Soft X-ray Core Level and UV-Vis Spectroscopies Applied to Organic Molecules in the Solid State and in Solution

A thesis submitted to The University of Manchester for the degree of

Doctor of Philosophy

in the Faculty of Engineering and Physical Sciences

2016

Adrian Gainar

School of Chemical Engineering and Analytical Science

The University of Manchester

Table of Contents

TOTAL WORD COUNT: 43835

List of Figures

List of Tables

List of Equations

List of Abbreviations

ATP	Adenosine triphosphate
BNL	Brookhaven National Laboratory
CCD	Charge-coupled device
CNT	Classical nucleation theory
DFT	Density functional theory
DMSO	Dimethylsulfoxide
DOS	Density of states
EPSR	Empirical potential structure refinement
EXAFS	Extended X-ray absorption fine structure
FMS	Full multiple scattering
FTIR	Fourier transformed infrared spectroscopy
FY	Fluorescence yield
HMT	Hexamethylenetetramine
HOMO	Highest occupied molecular orbital
HPLC	High-performance liquid chromatography
HZB	Helmholtz Zentrum Berlin
INA	Isonicotinamide
IP	Ionisation potential
IR	Infrared spectroscopy
IUPAC	International Union of Pure and Applied Chemistry
LDA	Local density approximation
LUMO	Lowest unoccupied molecular orbital
MCP	Microchannel plate
MO	Molecular orbital
NEXAFS	Near-edge X-ray absorption fine structure
NMR	Nuclear magnetic resonance
NSLS	National Synchrotron Light Source
PABA	*Para*-aminobenzoic acid
PCM	Polarisable continuum model
PEY	Partial electron yield
PFY	Partial fluorescence yield
RIXS	Resonant inelastic X-ray scattering
RNA	Ribonucleic acid
RXES	Resonant X-ray emission spectroscopy
SAXS	Small-angle X-ray scattering
SCF	Self-consistent field
SCRF	Self-consistent reaction field
TD-DFT	Time-dependent density functional theory
TEY	Total electron yield
TS	Total spectrum
UV-Vis	Ultraviolet-visible spectroscopy
VLS	Varied line spacing
WAXS	Wide-angle X-ray scattering
XANES	X-ray absorption near-edge structure
XAS	X-ray absorption spectroscopy
XES	X-ray emission spectroscopy

XPS	X-ray photoelectron spectroscopy
XRD	X-ray diffraction
XRPD	X-ray powder diffraction

ABSTRACT

The University of Manchester
Adrian Gainar
For the degree of Doctor of Philosophy

Soft X-ray Core Level and UV-Vis Spectroscopies Applied to Organic Molecules in the Solid State and in Solution

March 2016

The process of homogeneous nucleation from solution involves an incipient stage preceding spontaneous crystal growth, in which an embryo-like cluster is formed in a supersaturated homogeneous phase. Elucidating the physicochemical processes that govern nucleation (*e.g.* solute speciation, solute-solvent interactions/solvation, solute-solute 'self'-association, as well as charge density and bond length variations in the solute molecules) is often referred to as one of the remaining grand challenges in the physical sciences. Laboratory analytical techniques such as IR, Raman, NMR spectroscopy are in principle sensitive to the relevant molecular level changes in solution, but the type of information they provide is often only indirectly related to molecular structure. This dissertation shows that core-level X-ray spectroscopies (near-edge X-ray absorption fine-structure, NEXAFS, and resonant inelastic X-ray scattering, RIXS) are emerging as versatile techniques that provide more direct insight into the electronic and geometrical structure of molecular species in solution and in the solid state. Another idea explored in this dissertation is that laboratory UV-Vis spectroscopy provides complementary information to NEXAFS and RIXS on the occupied and unoccupied states in the valence region of the solutes. The dissertation reports these studies through a series of work packages, as follows.

First, a study of a series of crystalline saccharides was carried out to examine the hypothesis that even subtle changes in intramolecular bond lengths should manifest through detectable shifts in the shape resonances in the NEXAFS spectra even for the solid state of relatively complex organic molecules. Detection of C–O bond length variations in the saccharides, including non-crystalline solids without long-range order, demonstrates the applicability of the 'bond length with a ruler' approach and thereby extends its use for gas-phase and adsorbed molecule NEXAFS to the organic solid state.

Second, it is shown in two chapters how pH changes and the nature of the solvent induce chemical state and local electronic structure variations in solute molecules, including the formation of cationic and anionic species. The high sensitivity to bond length changes in dissolved molecular species is shown for *para*-aminobenzoic acid (PABA) through comparison with reference NEXAFS data for the solid state. This technique thus provides an avenue to computational predictions of solute properties and solute-solvent interactions.

Finally, combining NEXAFS and RIXS PABA in solution with UV-Vis spectra demonstrates that UV absorption peak shifts associated with changes in the electronic structure triggered by pH variation, are compatible with the electronic structure changes observed by NEXAFS/RIXS. Moreover, a shift of the main UV absorption band to shorter wavelengths as a function of solvent polarity is noticed, which further contributes to an understanding of solute-solvent interactions in solution.

DECLARATION

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Adrian Gainar

The University of Manchester

COPYRIGHT STATEMENT

ACKOWLEDGEMENTS

I would like to thank my supervisor, Prof Sven Schroeder, for being the guide towards understanding the fascinating ‘world’ of X-ray spectroscopies.

I need to thank Dr Joanna Stevens, who has provided enormous support during my PhD, dedicating her time to advise me on data analysis, beamline experiments, and closely collaborating throughout the entire project.

My personal financial support throughout the PhD programme has been provided by EPSRC to whom I am extremely thankful.

Special thanks to our collaborators from the NSLS (Brookhaven, U.S.): Dr Cherno Jaye, Daniel Fischer, and from BESSYII (Berlin, Germany): Dr Jie Xiao, Dr Edlira Suljoti, Dr Mikhail Yablonskikh, Ronny Golnak, Ervis Suljoti, Tim Brandenburg and Prof Emad Aziz, for providing technical support and advice during and after beamline experiments.

My thanks go to my co-supervisor, Prof Roger Davey, who helped me with pieces of advice whenever crystallisation knowledge was involved, to the team of researchers from University of Leeds I have worked with on this project: Prof Kevin Roberts, Dr Robert Hammond, Dr Andrew Scott, Dr Xiaojun Lai, Dr Che Seabourne, Dr Dimitrios Toroz, Dr Thomas Turner and Dr Ian Rosbottom.

Also, I am very grateful to my colleagues from D floor: Dr Sara Gilks who will always be remembered, Dr Ghazala Sadiq, Dr Rachel Sullivan, Sharlinda Salim Sachithananthan, Till Koellges, Lauren Newton, Dr Adrian Hutchinson, Loris Doyle for support and assistance, to my colleagues on the F floor: Dr Elizabeth Willneff, Dr Joshua Goldberg, Dr Devon Indar and Amani Musharah for being excellent team mates.

Lastly but not last to thank are my mum, dad, and relatives, who morally sustained me even if I was miles away from home, and empowered me with pieces of advice even in delicate situations.

DEDICATION

This thesis is dedicated to the memory of my beloved father, Ioan Gainar, who passed away in October 2014. He offered me strong moral support, supported me with every bit of his heart for the chemistry contests and olympiads I participated to when I was younger, and although I miss him every single day, I am content that he has got to know I pursue a dedicated work towards fulfilling my PhD goals and finding a suitable job in academia or industry.

CHAPTER 1

INTRODUCTION

1.1. Nucleation

Materials crystallised from solutions are of great importance to industry and academic environments, and the form produced impacts on chemical and physical properties that are relevant to industrial processability and handling. Control on the self-assembling manner into which molecules form well-defined crystals, also known as 'crystal engineering', requires in-depth understanding of the mechanisms involved in solution prior to crystallisation.[1]

Nucleation is the stage at which an embryo-like cluster results from a supersaturated homogeneous phase, and precedes the crystal formation process. There are currently two main nucleation theories, both assuming that some form of molecular self-aggregation constitutes an intermediate step in the nucleation process. The classical nucleation theory (CNT)[2] takes into account that in a solution system, parallel density and degree order variations lead to formation of clusters with a molecular packing that is identical to a defined crystal.[3] The alternative view is that the crystal packing is preceded by a liquid-like cluster generating crystalline nuclei of higher order that will consequently lead to the crystal structure.[4]

Research in the field of nucleation has evolved rapidly owing to the advanced analytical techniques and molecular simulations available. UV-Vis spectroscopy, Raman, *in situ* FTIR, and NMR spectroscopy have all contributed to the overall advances in the knowledge of solution chemistry. However, the complex process of nucleation is still not understood in detail. Employing additional experimental techniques, supported by computational methods to predict spectra and create molecular simulations, allows the local chemical and electronic environment and structure to be probed, leading to a more detailed understanding of the nature and interactions of solution species and thereby potential evidence for the nature of nucleation.

