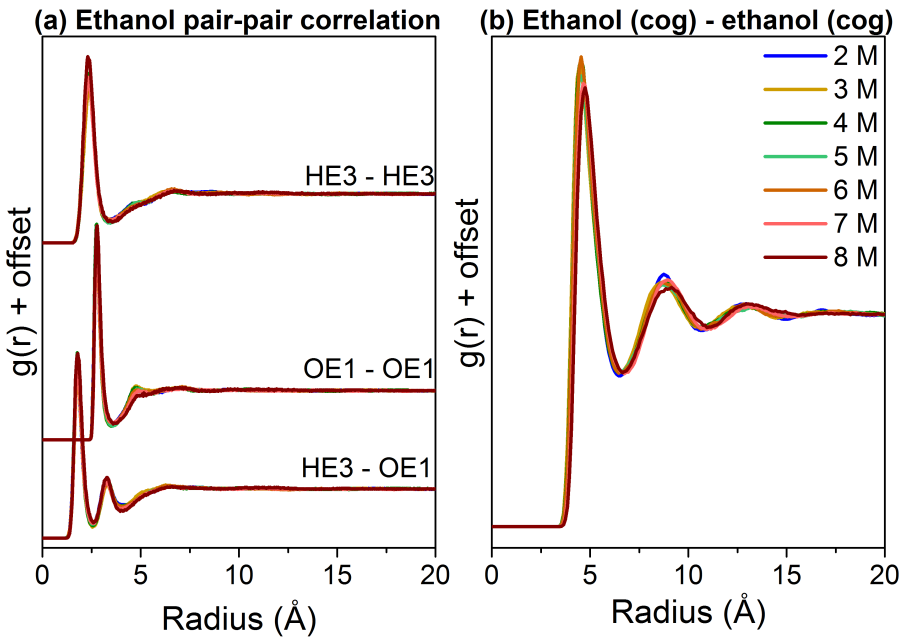

Figure 7.21. (a) The radial distribution functions (RDFs) of the atomic pair-pair correlations in ethanol (b) The radial distribution function (RDF) of the centre of geometry (cog) of ethanol around cog of ethanol.

Table 7.9. Coordination numbers of atomic pairs in ethanol

Central atom-surrounding atom	cut-off /Å	2 M	3 M	4 M	5 M	6 M	7 M	8 M
OE1 – HE3	2.7	0.8	0.8	0.7	0.6	0.6	0.4	0.3
OE1 – OE1	3.7	1.8	1.6	1.5	1.2	1.2	0.9	0.5
Ethanol (cog) – ethanol (cog)	6.8 10.8	12.2 51.3	11.0 46.4	10.0 42.0	9.3 39	8.4 34.9	6 25.4	3.3 14.1

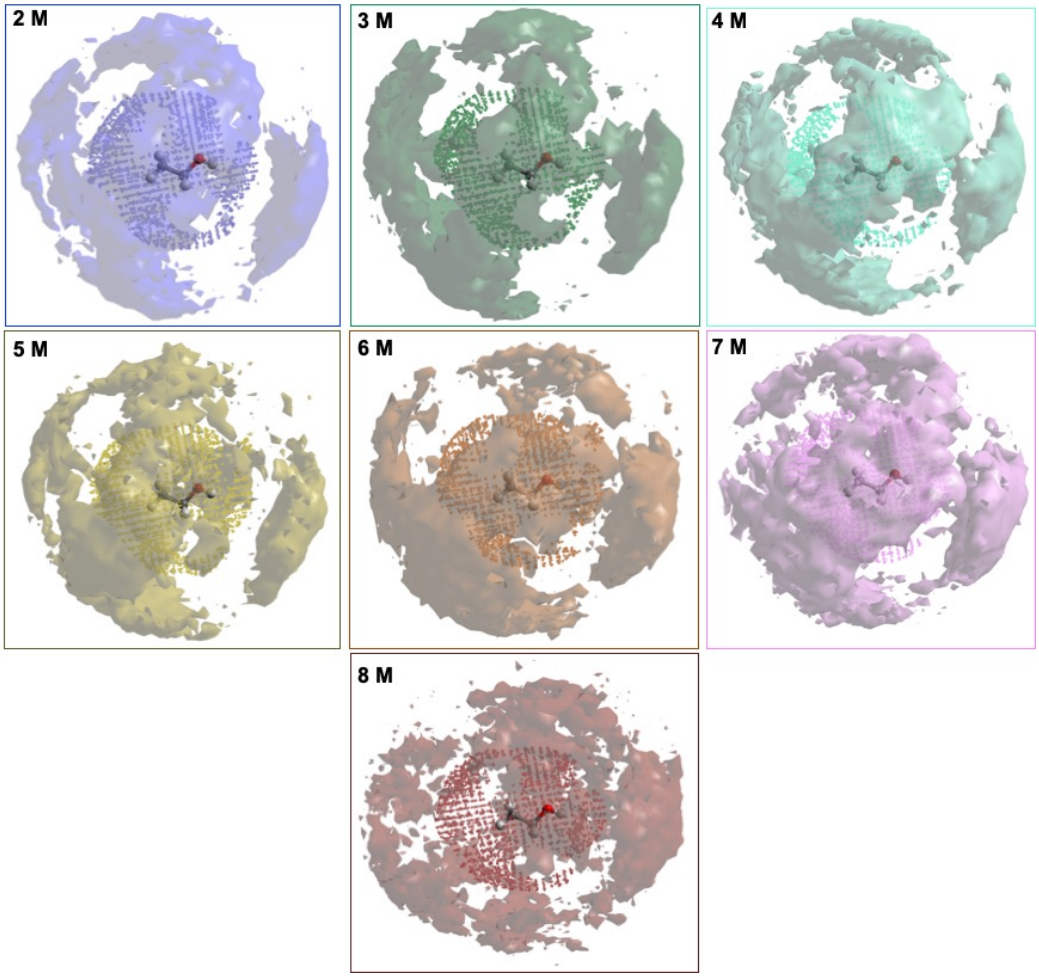

Figure 7.22. Spatial density function (SDF) of ethanol around ethanol for distances up to: ~ 6.8 Å where 30% of the ethanol molecules are shown (dots) and ~ 10.8 Å where 15 % of the ethanol molecules are shown (solid surface)

Imidazole – Imidazole interactions

Linear N1-H1…N3 bonds in the ethanolic solutions of imidazole are inferred from the 1 Å difference between the first peak in the N1…N3 RDF and the N3…H1 RDF (Figure 7.23(a)). However, the probability of finding such species is less than 1 as seen from the coordination numbers of H1 around N3 and N3 around N1 at 2.9 Å and 3.5 Å respectively (Table 7.10). The probability of finding imidazole molecules coordinating a central

molecule (regardless of the direction) increases with increasing the concentration of imidazole in ethanolic solutions (Table 7.10). This is further illustrated by the SDF of the imidazole around imidazole (Figure 7.24), which shows the presence of imidazole molecules coordinating the N1-H side of the imidazole ring as well as above and below the plane of the imidazole ring.

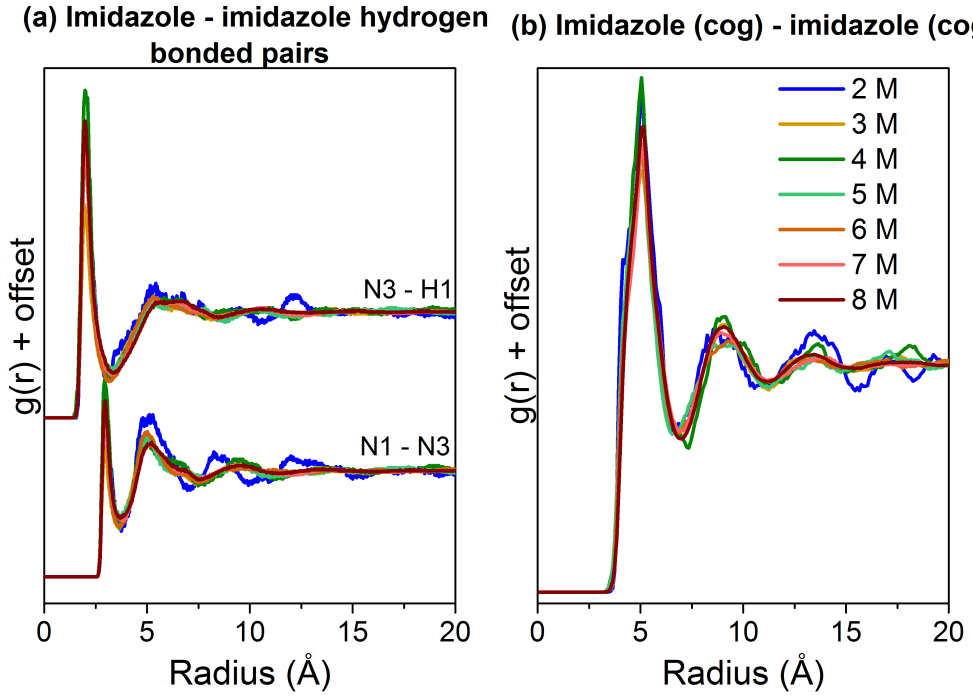

Figure 7.23. (a) The radial distribution functions (RDFs) of hydrogen-bonded pairs in imidazole (b) The radial distribution function (RDF) of the centre of geometry (cog) of imidazole around the cog of imidazole.

Table 7.10. Coordination numbers of hydrogen-bonded imidazole pairs

Central atom-surrounding atom	cut-off /Å	2 M	3 M	4 M	5 M	6 M	7 M	8 M
N3 – H1	3.1	0.1	0.2	0.3	0.3	0.4	0.5	0.6
N1 – N3	3.7	0.1	0.2	0.3	0.3	0.4	0.5	0.6
Imidazole (cog) – imidazole (cog)	7	1.8	2.6	3.3	4.1	5.0	5.7	6.7

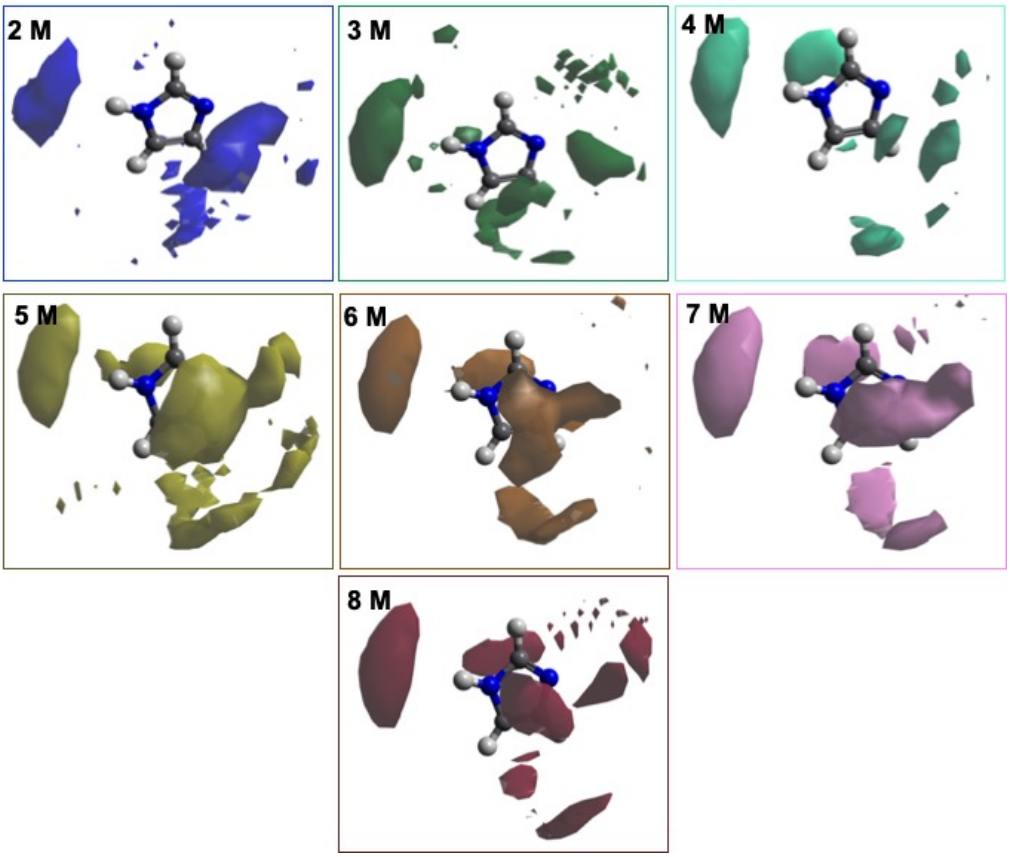

Figure 7.24. Spatial density function (SDF) of imidazole around imidazole for distances up to: ~7 Å where 30% of the ethanol molecules are shown (dots) and ~10.8 where 15 % of the ethanol molecules are shown (solid surface).

7.3.3 Structural Speciation of Ethanolic Solutions of Imidazole during Cooling Crystallisation

Total X-ray scattering data was collected for a 7 M ethanolic solution of imidazole during cooling crystallisation until the onset of crystal formation. The experiment was repeated two times and in the second time total X-ray scattering data was collected during crystallisation and dissolution. The X-ray pair distribution function (XPDF) patterns of the 7 M solution was generated through Fourier transformation of the Total X-ray scattering patterns. These are presented in Figures (7.25, and 7.29).

7.3.3.1 First Cooling Crystallisation Experiment

Total X-ray scattering data was collected for the 7 M ethanolic solution of imidazole as it was cooled from 25 °C to 13.5 °C and the XPDF patterns from the experiment are shown in Figure 7.25. Upon crystallisation, the XPDF patterns develop features in the longer radial range (10 Å – 50 Å) (blue patterns (scans 126 – 130) in Figure 7.25 (c -2)) in

a similar manner to crystallisation of imidazole from aqueous solutions (chapter 6, section 6.3.2). This coincide with a loss of the laser signal (Figure 7.25 (e)) and an increase in temperature (Figure 7.25 (d)), which is expected since crystallisation is an exothermic process. The similarity index between the XPDF patterns of the solution after crystallisation (scans 126 – 130) is ~ 1 (Figure 7.26 (a) and (b)).

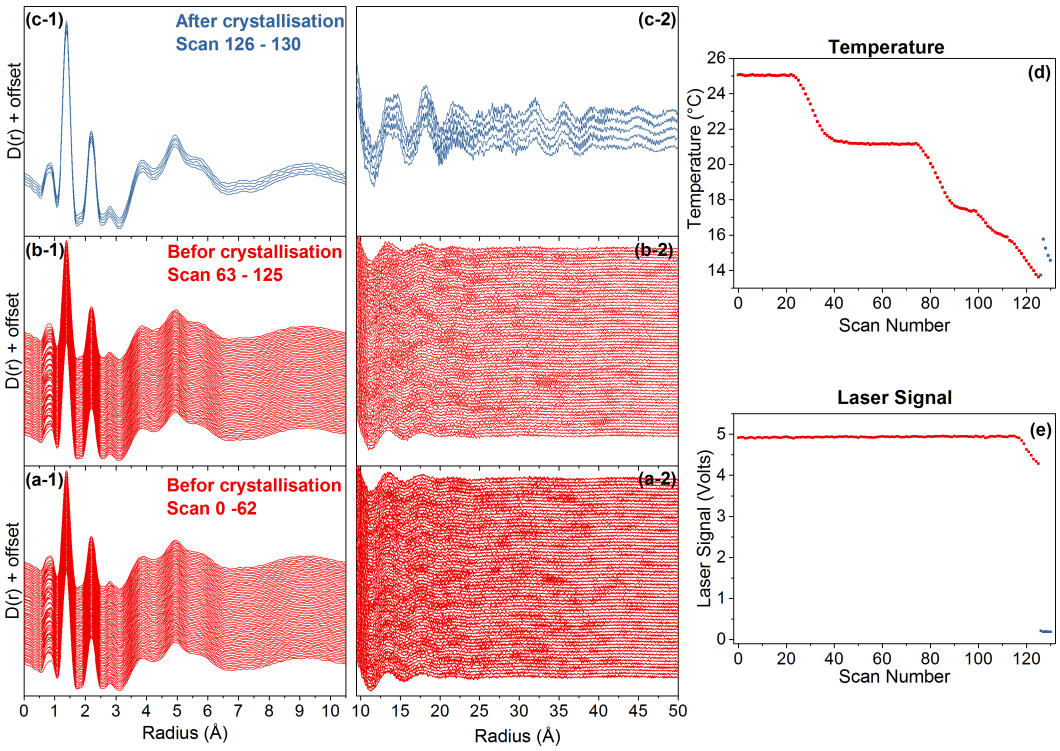

Figure 7.25. XPDF patterns of 7 M ethanolic solution of imidazole during the first cooling crystallisation experiment (a) and (b) before crystallisation and (c) after crystallisation. (d) The temperature of the solution and (e) The laser signal during the cooling crystallisation experiment.

The developed features in the XPDF patterns of the solution after crystallisation (in the 10 Å – 50 Å radial range) are highly similar to those in the XPDF pattern of solid imidazole as indicated by the high similarity index of ~0.9 (Figure 7. 26(d)). This is further illustrated by the contour maps of the XPDF patterns of the solution from the first cooling cycle shown in Figure 7.27 (c) and (d). The pink dotted lines highlight regions where the intensity of features in the XPDF patterns increase as a result of crystallisation while the red dotted lines highlight regions of decreased intensity. the features centred at 3.9 Å, 7.2 Å, 14.4 Å and 18.3 Å increase in intensity and they correspond to peaks in the XPDF pattern of solid imidazole while the features centred at 13.5 Å, 16.2 Å and 20.3 Å

decrease in intensity and they resemble troughs in the XPDF pattern of the solid (Figure 7.27).

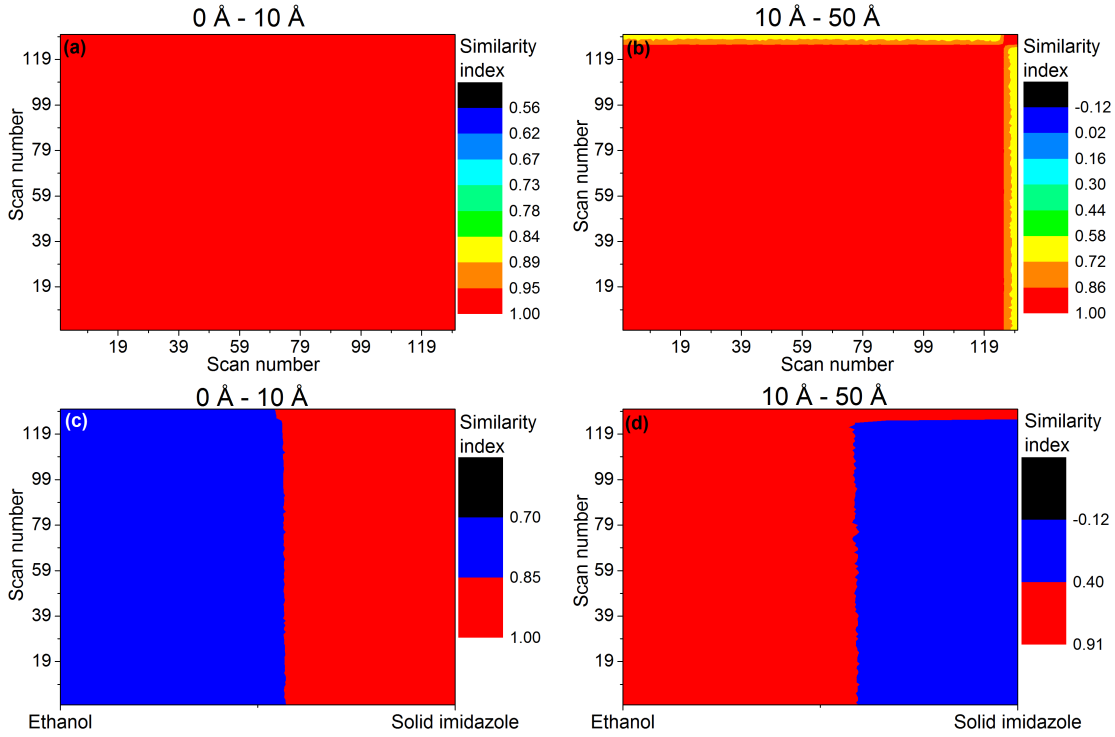

Figure 7.26. Correlation matrix of the XPDF patterns of the 7 M ethanolic solution of imidazole collected during the first cooling crystallisation experiment in the (a) 0 Å – 10 Å range and (b) 10 Å – 50 Å range. Correlation between the XPDF patterns of the ethanolic solution of imidazole and ethanol and solid imidazole in the (c) 0 Å – 10 Å range and (d) 10 Å – 50 Å range.

It is worth noting here that the feature at 11 Å in the XPDF pattern of solid imidazole is not observed in the XPDF pattern of the ethanolic solution of imidazole after crystallisation unlike what was observed for the aqueous imidazole solution (chapter 6, section 6.3.2). This could explain the slightly lower similarity index between the XPDF pattern of the ethanolic solution of imidazole after crystallisation and solid (~0.9) (Figure 7.26 (d)) when compared to the similarity index of the XPDF pattern of the aqueous imidazole solution after crystallisation and solid imidazole (~1) (chapter 6, section 6.3.2). The observation above could imply that the structure of the crystals forming from ethanol solutions differs from the structure of crystalline imidazole or crystals forming from aqueous solutions.

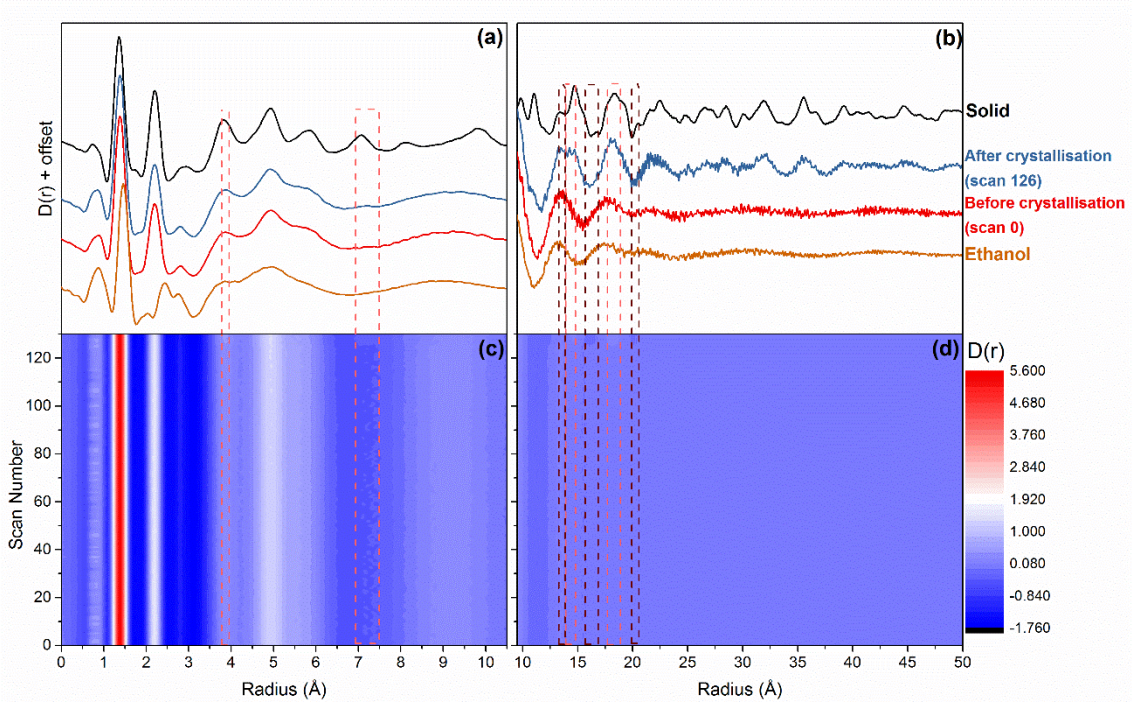

Figure 7.27. XPDF patterns of solid imidazole, ethanol and the 7 M ethanolic solution of imidazole before and after crystallisation from the first cooling cycle in the range: (a) 0 – 10 Å and (b) 10 – 50 Å. Contour plot of the XPDF patterns of the 7 M ethanolic solution of imidazole collected during the first cooling crystallisation experiment in the range: (c) 0 – 10 Å and (d) 10 – 50 Å.