1.2. Use of core-level spectroscopies

Detection of defined nuclei can be experimentally challenging due to their reduced sizes and generally short lifetimes. Various analytical techniques such as NMR, FTIR, Raman, UV-Vis, and neutron scattering are employed to reveal information on solute-solute and solute-solvent associations in solutions, and simultaneous small-angle and wide-angle X-ray spectroscopies (SAXS-WAXS) have been used in order to gain structural information on the nucleus phase.[5-7] Valuable levels of details that are otherwise difficult to attain by conventional techniques[7-9] can be achieved with core-level X-ray spectroscopies (*e.g.* X-ray photoelectron spectroscopy, XPS, and X-ray absorption spectroscopy, XAS).

The XAS, near-edge X-ray absorption fine structure (NEXAFS), probes the local chemical and electronic structure, being specific to the type of element and to the local environment of the X-ray absorbing atoms. This should be ideal to probe the stage preceding the nucleus formation, as the structural information in the nucleation germ depends on the neighbouring environment, allowing changes in the local environment as a function of time to be monitored using NEXAFS.[5]

An important aspect of the NEXAFS technique is that experimentally detectable energy shifts can be interpreted using specific codes[10,11] to simulate the spectra and provide a guiding analysis of the electronic transitions involved in the experiments.

1.3. Aims and objectives

The PhD programme constitutes a part of the overall project, which is a collaboration between partners from The University of Manchester and University of Leeds titled 'Molecules, Clusters and Crystals: A Multi-Scale Approach to Understanding Kinetic Pathways in Crystal Nucleation from Solution'. The main objective is to build a more complete understanding of the mechanisms encountered in the solid and solution state of organic systems across molecular, mesoscopic and macroscopic scales through structural aggregation from individual molecules to macroscopic crystals. The novelty of the project

consists in combining for the first time experimental and computational methods with *in situ* synchrotron techniques that probe local electronic and structural changes.

The methodology plan integrates 6 complex work programmes, described as below:

- Work programme 1 consists in managerial and administrative support for the project in terms of organising project meetings, semi-annual reports and annual meetings with the Advisory Board. IT support will be provided to facilitate data acquisition and storage.
- Work programme 2 aims to define the characteristic thermodynamic data, to experimentally verify the presence of dimers/higher-order aggregates, to model molecular association in solution, and to quantify the nucleation rates/induction time measurements.
- Work programme 3 refers to single crystal growth (influence of solvent, supersaturation and temperature on crystal morphology) and modelling of growth interface
- Work programme 4 consists in synchrotron data collection through the available infrastructure located at BESSY (liquid microjet NEXAFS and RIXS) and NSLS (solid-state NEXAFS), as well as in calculations that are consistent with the experimental data
- Work programme 5 aims to develop optimised cluster models corresponding to the measured bulk systems in the working programmes 2 and 3
- Work programme 6 has the objective of integrating the project results and presenting them at conferences or in scientific journals to generate a high impact within the international research community

In the context of the overall project, the applications of soft X-ray spectroscopies under experimental conditions to systems where nucleation is taking place constitutes a *'Grand Challenge'* in that it necessitates experimentally difficult measurements of solution state condensed matter within fairly reduced timescales due to the short lifetime of nuclei. The windowless microjet setup developed by groups from BESSY[12-14] applied to dilute and concentrated solutions represents a revolutionary advantage as windows normally limit the time resolution and photon depth for the measurements. Consequently, the laminar flow region of the microjet allows measurements within the timescale of less than one second.

While the long-term objective is to observe nucleation or cluster formation, the specific aims of the PhD are centred on aspects that develop the fundamental understanding of the speciation, interactions and importance of local environment in condensed matter:

- To apply synchrotron solid and solution-state X-ray core level spectroscopies to organic materials (*i.e.* saccharides, *para*-aminobenzoic acid) for elucidating local electronic changes in the context of bond lengths, pH, concentration or solvent dependence
- To validate data acquired through the use of a density functional theory (DFT) or full multiple scattering (FMS) code for assisting with the interpretation of spectral features
- To establish a correlation between synchrotron NEXAFS-RIXS spectroscopies and home laboratory UV-Vis as pH is varied in aqueous solutions by combining experimental data with DFT and time-dependent density functional theory (TD-DFT) calculations
- To examine the level of solute self-associations in solutions through plots of UV-Vis absorption as a function of concentration for probing the predominance of monomers in aqueous and alcohol solutions.

1.4. Thesis outline by chapters

Chapter 2 (Literature Review) aims to highlight the general background of nucleation/crystallisation aspects, focussing especially on alternative techniques in the study of local electronic/structure changes such as NEXAFS.

Chapter 3 (Methodology) describes general aspects on the X-ray absorption/emission and UV-Vis spectroscopies, along with other supporting experimental techniques and theoretical approaches employed.

Chapter 4 (Solid-State NEXAFS of Saccharides) outlines the applicability of NEXAFS in detecting minute variations in the C–OH average bond lengths of saccharides, thus extending the applicability of the 'bond length with a ruler' concept to complex organic molecules, including non-crystalline systems without long-range order such as amorphous phases.

The work reflected in this chapter was published as *'NEXAFS Sensitivity to Bond Lengths in Complex Molecular Materials: A Study of Crystalline Saccharides'*, A. Gainar, J. S. Stevens, C. Jaye, D. A. Fisher, S. L. M. Schroeder, Journal of Physical Chemistry B, 2015, **119**, 14373.

Chapter 5 (Liquid-State NEXAFS/RIXS of PABA in Water at High/Low pH and in Methanol Solution) presents how NEXAFS and resonant inelastic X-ray spectroscopy (RIXS) microjet studies can be combined with density functional theory (DFT) calculations to elucidate the pH- and solvent-dependent changes in the local chemical and electronic environment of *para*-aminobenzoic acid (PABA).

This work was published as *'Chemical Speciation and Bond Lengths of Organic Solutes by Core-Level Spectroscopy: pH and Solvent Influence on p-Aminobenzoic Acid'*, J. S. Stevens[§], A. Gainar[§], E. Suljoti, J. Xiao, R. Golnak, E. F. Aziz, S. L. M. Schroeder, Chemistry – A European Journal, 2015, **21**, 7256.

([§]Equal contributions from the underlined authors)

Chapter 6 (NEXAFS of neutral PABA species in Water and in Alcohol Solutions) examines how pH variations influence the PABA speciation composition in water, and studies the solvent effect in the alcohol and aqueous solutions in terms of solute-solute and solute-solvent interactions.

Chapter 7 (UV-Vis of PABA in Water and Alcohols) reflects the effect of pH in PABA aqueous solutions and shows how laboratory UV-Vis and synchrotron NEXAFS-RIXS spectroscopies complement each other in providing access to the PABA speciation information. The solvent influence on the UV-Vis spectra of PABA solutions is also presented here.

Part of this work was published as *'The Structure of p-Aminobenzoic Acid in Water: Studies Combining UV-Vis, NEXAFS and RIXS Spectroscopies'*, A. Gainar, J. S. Stevens, E. Suljoti, J. Xiao, R. Golnak, E. F. Aziz, S. L. M. Schroeder, Journal of Physics: Conference Series, 2016, **712**, 012034.

Chapter 8 (Conclusions) highlights the important conclusions arising from the presented results in the context of the overall research pursued within the PhD programme, and proposes views on the future work that could continue from the results obtained.

Appendices show the crystal structures of the saccharides studied in Chapter 4, of the PABA polymorphs and of the PABA salts presented in Chapter 5, the X-ray powder diffraction and FEFF calculations for the saccharides from Chapter 4, personal Curriculum Vitae, participations to conferences, mettings, and workshops, as well as peer-reviewed publications.