Principal component analysis (PCA) was performed for the XPDF patterns of the 7 M ethanolic imidazole solution collected during the first cooling crystallisation experiment. Three principal components were identified from the analysis (PC1 – PC3) (Figure 7.28 (a)).

- PC1 has the highest eigenvalue and it represents the XPDF pattern of the solution before crystallisation (Figure 7.28 (b-1)). The contribution from the different XPDF patterns to this principal component is almost equal with an insignificant decrease (less than 0.001) from the XPDF patterns of the solution after crystallisation (Figure 7.28 (b-2)).
- PC2 is very noisy with some features corresponding to the solution after crystallisation at 18 Å, 32.5 Å and 36 Å (Figure 7.28 (c-1)). It is difficult to find a trend for the contribution to PC2 from the different XPDF patterns (Figure 7.28 (c-2)).
- PC3 has the lowest eigenvalue and it represents features in the XPDF pattern of solid imidazole (Figure 7.28 (d-1)). The contribution to PC3 from the XPDF

patterns of the solution before crystallisation is null and it increases upon crystallisation (Figure 7.28 (d-2)); indicating that the structural properties of the precipitated crystals highly resembles those of solid imidazole.

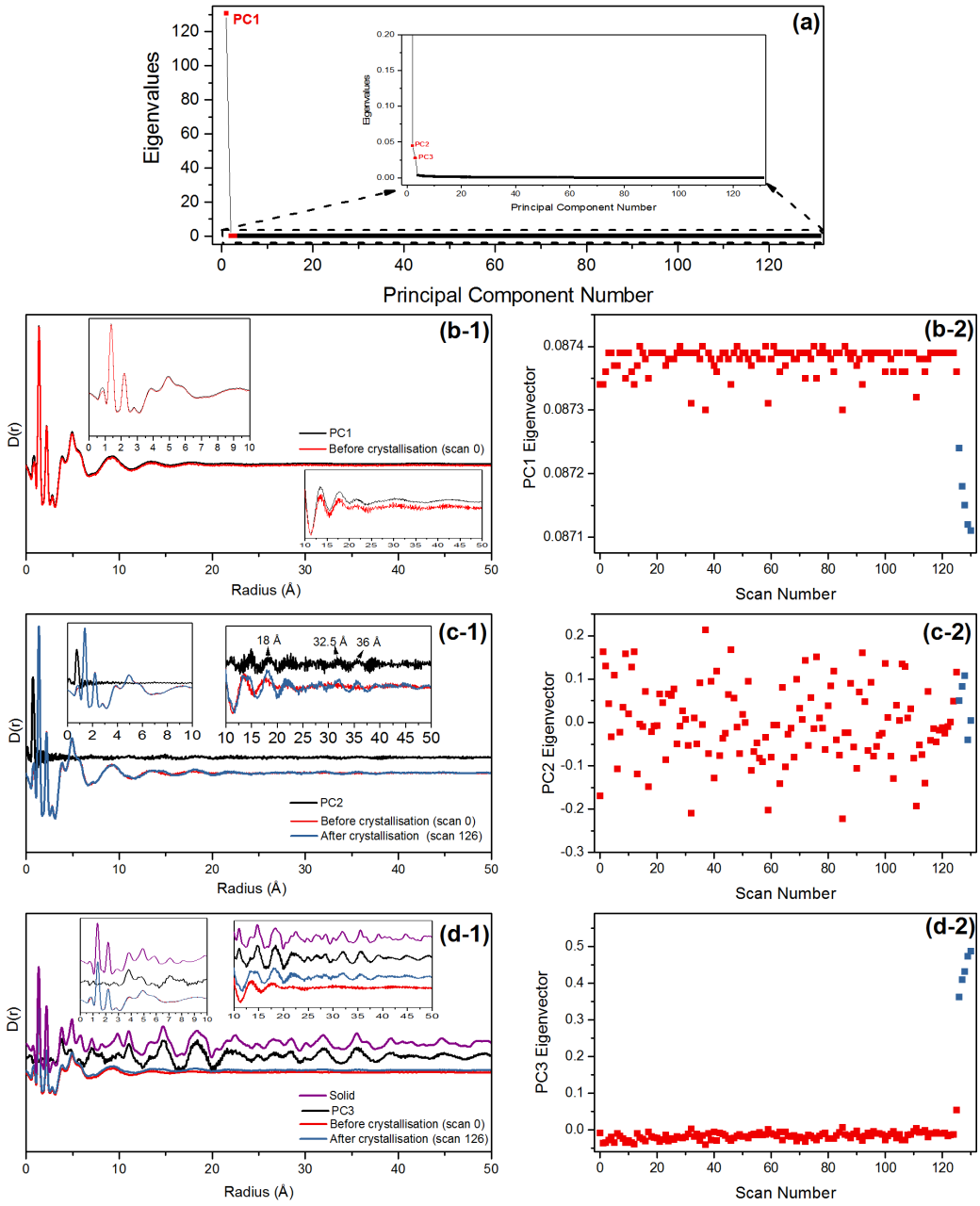

Figure 7.28. (a) Eigenvalues of the principal components from the PCA of the XPDF patterns of the 7 M ethanolic solution of imidazole from the first cooling crystallisation experiment. (b-1) (c-1) (d-1) comparison between the patterns of PC1 – PC3 and the XPDF patterns of solid imidazole, ethanolic solution of imidazole and ethanol (b-2) (c-2) (d-2) eigenvectors of PC1 – PC3.

7.3.3.2 Second Cooling Crystallisation Experiment

Similar to the XPDF patterns of the previous cycle, features in the longer radial range (10 Å – 50 Å) develop as a result of crystal formation (blue patterns in Figure 7.29 (b-2)). This coincides with an increase in temperature (Figure 7.29 (f)) and a loss in the laser signal (Figure 7.29 (g)). During dissolution (as the temperature of the solution is increased), the features in the longer radial range in the XPDF patterns (green patterns in Figure 7.29 (d-2)) disappear. The laser signal does not recover until the features in the longer radial range have completely disappeared (yellow patterns in Figure 7.29 (e-2)).

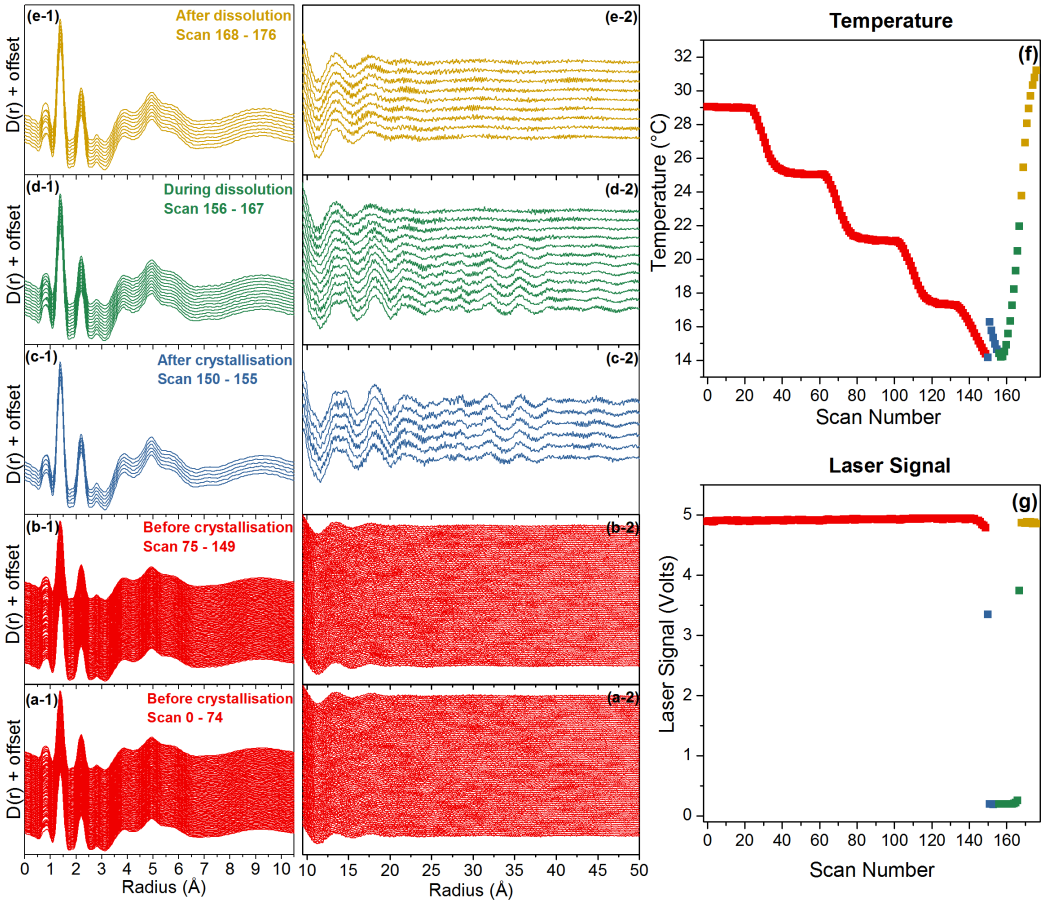

Figure 7.29. XPDF patterns of 7 M ethanolic solution of imidazole from the second cooling crystallisation experiment (a) and (b) before crystallisation (c) after crystallisation (d) during dissolution and (e) after dissolution. (f) The temperature of the solution and (g) The laser signal during the cooling crystallisation experiment.

The features in the 0 Å – 10 Å radial range in the XPDF pattern of the solution are highly similar throughout the crystallisation experiment as indicated by the correlation matrix in Figure 7.30 (a) and these features resemble those of solid imidazole (Figure 7.30 (c)).

This could be attributed to the fact that the most intense features in the 0 Å – 10 Å radial range correspond to intramolecular interactions in imidazole.

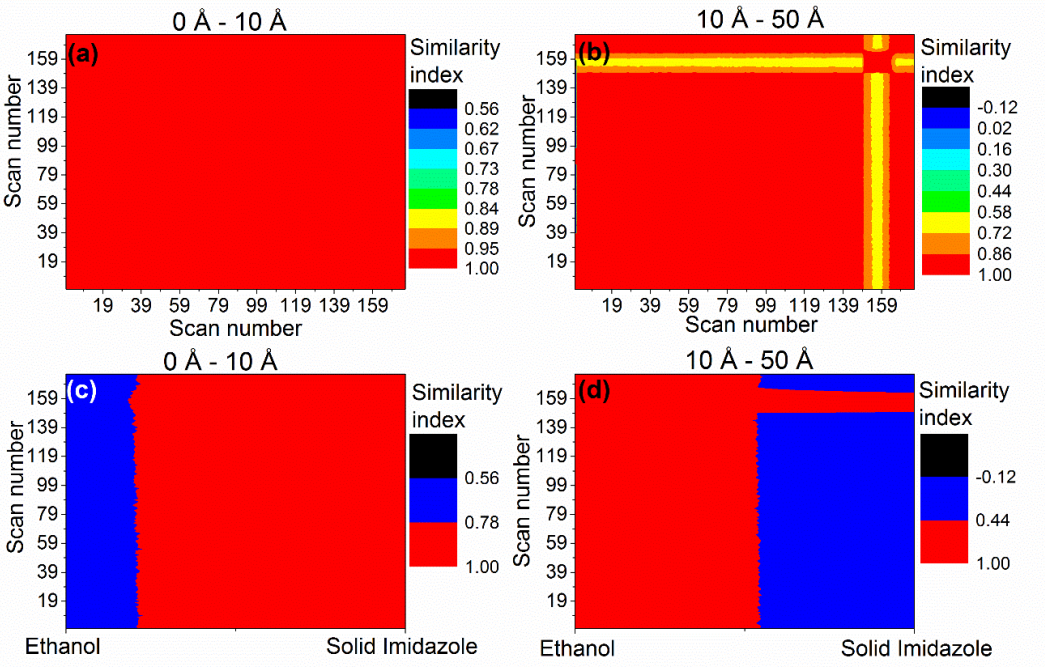

Figure 7.30. Correlation matrix of the XPDF patterns of the 7 M ethanolic solution of imidazole collected during the second cooling crystallisation experiment in the (a) 0 Å – 10 Å range and (b) 10 Å – 50 Å range. Correlation between the XPDF patterns of the ethanolic solution of imidazole and ethanol and solid imidazole in the (c) 0 Å – 10 Å range and (d) 10 Å – 50 Å range.

The features in the 10 Å – 50 Å radial range change during the cooling crystallisation experiment. They are comparable for the solution prior to crystallisation and after dissolution with a similarity index of ~1 (Figure 7.30 (b)) and resemble those of ethanol (Figure 7.30 (d)). This indicates that the structural properties of the solution prior to crystallisation and after dissolution are similar with no evidence for the presence of intermediate structures prior to crystallisation.

The XPDF features for the solution after crystallisation in the 10 Å – 50 Å radial range are highly comparable (Figure 7.30 (b)) and they resemble those of solid imidazole (Figure 7.30 (d)). The contour maps (Figure 7.31) highlight that XPDF features centred at 3.9 Å, 7.1 Å, 14.5 Å and 18.4 witness an increase in intensity as a result of crystallisation and they correspond to peaks in the XPDF pattern of solid imidazole. Similarly features with decreased intensity (centred at 12.5 Å, 16.3 Å and 20.1 Å) correspond to troughs in the XPDF pattern of the solid. This indicates that the structure of the precipitated crystals is very similar to that of solid imidazole.

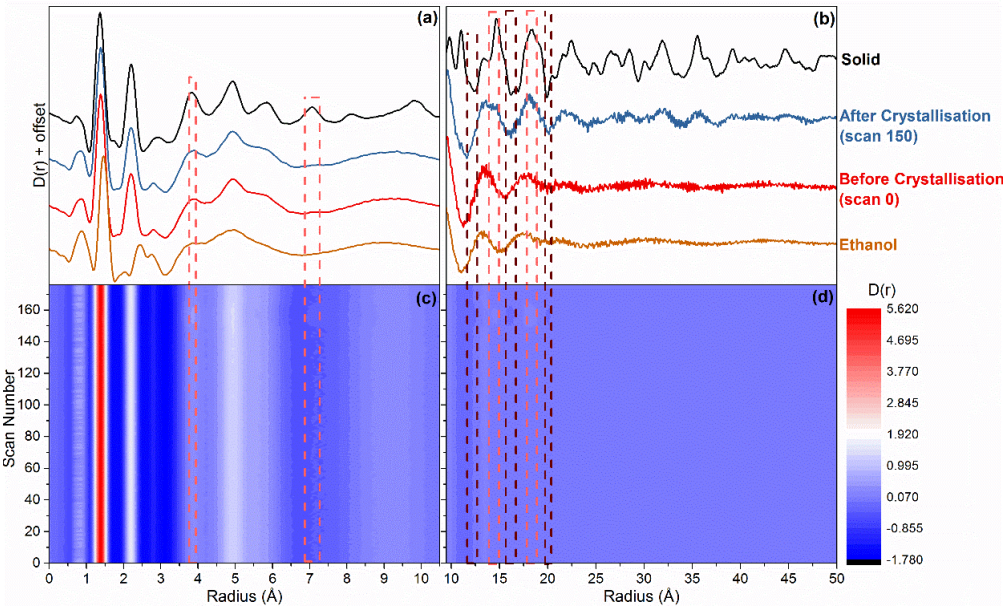

Figure 7.31. XPDF patterns of solid imidazole, ethanol and the 7 M ethanolic solution of imidazole before and after crystallisation from the second cooling cycle in the range: (a) 0 – 10 Å and (b) 10 – 50 Å. Contour plot of the XPDF patterns of the 7 M ethanolic solution of imidazole collected during the second cooling crystallisation experiment in the range: (c) 0 – 10 Å and (d) 10 – 50 Å.

Principal component analysis (PCA) was performed on the XPDF patterns collected during the second cooling crystallisation experiment. Three principal components were derived from the analysis (PC1 – PC3) (Figure 7.32 (a)).

- PC1 has the highest eigenvalue and it represents the XPDF pattern of the solution before crystallisation (Figure 7.32 (b.1)). The contributions from the different XPDF patterns to this principal component are almost equal with a slight reduction in contribution as a result of crystal formation (Figure 7.32 (b.2)). During dissolution, eigenvector values from the different XPDF to PC1 increases and continues to until these values are equal to those of the solution prior to crystallisation; this marks complete dissolution of the crystals (Figure 7.32 (b.2)).
- PC2 and PC3 represents the XPDF pattern of the solid in the 10 Å – 50 Å range with some features in the 0 Å – 10 Å range corresponding to the solid (Figure 7.32 (c.1) and (d.1)). The contributions to PC2 and PC3 increase after crystallisation (Figure 7.32 (c.2) and (d.2)). During dissolution, the contributions to PC2 and PC3 reduce until they are ~ 0 (Figure 7.32 (c.2) and (d.2)). This implies complete dissolution of the crystals.

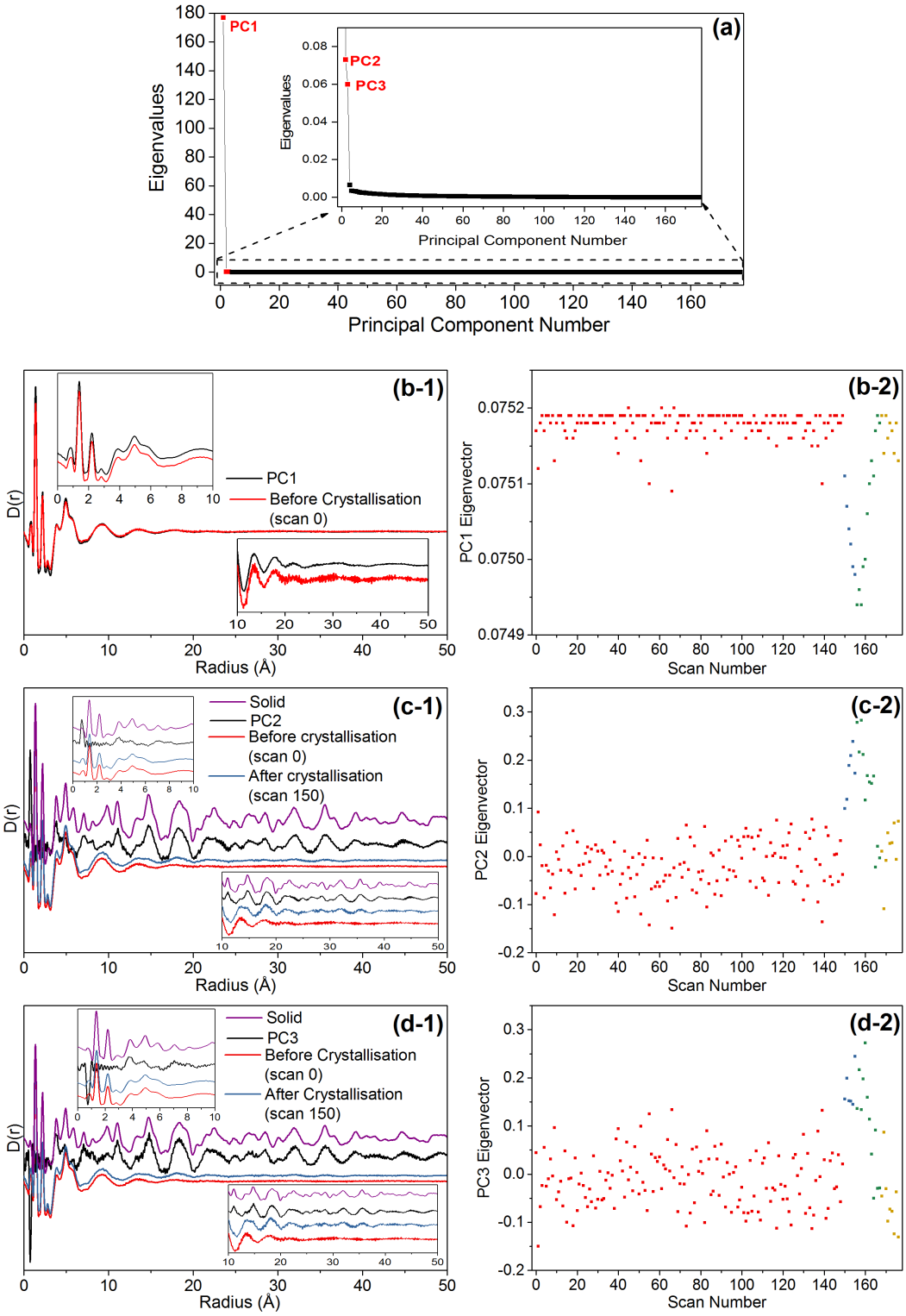

Figure 7.32. (a) Eigenvalues of the principal components from the PCA of the XPDF patterns of the 7 M ethanolic imidazole solution collected during the second cooling crystallisation experiment. (b-1) (c-1) (d-1) comparison between the patterns of PC1 – PC3 and the XPDF patterns of solid imidazole, ethanol and the ethanolic solution of imidazole before and after crystallisation. (b-2) (c-2) (d-2) eigenvectors of PC1 – PC3.

7.4 Conclusions

Direct imidazole-imidazole interactions are anticipated in ethanolic solutions of imidazole. These intermolecular interactions are different from the ones predicted for aqueous imidazole solutions. The FTIR and N K-edge spectra illustrated that the bonding environment in ethanolic solutions of imidazole is similar to that of the hydrogen-bonded solid than the aqueous solution. This is further illustrated by the RDF and the SDF of imidazole around imidazole in ethanolic solution, which indicated directional imidazole-imidazole interactions. The change in the solvation structure of ethanolic solutions of imidazole when compared to aqueous solutions can be attributed to the reduction in the number of hydrogen bonds formed by ethanol which is half of those formed by water. This reduction in the number of hydrogen bonds could enable direct imidazole – imidazole interactions.

FTIR spectroscopy and total X-ray scattering were utilised to study the effect of increasing the solute concentration on the solvation structure of ethanolic solutions of imidazole. The features in FTIR spectra and XPDF patterns of ethanolic solutions of imidazole were broadly similar across a range of concentrations, which indicates that the local solvation structure does not change massively in response to the concentration changes of this magnitude. The similarities in the local bonding environment around the imidazole molecules have been confirmed by the radial and spatial distribution functions (RDFs and SDFs) of ethanol around imidazole.