1.5. Synchrotron experiments

The work in this thesis required experiments overseas at synchrotron X-ray sources:

- Liquid microjet NEXAFS measurements at the U41-PGM beamline from BESSY II, Helmholtz Zentrum Berlin (HZB), Berlin, Germany, 23rd-30th June 2014
- Solid-state NEXAFS measurements at the U7A beamline from National Synchrotron Light Source (NSLS), Brookhaven National Laboratories (BNL), New York, US, 24th-29th July 2013
- Liquid microjet NEXAFS measurements at the U41-PGM beamline from BESSY II, Helmholtz Zentrum Berlin (HZB), Berlin, Germany, 17th-23rd May 2013
- Solid-state NEXAFS measurements at the U7A beamline from National Synchrotron Light Source (NSLS), Brookhaven National Laboratories (BNL), New York, US, 29th September-4th October 2012
- Solid-state NEXAFS measurements at the U7A beamline from National Synchrotron Light Source (NSLS), Brookhaven National Laboratories (BNL), New York, US, 2nd-6th of August 2012

1.6. List of publications

1. J. S. Stevens[§], A. Gainar[§], E. Suljoti, J. Xiao, R. Golnak, E. F. Aziz, S. L. M. Schroeder, *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*, 7256-7263, DOI: 10.1002/chem.201405635.
   [§]Equal contributions from the underlined authors

2. A. Gainar, J. S. Stevens, C. Jaye, D. A. Fischer, S. L. M. Schroeder, *NEXAFS Sensitivity to Bond Lengths in Complex Molecular Materials: A Study of Crystalline Saccharides*, Journal of Physical Chemistry B, (2015), *119*, 14373-14381, DOI:10.1021/acs.jpcb.5b07159.
3. R. Golnak, J. Xiao, K. Atak, J. S. Stevens, A. Gainar, S. L. M. Schroeder, E. F. Aziz, *Intermolecular Bonding of Hemin in Solution and in the Solid State Probed by N K-edge Core Level Spectroscopies*, Physical Chemistry Chemical Physics, (2015), *17*, 29000-29006, DOI: 10.1039/C5CP04529K
4. A. Gainar, J. S. Stevens, E. Suljoti, J. Xiao, R. Golnak, E. F. Aziz, S. L. M. Schroeder, *The Structure of p-Aminobenzoic Acid in Solution: Studies Combining NEXAFS and UV-Vis Spectroscopy*, Journal of Physics: Conference Series, (2016), *712*, 012034, (pages 1-4), DOI: 10.1088/1742-6596/712/1/012034.
5. J. S. Stevens, A. Gainar, C. Jaye, D. A. Fischer, S. L. M. Schroeder, *NEXAFS and XPS of p-Aminobenzoic Acid Polymorphs: The Influence of Local Environment*, Journal of Physics: Conference Series, (2016), *712*, 012133 (pages 1-4), DOI: 10.1088/1742-6596/712/1/012133.
6. J. S. Stevens, A. Gainar, E. Suljoti, J. Xiao, R. Golnak, E. F. Aziz, S. L. M. Schroeder, *NEXAFS Chemical State and Bond Lengths of p-Aminobenzoic Acid in Solution and Solid State*, Journal of Physics: Conference Series, (2016), *712*, 012136 (pages 1-4), DOI: 10.1088/1742-6596/712/1/012136.
7. A. Gainar, M. C. Tzeng, B. Donnio, D. W. Bruce, *Control of Mesophase Structure Using Polyphilic Mesogens*, submitted June 2016.
8. Acknowledged for offering technical support for the NEXAFS data included in the paper: E. A. Willneff, S. L. M. Schroeder, B. A. Ormsby, *Spectroscopic Techniques and the Conservation of Artists' Acrylic Emulsion Paints*, Heritage Science, (2014), *2*, pages 1-10, DOI: 10.1186/s40494-014-0025-y.
9. A. Gainar, J. S. Stevens, E. Suljoti, J. Xiao, R. Golnak, E. F. Aziz, S. L. M. Schroeder, *The Influence of pH and Solvent on the Structure of p-Aminobenzoic Acid*, estimated date of submission July 2016.

References

[1] J. T. Carstensen, *Pharmaceutical Principles of Solid Dosage Forms*, Technomic Publishing Co., 1993.
[2] M. Volmer, *Kinetics of Phase Formation*, Steinkopff, Leipzig, 1939.
[3] I. Weissbuch, M. Lahav, L. Leiserowitz, *Cryst. Growth Des.*, 2003, **3**, 125.
[4] P. G. Vekilov, *Cryst. Growth Des.*, 2010, **10**, 5007.
[5] N. Pienack, W. Bensch, *Angew. Chem. Int. Ed.*, 2011, **50**, 2014.
[6] H. G. Alison, R. J. Davey, J. Garside, M. J. Quayle, G. J. T. Tiddy, D. T. Clarke, G. R. Jones, *Phys. Chem. Chem. Phys.*, 2003, **5**, 4998.
[7] A. M. Beale, A. M. J. van der Eerden, S. D. M. Jacques, O. Leynaud, M. G. O'Brien, F. Meneau, S. Nikitenko, W. Bras, B. M.Weckhuysen, *J. Am. Chem. Soc.*, 2006, **128**, 12386.
[8] J. Polte, R. Kraehnert, M. Radtke, U. Reinholz, H. Riesemeier, A. F. Thenemann, F. Emmerling, *J. Phys. Conf. Ser.*, 2010, **247**, 012051.
[9] S. Calvin, E. E. Carpenter, V. Cestone, L. K. Kurihara, V. G. Harris, E. C. Brown, *Rev. Sci. Instrum.*, 2005, **76**, 016103.
[10] R. G. Parr, W. Yang, *Density-Functional Theory of Atoms and Molecules*, New York, Oxford University Press, 1989.
[11] D. Cabaret, P. Sainctavit, P. Ildefonse, A. M. Flank, *Am. Mineral.*, 1998, **83**, 300.
[12] K. Atak, S. I. Bokarev, M. Gotz, R. Golnak, K. M. Lange, N. Engel, M. Dantz, E. Suljoti, Oliver Kühn, E. F. Aziz, *J. Phys. Chem. B*, 2013, **117**, 12613.
[13] R. Golnak, K. Atak, E. Suljoti, K. F. Hodeck, K. M. Lange, M. A. Soldatov, N. Engel, E. F. Aziz, *Phys. Chem. Chem. Phys.*, 2013, **15**, 8046.
[14] K. Atak, N. Engel, K. M. Lange, R. Golnak, M. Gotz, M. Soldatov, J.-E. Rubensson, N. Kosugi, E. F. Aziz, *Chem. Phys. Chem.*, 2012, **13**, 3106.

CHAPTER 2

LITERATURE REVIEW

2.1. General aspects on crystallisation

Crystallisation is a process during which a highly-ordered solid results from a saturated solution, melt or material deposition from a gas phase. Crystallisation from solutions has many applications in industry such as the pharmaceutical, material or dairy production.[1,2] It is often used as a powerful purification tool due to the self-assemblage of solute molecules forming a defined crystal packing.

Generally, crystallisation comprises three steps:

i. supersaturation
ii. nucleation
iii. crystal growth

2.1.1. Supersaturation

Supersaturation is the state above the saturation point of a solution when the solute is dissolved in a higher quantity than the solvent can normally take, and represents the principal driving force for the crystallisation to occur.[3]

There are several procedures for the supersaturation to be generated:

i. solution cooling, with the immediate consequence of an increase in the amount of undissolved solute particles as the system is pushed towards the supersaturated region
ii. solvent evaporation, which is preferred when the solvent is relatively volatile
iii. anti-solvent method, when a second solvent is added to the saturated solution, creating a three-component system in which the solubility is lower than in the initial saturated solution; therefore, the solute nucleates out and crystallisation takes place

2.1.2. Nucleation

Nucleation can take two different forms depending on the nature of the system – it can be either homogenous or heterogeneous.

Homogeneous nucleation takes place spontaneously in a solution with no other foreign particles. In supersaturated solutions, the solute molecules self-aggregate under the form of nuclei, many of which are not usually able to reach the critical size and immediately tend to dissolve back into the system. Thus, a resulting nucleus of spherical geometry presents two alternatives for the free energy of the system to decrease: either to grow or dissolve.[4] A way to quantify the nucleation process is to consider the induction time at a specific temperature as the time when the first visible crystals appear.

Heterogeneous nucleation, as opposed to homogeneous nucleation, can occur at much lower supersaturations due to formation of a critical nucleus size involving lower free energies, and is dictated by the similarity between the foreign particles and the solute molecules. Secondary nucleation is a particular type of heterogeneous nucleation when the system is seeded with solute crystals and can take place even at lower supersaturations than the standard heterogeneous nucleation. Alternatively, the seed crystals can form when the solute particles collide with the walls of the vessel.[3]

2.1.3. Crystal growth

Crystal growth is the next phase after the formation of the stable nucleus and influences the crystal shape, product quality and purity. This complex process involves diffusional and interfacial aspects through which molecules adhere to crystals via either adsorption or bonding processes, and is dependent on the number of interactions or the nature of the bonds between the crystal faces and the molecules adhering to these, thus influencing the growth rate. Therefore, a classification of the crystal faces can be made as follows:

i. kinked faces are surfaces where three molecule-crystal bonds can form
ii. stepped faces are surfaces involving two bonds
iii. flat faces are faces where only one bond can form

Additionally, a molecule can adhere to the crystal surface in three different possible ways through continuous growth, surface nucleation and spiral growth. On the other hand, mass

transfer plays an important role in the crystal growth rate, as solute molecules are transported from the solution onto the crystal surface. Generally, the local concentration at the interface between the growing crystal and solution is lower than the rest due to the resulting mass transfer gradient.[3]

2.2. Analytical techniques for the study of nucleation

In order to study the solution-state behaviour and the experimental thermodynamics of a solute, conventional analytical methods can be normally employed in standard analytical laboratories, using appropriate pieces of equipment.