Finally, the structural evolution of ethanolic solutions of imidazole during cooling crystallisation were followed by *in-situ* X-ray pair distribution function (XPDF). Upon crystal formation, features developed in the longer radial range (10 Å – 50 Å) of the XPDF patterns of the solution. The intensity of the features in the XPDF pattern of the solution in the metastable zone are similar to those of the solution outside the metastable zone and after dissolution. This implies that the nucleation of imidazole from ethanol solutions is dominated by local hydrogen-bonding to ethanol in the inner solvation shell around the imidazole molecules. The breakage of these solute-solvent hydrogen bonds during desolvation is likely to be the key step in the nucleation of imidazole crystals from ethanol solutions.

Chapter 8 Additional Work - Automated X-ray Absorption Spectroscopy Analysis

8.1 Introduction

As mentioned through this dissertation, understanding the electronic structures and the intermolecular bonding in materials is crucial to predict their macroscopic structures and surface properties which will subsequently control their chemical interaction with the surrounding environment. X-ray absorption spectroscopy (XAS) and other core-level spectroscopic techniques, such as electron energy loss spectroscopy (EELS) and X-ray Raman scattering (XRS), are powerful and versatile techniques to probe the electronic structures of materials since they carry element specific information on the local geometrical structures and bonding environment. This enabled utilising core-level spectroscopy across a wide range of fields; including but not limited to materials research, electrochemistry, pharmacy, biological sciences, nuclear fuel, waste processing and heritage science.

The X-ray absorption spectroscopy is divided into two regions: near-edge X-ray absorption fine structure (NEXAFS) spectroscopy and extended X-ray absorption fine structure (EXAFS) spectroscopy. NEXAFS usually probes transitions from the deepest core shell of atomic species (the K-edge) to intra-molecular and extra-molecular neighbours [168]. Although NEXAFS spectra are generated conjunctively with EXAFS spectra, the region is less well understood compared to EXAFS and NEXAFS analysis is considered to be challenging and time consuming.

The analysis of NEXAFS has been mainly qualitative through fingerprinting, linear combination fitting (LCF) and principal component analysis (PCA) [270]. Fingerprinting is the most straightforward method in which experimental spectra are compared to calculated spectra (theoretical standard) or known spectra of the sample (empirical standard) to assign different features that appear in the experimental spectra [270]. For more accurate fingerprinting analysis, small changes in spectra and precise position of peaks should be defined which could be achieved through peak fitting. Peak fitting is a widely used method performed by approximating the edge-step as an arctangent function and the other spectral features as Gaussian, Lorentzian, Voigt (which is the

convolution of Gaussian with Lorentzian) or Pseudo-Voigt (which is the sum of Gaussian and Lorentzian) functions [164, 270]. The peak fitting process could be conducted using XAS specialised programs such as Athena [239] and WINXAS [271] or non XAS specialised programs such as Fityk [244]. Recently, an automated peak fitting algorithm was developed utilising MATLAB [272] as a first step towards automating the analysis of NEXAFS data.

Linear combination fitting (LCF) and principal component analysis (PCA) are used to qualitatively analyse the composition of mixtures in the probed sample. In LCF, a library of standards is assembled and a computer is used to sum them in various linear combinations reporting those that match the experimental data [164, 270]. If LCF fails then it suggests that there is a constituent not available in the library of standards [270]. This presents a weakness in the LCF method since a standard for every constituent in the sample is required. PCA estimates the number of distinct species in a series of spectra (determine the basis vectors) hence minimising the number of basis vectors that are required to generate the spectra adequately [164, 270]. PCA does not require any standards and it can define similar and different spectra. However, there is no guarantee that the generated basis vectors represent any spectral features of interest. PCA is usually used as a first stage in the analysis process for unknown systems to understand the intrinsic complexity of the spectrum and is usually followed by fingerprinting, LCF or fitting to calculated spectra.

So far, the quantitative analysis of NEXAFS data is performed in separate software packages and is not fully automated. It requires processing raw data through data reduction, normalisation and peak fitting followed by calculating theoretical NEXAFS spectra to be compared to the experimental spectra. Data reduction and peak fitting of the experimental spectra can be performed utilising Athena [239], WINXAS [271] and Fityk [244]. And as mentioned previously, a MATLAB algorithm was developed to automate the peak fitting stage [272]. Calculating theoretical NEXAFS spectra is currently conducted in a number of software packages utilising different levels of theory; including multiple scattering theory [273-275] and molecular orbital theory [168].

Multiple-scattering theory method starts by separating the transition potential to contributions from scattering potentials. The scattering potentials correspond to the paths the excited photoelectron takes as it scatters off of neighbouring atoms before scattering back to the absorbing atom [273]. NEXAFS photoelectron has low kinetic energy and long inelastic electron mean free path, which increases the probability of scattering off of each neighbouring atom [270, 276]. Thus, considering a few important multiple scattering paths is not sufficient for NEXAFS to converge (unlike EXAFS) and a full multiple scattering path method should be utilised to generate the theoretical NEXAFS spectra. A number of software packages allow modelling of NEXAFS using full multiple scattering theory; such as FDMNES [277], MXAN [278] and FitIT [279]; the latter is based on FEFF calculations [274]. These codes would compute the spectrum for one structure and then modify the structure and compute the spectrum again and repeat the process until the calculated spectrum is in good agreement with the experimental spectrum.

In molecular orbital theory, the positions of bonded electrons are approximated as linear combinations of atomic orbitals rather than assigning electrons to certain bonds between atoms [168]. X-rays would then interact with electrons in the core-level orbitals exciting them to the lowest unoccupied molecular orbitals (LUMO) or to the continuum. The theoretical NEXAFS spectra are generated by accurately calculating the ground states and final states of the investigated system. Then the transition energy is approximated, and the theoretical spectra is generated. Molecular orbital theory utilises Hartree-Fock (HF) theory, Density Functional Theory (DFT) and other empirical approximations. Examples of software packages utilising molecular orbital theory to quantitatively analyse NEXAFS spectra include (1) Gaussian, ORCA and StoBe, which are quantum chemistry software packages (2) WIEN2K which is a band structure code and (3) CASTEP and PARATEC, which are plane wave pseudopotential packages [274]. These packages are based on ground-state density functional theory and are well developed in treating the electronic structures in the ground-state. They can also perform calculation to determine low energy excited-state properties.

The molecular orbital theory will be better at interpreting NEXAFS spectra since they carry information on chemical bonding and local electronic structure. Adding to that, the molecular orbital theory avoids the poor muffin-tin approximations and the complex

non-muffin tin approximations used in the multiple scattering theory. However, the calculated NEXAFS peaks positions and amplitudes by the molecular orbital theory are usually shifted from the experimental values since this method does not account for self-energy effects [274] and corrections need to be applied.

Overall, the analysis of NEXAFS data is currently conducted in separate software packages and is challenging. Automating the analysis of NEXAFS data will save the researchers' time which will probably encourage the use of NEXAFS as a powerful spectroscopic technique. Moreover, automated analysis has also become a necessity because the current-trend is to develop high-throughput synchrotron radiation facilities [280-283] and bench-top machines [284], which will enable providing hundreds of NEXAFS spectra within minutes to hours. Hence, it will be useful to have access to an analysis code that permits initial assessment of such datasets during experiments. This chapter explores the possibilities of an automated XAS analysis code for the parameterisation of the spectral features, conducting excited-state calculations and assigning the experimental features based on the theoretical calculations.

8.2 Code Architecture

The automated XAS analysis code presented here is not a definite general analysis package but a start of automating the computational workflow for XAS analysis and it works on three axes:

1. The parameterisation of the NEXAFS region based on the determination of energies, intensities and widths of the most prominent features.
2. The calculation of the theoretical X-ray absorption (XA) spectra of the probed species.
3. The comparison of the experimental and theoretical XA spectra to assign the spectral features to electronic transitions.

The code was developed with Python (version 3.5.2) on a windows operating system using Anaconda [285] and it has been tested on Linux and Mac operating systems. Python was the chosen programming language because the code aims to integrate domain-specific knowledge on the analysis of X-ray absorption spectroscopy (XAS) data into the larger ecosystem of Python data analysis. Python programming language already has a rich suite of tools for data science, which is heavily used by the data science community. These libraries include NumPy and SciPy, which provide a suite of high-

performance numerical methods. The automated XAS analysis code aims to allow users to utilise these professional data science libraries for XAS analysis.

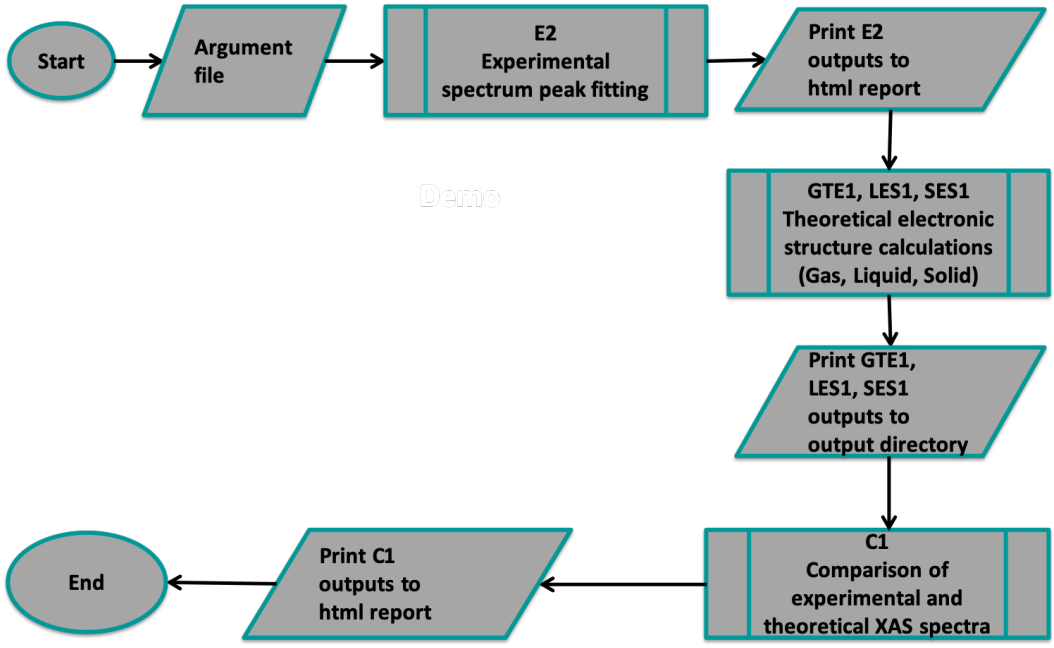

Figure 8.1. Algorithm of automated XAS analysis code

Table 8.1. Scripts of the automated XAS analysis code

Script name	Description
E1	Background noise subtraction and normalisation of the experimental spectra. This is still to be implemented as a deep learning algorithm.
E2	Experimental spectra peak fitting
GTE1	Theoretical electronic structure calculation for a gas
LES1	Theoretical excited state calculation for a liquid
SES1	Theoretical excited state calculation for a solid
C1	Compares experimental spectra with calculated spectra
GW1	Wrapper Script to run E2, GTE1 and C1 for a gas
LW1	Wrapper Script to run E2, LES1 and C1 for a liquid
SW1	Wrapper Script to run E2, SES1 and C1 for a solid

The main algorithm of the code (Figure 8.1) consists of three components which address the three axes mentioned above. The second component, which related to executing the calculations to generate the theoretical XA spectra, has different Python scripts for each state of matter (i.e. gas, liquid and solid). These scripts were named using the initials of the named state of matter (G for gas, L for liquid and S for solid). The other two components of the code, which perform the peak fitting of the experimental spectra and the comparison between the experimental and theoretical spectra, are the same

for the different states of matter. Wrapper scripts are used to run the calculations and generates the results. The names and descriptions of all the scripts are listed in Table 8.1.

8.3 Components of the Code

The parent directory of the code has sub-directories for each of the scripts listed in Table 8.1 alongside with the edge data file, which is a text file containing the natural line width table; obtained from reference [286]. The edge data file is stored in the parent directory of the code because it is used by more than one script. The main functions of the scripts of the automated XAS analysis code are described below.

8.3.1 Experimental Spectra Peak Fitting (E2)

The first script in the pipeline for automating the analysis of X-ray absorption spectroscopy (XAS) data focuses on parameterising the near-edge X-ray absorption fine structure (NEXAFS) region based on the determination of energies, intensities and widths of the most prominent features. Any absorption features below the ionisation potential (IP) were simulated with Gaussian functions [168]. This fitting procedure is based on findings that Gaussian functions fitted most reference spectra well [90, 217, 272]. A single arctangent step function was fitted to resemble the absorption edge (i.e. IP) and this was done to increase the reliability of the fits in the NEXAFS region and not to analyse or parameterise the ionisation potentials.

The peak fitting procedure was conducted with the aid of LmFit Python package [287], which is a high-level interface to curve fitting and non-linear optimisation. It builds on the optimisation methods of SciPy.Optimize by providing a number of useful enhancements; such as, using parameter objects as variables of the fitted functions instead of plain floats. Each parameter has a value that could be varied during the fit or fixed to a certain value and it can have an upper and lower bound. For E2 script, the parameters to be optimised by LmFit package were the number of Gaussian functions fitted, the amplitude (which represents the overall height or area of the function), the centre (which represents the peak centroid position) and sigma (which gives the characteristic width of the peak). Other peak parameters; such as, the full-width-at-half-maximum (FWHM) and the height are constrained by at least one of the parameters previously mentioned.

Energy boundaries and constraints were imposed to ensure that the automated fitting results retained physical meaning. In order to determine the centroid of arctan functions, an initial guess and boundary limits were obtained from tabulated natural line width data (i.e from the edge data file taken from reference [286]). It was assumed that any absorption maximum located at a lower energy than the step function modelling the IP must correspond to the excitation of a core electron to an unoccupied bound electronic state and these features were fitted with Gaussian functions. The fitted functions must be separated from each other by a threshold value to prevent overfitting; 0.8 eV gave good results. The amplitude and sigma of the IP arctan functions were initially fixed at 0.5, which gave reasonable results.

In the first optimisation loop, the number of peaks to be fitted and their centroids are guessed from the first derivative of the experimental spectrum. A peak centroid is then assigned to each maximum in the spectrum. The initial guess for the IP is the inflection point of the experimental spectrum (*E*0) + 6 eV. In the subsequent minimisation loop the difference between the experimental spectrum and the fitted line shapes is examined, and Gaussians are added where the deviation between experiment and fit is larger than the threshold value of 0.8 eV. The algorithm for E2 script is summarised in Figure 8.2. Finally, the fitting procedure is attempted with smooth and non-smooth data and the best fit is chosen based on the goodness of the fit value.

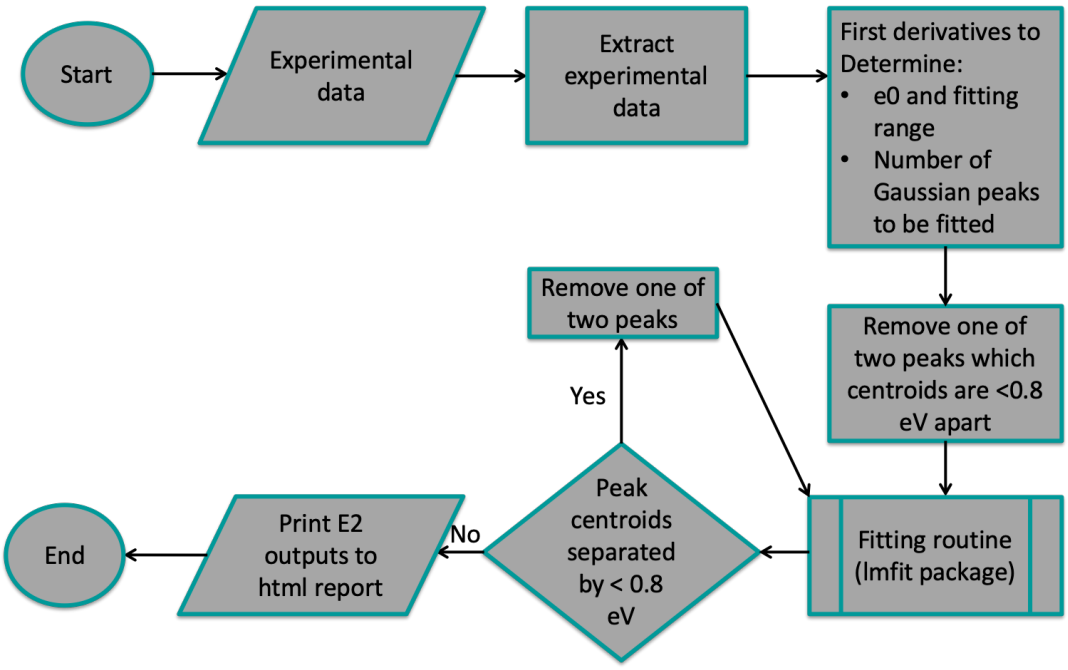

Figure 8.2. Algorithm of E2 script

8.3.2 Theoretical Calculations

As mentioned previously there are three scripts to perform the excited state calculations to generate the theoretical X-ray absorption (XA) spectra for each state of matter (i.e. gas, liquid and solid). The three scripts were developed to work with the software ORCA version 4.0.1.2.[208] The computational procedure is different for each script (which will be discussed further) and the success of the execution of these scripts relies on completing each step of the computational procedure successfully. A variety of computational strategies can be applied either in a pre-defined order within the code or through the user input. If the script terminates early without completing all the computational steps successfully, there will be an error message printed.

8.3.2.1 Theoretical Electronic Structure Calculations for Gas-Phase (GTE1)

GTE1 is the script used for calculating the theoretical XA spectra of monomeric molecules or samples in the gaseous phase. The default ORCA parameters used for the calculations are the B3LYP functional [25] and the 6-31G* basis set with TIGHTSCF and Grid3 for SCF iterations and Grid5 for the final energy. The computational algorithm (shown in Figure 8.3) runs in three steps:

- Geometry optimisation where the intramolecular structure of the studied molecule is optimised. The script will check the geometry optimisation output file to ensure that the calculation was successfully completed before running the next step.
- Frequency calculation is executed to ensure that the geometry optimised structure corresponds to the global energy minimum.
- Excited-stated calculation or TDDFT (time-dependent density functional theory) calculation is performed to generate the theoretical XA spectra.

If errors are encountered at any stage of the calculation, the script will terminate printing out an error message, which is also printed in the log file for the script.

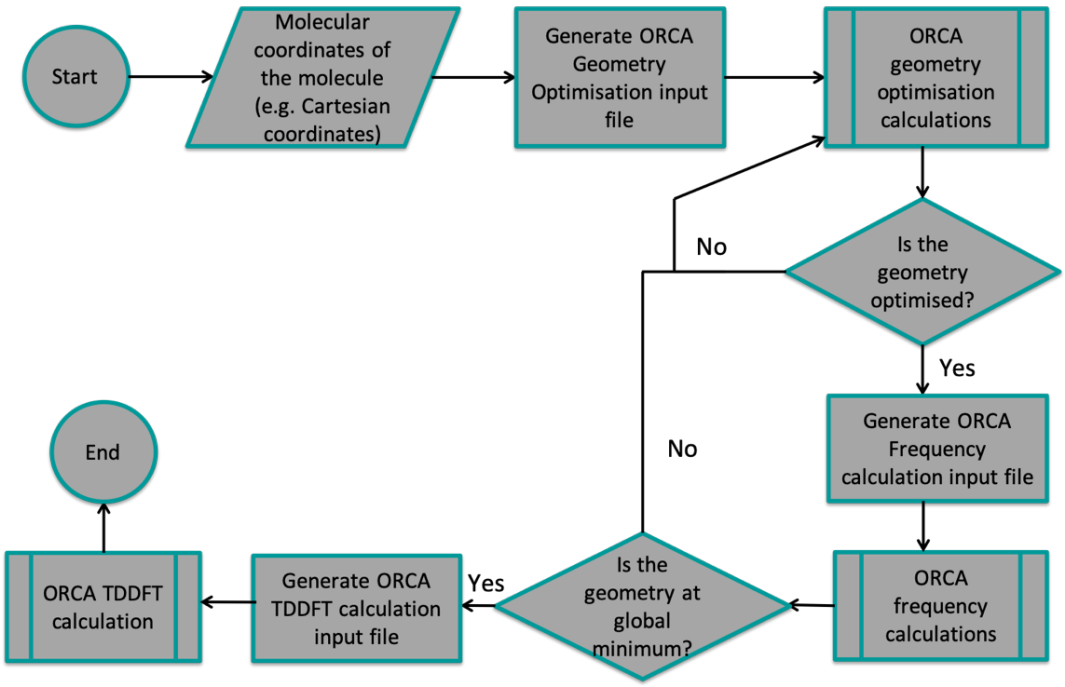

Figure 8.3. Algorithm of GTE1 script

8.3.2.2 (Theoretical) Excited-State Calculations for Liquid-Phase (LES1)

LES1 is the second step in the computational workflow for the calculation of the X-ray absorption (XA) spectra of samples in the liquid-phase. The default ORCA parameters used in this script are B3LYP functional [25] and minimally augmented def2-SVP [26] basis set with D3 dispersion correction [27], TIGHTSCF, Grid3 for SCF iterations and Grid5 for the final energy. It is assumed that the geometrical structure provided by the user is optimised; hence the computations performed by LES1 algorithm (shown in Figure 8.4) do not include a geometry optimisation step (similar to GTE1) and they run in two steps:

- Single point (SP) energy calculation to ensure that the provided geometry files and the ORCA parameters used are error-free. This calculation step also provides information on the molecular orbitals and their energies, which is needed for the excited-state calculation step.
- Excited-stated or TDDFT calculation is performed to generate the theoretical X-ray absorption (XA) spectra for the studied system.

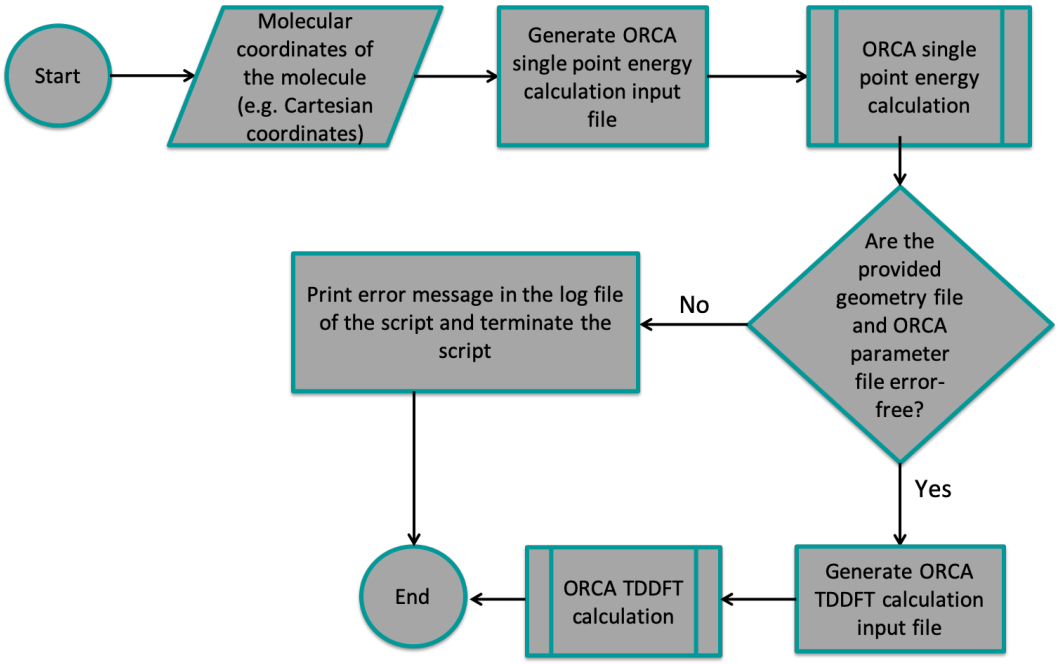

Figure 8.4. Algorithm of LES1 script

8.3.2.3 (Theoretical) Excited-State Calculation for Solid-State (SES1)

SES1 (algorithm shown in Figure 8.5) is used for the calculation of the X-ray absorption (XA) spectra of solid samples and it runs in two steps, which are similar to those of LES1. An extra feature provided by SES1 is the capability to perform geometry optimisation calculations for the positions of the hydrogen atoms. This might be desirable since the crystal structure of solid samples is mainly obtained from X-ray diffraction data, which does not accurately define the position of the hydrogen atoms. This option needs to be specified by the user in the argument file (S_args.txt).