As a common technique, UV-Vis spectroscopy is suitable for probing solute-solute associations in solutions[5] and allows distinguishing between dimer formation and higher-order associations. Previous literature examples present the self-association of organic molecules such as imidazole,[6] purine,[7] pyridine[8,9] in solution by studying the deviations (inflection points) from the Beer-Lambert law[10] as concentration is increased over a limited range.

Other conventional techniques such as solution-state IR and Raman spectroscopies are able to identify the stretching frequencies (*e.g.* for N–H and C=O from *para*-aminobenzoic acid, PABA), as well as bending frequencies (*e.g.* for O–H from PABA or benzoic acid), in order to probe structural changes induced by self-associations in solution or to identify different hydrogen-bonded motifs.[11,12] A specific published study presents the solvent effect on the formation of isonicotinamide (INA) polymorphs through the use of IR and Raman spectroscopies; thus, solvents with strong hydrogen-bond acceptor sites determine INA to self-associate as head-to-tail chains, while solvents with strong hydrogen-bond donor sites lead to head-to-head dimers.[13]

A manner to quantify the magnitude of the thermodynamic driving force is to analyse nucleation rates for estimating interfacial tensions. Moreover, in a recent review[14] it was discussed that the fundamental kinetic processes can be examined through determination of nucleation rates. When combined with solution IR, computational simulations and solubility data, the overall nucleation process can be successfully determined through solute dimerisation and desolvation.[11] However, although the use of nucleation rates leads

to increasing the volume of kinetic data, the full interpretation of interactions on a molecular scale still remains unattainable. One significant downside represents the lack of precise control of the use of global parameters, notably supersaturation, temperature and cooling rates that are often irreproducible.

In general, it is possible to determine association constants and to confirm the presence of hydrogen bonds through NMR measurements.[15] The case of *para*-acetanisidide in chloroform reveals that solute molecules were N–H$\cdots$O=C hydrogen-bonded similarly to the crystal structure motifs.[16] From the nuclear Overhauser effect it was indicated here that the solute molecules self-aggregate in pre-nucleation clusters, albeit without having the same packing as the crystal structure.

Alternatively, neutron scattering experiments (normally conducted at neutron sources) enhanced by empirical potential structure refinement (EPSR) method describe in detail the solution structure by determining the radial distribution functions of both the solute and solvent.[17,18] A specific example for a mixture methane-water, the order of the hydration shell around the methane was proved to be reduced, while the methane ordering increased after crystallisation.[19] Another case represents the application of hexamethylenetetramine (HMT) in aqueous solution, where it was possible to observe similarities between solute-solute associations and interactions in the anhydrous crystalline HMT, as well as between solute-solvent associations and the crystal structure of the HMT hexahydrate.[17]

However, the acquisition time for these techniques is generally longer than the lifetime of the nuclei, while these consequently fail in defining the bigger picture on the correlation between solution chemistry and achievement of the crystallinity through a precise structural model of the global process. The synchrotron radiation techniques present important advantages for the study of nucleation/crystal growth, with the high photon intensity enabling these processes to be examined *in situ* starting even from the early stages. The wide spectral range employed in near-edge X-ray absorption fine structure (NEXAFS) or X-ray photoelectron spectroscopy (XPS) allows the local environment and chemical state to be probed within seconds or even tenths of a second. However, both of these techniques normally require ultra-high vacuum experimental conditions.

Previous work at the BESSYII synchrotron (Berlin, Germany) showed that solutions can be investigated by soft X-ray absorption spectroscopies (XAS) using an enclosed flow cell in specially designed high vacuum chambers,[20,21] providing information on the local environment, geometry and chemical state of the species in solution. The measurements revealed changes in the local structure of imidazole in solution as a function of concentration caused by an increase in solute self-associations.[22]

Resonant inelastic X-ray scattering (RIXS) is complementary to NEXAFS to provide a more complete picture of the electronic structure, yet data acquisition time is longer (order of minutes). Small angle X-ray scattering (SAXS) and wide-angle X-ray scattering (WAXS) techniques probe a longer range order than NEXAFS does. SAXS is suitable for investigating solute-solute interactions that lead to the formation of clusters/nano-crystals, whereas WAXS examines the growth of nuclei to the final defined stage of crystals. For a small organic molecule of 2,6-dibromo-4-nitroaniline, simultaneous SAXS/WAXS measurements indicated that the initial phase detected by SAXS leads to a crystalline structure in several tenths of a second.[23]

2.3. X-ray core level spectroscopies for liquids

Originally developed by Faubel,[24] the use of liquid microjets for core level X-ray spectroscopies succeeds in avoiding problems caused by radiation damage owing to the continuous renewal of the sample.[25,26]

The mechanisms present in solution are quite complex as the solvation process is dynamic, each solute molecule being surrounded by solvation shells. The key factors for the distribution of free energy throughout the system represent solute-solute or solute-solvent intermolecular interactions (*e.g.* π-π stacking, van der Waals, proton transfer, hydrogen bonding).

In the case of imidazole in aqueous solutions, there has been an increasing interest in explaining the molecular self-association for a long period of time, initially being suggested that π-π stacking are the main driving forces.[27] Comparing the NEXAFS spectra of aqueous solutions over a large concentration range, it was suggested that the local solvation environments remain similar, without significant variations. In this study,

two possible mechanisms of solute self-association were proposed: one which assumes the presence of interactions between solvated imidazole molecules that do not massively impact on the measured spectra, and another one where interactions occur mostly at the lower end of the concentration range investigated. The spectrum of imidazole aqueous solution presents two pre-edge peaks corresponding to transitions 1s→1π* of the N1 and N3 moieties in the imidazole ring (Figure 2.1), separated by a 1.7 eV energy difference that is slightly larger, albeit closer to the solid-state imidazole (1.5 eV), and smaller than the gas phase monomer (2.4 eV). This indicates that hydrogen bonding takes place in solution with water or imidazole molecules, similarly to the crystalline solid state.

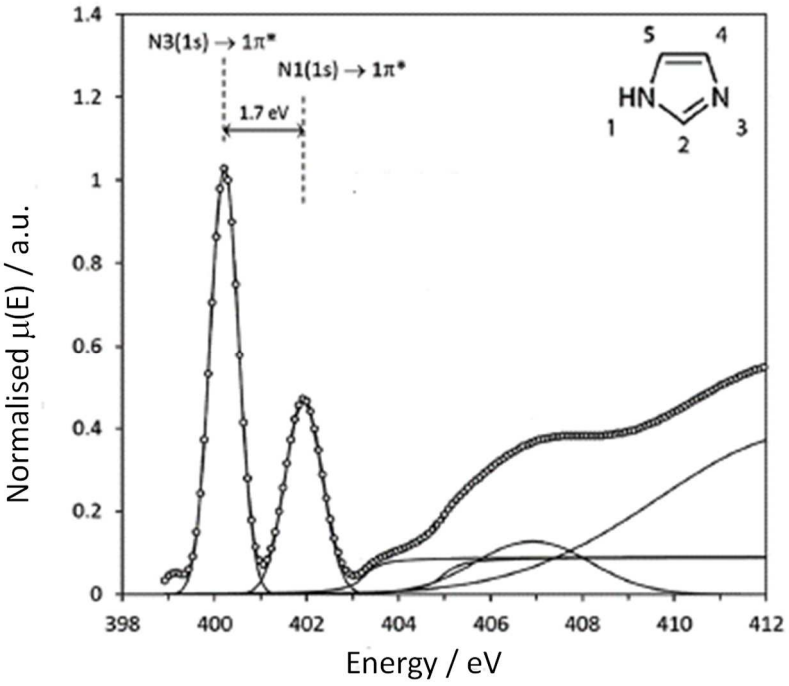


Figure 2.1. N K NEXAFS of imidazole aqueous solution showing the two pre-edge peaks separated by a 1.7 eV energy difference. The top right corner presents the imidazole molecule with atoms numbered according to IUPAC. Figure adapted from reference [22].