The default ORCA parameters used in this script are similar to those used in LES1 script. B3LYP functional [25] is utilised with minimally augmented def2-SVP [26] basis set as well as D3 dispersion correction [27], TIGHTSCF, Grid3 for SCF iterations and Grid5 for the final energy. In addition to these parameters, RIJCOSX approximation [212] and def2/J auxiliary set were applied to speed up the calculations of the solid-state.

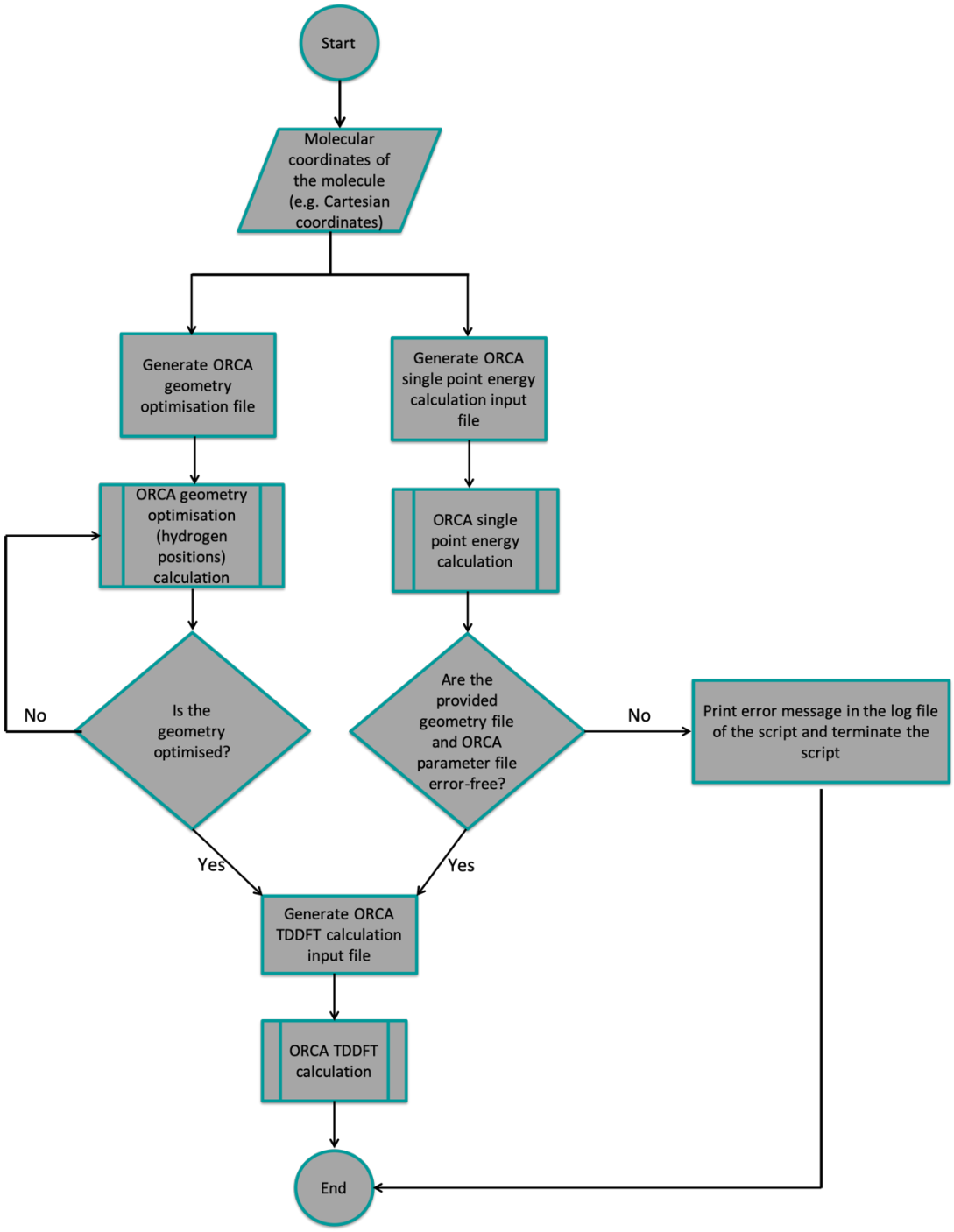

Figure 8.5. Algorithm of SES1 script

8.3.3 Comparison of Experimental and Theoretical Spectra (C1)

C1 (algorithm shown in Figure 8.6) is the last script in the pipeline for automating the analysis of X-ray absorption spectroscopy (XAS) data. It compares the experimental and theoretical X-ray absorption (XA) spectra of the studied system and assigns the features in the experimental spectra to specific electronic transitions. The output file from the excited-state calculations (TDDFT.out file) is copied to the output directory of the C1 script and the information on excited-states and population analysis of Loewdin orbitals

are extracted from the copied output file. Also, the TDDFT.out file is used to calculate the theoretical X-ray absorption spectrum using mapspc routine in ORCA [208]. Then, the peaks in the theoretical spectra are defined and associated with the corresponding excited states. The final step is to translate the energy scale of the theoretical spectrum so that the first peak is aligned with the first peak in the experimental spectrum, which enables visual comparison between the experimental and theoretical XA spectra.

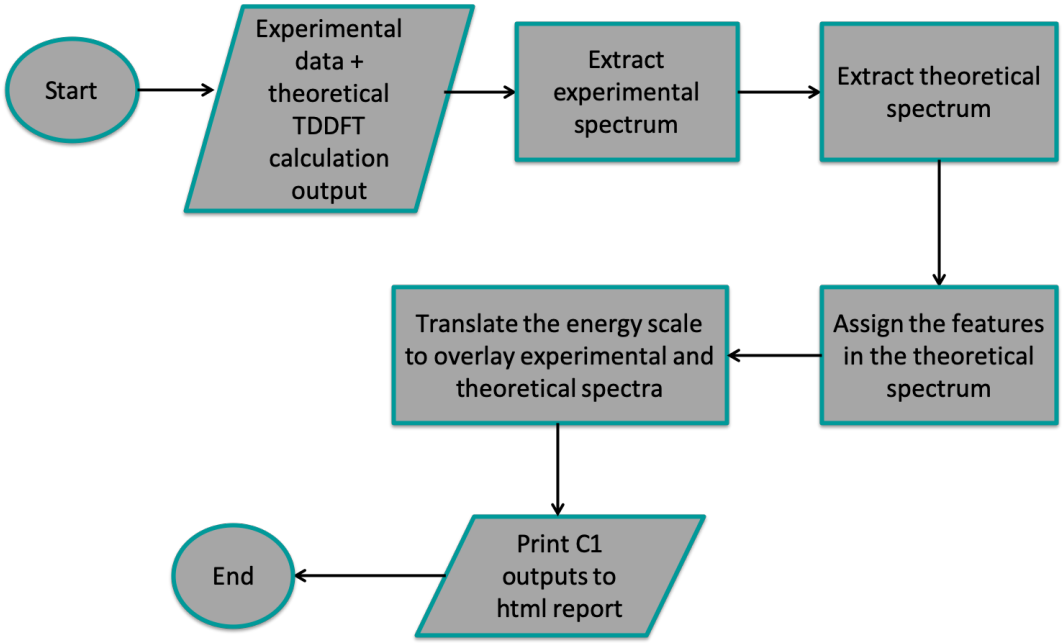

Figure 8.6. Algorithm of C1 script

8.3.4 Wrapper Scripts

The wrapper scripts (GW1, LW1 or SW1) were written to ease the use of the automated XAS analysis code. They run the scripts needed for gas-, liquid- or solid- phase calculations. GW1 is the wrapper script for gas-phase calculations and the computation runs in 3 steps; running E2 to perform the peak fitting of the experimental spectrum, followed by GTE1 to perform the electronic structure and excited-state calculations needed for the generation of the theoretical XA spectrum and finally C1 which compares the experimental and theoretical spectra to assign the spectral features to electronic transitions.

LW1 is the wrapper script for liquid-phase calculations and the computation starts by running E2 script to perform curve fitting of the experimental spectrum. This is followed by running a loop of LES1 and C1 scripts until theoretical and experimental data compare

well. This way a number of geometrical files representing the liquid can be used in the calculation; for example, files from a molecular dynamics (MD) simulations.

SW1 is the wrapper script for solid-phase calculations and the computation starts with running E2 script like the previous two wrapper scripts to perform peak fitting of the experimental spectrum. This is followed by running the SES1 script to perform the excited-state calculations and finally ending with C1 script to compare the experimental and theoretical spectra.

8.4 Operation of the Code

The code can be downloaded from GitHub [288]; the parent directory contains sub-directories for each Python script used in the code (these scripts were discussed in section 8.3) and the edge-data file (containing natural line width data) which is used by more than one of these scripts. The code is operated by modifying the parameters in the argument files located in the directory of the wrapper scripts; there are three wrapper scripts for each state of matter as mentioned previously. The parameters needed by each wrapper script are listed in Table 8.2. The code is operated by typing in the following command in the terminal:

```
Python ../../XAS_analysis/GW1/GW1.py../../XAS_analysis/GW1/G_args.txt
```

The command above is for gas-phase calculation and a video tutorial that goes through the steps of operating the code is available [289].

The output of E2 and C1 scripts are html reports that are created automatically when these scripts successfully run. Images in these scripts are created and are referenced in the html report. The images are also available in the output directories for each of these scripts. The resolution of these images is determined by the graphics settings of the system that creates them. These are scaled in the html to 100% so that high resolution images will be visible on the screen (it will also increase the size of low-resolution images). At the time of writing the code, the html 5 format was new and was not fully implemented in all browsers. Also, the html 5 standard had new ways to handle images that did not include scaling, so the html template is not compatible with html 5. Adding to that, all scripts create log files that provide meta data when each script is executed. These are plain text files and do not use the log file handler. This is because at the time

of development the argparse library (used in the code) was not compatible with the log file handler.

Table 8.2. The Parameters required by wrapper scripts of the automated XAS analysis code and their description

Parameter	Description	GW1	LW1	SW1
geom_file_name	Name and path of the geometry file for the studied system	✓		
geom_directory	Path to the directory containing the geometry files		✓	✓
orca_param	Name and path of the file containing the parameters for the theoretical calculations (If this is left blank then the default parameters will be used)	✓	✓	✓
orca_executable	The path to ORCA executable (if left empty then the default path "C:\Orca\orca.exe" will be used)	✓	✓	✓
experimental_spectra	Name and path of the file containing the experimental X-ray absorption spectra	✓	✓	✓
experimental_energy_column_number	The number of the column in the experimental data file which contains the absorption values	✓	✓	✓
experimental_intensity_column_number	The total number of columns in the experimental data file	✓	✓	✓
experimental_number_columns	The number of the column in the experimental data file which contains the energy values	✓	✓	✓
experimental_header_skip	The total number of rows to skip (headers) in the experimental data file	✓	✓	✓
element_calculate	The symbol of the element to calculate the theoretical core-level XA spectra for (e.g. C for carbon K-edge)	✓	✓	✓
number_of_processors	Number of processors if running ORCA calculations in parallel		✓	✓
hydrogen_opt	Run geometrical optimisation calculation of the hydrogen position			✓

8.5 Examples of Using the Automated XAS Analysis Code

8.5.1 Gas-Phase Calculations

The automated XAS analysis code was tested on 35 monomeric molecules. The experimental data for these molecules was ISEELS (inner-shell electronic energy loss spectroscopy) data obtained from the gas-phase core excitation database (Hitchcock database) [290]. The timescale for the calculations varied between few seconds to few minutes depending on the size of the molecule studied. The code was tested on Linux, Windows and Mac operating system and the calculations for the 35 molecules were run on a Mac operating system. The outcome of these calculations is deposited on the GitHub repository of the code under "examples" directory.

The peak fitting and assignment results for monomeric imidazole, which were generated by the code, are compared with manual peak fitting and assignment which were previously reported in chapter 4. The manual peak fitting was performed using Fityk software [244] and the peak assignment procedure was done by manually comparing the experimental spectrum to the calculated theoretical spectrum. Good agreement was achieved between the manual and automated procedures as illustrated in Figures (8.7) (8.8) and Tables (8.3) and (8.4).

Table 8.3. Peak position and assignment of the experimental ISEELS and calculated XA C K-edge spectra using manual and automated fitting procedures.

		Manual fitting		Automated fitting	
	Peak number	Exp/ eV	Calc/ eV *	Exp/eV	Calc/eV**
E (C 1s–1π*)	1	286.3	276.3 (C4)	286.5	276.3 (C4)
	1	286.8	276.5 (C5)	287.1	276.5 (C5)
			276.9 (C2)		276.9 (C2)
E (C 1s–2π*)	1	286.8	277.1 (C4)	288.9	277.1 (C4)
	2	288.0	278.1 (C5)		278.1 (C5)
	3	289.3	278.9 (C2)		278.9 (C2)

*** The energy shift of 10.1 eV has not been applied to the calculated transitions**
**** The energy shift of 9.8 eV has not been applied to the calculated transitions**

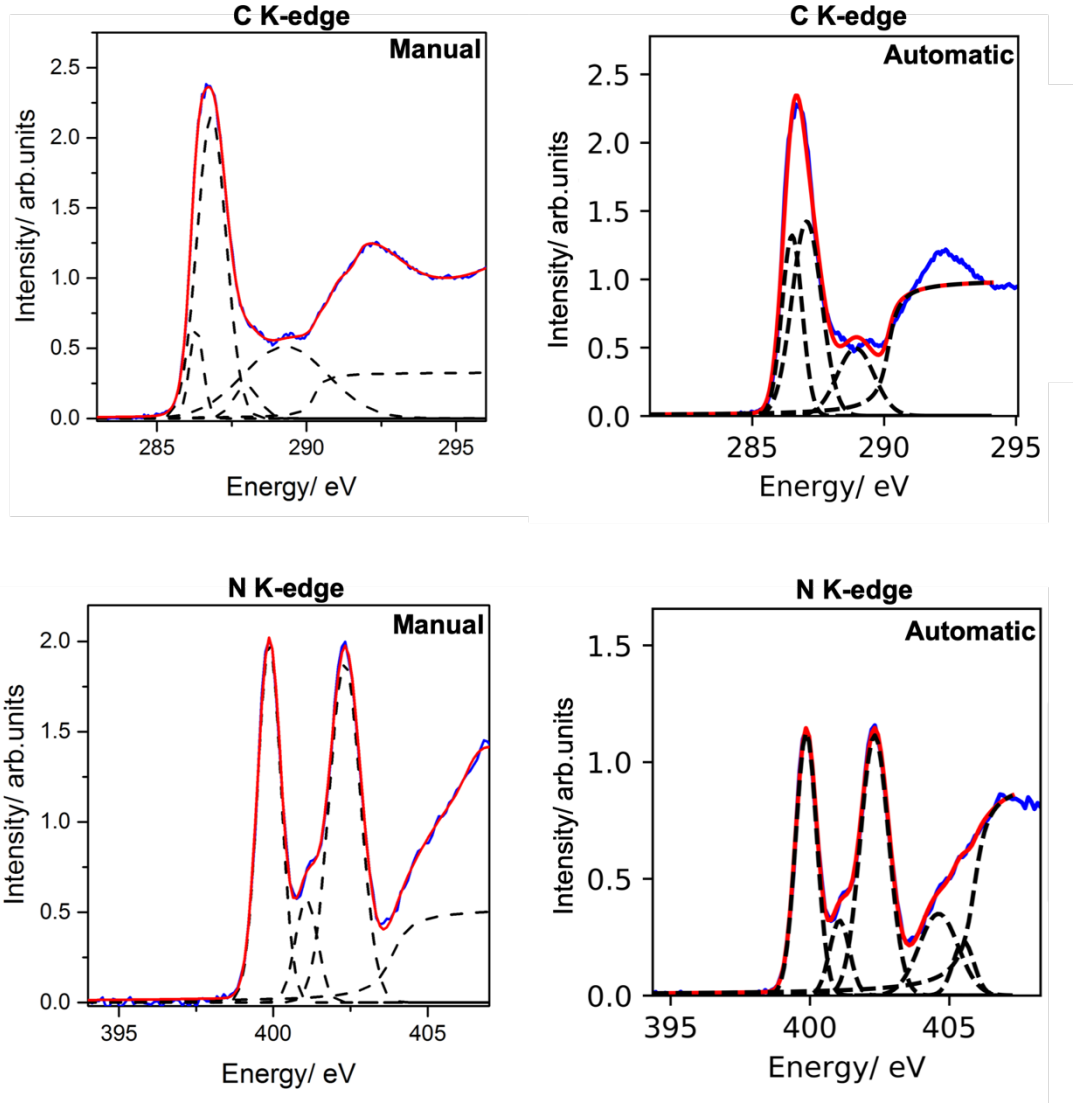

Figure 8.7. Fitted C (top) and N (bottom) K-edge ISEELS spectra of imidazole in the gas-phase utilising manual and automated fitting procedures.

Table 8.4. Peak position and assignment of the experimental ISEELS and calculated XA N K-edge spectra using manual and automated fitting procedures.

		Manual fitting		Automated fitting	
	Peak number	Exp/ eV	Calc/ eV*	Exp/ eV	Calc/ eV**
E (N3 1s - 1π*)	1	399.9	388.0	399.9	388.0
E (N3 1s - 2π*)	2	401.1	389.2	401.1	389.2
E (N1 1s - 1π*)	3	402.3	390.1	402.3	390.1
*** The energy shift of 11.8 eV has not been applied to the calculated transitions**					
**** The energy shift of 11.8 eV has not been applied to the calculated transitions**					

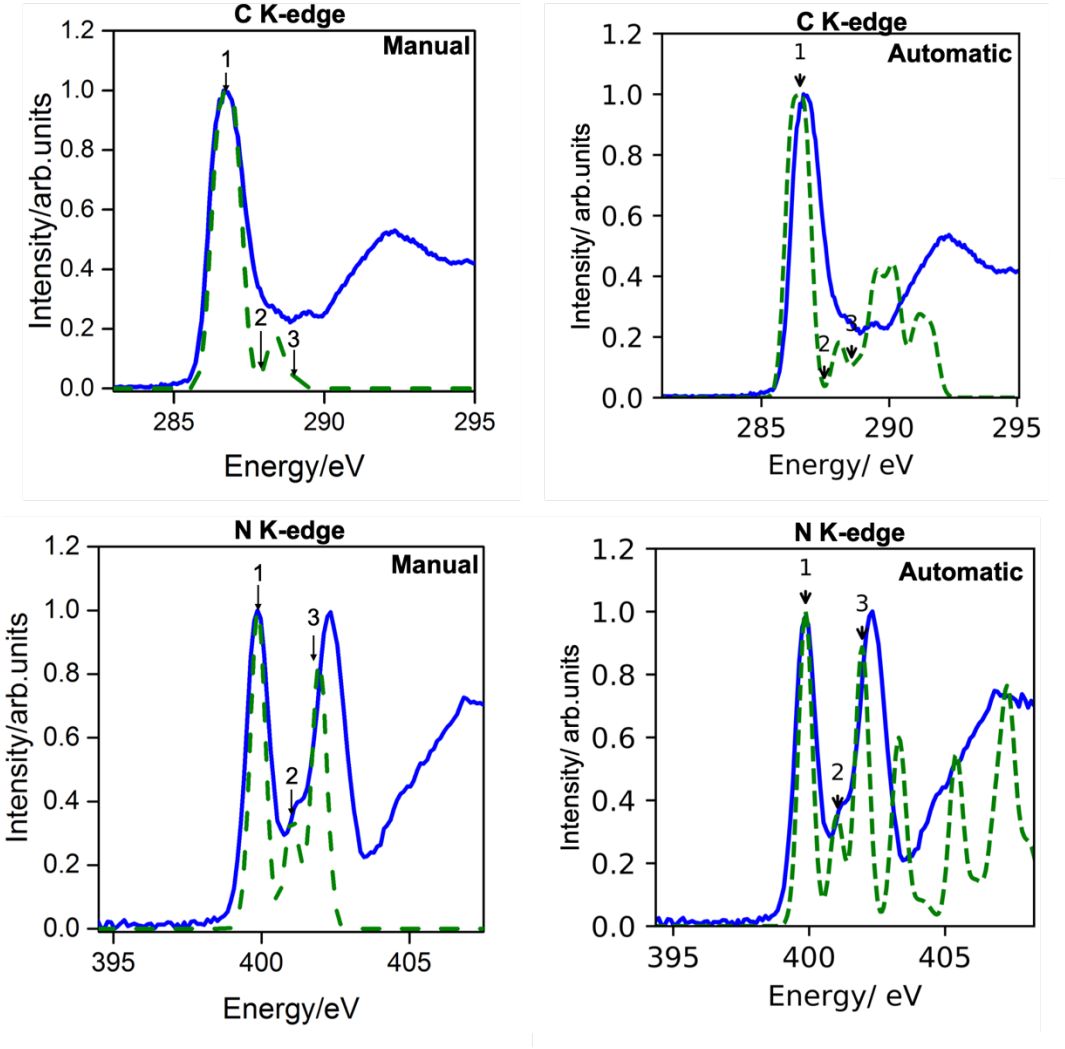

Figure 8.8. Comparison of the experimental ISEELS and calculated C (top) and N (bottom) K-edge spectra of imidazole in the gas-phase utilising manual and automated fitting procedures.

8.5.2 Solution-Phase Calculations using Simplified Structural Model of Aqueous Solutions of Imidazole

The automated XAS analysis code was tested on an explicit solvation model of aqueous imidazole solutions where a water molecule was added at each side of the imidazole ring. This solvation model has been presented in Chapter 4 (Model 4). The experimental XRS data for the 10.6 M aqueous imidazole solution reported in Chapter 6 was used here. The peak fittings and assignments, which were generated by the code, are compared with manual peak fittings and assignments which were previously reported in Chapters 4 and 6. The manual peak fitting was performed using Fityk software [244] and the peak assignment procedure was done by manually comparing the experimental spectrum to the calculated theoretical spectrum (details of the calculation procedure is provided in Chapter 4). Good agreement was achieved between the manual and

automated procedures as illustrated in Figures (8.9) and (8.10) and Tables (8.5) and (8.6).