A previous X-ray scattering analysis of aqueous imidazole solutions suggests that the solute molecules self-associate in stacks through water molecules.[27] To verify this aspect, calculations using a geometry-optimised stack structure of three imidazole molecules were employed, returning a result of 1.5 eV between the N 1s→1π* resonances, which is

closer to the experimental value than geometry-optimised microsolvated clusters with only one single imidazole molecule. This reflects that the mechanism of self-association can be intuitively understood by hydrogen-bonded π-π stacks in competition with chain-like hydrogen-bonded imidazole molecules.

Another process that may occur in solution is solvolysis which involves dissociation of solvent molecules into ionic species. For the water molecule, H_2O, dissociation leads to H_3O^+ and HO^- aqueous species (Figure 2.2), which can further interact with the solute molecules that have either an acidic or basic character (Figure 2.3).

Figure 2.2. Dissociation of water molecules into ionic species, H_3O^+ and HO^-.

Figure 2.3. Mechanism of proton transfer between a base B and H_3O^+ species (a), and between an acid A–H and HO^- species (b) in an aqueous solution.

The proton transfer equilibria presented in Figures 2.2 and 2.3 can be affected if in the aqueous system the pH value is changed when introducing quantities of acids or bases in the solution. For a low pH value, the system is acidic and formation of the B^+–H species is favoured. The situation is different at high pH values, when the system is basic and formation of A^- is preferred.

In general, pH-dependent NEXAFS measurements in solution can be sensitive enough to allow probing the local environment of the functional groups involved in the proton transfer by indicating the preferred charge state. The electronic structure of glycine was previously investigated in water keeping the solution concentration constant and systematically varying only the pH.[28] Glycine is an amino acid, and therefore presents both $–NH_2$ and –COOH functional groups in the molecular structure; the $–NH_2$ group can accept a proton, while –COOH can donate a proton in an aqueous environment, allowing

glycine to behave either as an acid or base (amphoteric character). Moreover, the process of proton transfer can happen within the same molecule of glycine, leading to a zwitterionic form, in which $–NH_3^+$ and $–COO^-$ co-exist (Figure 2.4).

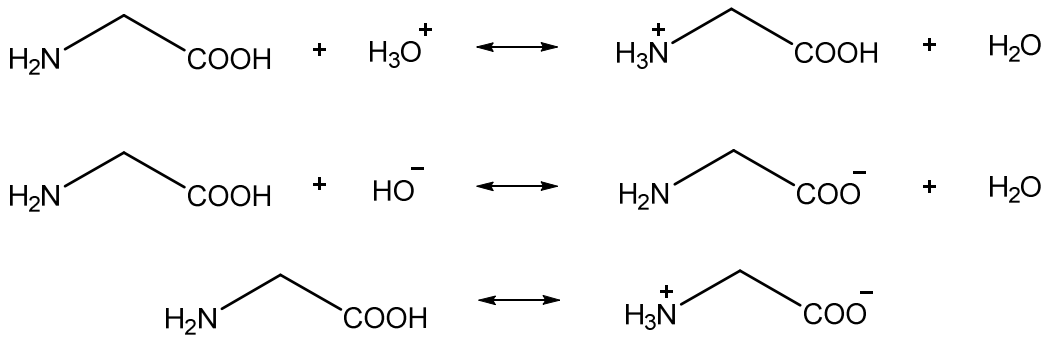


Figure 2.4. Protonation of glycine in acidic aqueous solution (top), deprotonation of glycine in basic aqueous solution (middle), and intramolecular proton transfer involving the $–NH_2$ and –COOH groups (bottom).

Studying the dominant form of glycine in water at various pH values provides significant information with direct relevance to peptides and proteins, as changes in the local environment can modify its reactivity and can have a high impact on its biological functionality.[28,29] Donating or accepting protons modifies the local environment surrounding the nitrogen atom from the $–NH_2$ moiety or the oxygen atoms from the –COOH group in the molecular structure. Consequently, changes in pH can imply shifts in the energy levels of atomic and molecular orbitals involved in the $–NH_2$ or –COOH functionalities.[30,31] Unlike the $–NH_2$ and –COOH groups, the carbon-containing backbone, $–CH_2–$, is not directly affected by pH changes (Figure 2.4). As a consequence, in the C K-edge spectra it is expected there will be negligible changes in the electronic structure of the carbon-containing moieties, as presented in Figure 2.5 – only a small shift of ~ 0.15 eV is noticed for the first sharp peak around 287 eV for pH 1, and a redistribution of amplitudes and widths occurs for the broad resonances between 290 and 305 eV.[28] For the N K-edge (Figure 2.6), the anionic species (at pH 12) presents distinct features compared to the other two species, neutral (at pH 6) and cationic (at pH 1) glycine. A large energy shift of ~ 1.3 eV of the main peak (~ 406 eV) is also visible for the pH 12 solution as the local environment of the nitrogen atom is changed from the planar

configuration of $–NH_2$ to tetrahedral $–NH_3^+$ via protonation.[28] On the other hand, as the three spectra were recorded at the same glycine concentration, the glycine-glycine interactions are quasi-constant in magnitude and do not cause noticeable changes in the spectra, unlike the charge state on the nitrogen moiety.

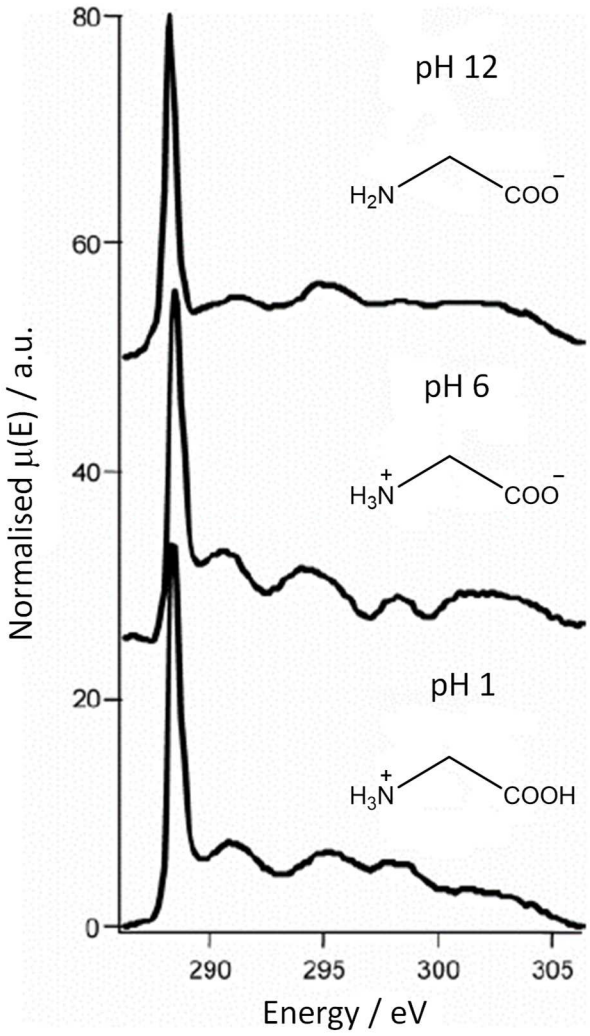


Figure 2.5. C K NEXAFS spectra of glycine aqueous solution at pH 12 (top), 6 (middle), and 1 (bottom). Figure adapted from reference [28].

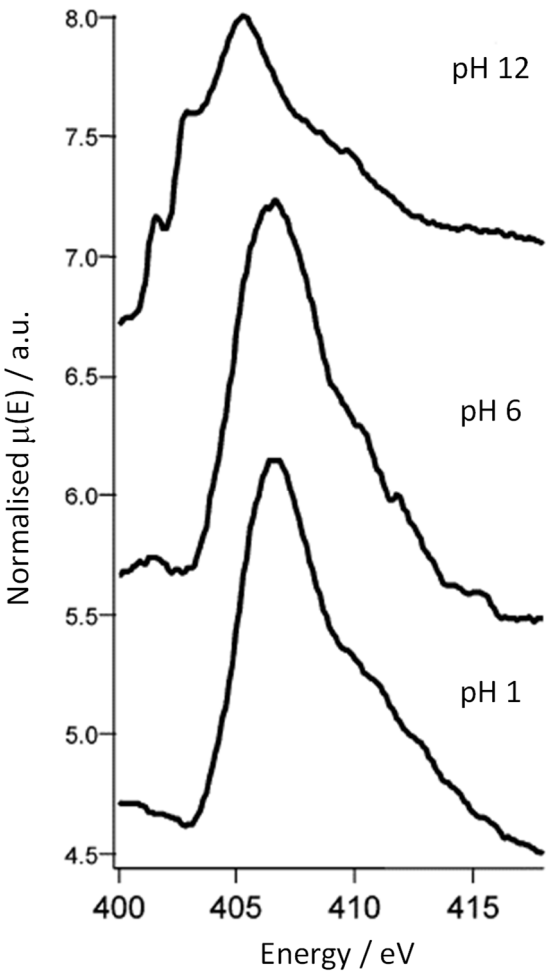


Figure 2.6. N K NEXAFS spectra of glycine aqueous solution at pH 12 (top), 6 (middle), and 1 (bottom). Figure adapted from reference [28].