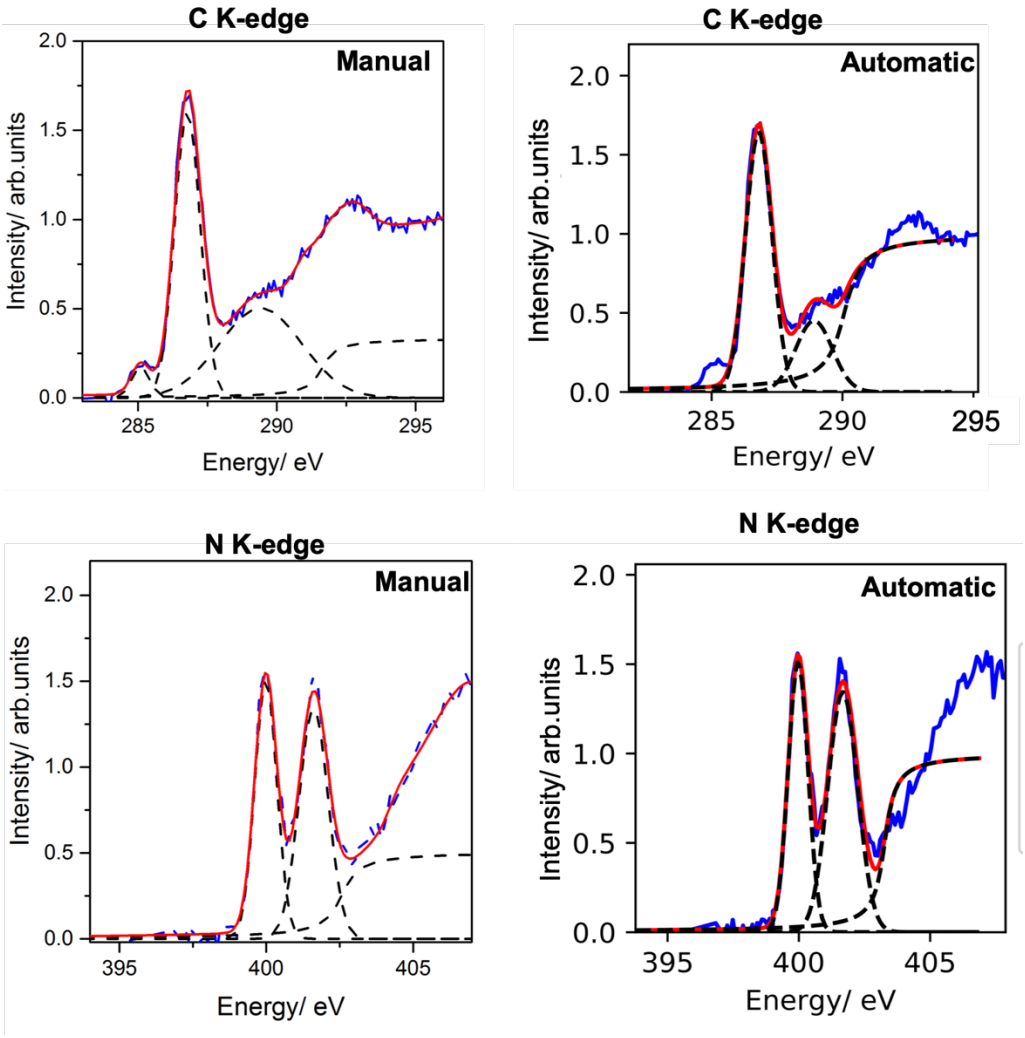

Figure 8.9. Fitted C (top) and N (bottom) K-edge XRS spectra of imidazole in aqueous solution utilising manual and automated fitting procedures.

Table 8.5 Peak position and assignment of the experimental XRS and calculated XA N K-edge spectra using manual and automated fitting procedures.

		Manual fitting		Automated fitting	
	Peak number	Exp/ eV	Calc/ eV*	Exp/ eV	Calc/ eV**
E (N3 1s - 1π*)	1	400.0	387.5	400.0	387.8
E (N3 1s - 2π*)	2	-	388.6	-	388.9
E (N1 1s - 1π*)	3	401.6	389.2	401.7	389.5
*** The energy shift of 12.5 eV has not been applied to the calculated transitions**					
**** The energy shift of 12.2 eV has not been applied to the calculated transitions**					

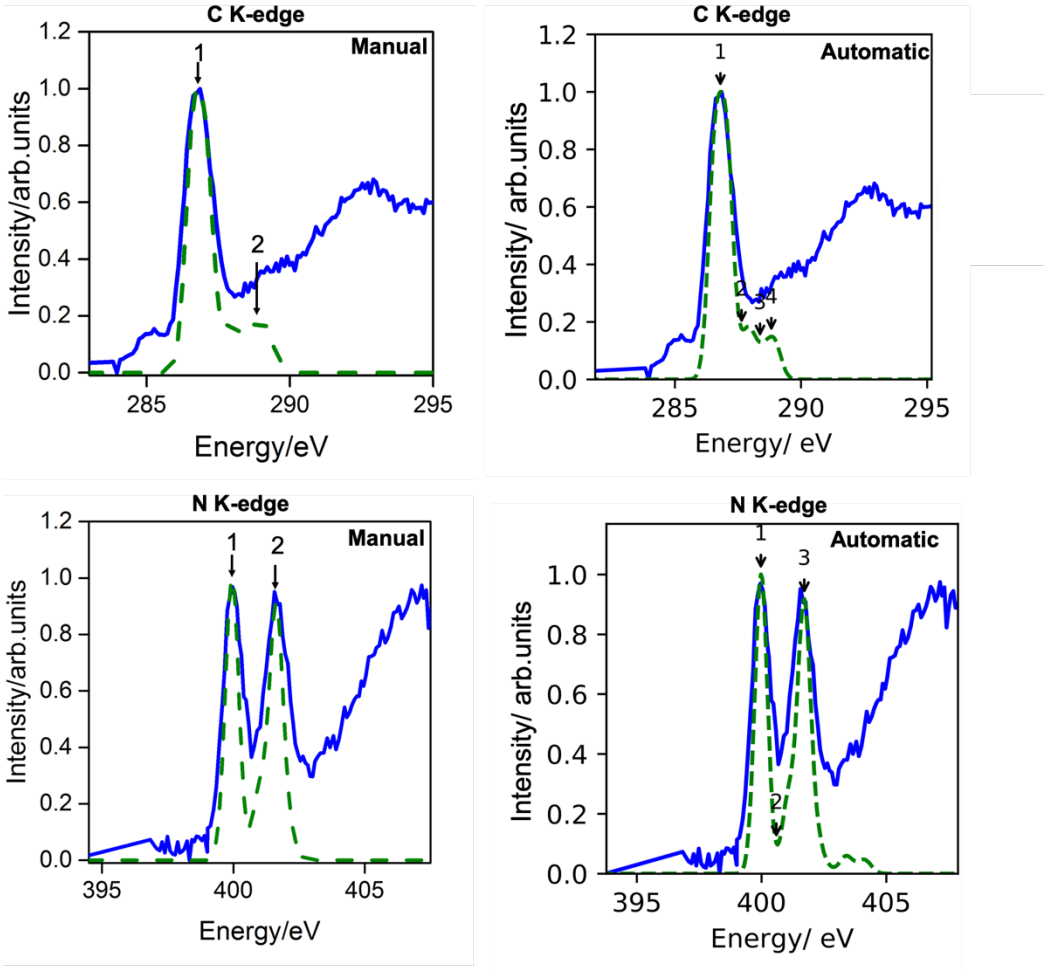

Figure 8.10. Comparison of the experimental XRS and calculated C (top) and N (bottom) K-edge XRS spectra of imidazole in aqueous solution utilising manual and automated fitting procedures.

Table 8.6. Peak position and assignment of the experimental XRS and calculated XA C K-edge spectra using manual and automated fitting procedures.

		Manual fitting		Automated fitting	
	Peak number	Exp/ eV	Calc/ eV *	Exp/eV	Calc/eV**
E (C 1s–1π*)	1	285.2	275.8 (C4)	286.8	276.2 (C4)
			275.9 (C5)		276.3 (C5)
	1	286.8	276.3 (C2)		276.7 (C2)
E (C 1s–2π*)	1	288.9	276.3(C4)	288.9	276.7 (C4)
	2		277.1(C5)		277.6(C5)
	3		278.0(C2)		278.4 (C2)

*** The energy shift of 10.8 eV has not been applied to the calculated transitions**
**** The energy shift of 10.3 eV has not been applied to the calculated transitions**

8.5.3 Solid-Phase Calculations using Crystalline Imidazole Structure

The automated XAS analysis code was tested on a cluster of 4 imidazole molecules taken from the crystal structure of imidazole. The experimental XRS data for the crystalline imidazole and the manual peak fitting and assignment results were reported in Chapter 4. In terms of peak fitting, good agreement was achieved between the manual and automated procedures as illustrated in Figures (8.11) and Tables (8.7) and (8.8). Assignment of the spectral features to electronic transitions was not achieved through the automated analysis route although the theoretical X-ray absorption spectra was successfully calculated as seen in Figure 8.12. This is attributed to how the electronic transitions are assigned using the Loewdin population analysis. Further refinement is required for the C1 script to accommodate the requirements of the solid-state calculations.

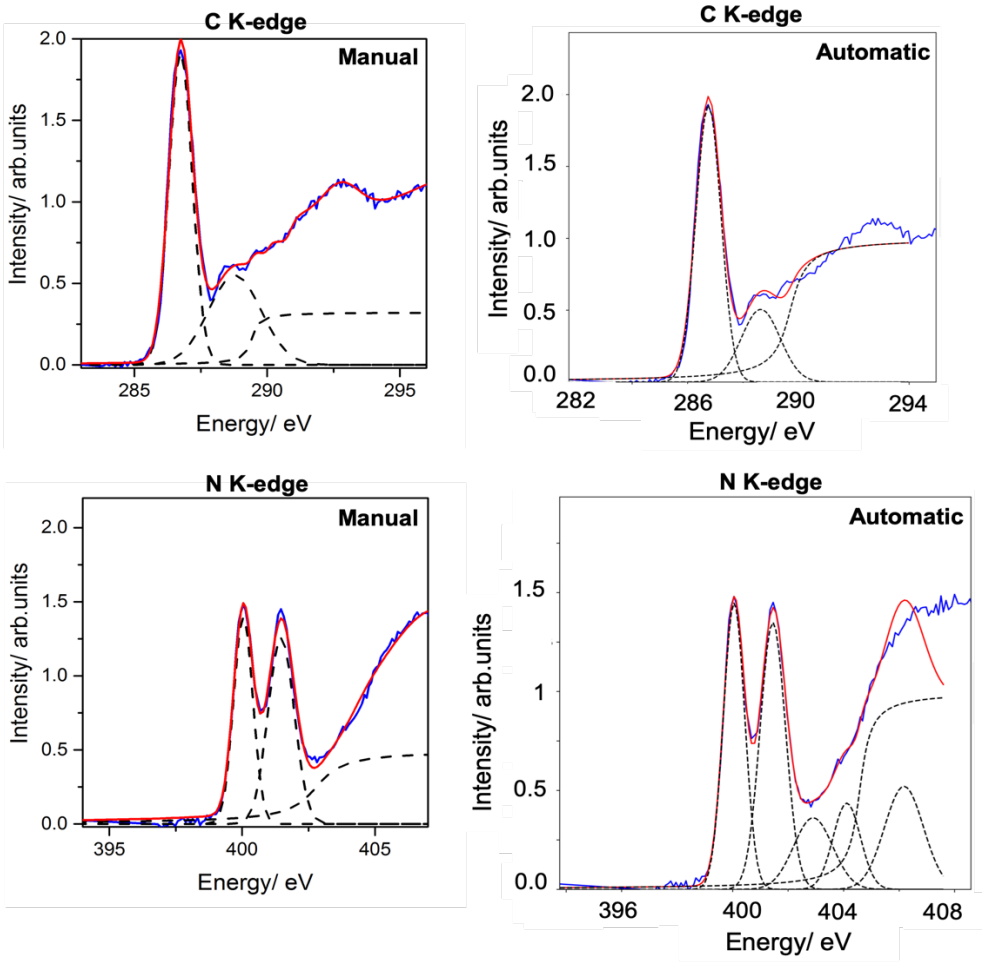

Figure 8.11. Fitted C (top) and N (bottom) K-edge XRS spectra of crystalline imidazole utilising manual and automated fitting procedures.

Table 8.7. Peak position and assignment of the experimental XRS and calculated XA C K-edge spectra using manual and automated fitting procedures.

		Manual fitting		Automated fitting	
	Peak number	Exp/ eV	Calc/ eV *	Exp/eV	Calc/eV**
E (C 1s–1π*)	1	286.7	275.8 (C4)	286.5	
	1		275.9 (C5)		
			276.3 (C2)		
E (C 1s–2π*)	1	288.8	276.0(C4)	286.5	
	2		277.0(C5)	287.1	
	3		277.7(C2)	288.9	

*** The energy shift of 10.8 eV has not been applied to the calculated transitions**
**** The energy shift of 10.7 eV has not been applied to the calculated transitions**

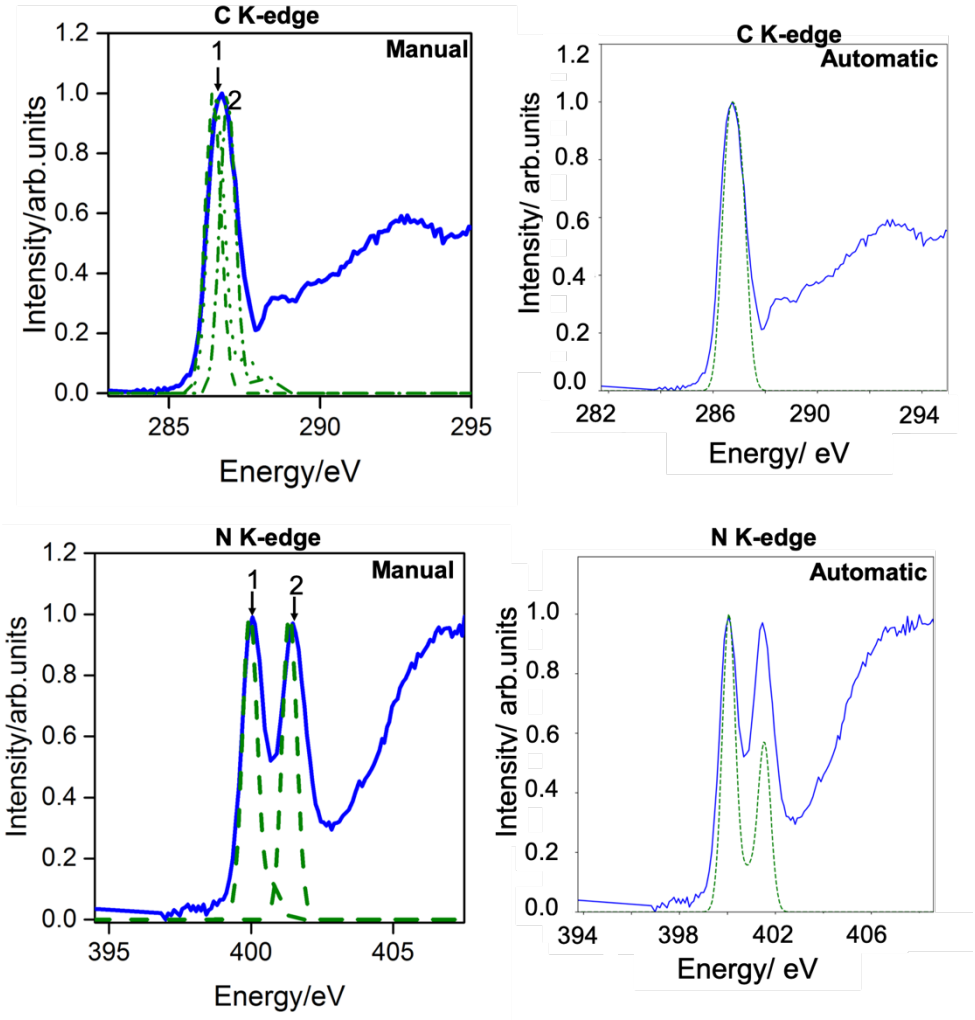

Figure 8.12. Comparison of the experimental XRS and calculated C (top) and N (bottom) K-edge XRS spectra of crystalline imidazole utilising manual and automated fitting procedures.

Table 8.8. Peak position and assignment of the experimental XRS and calculated XA N K-edge spectra using manual and automated fitting procedures.

		Manual fitting		Automated fitting	
	Peak number	Exp/ eV	Calc/ eV*	Exp/ eV	Calc/ eV**
E (N3 1s - 1π*)	1	400.0	387.6	400.1	
E (N3 1s - 2π*)	2	-	388.5	-	
E (N1 1s - 1π*)	3	401.5	389.0	401.5	
*** The energy shift of 12.5 eV has not been applied to the calculated transitions**					
**** The energy shift of 12.6 eV has not been applied to the calculated transitions**					

8.6 Conclusions

The automated XAS analysis procedure parameterises most of the essential electronic structure calculations and chemical state information at the K-edge spectra of organic molecules with minimal user input. The analysis of the experimental ISEELS spectra of imidazole and other organic molecules as well as the experimental XRS spectra of aqueous imidazole solution was adequate using the automated code. The results were quantitatively comparable to the manual fits. Further refinements are required for the script dealing with the comparison of the experimental and calculated data to enable assigning the spectral features in the spectra of the solid-state. Such simple automation procedures will be very useful for the initial stages of data analysis for large bodies of spectra and for users with beginner knowledge of the analysis procedure for XA spectra.

Chapter 9 Conclusions and Future Work

The research reported in this dissertation explored the combination of different analytical techniques; namely, vibrational spectroscopy (FTIR), core-level X-ray spectroscopy (X-ray Raman scattering) and neutron and X-ray scattering to elucidate the nature of intermolecular interactions in imidazole solutions. A bottom-up strategy was followed, starting by understanding the solvation structure of aqueous solutions of imidazole (chapter 3), followed by independent validation of the obtained structure by core-level spectroscopy (chapter 4). Next, the effect of changing the supersaturation of the solution on the solvation structure predicted in chapter 3 was studied. This was performed in two stages; (1) by changing the concentration of imidazole in aqueous solutions (chapter 5), and (2) through changing the temperature of the solution; i.e. by performing cooling crystallisation experiments (chapter 6). The crystallisers with *in-situ* probing were novel in themselves, providing experimental insight into the structural evolution of crystallising solutions prior to nucleation. The effect of solvent on the solvation structure of imidazole in solutions was examined by probing ethanolic solutions of imidazole and the effect of imidazole concentration on the structural properties of these solutions.

The most dominant structural motifs in concentrated aqueous solutions of imidazole are hydrated imidazole molecules with minimal direct solute-solute interactions, as shown by neutron diffraction combined with total X-ray scattering and empirical potential structure refinement (EPSR) simulations. The nearest imidazole neighbours were found to prefer parallel orientation (π-π stacking) at shorter radial distances (up to ~ 3.5-3.8 Å) and both parallel and edge-to-face orientations at longer radial distances (up to ~ 5 Å). The structural model predicted by neutron diffraction was validated by conducting excited-state density functional theory (DFT) calculations to generate the C and N K-edge spectra, which were in good quantitative agreement with the experimental X-ray Raman scattering (XRS) spectra.

The local solvation structure in aqueous solutions of imidazole, which is dominated by hydrogen-bonding to water, does not seem to be disturbed by changing the concentration of imidazole in the solution as indicated by FTIR spectroscopy, XRS and total X-ray scattering and the associated X-ray pair distribution function (XPDF) analysis.

These results were in agreement with the results of the cooling crystallisation experiments, which utilised XRS and XPDF to probe the structural evolution of imidazole in aqueous solutions from undersaturation through the metastable zone and until the onset of crystal formation. It was concluded that the breakage of solute-solvent hydrogen bonds is the key step in the nucleation of imidazole from the aqueous phase.

Direct solute-solute interactions are predicted in ethanolic solutions of imidazole. FTIR, N K-edge XRS and X-ray pair distribution function (XPDF) analysis indicated that the bonding environment in the ethanolic solutions is similar to that of the hydrogen-bonded solid than the aqueous solution. However, similar to the aqueous solutions of imidazole, the solvation structure does not seem to be disturbed by varying the supersaturation level of the ethanolic solutions. This reinforces that breakage of imidazole-solvent hydrogen bonds during desolvation is the key step in the nucleation of imidazole crystals from solutions.

The results presented in this thesis suggest the importance of studying solvation and nucleation from the molecular level. They also highlight that spectroscopic and scattering techniques are a promising combination for obtaining information about the local structure as well as the short- and long-range ordering of molecular solutes in the solution-phase. In fact, combining spectroscopic and scattering techniques with excited-state DFT calculations and EPSR simulations is the most novel feature of this thesis as this combination has not been applied to study solvation or nucleation of organic molecules from solutions.

There are a number of interesting ideas for future work that could develop from this thesis. Combining X-ray spectroscopic and scattering techniques with electronic structure calculations and reverse Monte Carlo simulations could be applied to probe other forms of crystallisation; such as anti-solvent crystallisation. The combination of above-mentioned analytical and computational techniques could also provide valuable information on the changes in the local structures of solutions during protonation, which is of high relevancy to understanding salts and co-crystals of active pharmaceutical ingredients (APIs).

Apart from further work using the techniques presented in this thesis; future work on imidazole and imidazole derivatives based utilising additional analytical techniques

could be beneficial. Some of these techniques include X-ray photoelectron spectroscopy of imidazole in solid- and solution-phase and NMR studies. Phase-contrast X-ray imaging is another possibility for obtaining more information on the structural evolution during nucleation due to the high spatial resolution of the technique. Dynamic light scattering, which provides a size distribution profile of small particles in suspension in solutions, may also supply useful pointers on self-association prior to nucleation.

Finally, the work in Chapter 8 is illustrative of what the future may hold for core-level X-ray spectroscopy. X-ray absorption spectroscopy (XAS) and other core-level spectroscopic techniques, such as electron energy loss spectroscopy (EELS), and X-ray Raman scattering (XRS), are powerful techniques to probe the local geometrical structures and bonding environments. However, the generated spectra are not well understood, and their analysis is considered to be challenging and time consuming. The quantitative analysis of core-level spectroscopy data is currently performed in separate software packages and is not fully automated. In Chapter 8, automating the analysis of core-level spectroscopy data was attempted. This will potentially save the time currently spent on the analysis process and will consequently encourage the use of core-level spectroscopy as a powerful spectroscopic technique. Testing the code on other systems will be useful to ensure the robustness of the algorithms utilised to conduct the analysis. Therefore, the code is publicly available to the scientific community through GitHub.