In the O K-edge spectra, the unique peak corresponds to a transition from an O 1s atomic orbital to an unoccupied π^* orbital in which the –COOH functionality is involved. The peak width and energy position are the same in the pH 12 (basic) and 6 (neutral) solutions because the local environment of the oxygen atoms is predominantly $–COO^-$, as opposed to –COOH in the pH 1 (acidic) solution where the peak exhibits a small shift of ~ 0.25 eV and a broadening (Figure 2.7).

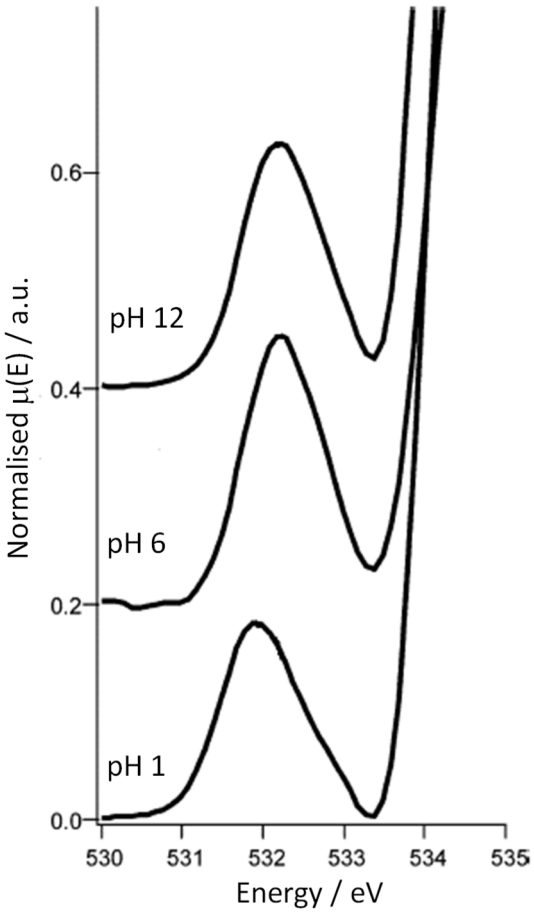


Figure 2.7. O K NEXAFS spectra of glycine aqueous solution at pH 12 (top), 6 (middle), and 1 (bottom), where due to strong water background, only the spectral region before the first background resonance is shown. Figure adapted from reference [28].

Another study presents a different amino acid, proline,[32] with the aim to compare the pH-induced spectral differences with the ones met in the case of glycine and to demonstrate the NEXAFS sensitivity to the local solvation environment of these two amino acids. The chemical structures of proline at neutral and basic pH are presented in Figure 2.8.

Figure 2.8. Chemical structures of proline in a neutral (left), and basic (right) aqueous solution.

Figure 2.9 shows the N K NEXAFS spectra of glycine and proline, each in aqueous solutions of pH 6 (neutral) and pH 12 (basic). At neutral pH, each spectrum presents a main feature at ~ 406.5 eV, whilst at basic pH this feature shows up at ~ 405 eV. Based on the previous study of glycine (Figure 2.6), this quasi-constant shift of ~ 1.5 eV is an indication that the protonation states of the nitrogen atom are the same for both amino acids when pH is the same, namely protonated (zwitterionic species) at neutral pH and deprotonated (anionic species) at basic pH.[28] Moreover, a broad resonance ~ 3 eV higher than the main feature was found in all four spectra, albeit more pronounced for proline.

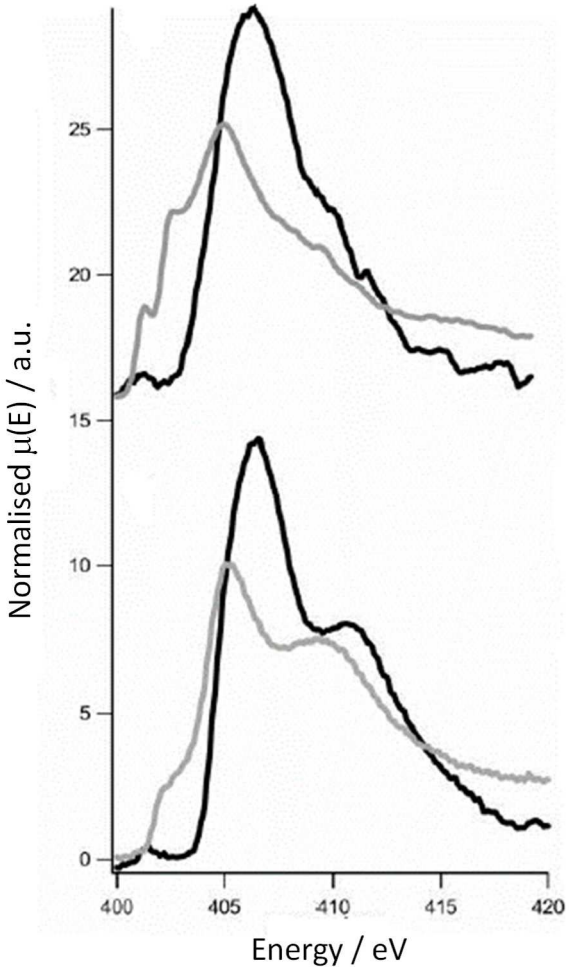


Figure 2.9. N K NEXAFS spectra of aqueous solutions of glycine (top) at pH 6 (black line) and pH 12 (grey line), and proline (bottom) at pH 6 (black line) and pH 12 (grey line). Figure adapted from reference [32].

In the pH-dependent study of glycine in solution,[28] changes in the protonation and conformational state of the molecule (*i.e.* specific rotation of the amino moiety around

the C–N bond) are expected to lead to differences in the N K-edge spectra (*i.e*. the small features of the pre-edge region undergo intensity and width variations). A distinctive feature of proline is its secondary amine nature, reflected by the 'locked' –NH– moiety that sterically hinders the C–N bond from freely rotating around its axis. Turning now to the small features noticeable in the spectra of anionic glycine at energies lower than ~ 405 eV, yet partially absent for anionic proline, calculations were implemented to conclude that the endocyclic C–N fixed bond of proline is the fundamental reason for these spectral differences.[32] The immediate consequence is that the hydrogen bonding acceptor-only conformation observed in the anionic glycine is primarily due to steric shielding from the water molecules induced by the rotation around the C–N bond, which is impossible for proline.

An example of using the liquid microjet uses NEXAFS spectroscopy as a tool to investigate the hydration of adenosine triphosphate (ATP) as a function of concentration and pH.[33] The concentration-dependent measurements did not outline noticeable spectral changes despite the reported π-stacking interactions for nucleotides at high concentrations.[34,35] In contrast, the pH-dependent measurements reflected structural changes due to the protonation of adenine from the structure of ATP.

As the ribose-monophosphate component of ATP does not significantly influence the π* transitions in the NEXAFS spectra,[36] the pre-edge regions of adenine in the solid and gaseous state, and ATP in aqueous solution (pH 7.5) can be directly compared. In Figure 2.10, two peaks at 286.5 eV and 287.3 eV in the pre-edge region of the C K-edge are in good agreement with the spectrum of solid phase adenine,[37] and gas phase adenine,[38] apart from an extra peak at 286.8 eV that is encountered only in the gas phase. These peaks correspond to two different carbon environments, *i.e*. a carbon atom bound to either one or two nitrogen atoms. Additionally, the relative peak intensities of the solid and aqueous state are similar, implying that the aqueous chemical state of ATP is more similar to the solid phase, rather than the gas phase.

Figure 2.11 shows the N K NEXAFS spectra for solid and gas phase adenine, and aqueous ATP solution at pH 7.5. The most intense peak is observed at 399.5 eV for adenine, whereas for ATP in water, this is blue shifted (higher energy) to 400.3 eV due to hydrogen bonding with the solvent molecules. These features correspond to transitions of core electrons of ring nitrogen atoms to the lowest unoccupied molecular orbital (LUMO).

Apart from the main peak, aqueous ATP presents a resonance at 401.5 eV (as a consequence of the LUMO orbital alteration due to hydration) and at 402.5 eV (which is attributed to transitions of ring nitrogen core level electrons). Overall, it is expected that lone pairs of electrons on the nitrogen atoms and N–H bonds are more involved in hydrogen bonds than carbon atoms, hence the more noticeable spectral differences in Figure 2.11.

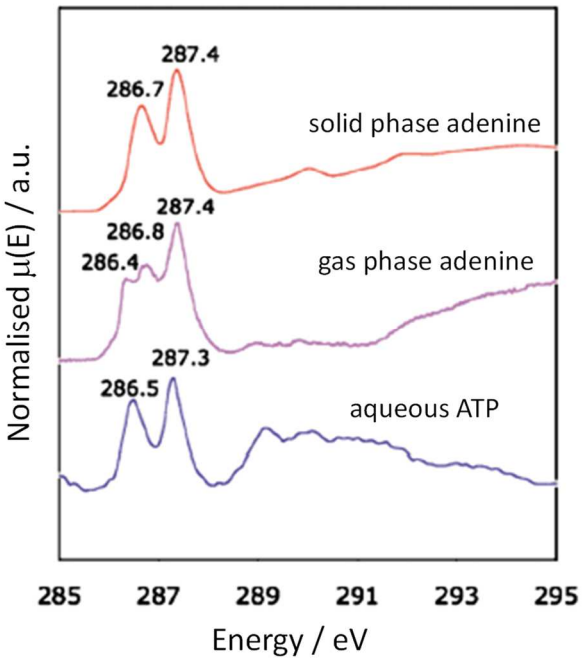


Figure 2.10. C K NEXAFS spectra of solid phase adenine (top), gas phase adenine (middle), and aqueous solution of ATP at pH 7.5 (bottom). Figure adapted from reference [33].

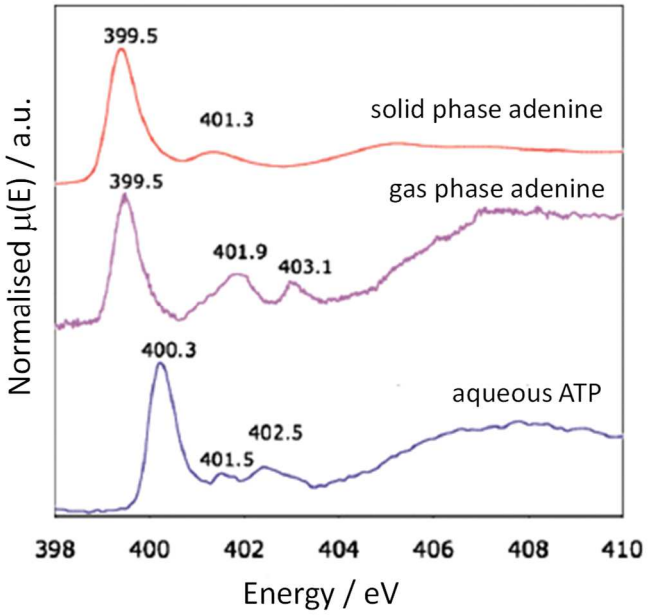


Figure 2.11. N K NEXAFS spectra of solid phase adenine (top), gas phase adenine (middle), and aqueous solution of ATP at pH 7.5 (bottom). Figure adapted from reference [33].

Variations of the pH for ATP in aqueous solution are expected to trigger several significant electronic changes. The measured C K NEXAFS spectra for aqueous solutions are presented in Figure 2.12, where the peak encountered at 287.3 eV in a solution of pH 7.5 splits at pH 2.5 in two distinct peaks with energies 287 eV and 287.7 eV, while the peak at 286.5 eV remains at the same energy position in both spectra. In the study of NEXAFS spectra of solid state adenine,[37] the peak at ~ 287.3 eV corresponds to the four carbon atoms in the adenine ring that are bound to two nitrogen atoms (Figure 2.13). The splitting of this peak in two features at pH 2.5 is explained by the nitrogen that becomes protonated and generates two types of carbon chemical environments on the adenine ring as this directly affects the states of the carbons that are attached to it, and also has secondary effects on the other carbons of the adenine ring.[33]

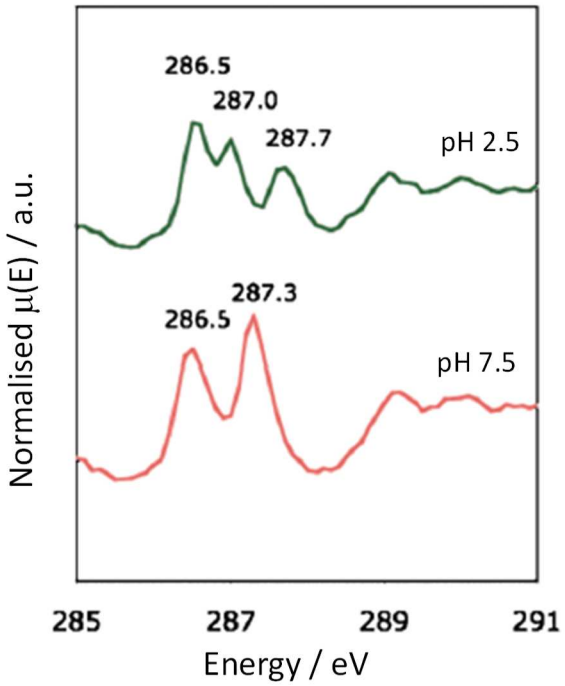


Figure 2.12. C K NEXAFS spectra of aqueous ATP at pH 2.5 (top) and 7.5 (bottom). Figure adapted from reference [33].

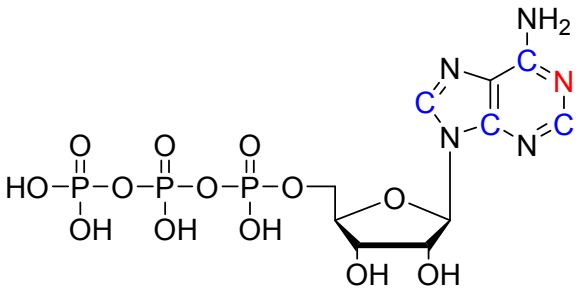


Figure 2.13. The structure of ATP, where the carbon atoms bound to two nitrogen atoms are marked in blue, and the preferred nitrogen atom prone to protonation is marked in red.

Turning to the N K-edge spectra presented in Figure 2.14, in the acidic solution (pH 2.5), all the features seem to be red shifted (lower energy) with respect to the peaks from the neutral solution (pH 7.5). At pH 2.5, the dominant peaks at 399.9 eV and 401.8 eV match the gas-phase adenine data.[36]

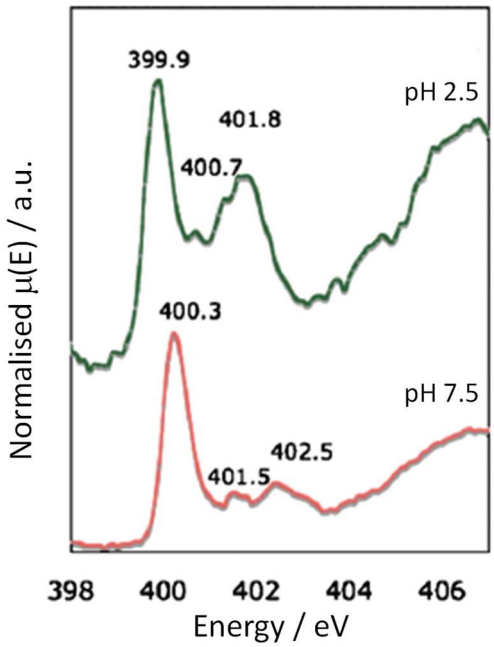


Figure 2.14. N K NEXAFS spectra of aqueous ATP at pH 2.5 (top) and pH 7.5 (bottom). Figure adapted from reference [33].

Another applicability of NEXAFS to liquids is that it can probe the interaction strength between charged molecules and their counter-ions.[39-42] As part of the bigger picture of the processes occurring in solution, a few key factors shall be considered, such as dynamics of the solute, intramolecular interactions within the solute molecule (*e.g.* between the charged $–NH_3^+$ and $–COO^-$ groups), and intermolecular solute-solute and solute-solvent interactions.