References

1. Fleming, G.R. and P.G. Wolynes, *Chemical Dynamics in Solution.* Physics Today, 1990. **43**(5).
2. Barbara, P.F. and W. Jarzeba, *Dynamic solvent effects on polar and nonpolar isomerizations.* Accounts of Chemical Research, 1988. **21**(5): p. 195-199.
3. Maroncelli, M., J. Macinnis, and G.R. Fleming, *Polar Solvent Dynamics and Electron-Transfer Reactions.* Science, 1989. **243**(4899): p. 1674.
4. Bagchi, B., *Dynamics of Solvation and Charge Transfer Reactions in Dipolar Liquids.* Annual Review of Physical Chemistry, 1989. **40**(1): p. 115-141.
5. Rossky, P.J. and J.D. Simon, *Dynamics of chemical processes in polar solvents.* Nature, 1994. **370**(6487): p. 263-269.
6. Fleming, G.R. and M. Cho, *CHROMOPHORE-SOLVENT DYNAMICS.* Annual Review of Physical Chemistry, 1996. **47**(1): p. 109-134.
7. Fleming, G.R., T. Joo, and M. Cho, *Femtosecond Chemical Dynamics in Condensed Phases.* Advances in Chemical Physics, 1997.
8. Pal, S.K., et al., *Biological Water: Femtosecond Dynamics of Macromolecular Hydration.* The Journal of Physical Chemistry B, 2002. **106**(48): p. 12376-12395.
9. Bagchi, B. and B. Jana, *Solvation dynamics in dipolar liquids.* Chemical Society Reviews, 2010. **39**(6): p. 1936-1954.
10. Hulliger, J., *Chemistry and Crystal Growth.* Angewandte Chemie International Edition in English, 1994. **33**(2): p. 143-162.
11. Cross, W.I., et al., *A Whole Output Strategy for Polymorph Screening: Combining Crystal Structure Prediction, Graph Set Analysis, and Targeted Crystallization Experiments in the Case of Diflunisal.* Crystal Growth & Design, 2003. **3**(2): p. 151-158.
12. Miller, J.M., et al., *Solvent Systems for Crystallization and Polymorph Selection*. 2007.
13. Cannon, C.G., *The nature of hydrogen bonding: A review of published work and discussion.* Spectrochimica Acta, 1958. **10**(4): p. 341-368.
14. Arunan, E., et al., *Definition of the hydrogen bond (IUPAC Recommendations 2011)*, in *Pure and Applied Chemistry*. 2011. p. 1637.
15. Buckingham, A.D., J.E. Del Bene, and S.A.C. McDowell, *The hydrogen bond.* Chemical Physics Letters, 2008. **463**(1): p. 1-10.
16. Steiner, T., *The Hydrogen Bond in the Solid State.* Angewandte Chemie International Edition, 2002. **41**(1): p. 48-76.
17. Lewis, G.N., *Valence and the structure of atoms and molecules*. 1923, New York: Chemical Catalog Co.
18. Latimer, W.M. and W.H. Rodebush, *POLARITY AND IONIZATION FROM THE STANDPOINT OF THE LEWIS THEORY OF VALENCE.* Journal of the American Chemical Society, 1920. **42**(7): p. 1419-1433.
19. Pauling, L., *The Nature of the Chemical Bond ... Third edition*.
20. Pimentel, G.C. and A.L. McClellan, *The hydrogen bond*. 1960, San Francisco ; London: W.H. Freeman.
21. Hamilton, W.C. and J.A. Ibers, *Hydrogen bonding in solids*. New York.
22. Jeffrey, G.A. and W. Saenger, *Hydrogen bonding in biological structures*. 1991, Berlin: Springer.
23. Smith, D.A.E., *Modeling the hydrogen bond : Symposium : 206th National meeting : Papers*. 1994: American Chemical Society.
24. Hadzi, D.C., *Theoretical treatments of hydrogen bonding*. 1997, Chichester: Wiley.
25. Jeffrey, G.A., *An introduction to hydrogen bonding*. 1997, New York ; Oxford: Oxford University Press.
26. Scheiner, S., *Hydrogen bonding : a theoretical perspective*. 1997, New York ; Oxford: Oxford University Press.

27. Desiraju, G.R. and T. Steiner, *The weak hydrogen bond in structural chemistry and biology*. 1999, Oxford: Oxford University Press.
28. Morokuma, K., *Why do molecules interact? The origin of electron donor-acceptor complexes, hydrogen bonding and proton affinity.* Accounts of Chemical Research, 1977. **10**(8): p. 294-300.
29. Baker, W., et al., *Organic chemistry.* Annual Reports on the Progress of Chemistry, 1946. **43**(0): p. 138-261.
30. Donohue, J., *The Hydrogen Bond in Organic Crystals.* The Journal of Physical Chemistry, 1952. **56**(4): p. 502-510.
31. Gilli, P., et al., *Predicting Hydrogen-Bond Strengths from Acid–Base Molecular Properties. The pKa Slide Rule: Toward the Solution of a Long-Lasting Problem.* Accounts of Chemical Research, 2009. **42**(1): p. 33-44.
32. Lakshmi, B., et al., *Is there a hydrogen bond radius? Evidence from microwave spectroscopy, neutron scattering and X-ray diffraction results.* New Journal of Chemistry, 2005. **29**(2): p. 371-377.
33. Hobza, P. and Z. Havlas, *Blue-Shifting Hydrogen Bonds.* Chemical Reviews, 2000. **100**(11): p. 4253-4264.
34. Hobza, P. and Z. Havlas, *Improper, blue-shifting hydrogen bond.* Theoretical Chemistry Accounts, 2002. **108**: p. 325 - 334.
35. Barnes, A.J., *Blue-shifting hydrogen bonds—are they improper or proper?* Journal of Molecular Structure, 2004. **704**(1): p. 3-9.
36. Bertolasi, V., et al., *Intramolecular O–H · · · O hydrogen bonds assisted by resonance. Correlation between crystallographic data and 1H NMR chemical shifts.* Journal of the Chemical Society, Perkin Transactions 2, 1997(5): p. 945-952.
37. Harris, T.K., Q. Zhao, and A.S. Mildvan, *NMR studies of strong hydrogen bonds in enzymes and in a model compound.* Journal of Molecular Structure, 2000. **552**(1): p. 97-109.
38. Stevens, J.S., S.J. Byard, and S.L.M. Schroeder, *Characterization of Proton Transfer in Co-Crystals by X-ray Photoelectron Spectroscopy (XPS).* Crystal Growth & Design, 2010. **10**(3): p. 1435-1442.
39. Stevens, J.S., et al., *NEXAFS Chemical State and Bond Lengths of p-Aminobenzoic Acid in Solution and Solid State.* 16th International Conference on X-Ray Absorption Fine Structure (Xafs16), 2016. **712**.
40. Naslund, L.A., et al., *X-ray absorption spectroscopy study of the hydrogen bond network in the bulk water of aqueous solutions.* Journal of Physical Chemistry A, 2005. **109**(27): p. 5995-6002.
41. Wernet, P., et al., *The Structure of the First Coordination Shell in Liquid Water.* Science, 2004. **304**(5673): p. 995.
42. Myneni, S., et al., *Spectroscopic probing of local hydrogen-bonding structures in liquid water.* Journal of Physics: Condensed Matter, 2002. **14**(8): p. L213-L219.
43. Tamenori, Y., et al., *Hydrogen bonding in methanol clusters probed by inner-shell photoabsorption spectroscopy in the carbon and oxygen K-edge regions.* The Journal of Chemical Physics, 2008. **128**(12): p. 124321.
44. Nilsson, A., et al., *The hydrogen bond in ice probed by soft x-ray spectroscopy and density functional theory.* The Journal of Chemical Physics, 2005. **122**(15): p. 154505.
45. Bergmann, U., et al., *X-ray Raman spectroscopy at the oxygen K edge of water and ice: Implications on local structure models.* Physical Review B, 2002. **66**(9).
46. Rappe, A.K. and C.J. Casewit, *Molecular mechanics across chemistry*. 1997, Sausalito, Calif.: University Science Books.
47. Rappé, A.K. and E.R. Bernstein, *Ab Initio Calculation of Nonbonded Interactions: Are We There Yet?* The Journal of Physical Chemistry A, 2000. **104**(26): p. 6117-6128.
48. Müller-Dethlefs, K. and P. Hobza, *Noncovalent Interactions: A Challenge for Experiment and Theory.* Chemical Reviews, 2000. **100**(1): p. 143-168.

49. Arrheniu, S., *Development of the Theory of Electrolytic Dissociation*, in *Noble Lecture*. 1903.
50. Brönsted, J.N., *Part III. Neutral salt and activity effects. The theory of acid and basic catalysis.* Transactions of the Faraday Society, 1928. **24**(0): p. 630-640.
51. Lowry, T.M. and I.J. Faulkner, *CCCXCIX.—Studies of dynamic isomerism. Part XX. Amphoteric solvents as catalysts for the mutarotation of the sugars.* Journal of the Chemical Society, Transactions, 1925. **127**(0): p. 2883-2887.
52. Lewis, G.N., *Acids and bases.* Journal of the Franklin Institute, 1938. **226**(3): p. 293-313.
53. Chandler, D., *From 50 Years Ago, the Birth of Modern Liquid-State Science.* Annual Review of Physical Chemistry, 2017. **68**(1): p. 19-38.
54. Sun, C.Q., *Aqueous charge injection: solvation bonding dynamics, molecular nonbond interactions, and extraordinary solute capabilities.* International Reviews in Physical Chemistry, 2018. **37**(3-4): p. 363-558.
55. Miller, J., M., et al., *Solvent Systems for Crystallization and Polymorph Selection*, in *Solvent Systems and Their Selection in Pharmaceutics and Biopharmaceutics , ed.*, P. Augustijns and M.E. Brewster, Editors. 2007, Springer: New York. p. 53-109.
56. Schein, C.H., *Solubility as a function of protein structure and solvent components.* Biotechnology (N Y), 1990. **8**(4): p. 308-17.
57. Eggers, D.K., H.W. Blanch, and J.M. Prausnitz, *Extractive catalysis: Solvent effects on equilibria of enzymatic reactions in two-phase systems.* Enzyme and Microbial Technology, 1989. **11**(2): p. 84-89.
58. Gilbert, L. and C. Mercier, *Solvent effects in heterogeneous catalysis : Application to the synthesis of fine chemicals*, in *Studies in Surface Science and Catalysis*, M. Guisnet, et al., Editors. 1993, Elsevier. p. 51-66.
59. Dyson, P.J. and P.G. Jessop, *Solvent effects in catalysis: rational improvements of catalysts via manipulation of solvent interactions.* Catalysis Science & Technology, 2016. **6**(10): p. 3302-3316.
60. Hynes, J.T., *Chemical Reaction Dynamics in Solution.* Annual Review of Physical Chemistry, 1985. **36**(1): p. 573-597.
61. Balakrishnan, G., et al., *Understanding solvent effects on structure and reactivity of organic intermediates: a Raman study.* Faraday Discussions, 2010. **145**(0): p. 443-466.
62. Cukier, R.I. and D.G. Nocera, *On the separation of static and dynamic solvent effects for electron transfer reactions.* The Journal of Chemical Physics, 1992. **97**(10): p. 7371-7376.
63. Tachiya, M., *Static and dynamical solvent effects on electron transfer rates.* Radiation Physics and Chemistry, 1996. **47**(1): p. 43-46.
64. *Asano, T., Kinetics in highly viscous solutions: dynamic solvent effects in 'slow' reactions.* Pure and Applied Chemistry, 1999. **71**: p. 1691-1704.
65. Wagner, C., *Kinetik der Phasenbildung. Von Prof. Dr. M. Volmer. (Bd. IV der Sammlung „Die chemische Reaktion", herausgegeben von K. F. Bonhoeffer.) XII und 220 S. Verlag Th. Steinkopff, Dresden und Leipzig 1939. Preis geh. RM. 19,—, geb. RM. 20,—.* Angewandte Chemie, 1939. **52**(30): p. 503-504.
66. Gibbs, J.W., Trans. Connect. Acad. Sci, 1876. **3**: p. 108-248.
67. Gibbs, J.W., Trans. Connect. Acad. Sci, 1877. **3**: p. 343-524.
68. Erdemir, D., A.Y. Lee, and A.S. Myerson, *Nucleation of Crystals from Solution: Classical and Two-Step Models.* Acc Chem Res, 2009. **42**(5): p. 621-9.
69. Oxtoby, D.W., *Nucleation of first-order phase transitions.* Accounts of Chemical Research, 1998. **31**(2): p. 91-97.
70. Chiarella, R.A., et al., *The Nucleation of Inosine: the Impact of Solution Chemistry on the Appearance of Polymorphic and Hydrated Crystal Forms.* Faraday Discussions, 2007. **136**(0): p. 179-193.

71. Davey, R.J., et al., *Concerning the Relationship between Structural and Growth Synthons in Crystal Nucleation: Solution and Crystal Chemistry of Carboxylic Acids As Revealed through IR Spectroscopy.* Crystal Growth & Design, 2006. **6**(8): p. 1788-1796.
72. Burton, R.C., et al., *The Relationship between Solution Structure and Crystal Nucleation: A Neutron Scattering Study of Supersaturated Methanolic Solutions of Benzoic Acid.* The Journal of Physical Chemistry B, 2010. **114**(26): p. 8807-8816.
73. Vekilov, P.G., *Nucleation.* Crystal Growth & Design, 2010. **10**(12): p. 5007-5019.
74. Gebauer, D. and H. Colfen, *Prenucleation clusters and non-classical nucleation.* Nano Today, 2011. **6**(6): p. 564-584.
75. tenWolde, P.R. and D. Frenkel, *Enhancement of protein crystal nucleation by critical density fluctuations.* Science, 1997. **277**(5334): p. 1975-1978.
76. Talanquer, V. and D.W. Oxtoby, *Crystal nucleation in the presence of a metastable critical point.* Journal of Chemical Physics, 1998. **109**(1): p. 223-227.
77. Bonnett, P.E., et al., *Solution crystallisation via a submerged liquid-liquid phase boundary: oiling out.* Chemical Communications, 2003(6): p. 698-699.
78. Davey, R.J., S.L.M. Schroeder, and J.H. ter Horst, *Nucleation of Organic CrystalsA Molecular Perspective.* Angewandte Chemie-International Edition, 2013. **52**(8): p. 2166-2179.
79. Parveen, S., et al., *Linking Solution Chemistry to Crystal Nucleation: the Case of Tetrolic Acid.* Chemical Communications, 2005(12): p. 1531-1533.
80. Hunter, C.A., J.F. McCabe, and A. Spitaleri, *Solvent Effects of the Structures of Prenucleation Aggregates of Carbamazepine.* CrystEngComm, 2012. **14**(21): p. 7115-7117.
81. Kulkarni, S.A., et al., *Isonicotinamide Self-Association: the Link between Solvent and Polymorph Nucleation.* Chemical Communications, 2012. **48**(41): p. 4983-4985.
82. Hamad, S., et al., *Kinetic Insights into the Role of the Solvent in the Polymorphism of 5-Fluorouracil from Molecular Dynamics Simulations.* The Journal of Physical Chemistry B, 2006. **110**(7): p. 3323-3329.
83. Chen, J. and B.L. Trout, *Computational Study of Solvent Effects on the Molecular Self-Assembly of Tetrolic Acid in Solution and Implications for the Polymorph Formed from Crystallization.* The Journal of Physical Chemistry B, 2008. **112**(26): p. 7794-7802.
84. Spitaleri, A., et al., *A 1H NMR Study of Crystal Nucleation in Solution.* CrystEngComm, 2004. **6**(80): p. 490-493.
85. Chadwick, K., et al., *Cocrystallization: A Solution Chemistry Perspective and the Case of Benzophenone and Diphenylamine.* Crystal Growth & Design, 2009. **9**(4): p. 1990-1999.
86. Peral, F. and E. Gallego, *Self-association of imidazole and its methyl derivatives in aqueous solution. A study by ultraviolet spectroscopy.* Journal of Molecular Structure, 1997. **415**(1): p. 187-196.
87. Williams, D.H. and I. Fleming, *Spectroscopic methods in organic chemistry*. 6th ed. 2008, London ; New York: McGraw-Hill. xiii, 291 p.
88. Thomason, M.J., et al., *Self-Association of Organic Solutes in Solution: a NEXAFS Study of Aqueous Imidazole.* Faraday Discussions, 2015. **179**: p. 269-289.
89. Golnak, R., et al., *Intermolecular bonding of hemin in solution and in solid state probed by N K-edge X-ray spectroscopies.* Physical Chemistry Chemical Physics, 2015. **17**(43): p. 29000-29006.
90. Gainar, A., et al., *NEXAFS Sensitivity to Bond Lengths in Complex Molecular Materials: A Study of Crystalline Saccharides.* The Journal of Physical Chemistry B, 2015. **119**(45): p. 14373-14381.
91. Nolting, D., et al., *Pseudoequivalent Nitrogen Atoms in Aqueous Imidazole Distinguished by Chemical Shifts in Photoelectron Spectroscopy.* Journal of the American Chemical Society, 2008. **130**(26): p. 8150-8151.

92. Jagoda-Cwiklik, B., et al., *Ionization of Imidazole in the Gas Phase, Microhydrated Environments, and in Aqueous Solution.* The Journal of Physical Chemistry A, 2008. **112**(16): p. 3499-3505.
93. Jagoda-Cwiklik, B., et al., *Ionization of Aqueous Cations: Photoelectron Spectroscopy and ab Initio Calculations of Protonated Imidazole.* The Journal of Physical Chemistry B, 2008. **112**(25): p. 7355-7358.
94. Chapman, H.N., et al., *Femtosecond X-ray protein nanocrystallography.* Nature, 2011. **470**(7332): p. 73-77.
95. Wilson, K.R., et al., *Investigation of volatile liquid surfaces by synchrotron x-ray spectroscopy of liquid microjets.* Review of Scientific Instruments, 2004. **75**(3): p. 725-736.
96. Faubel, M., S. Schlemmer, and J.P. Toennies, *A Molecular-Beam Study of the Evaporation of Water from a Liquid Jet.* Zeitschrift Fur Physik D-Atoms Molecules and Clusters, 1988. **10**(2-3): p. 269-277.
97. Winter, B. and M. Faubel, *Photoemission from Liquid Aqueous Solutions.* Chemical Reviews, 2006. **106**(4): p. 1176-1211.
98. Wilson, K.R., et al., *X-ray Spectroscopy of Liquid Water Microjets.* The Journal of Physical Chemistry B, 2001. **105**(17): p. 3346-3349.
99. Alison, H.G., et al., *Using a novel plug flow reactor for the in situ, simultaneous, monitoring of SAXS and WAXD during crystallisation from solution.* Physical Chemistry Chemical Physics, 2003. **5**(22): p. 4998-5000.
100. Wu, S., M. Li, and Y. Sun, *In Situ Synchrotron X-ray Characterization Shining Light on the Nucleation and Growth Kinetics of Colloidal Nanoparticles.* Angewandte Chemie International Edition, 2019. **58**(27): p. 8987-8995.
101. Abécassis, B., et al., *Probing in situ the Nucleation and Growth of Gold Nanoparticles by Small-Angle X-ray Scattering.* Nano Letters, 2007. **7**(6): p. 1723-1727.
102. Hulme, E.C., et al., *The hydration of the neurotransmitter acetylcholine in aqueous solution.* Biophysical Journal, 2006. **91**(6): p. 2371-2380.
103. Callear, S.K., et al., *Conformation and interactions of dopamine hydrochloride in solution.* The Journal of Chemical Physics, 2015. **142**(1): p. 014502.
104. McLain, S.E., A.K. Soper, and A. Luzar, *Investigations on the structure of dimethyl sulfoxide and acetone in aqueous solution.* The Journal of Chemical Physics, 2007. **127**(17): p. 174515.
105. McLain, S.E., et al., *Structure and Hydration of l-Proline in Aqueous Solutions.* The Journal of Physical Chemistry B, 2007. **111**(17): p. 4568-4580.
106. Rhys, N.H., et al., *On the structure of an aqueous propylene glycol solution.* JOURNAL OF CHEMICAL PHYSICS, 2016. **145**(22).
107. Soper, A.K., E.W. Castner, and A. Luzar, *Impact of urea on water structure: a clue to its properties as a denaturant?* Biophysical Chemistry, 2003. **105**(2): p. 649-666.
108. Meersman, F., et al., *An X-ray and neutron scattering study of the equilibrium between trimethylamine N-oxide and urea in aqueous solution.* Physical Chemistry Chemical Physics, 2011. **13**(30): p. 13765-13771.
109. Burton, R.C., et al., *The Structure of a Supersaturated Solution: A Neutron Scattering Study of Aqueous Urea.* Crystal Growth & Design, 2008. **8**(5): p. 1559-1565.
110. Duboué-Dijon, E., et al., *Changes in the hydration structure of imidazole upon protonation: Neutron scattering and molecular simulations.* The Journal of Chemical Physics, 2017. **146**(18): p. 185102.
111. Ramondo, F., et al., *Hydration of diazoles in water solution: pyrazole. A theoretical and X-ray diffraction study.* Physical Chemistry Chemical Physics, 2009. **11**(41): p. 9431-9439.
112. Mason, P.E., et al., *Observation of Pyridine Aggregation in Aqueous Solution Using Neutron Scattering Experiments and MD Simulations.* The Journal of Physical Chemistry B, 2010. **114**(16): p. 5412-5419.

113. Naslund, L.A., et al., *X-ray absorption spectroscopy measurements of liquid water.* Journal of Physical Chemistry B, 2005. **109**(28): p. 13835-13839.
114. Bergmann, U., et al., *X-ray Raman spectroscopy - A new tool to study local structure of aromatic hydrocarbons and asphaltenes.* Petroleum Science and Technology, 2004. **22**(7-8): p. 863-875.
115. Sternemann, C. and M. Wilke, *Spectroscopy of low and intermediate Z elements at extreme conditions: in situ studies of Earth materials at pressure and temperature via X-ray Raman scattering.* High Pressure Research, 2016. **36**(3): p. 275-292.
116. Tse, J.S., et al., *X-ray Raman spectroscopic study of water in the condensed phases.* Physical Review Letters, 2008. **100**(9).
117. Bergmann, U., et al., *Isotope effects in liquid water probed by x-ray Raman spectroscopy.* Physical Review B, 2007. **76**(2).
118. Sahle, C.J., et al., *Influence of TMAO and urea on the structure of water studied by inelastic X-ray scattering.* Physical Chemistry Chemical Physics, 2016. **18**(24): p. 16518-16526.
119. Xu, H., et al., *The Chemistry of Nucleation: In Situ Pair Distribution Function Analysis of Secondary Building Units During UiO-66 MOF Formation.* Chemistry – A European Journal, 2019. **25**(8): p. 2051-2058.
120. Marcial, J. and J. McCloy, *Role of short range order on crystallization of tectosilicate glasses: A diffraction study.* Journal of Non-Crystalline Solids, 2019. **505**: p. 131-143.
121. Zobel, M., R.B. Neder, and S.A. Kimber, *Universal solvent restructuring induced by colloidal nanoparticles.* Science, 2015. **347**(6219): p. 292-4.
122. Bøjesen, E.D. and B.B. Iversen, *The chemistry of nucleation.* CrystEngComm, 2016. **18**(43): p. 8332-8353.
123. Will, G., *Crystal Structure of Imidazol.* Nature, 1963. **198**(4880): p. 575-575.
124. Martinez-Carrera, S., *The Crystal Structure of Imidazole at -150degreesC.* Acta Crystallographica, 1966. **20**(6): p. 783-789.
125. Craven, B.M., et al., *The crystal structure of imidazole by neutron diffraction at 20degreesC and -150degreesC.* Acta Crystallographica Section B, 1977. **33**(8): p. 2585-2589.
126. Colombo, L., *Low-Frequency Raman Spectrum of Imidazole Single Crystals.* The Journal of Chemical Physics, 1968. **49**(10): p. 4688-4695.
127. Apen, E., A.P. Hitchcock, and J.L. Gland, *Experimental Studies of the Core Excitation of Imidazole, 4,5-dicyanoimidazole, and s-triazine.* The Journal of Physical Chemistry, 1993. **97**(26): p. 6859-6866.
128. Holliday, G.L., J.B.O. Mitchell, and J.M. Thornton, *Understanding the Functional Roles of Amino Acid Residues in Enzyme Catalysis.* Journal of Molecular Biology, 2009. **390**(3): p. 560-577.
129. Jeuken, L.J.C., et al., *Role of the surface-exposed and copper-coordinating histidine in blue copper proteins: The electron-transfer and redox-coupled ligand binding properties of His117Gly azurin.* Journal of the American Chemical Society, 2000. **122**(49): p. 12186-12194.
130. Smith, S.R., et al., *Investigation of the copper binding site and the role of histidine as a ligand in riboflavin binding protein.* Inorganic Chemistry, 2008. **47**(15): p. 6867-6872.
131. Haydon, M.J. and C.S. Cobbett, *Transporters of ligands for essential metal ions in plants.* New Phytologist, 2007. **174**(3): p. 499-506.
132. Callahan, D.L., et al., *Metal ion ligands in hyperaccumulating plants.* Journal of Biological Inorganic Chemistry, 2006. **11**(1): p. 2-12.
133. Alberts, I.L., K. Nadassy, and S.J. Wodak, *Analysis of zinc binding sites in protein crystal structures.* Protein Science, 1998. **7**(8): p. 1700-1716.
134. Sankararamakrishnan, R., S. Verma, and S. Kumar, *ATCUN-like metal-binding motifs in proteins: Identification and characterization by crystal structure and sequence analysis.* Proteins-Structure Function and Bioinformatics, 2005. **58**(1): p. 211-221.

135. Li, H.R., et al., *Histidine 103 in Fra2 Is an Iron-Sulfur Cluster Ligand in the [2Fe-2S] Fra2-Grx3 Complex and Is Required for in Vivo Iron Signaling in Yeast.* Journal of Biological Chemistry, 2011. **286**(1): p. 867-876.
136. Tsud, N., et al., *Bonding of Histidine to Cerium Oxide.* Journal of Physical Chemistry B, 2013. **117**(31): p. 9182-9193.
137. Feyer, V., et al., *Adsorption of Histidine and Histidine-Containing Peptides on Au(111).* Langmuir, 2010. **26**(11): p. 8606-8613.
138. Feyer, V., et al., *The Electronic Structure and Adsorption Geometry of L-Histidine on Cu(110).* Journal of Physical Chemistry B, 2008. **112**(43): p. 13655-13660.
139. Zubavichus, Y., et al., *X-ray Absorption Spectroscopy of the Nucleotide Bases at the Carbon, Nitrogen, and Oxygen K-Edges.* Journal of Physical Chemistry B, 2008. **112**(44): p. 13711-13716.
140. De Luca, L., *Naturally occurring and synthetic imidazoles: Their chemistry and their biological activities.* Current Medicinal Chemistry, 2006. **13**(1): p. 1-23.
141. Katritzky, A.R. and A.F. Pozharskii, *Handbook of heterocyclic chemistry*. 2000, Amsterdam; Lausanne; New York: Pergamon Press.
142. Zhao, G. and H. Lin, *Metal complexes with aromatic N-containing ligands as potential agents in cancer treatment.* Curr Med Chem Anticancer Agents, 2005. **5**(2): p. 137-47.
143. Balakrishnan, G., et al., *Mode Recognition in UV Resonance Raman Spectra of Imidazole: Histidine Monitoring in Proteins.* Journal of Physical Chemistry B, 2012. **116**(31): p. 9387-9395.
144. Lehmann, J., et al., *Synchrotron-Based Near-Edge X-Ray Spectroscopy of Natural Organic Matter in Soils and Sediments*, in *Biophysico-Chemical Processes Involving Natural Nonliving Organic Matter in Environmental Systems*. 2009, John Wiley & Sons, Inc. p. 729-781.
145. Leinweber, P., et al., *Nitrogen K-edge XANES - An overview of reference compounds used to identify 'unknown' organic nitrogen in environmental samples.* Journal of Synchrotron Radiation, 2007. **14**: p. 500-511.
146. Chen, T.R., A.C. Yeh, and J.D. Chen, *A new imidazolylquinoline for organic thin film transistor.* Tetrahedron Letters, 2005. **46**(9): p. 1569-1571.
147. Mao, J., et al., *Imidazole, a New Tunable Reagent for Producing Nanocellulose, Part I: Xylan-Coated CNCs and CNFs.* Polymers, 2017. **9**(10): p. 473.
148. Domańska, U., M.K. Kozłowska, and M. Rogalski, *Solubilities, Partition Coefficients, Density, and Surface Tension for Imidazoles + Octan-1-ol or + Water or + n-Decane.* Journal of Chemical & Engineering Data, 2002. **47**(3): p. 456-466.
149. Hückel, W., J. Datow, and E. Simmersbach, *Physikalische eigenschaften von pyrazol, imidazol und 4-methylimidazol un ihrer Lösungen, besonders in benzol.* Z. Phys. Chem. Abt. A, 1940. **186**: p. 129.
150. Otting, W., Chemische Berichte Recueil, 1956. **89**: p. 2887.
151. Prałat, K. and G. Czechowski, *Investigation of imidazole self-association in non-polar solvents.* Journal of Molecular Liquids, 1988. **38**(3): p. 147-156.
152. Gontrani, L., et al., *Molecular aggregation phenomena in solution: an energy dispersive X-ray diffraction study of concentrated imidazole water solutions.* Chemical Physics Letters, 1999. **301**(1): p. 131-137.
153. Zischang, J., J.J. Lee, and M.A. Suhm, *Communication: Where does the first water molecule go in imidazole?* Journal of Chemical Physics, 2011. **135**(6).
154. Munowitz, M., et al., *Acid-Base and Tautomeric Equilibria in the Solid-State - N-15 Nmr-Spectroscopy of Histidine and Imidazole.* Journal of the American Chemical Society, 1982. **104**(5): p. 1192-1196.
155. Bachovchin, W.W. and J.D. Roberts, *N-15 Nuclear Magnetic-Resonance Spectroscopy - State of Histidine in Catalytic Triad of Alpha-Lytic Protease - Implications for Charge-Relay Mechanism of Peptide-Bond Cleavage by Serine Proteases.* Journal of the American Chemical Society, 1978. **100**(26): p. 8041-8047.

156. Hickman, B.S., et al., *Protonic conduction in imidazole: A solid-state N-15 NMR study.* Journal of the American Chemical Society, 1999. **121**(49): p. 11486-11490.
157. Nesmeyanov, A.N., et al., *H-1 and C-13 Nmr-Study of Tautomerism in Azoles .1. Proton-Transfer in Absence of Solvation.* Tetrahedron, 1975. **31**(11-1): p. 1461-1462.
158. Goward, G.R., et al., *High-resolution solid-state NMR studies of imidazole-based proton conductors: Structure motifs and chemical exchange from H-1 NMR.* Journal of Physical Chemistry B, 2002. **106**(36): p. 9322-9334.
159. Chang, K.C. and E. Grunwald, *Bifunctional Proton-Transfer of Conjugate Base of Uracil Compared with That of Imidazole in Aqueous-Solution.* Journal of the American Chemical Society, 1976. **98**(12): p. 3737-3738.
160. Al-Madhagi, L.H., et al., *X-ray Raman scattering: a new in situ probe of molecular structure during nucleation and crystallization from liquid solutions.* CrystEngComm, 2018. **20**(43): p. 6871-6884.
161. Meyer, F., et al., *Site-specific electronic structure of imidazole and imidazolium in aqueous solutions.* Physical Chemistry Chemical Physics, 2018. **20**(12): p. 8302-8310.
162. Liem, S.Y., M.S. Shaik, and P.L.A. Popelier, *Aqueous Imidazole Solutions: A Structural Perspective from Simulations with High-Rank Electrostatic Multipole Moments.* The Journal of Physical Chemistry B, 2011. **115**(39): p. 11389-11398.
163. Pagliai, M., et al., *Imidazole in Aqueous Solution: Hydrogen Bond Interactions and Structural Reorganization with Concentration.* The Journal of Physical Chemistry B, 2019. **123**(18): p. 4055-4064.
164. Bunker, G., *Introduction to XAFS: A Practical Guide to X-ray Absorption Fine Structure Spectroscopy.* 2010, Cambridge: Cambridge University Press.
165. Pynn, R., *Neutron Scattering: A Primer*. 1990, Los Alamos Science 19.
166. Smith, B.C., *Fundamentals of Fourier transform infrared spectroscopy*. 2nd ed. ed. 2011, Boca Raton, FL: CRC Press.
167. Lange, K.M. and E.F. Aziz, *Electronic Structure of Ions and Molecules in Solution: a View from Modern Soft X-ray Spectroscopies.* Chemical Society Reviews, 2013. **42**(16): p. 6840-6859.
168. Stöhr, J., *NEXAFS spectroscopy*. 1st ed. Springer series in surface sciences. 1996, Berlin ; New York: Springer. xv, 403 p.
169. Gudat, W. and C. Kunz, *Close Similarity between Photoelectric Yield and Photoabsorption Spectra in Soft-X-Ray Range.* Physical Review Letters, 1972. **29**(3): p. 169-172.
170. Jaklevic, J., et al., *Fluorescence Detection of Exafs - Sensitivity Enhancement for Dilute Species and Thin-Films.* Solid State Communications, 1977. **23**(9): p. 679-682.
171. Achkar, A.J., et al., *Bulk Sensitive X-ray Absorption Spectroscopy Free of Self-absorption Effects.* Physical Review B, 2011. **83**(8): p. 0811061 - 0811064.
172. Eisebitt, S., et al., *Determination of Absorption-Coefficients for Concentrated Samples by Fluorescence Detection.* Physical Review B, 1993. **47**(21): p. 14103-14109.
173. Troger, L., et al., *Full Correction of the Self-Absorption in Soft-Fluorescence Extended X-Ray-Absorption Fine-Structure.* Physical Review B, 1992. **46**(6): p. 3283-3289.
174. Mackova, A., et al., *Surface Analysis Techniques* in *Handbook of Spectroscopy* G.a.T.V.-D. Gauglitz, Editor. 2003, Wiley-VCH: Weinheim; Cambridge. p. 498 - 599.
175. Newville, M., *Fundamentals of XAFS*. 2004, University of Chicago: Chicago.
176. Sahle, C.J., et al., *Planning, Performing and Analyzing X-ray Raman Scattering Experiments.* Journal of Synchrotron Radiation, 2015. **22**(2): p. 400-409.
177. Schülke, W., *Electron Dynamics by Inelastic X-Ray Scattering*. 2007 ed. 2007, Oxford: Oxford Univ. Press.
178. Bernal, S.a., *A practical introduction to beam physics and particle accelerators*.
179. Wangler, T.P., *Introduction to Linear Accelerators*. 1993, Los Alamos National Laboratories: New Mexico. p. 171.

180. Balerna A., M.S., *Introduction to Synchrotron Radiation*, in *Synchrotron Radiation*, B.F. Mobilio S., Meneghini C., Editor. 2015, Springer: Berlin, Heidelberg.
181. Winick, H., *Synchrotron Radiation Sources: A Primer*. 1994, Singapore: London: World Scientific.
182. Martins, M.N. and T.F. Silva, *Electron accelerators: History, applications, and perspectives.* Radiation Physics and Chemistry, 2014. **95**: p. 78-85.
183. Huotari, S., et al., *A large-solid-angle X-ray Raman scattering spectrometer at ID20 of the European Synchrotron Radiation Facility.* Journal of Synchrotron Radiation, 2017. **24**: p. 521-530.
184. Fister, T.T., et al., *Multielement spectrometer for efficient measurement of the momentum transfer dependence of inelastic x-ray scattering.* Review of Scientific Instruments, 2006. **77**(6).
185. Seidler, G.T., et al., *The LERIX user facility.* X-Ray Absorption Fine Structure-Xafs13, 2007. **882**: p. 911-+.
186. Sumner, T., *Dazzling History.* Science, 2014. **343**(6175): p. 1092.
187. *Nature Milestones Crystallography*. 2014 [cited 2019 8th January]; Available from: https://www.nature.com/milestones/milecrystal/timeline.html.
188. Fischer, H.E., A.C. Barnes, and P.S. Salmon, *Neutron and x-ray diffraction studies of liquids and glasses.* Reports on Progress in Physics, 2006. **69**(1): p. 233-299.
189. Petkov, V., *Pair Distribution Functions Analysis.* Characterization of Materials, 2012: p. 1-14.
190. Billinge, S.J.L. and M.G. Kanatzidis, *Beyond crystallography: the study of disorder, nanocrystallinity and crystallographically challenged materials with pair distribution functions.* Chemical Communications, 2004(7): p. 749-760.
191. Wagner, C.N.J., *Structure of Amorphous Alloy Films.* Journal of Vacuum Science & Technology, 1969. **6**(4): p. 650-+.
192. Sutter, J.P., et al., *Three-Energy Focusing Laue Monochromator for the Diamond Light Source X-Ray Pair Distribution Function Beamline I15-1.* Proceedings of the 12th International Conference on Synchrotron Radiation Instrumentation (Sri2015), 2016. **1741**.
193. White, C. and K. Page, *Advanced Simulation Techniques for Total Scattering Data.* Neutron News, 2013. **24**(1): p. 13-14.
194. Billinge, C.L.F.a.P.J.a.J.W.L.a.D.B.a.E.S.B.a.J.B.a.T.P.a.S.J.L., *PDFfit2 and PDFgui: computer programs for studying nanostructure in crystals.* Journal of Physics: Condensed Matter, 2007. **19**(33): p. 335219.
195. Coelho, A.A., P.A. Chater, and A. Kern, *Fast synthesis and refinement of the atomic pair distribution function.* Journal of Applied Crystallography, 2015. **48**(3): p. 869-875.
196. Metropolis, N., et al., *Equation of State Calculations by Fast Computing Machines.* The Journal of Chemical Physics, 1953. **21**(6): p. 1087-1092.
197. McGreevy, R.L. and L. Pusztai, *Reverse Monte Carlo Simulation: A New Technique for the Determination of Disordered Structures.* Molecular Simulation, 1988. **1**(6): p. 359-367.
198. Soper, A.K., *Tests of the empirical potential structure refinement method and a new method of application to neutron diffraction data on water.* Molecular Physics, 2001. **99**(17): p. 1503-1516.
199. Soper, A.K., RAL Technical Reports, ((RAL-TR-2011-013, 2011)).
200. Youngs, T.G.A. *Dlputils*. 2017; Available from: https://www.projectaten.com/dlputils.
201. McGreevy, R.L., *Reverse Monte Carlo modelling.* Journal of Physics: Condensed Matter, 2001. **13**(46): p. R877.
202. Matthew, G.T.a.D.A.K.a.M.T.D.a.A.L.G.a.Q.H., *RMCProfile: reverse Monte Carlo for polycrystalline materials.* Journal of Physics: Condensed Matter, 2007. **19**(33): p. 335218.

203. Soper, A.K., *Joint structure refinement of x-ray and neutron diffraction data on disordered materials: application to liquid water.* Journal of Physics: Condensed Matter, 2007. **19**(33): p. 335206.
204. Soper, A.K., *The Radial Distribution Functions of Water as Derived from Radiation Total Scattering Experiments: Is There Anything We Can Say for Sure?* ISRN Physical Chemistry, 2013. **2013**: p. 67.
205. Soper, A.K., *The radial distribution functions of water and ice from 220 to 673 K and at pressures up to 400 MPa.* Chemical Physics, 2000. **258**(2): p. 121-137.
206. Dixit, S., et al., *Molecular segregation observed in a concentrated alcohol–water solution.* Nature, 2002. **416**: p. 829.
207. Dougan, L., et al., *Methanol-water solutions: A bi-percolating liquid mixture.* The Journal of Chemical Physics, 2004. **121**(13): p. 6456-6462.
208. Neese, F., *The ORCA program system.* Wiley Interdisciplinary Reviews: Computational Molecular Science, 2012. **2**(1): p. 73-78.
209. Stephens, P.J., et al., *Ab-Initio Calculation of Vibrational Absorption and Circular-Dichroism Spectra Using Density-Functional Force-Fields.* Journal of Physical Chemistry, 1994. **98**(45): p. 11623-11627.
210. Zheng, J.J., X.F. Xu, and D.G. Truhlar, *Minimally augmented Karlsruhe basis sets.* Theoretical Chemistry Accounts, 2011. **128**(3): p. 295-305.
211. Grimme, S., et al., *A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu.* Journal of Chemical Physics, 2010. **132**(15).
212. Neese, F., et al., *Efficient, approximate and parallel Hartree-Fock and hybrid DFT calculations. A 'chain-of-spheres' algorithm for the Hartree-Fock exchange.* Chemical Physics, 2009. **356**(1-3): p. 98-109.
213. Hanwell, M.D., et al., *Avogadro: an advanced semantic chemical editor, visualization, and analysis platform.* Journal of Cheminformatics, 2012. **4**.
214. Sahle, C.J., et al., *A miniature closed-circle flow cell for high photon flux X-ray scattering experiments.* Journal of Synchrotron Radiation, 2015. **22**: p. 1555-1558.
215. Juurinen, I., et al., *Saturation Behavior in X-ray Raman Scattering Spectra of Aqueous LiCl.* Journal of Physical Chemistry B, 2013. **117**(51): p. 16506-16511.
216. Pylkkanen, T., et al., *Universal Signature of Hydrogen Bonding in the Oxygen K-Edge Spectrum of Alcohols.* Journal of Physical Chemistry B, 2010. **114**(41): p. 13076-13083.
217. Thomason, M.J., *Soft X-ray Spectroscopy of Molecular Species in Solution: Studies of Imidazole and Imidazole/Water Systems*, in *Faculty of Engineering and Physical Sciences, School of Chemistry*. 2013, The University of Manchester: Manchester. p. 148.
218. Anwar, J. and D. Zahn, *Uncovering Molecular Processes in Crystal Nucleation and Growth by Using Molecular Simulation.* Angewandte Chemie-International Edition, 2011. **50**(9): p. 1996-2013.
219. Soper, A.K., *Partial structure factors from disordered materials diffraction data: An approach using empirical potential structure refinement.* Physical Review B, 2005. **72**(10).
220. Johnston, A.J., et al., *On the atomic structure of cocaine in solution.* Physical Chemistry Chemical Physics, 2016. **18**(2): p. 991-999.
221. Finney, J.L. and A.K. Soper, *Solvent Structure and Perturbations in Solutions of Chemical and Biological Importance.* Chemical Society Reviews, 1994. **23**(1): p. 1-10.
222. Enderby, J.E., *Ion Solvation Via Neutron-Scattering.* Chemical Society Reviews, 1995. **24**(3): p. 159-168.
223. Bowron, D.T., J.L. Finney, and A.K. Soper, *Structural investigation of solute-solute interactions in aqueous solutions of tertiary butanol.* Journal of Physical Chemistry B, 1998. **102**(18): p. 3551-3563.

224. Hannon, A.C., W.S. Howells, and A.K. Soper, *Atlas - a Suite of Programs for the Analysis of Time-of-Flight Neutron-Diffraction Data from Liquid and Amorphous Samples.* Institute of Physics Conference Series, 1990(107): p. 193-211.
225. Stewart, J.J.P., *Special Issue - Mopac - a Semiempirical Molecular-Orbital Program.* Journal of Computer-Aided Molecular Design, 1990. **4**(1): p. 1-45.
226. Berendsen, H.J.C., J.R. Grigera, and T.P. Straatsma, *The Missing Term in Effective Pair Potentials.* Journal of Physical Chemistry, 1987. **91**(24): p. 6269-6271.
227. Jorgensen, W.L., D.S. Maxwell, and J. TiradoRives, *Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids.* Journal of the American Chemical Society, 1996. **118**(45): p. 11225-11236.
228. Soper, A.K., *The excluded volume effect in confined fluids and liquid mixtures.* Journal of Physics-Condensed Matter, 1997. **9**(11): p. 2399-2410.
229. Luzar, A. and D. Chandler, *Structure and Hydrogen-Bond Dynamics of Water-Dimethyl Sulfoxide Mixtures by Computer-Simulations.* Journal of Chemical Physics, 1993. **98**(10): p. 8160-8173.
230. Leetmaa, M., et al., *Are recent water models obtained by fitting diffraction data consistent with infrared/Raman and x-ray absorption spectra?* Journal of Chemical Physics, 2006. **125**(24).
231. Verma, P.K., et al., *The Bend+Libration Combination Band Is an Intrinsic, Collective, and Strongly Solute-Dependent Reporter on the Hydrogen Bonding Network of Liquid Water.* The Journal of Physical Chemistry B, 2018. **122**(9): p. 2587-2599.
232. Will, G., *Die Kristallstruktur von Imidazol, C3N2H4, bei Zimmertemperatur*, in *Zeitschrift für Kristallographie - Crystalline Materials*. 1969. p. 211.
233. Headen, T.F., et al., *Structure of pi-pi Interactions in Aromatic Liquids.* Journal of the American Chemical Society, 2010. **132**(16): p. 5735-5742.
234. Chourasia, M., G.M. Sastry, and G.N. Sastry, *Aromatic-Aromatic Interactions Database, A(2)ID: An analysis of aromatic pi-networks in proteins.* International Journal of Biological Macromolecules, 2011. **48**(4): p. 540-552.
235. Turner, A.H., et al., *Applying neutron diffraction with isotopic substitution to the structure and proton-transport pathways in protic imidazolium bis{(trifluoromethyl)sulfonyl}imide ionic liquids.* Faraday Discussions, 2018. **206**: p. 247-263.
236. Stevens, J.S., et al., *NEXAFS and XPS of p-Aminobenzoic Acid Polymorphs: The Influence of Local Environment.* 16th International Conference on X-Ray Absorption Fine Structure (Xafs16), 2016. **712**.
237. Stevens, J.S., et al., *Chemical Speciation and Bond Lengths of Organic Solutes by Core-Level Spectroscopy: pH and Solvent Influence on p-Aminobenzoic Acid.* Chemistry-a European Journal, 2015. **21**(19): p. 7256-7263.
238. Stevens, J.S., et al., *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): p. 1776-1783.
239. Ravel, B. and M. Newville, *ATHENA, ARTEMIS, HEPHAESTUS: Data Analysis for X-ray Absorption Spectroscopy using IFEFFIT.* Journal of Synchrotron Radiation, 2005. **12**(4): p. 537-541.
240. Barone, V. and M. Cossi, *Quantum calculation of molecular energies and energy gradients in solution by a conductor solvent model.* Journal of Physical Chemistry A, 1998. **102**(11): p. 1995-2001.
241. Neese, F., *Prediction of molecular properties and molecular spectroscopy with density functional theory: From fundamental theory to exchange-coupling.* Coordination Chemistry Reviews, 2009. **253**(5-6): p. 526-563.
242. van Leeuwen, R., O.V. Gritsenko, and E.J. Baerends, *Analysis and modelling of atomic and molecular kohn-sham potentials*, in *Density Functional Theory I: Functionals and*

Effective Potentials, R.F. Nalewajski, Editor. 1996, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 107-167.
243. Sette, F., J. Stohr, and A.P. Hitchcock, *Determination of Intramolecular Bond Lengths in Gas-Phase Molecules from K-Shell Shape Resonances.* Journal of Chemical Physics, 1984. **81**(11): p. 4906-4914.
244. Wojdyr, M., *Fityk: a general-purpose peak fitting program.* Journal of Applied Crystallography, 2010. **43**(5 Part 1): p. 1126-1128.
245. Stevens, J.S., et al., *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.* Journal of Physical Chemistry B, 2010. **114**(44): p. 13961-13969.
246. Stevens, J.S., et al., *Quantitative analysis of complex amino acids and RGD peptides by X-ray photoelectron spectroscopy (XPS).* Surface and Interface Analysis, 2013. **45**(8): p. 1238-1246.
247. Soderholm, L., et al., *X-ray excited optical luminescence (XEOL) detection of x-ray absorption fine structure (XAFS).* Journal of Chemical Physics, 1998. **109**(16): p. 6745-6752.
248. Soper, A.K. and E.R. Barney, *Extracting the pair distribution function from white-beam X-ray total scattering data.* Journal of Applied Crystallography, 2011. **44**: p. 714-726.
249. Billes, F., H. Endredi, and G. Jalsovszky, *Vibrational spectroscopy of diazoles.* Journal of Molecular Structure-Theochem, 1999. **465**(2-3): p. 157-172.
250. Zimmermann, H., *Über den intermolekularen Protonenübergang und über die Zustände des Protons in der Wasserstoffbrückenbindung von Imidazol.* Zeitschrift für Elektrochemie, Berichte der Bunsengesellschaft für physikalische Chemie, 1961. **65**(10): p. 821-840.
251. Garfinkel, D. and J.T. Edsall, *Raman Spectra of Amino Acids and Related Compounds. VIII. Raman and Infrared Spectra of Imidazole, 4-Methylimidazole and Histidine1-3.* Journal of the American Chemical Society, 1958. **80**(15): p. 3807-3812.
252. Bellocq, A.-M., et al., *Spectres de vibration de l'imidazole, de l'imidazole (D)-1, de l'imidazole (D3)-2,4,5 et de l'imidazole (D4).* J. Chim. Phys., 1965. **62**: p. 1334-1343.
253. Perchard, C., A.-M. Bellocq, and A. Novak, *Spectres de vibration de l'imidazole, de l'imidazole (D) — 1, de l'imidazole (D3)-2,4,5 et de l'imidazole (D4).* J. Chim. Phys., 1965. **62**: p. 1344-1358.
254. Perchard, C. and A. Novak, *Far-Infrared Spectra and Hydrogen-Bond Frequencies of Imidazole.* Journal of Chemical Physics, 1968. **48**(7): p. 3079-&.
255. Majoube, M., *Raman and Ir-Spectra of the Nh and No Stretching Region in Polycrystalline Imidazole and Some of Its N-15 and D-Substituted Analogs.* Journal of Molecular Structure, 1980. **61**(Jan): p. 129-136.
256. Wolff, H. and H. Müller, *Structure of the NH stretching vibrational band of pyrazole. Multiple resonance of substances forming strong H or D bonds.* Spectrochimica Acta Part A: Molecular Spectroscopy, 1976. **32**(3): p. 581-585.
257. Sahle, C.J., et al., *Microscopic structure of water at elevated pressures and temperatures.* Proceedings of the National Academy of Sciences of the United States of America, 2013. **110**(16): p. 6301-6306.
258. Cai, Y.Q., et al., *Ordering of hydrogen bonds in high-pressure low-temperature H2O.* Physical Review Letters, 2005. **94**(2).
259. Pylkkanen, T., et al., *Role of Non-Hydrogen-Bonded Molecules in the Oxygen K-Edge Spectrum of Ice.* Journal of Physical Chemistry B, 2010. **114**(11): p. 3804-3808.
260. Cavalleri, M., et al., *The interpretation of X-ray absorption spectra of water and ice.* Chemical Physics Letters, 2002. **364**(3-4): p. 363-370.
261. Chen, W., X.F. Wu, and R. Car, *X-Ray Absorption Signatures of the Molecular Environment in Water and Ice.* Physical Review Letters, 2010. **105**(1).

262. Pylkkanen, T., et al., *Temperature Dependence of the Near-Edge Spectrum of Water.* Journal of Physical Chemistry B, 2011. **115**(49): p. 14544-14550.
263. Booth, S.G., et al., *In situ XAFS Study of Palladium Electrodeposition at the Liquid/Liquid Interface.* Electrochimica Acta, 2017. **235**: p. 251-261.
264. Chang, S.Y., et al., *Detection and characterisation of sub-critical nuclei during reactive Pd metal nucleation by X-ray absorption spectroscopy.* Crystengcomm, 2016. **18**(5): p. 674-682.
265. Chang, S.Y., et al., *Energy dispersive-EXAFS of Pd nucleation at a liquid/liquid interface.* 16th International Conference on X-Ray Absorption Fine Structure (Xafs16), 2016. **712**.
266. Zimmermann, N.E.R., et al., *Nucleation of NaCl from Aqueous Solution: Critical Sizes, Ion-Attachment Kinetics, and Rates.* Journal of the American Chemical Society, 2015. **137**(41): p. 13352-13361.
267. Anderson, D.M.W., J.L. Duncan, and F.J.C. Rossotti, *412. The hydrogen bonding of imidazole in carbon tetrachloride solution.* Journal of the Chemical Society (Resumed), 1961(0): p. 2165-2171.
268. Wang, S.-M., L.-Y. Lee, and J.-T. Chen, *Proton magnetic resonance studies on the self-association and hydrogen bonding of imidazole in chloroform solutions.* Spectrochimica Acta Part A: Molecular Spectroscopy, 1979. **35**(7): p. 765-771.
269. Lam, R.K., J.W. Smith, and R.J. Saykally, *Communication: Hydrogen bonding interactions in water-alcohol mixtures from X-ray absorption spectroscopy.* The Journal of Chemical Physics, 2016. **144**(19): p. 191103.
270. Calvin, S., *XAFS for Everyone*. 2013, Boca Raton:London:New York: CRC Press.
271. Ressler, T., *WinXAS: a Program for X-ray Absorption Spectroscopy Data Analysis under MS-Windows.* Journal of Synchrotron Radiation, 1998. **5**(2): p. 118-122.
272. Chang, S.Y., et al., *Automated analysis of XANES: A feasibility study of Au reference compounds.* Journal of Physics: Conference Series, 2016. **712**: p. 012070.
273. Rehr, J.J. and R.C. Albers, *Theoretical approaches to x-ray absorption fine structure.* Reviews of Modern Physics, 2000. **72**(3): p. 621-654.
274. Rehr, J.J. and A.L. Ankudinov, *Progress in the theory and interpretation of XANES.* Coordination Chemistry Reviews, 2005. **249**(1–2): p. 131-140.
275. Isao, T. and M. Teruyasu, *First-principles calculations of x-ray absorption near edge structure and energy loss near edge structure: present and future.* Journal of Physics: Condensed Matter, 2009. **21**(10): p. 104201.
276. Kas, J., *Toward Quantitative Calculation and Analysis of X-Ray Absorption Near Edge Spectra*. 2009, University of Washington: Washington.
277. Joly, Y., *X-ray absorption near-edge structure calculations beyond the muffin-tin approximation.* Physical Review B, 2001. **63**(12): p. 125120.
278. Benfatto, M. and S. Della Longa, *Geometrical fitting of experimental XANES spectra by a full multiple-scattering procedure.* J Synchrotron Radiat, 2001. **8**(4): p. 1087-94.
279. Smolentsev, G. and A.V. Soldatov, *FitIt: New software to extract structural information on the basis of XANES fitting.* Computational Materials Science, 2007. **39**(3): p. 569-574.
280. Scott, R.A., et al., *Bottlenecks and roadblocks in high-throughput XAS for structural genomics.* Journal of Synchrotron Radiation, 2005. **12**(1): p. 19-22.
281. Baudelet, F., et al., *ODE: a new beam line for high-pressure XAS and XMCD studies at SOLEIL.* High Pressure Research, 2011. **31**(1): p. 136-139.
282. Belin, S., et al., *SAMBA a New Beamline at SOLEIL for XRay Absorption Spectroscopy in the 440keV Energy Range.* Physica Scripta, 2005: p. 980.
283. Diaz-Moreno, S., et al., *I20; the Versatile X-ray Absorption spectroscopy beamline at Diamond Light Source.* Journal of Physics: Conference Series, 2009. **190**: p. 012038.
284. EasyXAFS. *EasyXAFS*. [cited 2019 27 November]; Available from: http://easyxafs.com/.
285. *Anaconda Python - Python version 3.5.2*. [cited 2018 17th May]; Available from: https://docs.anaconda.com/anaconda/packages/oldpkglists/.

286. Krause, M.O. and J.H. Oliver, *Natural widths of atomic K and L levels, Kα X-ray lines and several KLL Auger lines.* Journal of Physical and Chemical Reference Data, 1979. **8**(2): p. 329-338.
287. Matthew Newville, T.S.a.o. *LMFIT (Non-Linear Least-Square Minimization and Curve-Fitting for Python)*. 2019 [cited 2018 17th May]; Available from: https://lmfit.github.io/lmfit-py/.
288. Al-Madhagi, L.H.A., *XAS_analysis*. 2019, GitHub repository: https://github.com/LailaAlMadhagi/XAS_analysis.
289. Al-Madhagi, L.H.A., *Automated XAS Analysis Code*. 2018, YouTube: https://youtu.be/ZQvs84MoSkY.
290. Hitchcock, A.P., *Gas Phase Core Excitation Database*. 2003: http://unicorn.chemistry.mcmaster.ca/corex/cedb-title.html.

Appendix I: Individual C and N K-edge XRS spectra from Chapter 6

Figures A6.1-A6.5 show the individual N K-edge XRS spectra of imidazole solid, aqueous solution at 35 °C, aqueous solution at 20 °C, aqueous solution at 13 °C and aqueous solution after crystallisation. The individual spectra do not show a sign of time-dependent changes that could be attributed to radiation induced damage. This is expected because spinning of the solid sample and the large stock volume (~250 ml) of the aqueous solution accompined with constact circulation would minimise radiation damage. The same is noticed for the C K-edge individual spectra illustrated in Figures A6.6-A6.9

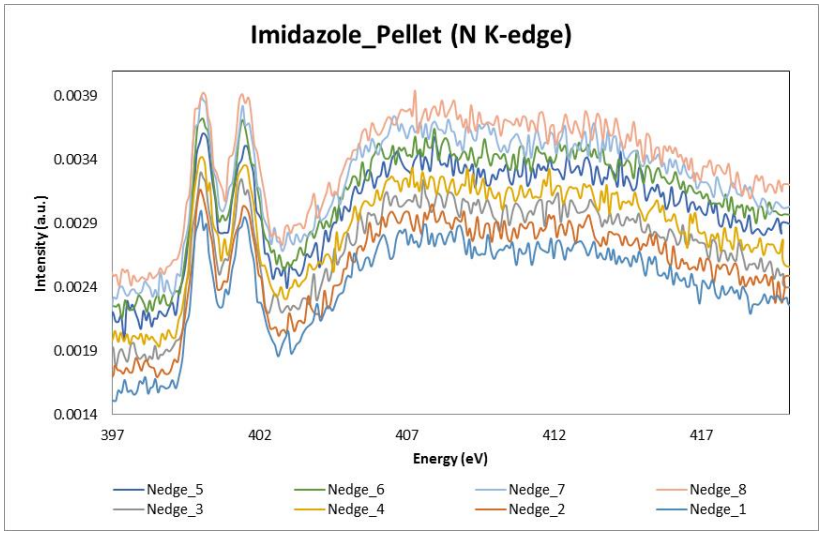

Figure A.1. Individual N K-edge XRS spectra of imidazole solid.

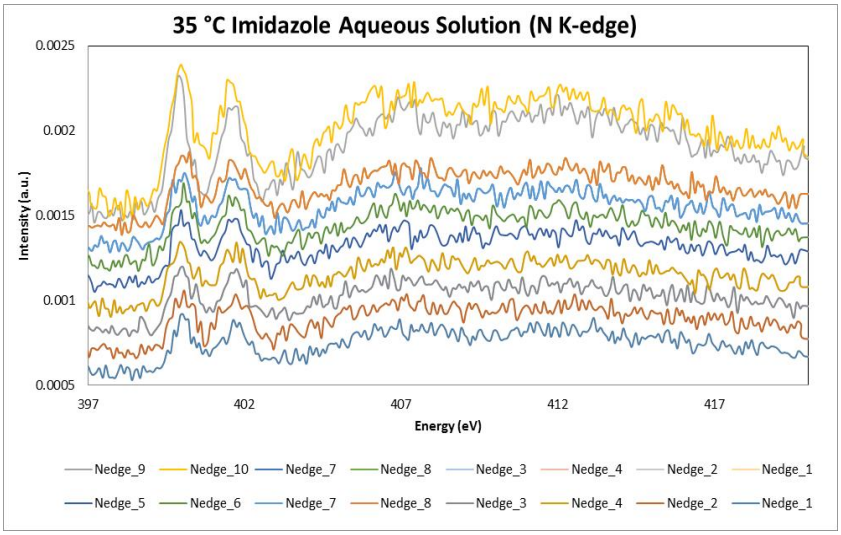

Figure A.2. Individual N K-edge XRS spectra of the 10.1 M aqueous imidazole solution at 35°C.

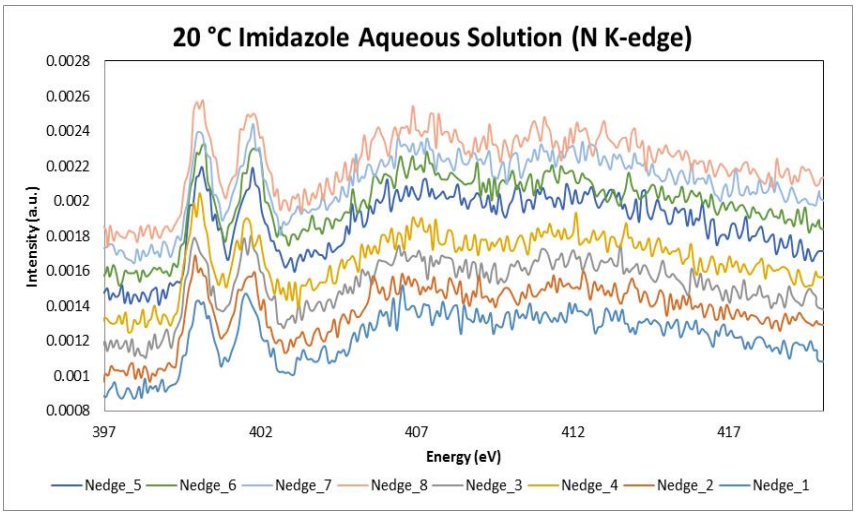

Figure A.3. Individual N K-edge XRS spectra of the 10.1 M aqueous imidazole solution 20°C.

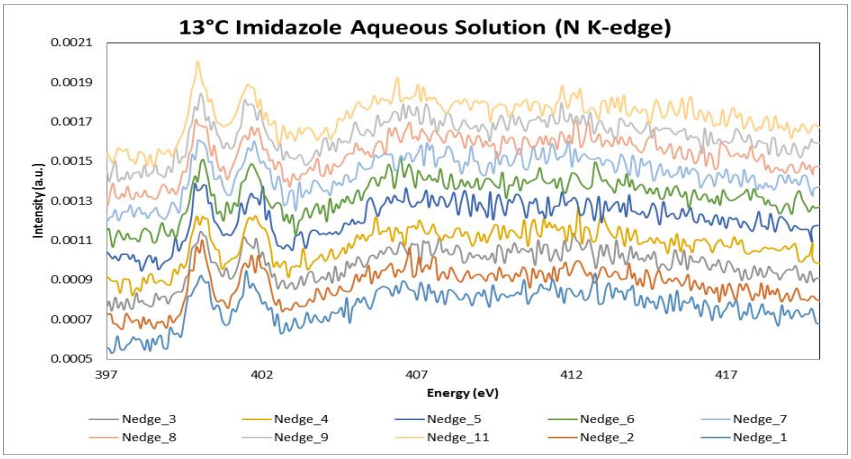

Figure A.4. Individual N K-edge XRS spectra of the 10.1 M aqueous imidazole solution at 13°C.

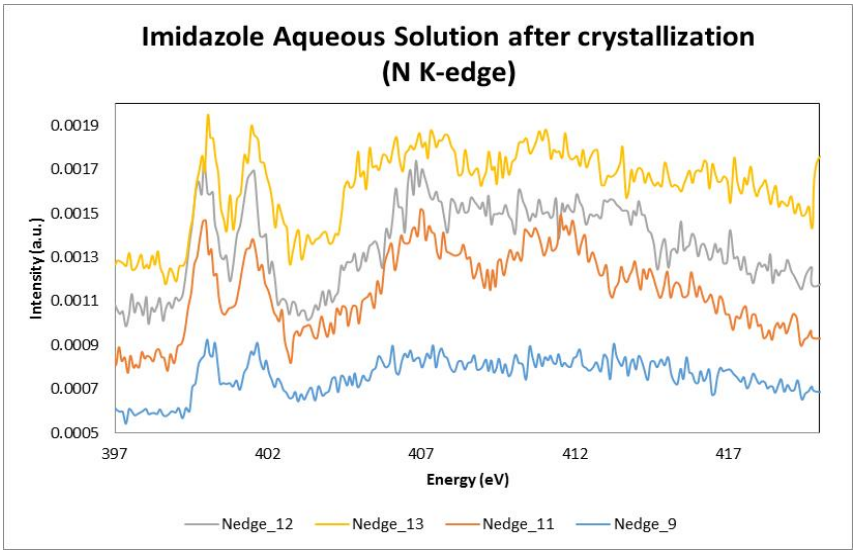

Figure A.5. Individual N K-edge XRS spectra of the 10.1 M aqueous imidazole solution after crystallisation.

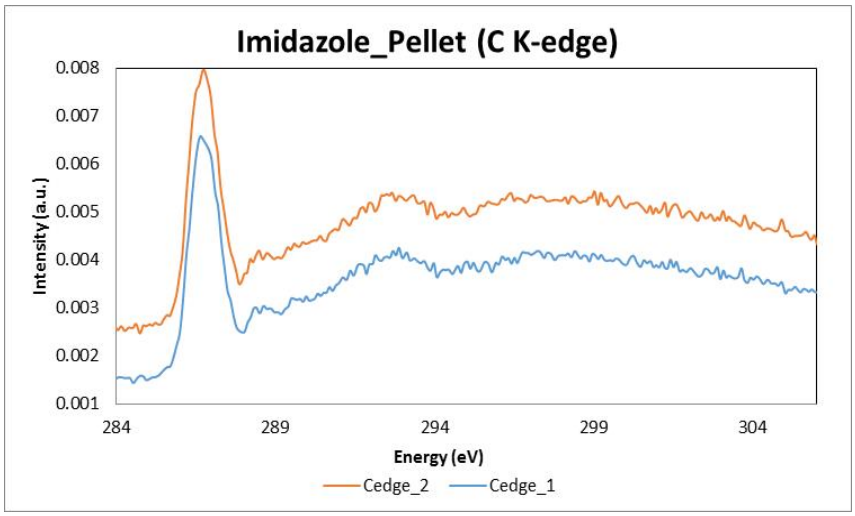

Figure A.6. Individual C K-edge XRS spectra of solid imidazole.

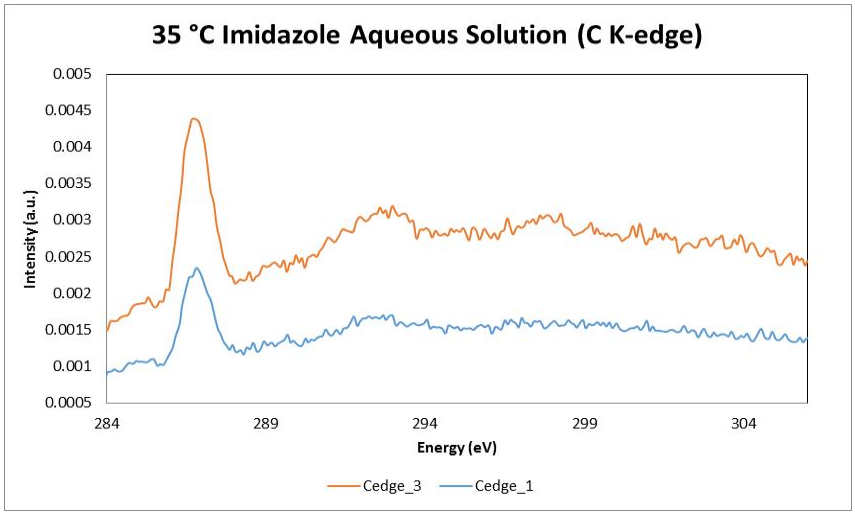

Figure A.7. Individual C K-edge XRS spectra of the 10.1 M aqueous imidazole solution at 35°C.

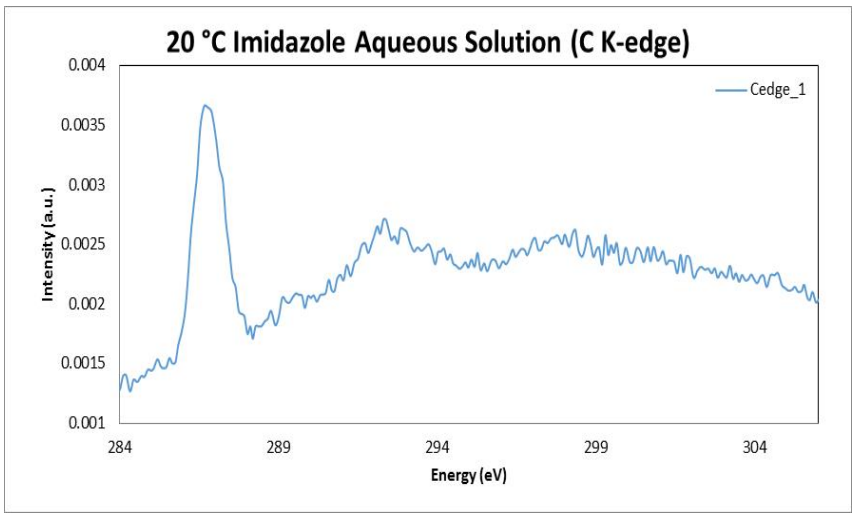

Figure A.8. Individual C K-edge XRS spectra of the 10.1 M aqueous imidazole solution at 20°C.

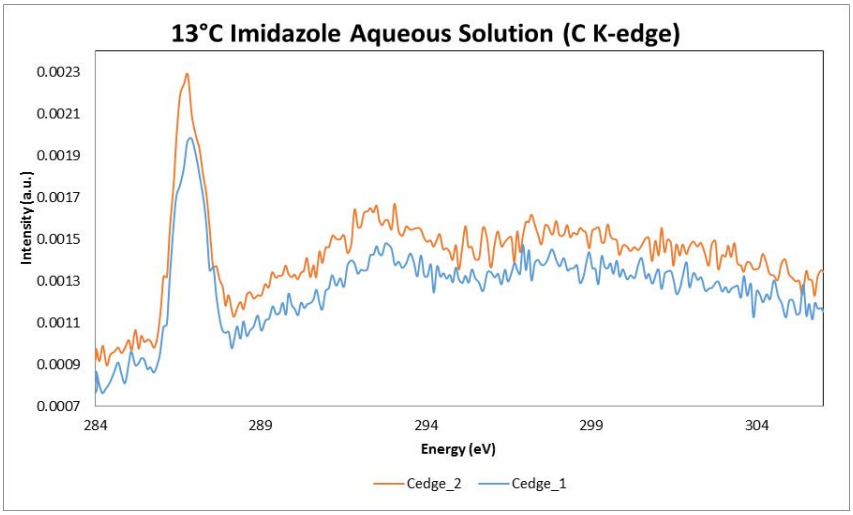

Figure A.9. Individual C K-edge XRS spectra of the 10.1 M aqueous imidazole solution at 13°C.