A particular example is represented by electrostatic interactions between cations and anions in water, which can be compared based on the electronic structure changes of the carboxylate moiety that is closest to the alkali counter-cation.[20] In this study, from intensity variations characteristic to the solvated $–COO^-$ in the O 1s XAS, it was possible to establish an order of the relative interaction strengths with various counter-cations, thus contributing to the elucidation of protein association and enzymatic activity mechanisms. Similarly, the aqueous zinc acetate, $Zn(CH_3COO)_2{\cdot}(H_2O)_n$, was studied to analyse the effect of the Zn^{2+} counter-cation on the carboxylate group, as well as of the acetate on water.[43] For this purpose, the spectra for zinc acetate in aqueous solution, aqueous acetic acid[44] and water[11] were collected and compared. The zinc acetate presents a peak at

~ 532.8 eV that is assigned to O 1s→$\pi^*_{C=O}$,[12,44,45] followed by another peak at ~ 535 eV that corresponds to the pre-edge region of water overlapped with the $O_{C=O}$ 1s→3p from the acetate ion. The spectrum of aqueous acetic acid at pH 13 is almost identical to the aqueous zinc acetate, suggesting that the anionic form is encountered in both cases. Comparing the RIXS spectra of aqueous zinc acetate and pure water at similar excitation energies, it was indicated that the hydrogen bonding network is reduced when Zn^{2+} and acetate ions are introduced in water. Continuing the series of examples with the aqueous ferrous (Fe^{2+}) cation, the effects on the electronic structure of water ligands were investigated.[46] In the RIXS spectra at the Fe L-edge, no intense ligand-to-metal charge transfer transitions were noticed as a consequence of the reduced orbital mixing between Fe^{2+} and water. Minor changes in the RIXS spectra at the O K-edge when compared to pure water indicate that the Fe^{2+} ions present a negligible disrupting effect on the hydrogen bonding and local electronic structure of water. It is also known that addition of ions can have a strong impact on the solubility of organic molecules. A specific example is illustrated by triglycine in water, where the effect of salt addition was investigated using NEXAFS.[47] Interesting observations were acquired from spectra comparison between aqueous solutions of pure triglycine, triglycine mixed with NaBr, and triglycine mixed with Na_2SO_3, keeping the same triglycine concentration in each case. The charge state of triglycine in water at neutral pH is the zwitterionic form that presents a protonated amino terminus, $–NH_3^+$, and a deprotonated carboxylic terminus, $–COO^-$, as shown in Figure 2.15. When dissolved in water, NaBr dissociates forming sodium cations, Na^+, and bromide anions, Br^-. Likewise, Na_2SO_3 dissociates in solution as Na^+ cations and SO_3^{2-} anions.

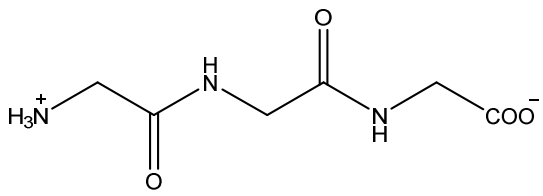


Figure 2.15. Zwitterionic form of triglycine in aqueous solution.

In the N K NEXAFS spectra from Figure 2.16, a distinct feature at ~ 403 eV is formed in the spectrum of triglycine aqueous solution with Na_2SO_3. On the other hand, the spectrum

of the triglycine solution containing Br^- anions presents negligible changes from the pure triglycine solution.[47] This observation can serve as a guide for evaluating the interaction strength between the SO_3^{2-}/Br^- anions and the protonated amino terminus $–NH_3^+$ of triglycine in water, which is stronger in the case of SO_3^{2-} considering the noticeable spectral differences. Another case is represented by electrostatic interactions between cations and anions in water, which can be compared based on the electronic structure changes of the carboxylate moiety that is closest to the alkali counter-cation.[5] In this study, from intensity variations characteristic to the solvated COO^- in the O K NEXAFS, it was possible to establish an order of the relative interaction strengths with various counter-cations, thus contributing to the elucidation of protein association and enzymatic activity mechanisms.

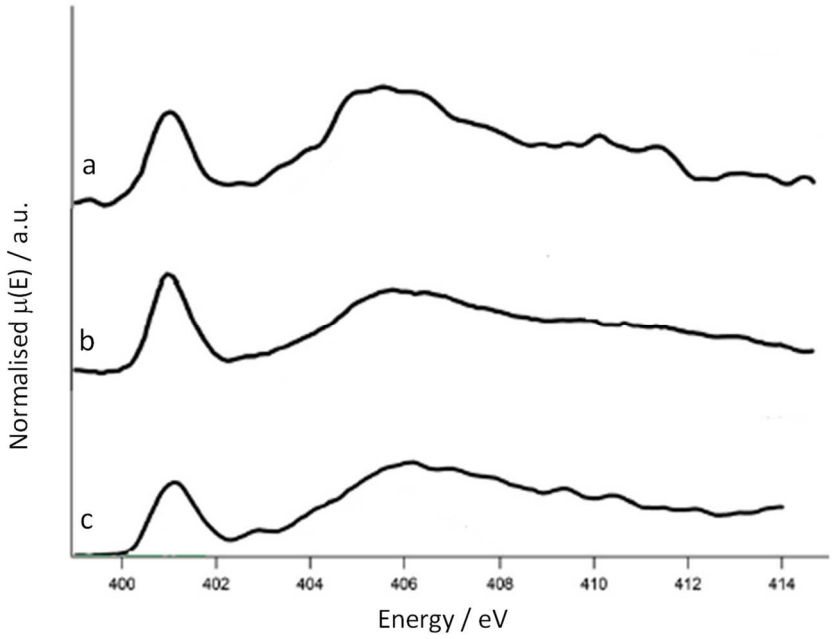


Figure 2.16. Compared N K NEXAFS spectra of pure triglycine in water (a), triglycine mixed with NaBr in water (b), and triglycine mixed with Na_2SO_3 in water (c). Figure adapted from reference [47].

A charge-transfer study on potassium ferrocyanide in water[48] analyses the local electronic structures of the CN^- ligands as well as the strong σ-donation and π-back-donation in the transitional metal complex, with the aim to provide a better grasp on the biological

catalytic processes based on the nature of the existent chemical bonds responsible for ligand exchange in solution. This constitutes a relevant example of an X-ray spectroscopic combination NEXAFS-RIXS, distinguishing between effects of bonding and back-bonding that cannot normally be analysed separately by conventional techniques. Here, the relative intensities of charge-transfer bands can be therefore associated with individual atomic orbital contributions for forming the molecular orbitals.

On a different note, a microscopic approach on the structure of hydrogen-bonded liquids aroused continuous interest due to the challenging theoretical and experimental conditions. Although NEXAFS is generally employed for detecting changes in bond lengths or angles, another significant use of NEXAFS in liquids is to probe the existence of hydrogen bonding that leads to cluster formation.[49-59] Acetone is an example of a solvent forming a weak hydrogen bonding network.[60] While in water and alcohols the –OH groups involved in these interactions can play both as H-donors and H-acceptors, in acetone clusters are formed on the basis of C=O$\cdots$H–C hydrogen bonds,[61-64] implying that the CH_3 functional group is the H-donor, while C=O is the H-acceptor. Acetone clusters were studied with NEXAFS to investigate the electronic structure of these associations.[49] The immediate observations arising from the density functional theory (DFT) calculations (that were used to substantiate the experiments) were that in the cluster, hydrogen bonding interactions C=O$\cdots$H–C cause the electron clouds around the non-donating H atoms to be pushed away from the hydrogen bond axis and the electron cloud on the C=O bond to be lost. These aspects explain the main difference in the C K-edge spectra, when the $C_{CH_3}1s\rightarrow 3p\pi(CH_3)$ resonance was strongly suppressed in the case of acetone clusters compared to isolated acetone.[49,65-67] On the other hand, the O K-edge spectra of the isolated acetone molecule and acetone clusters were almost identical, and the character of the π^* (C=O) orbital was not majorly influenced by the C=O$\cdots$H–C hydrogen bonding interaction, with just a small shift (0.13 eV) in the $O\ 1s\rightarrow\pi^*(C=O)$ resonance towards the higher energy side was observed for the clusters.[49]

The most common solvent, water, has been studied intensively for centuries to understand its anomalies triggered by the hydrogen-bonding network formed between molecules.[68,69] However, there are several aspects on the local geometric arrangement of pure liquid water molecules in clusters that were still not fully elucidated. In its O K NEXAFS spectrum three features are encountered: a pre-edge resonance at ~ 535 eV, a main resonance at ~ 537 eV, and a post-edge resonance at ~ 542 eV. In combination with

theoretical calculations, it was indicated that these features correspond to different types of hydrogen-bond coordinations: the pre-edge and main resonance are assigned to broken networks, while the post-edge resonance to tetrahedral motifs. A downside of these previous approaches was that the theoretical methods were not sufficiently advanced to corroborate with the experimental results, and the spectra were characteristic to bulk water, more precisely describing an average of hydrogen-bond interactions. To avoid these aspects, water was diluted with an oxygen-free solvent, acetonitrile, which is a hydrogen-bond acceptor. Consequently, the number of hydrogen bonds per water molecule is expected to reduce as the volumetric percentage of acetonitrile in the mixture is increased. The immediate observation from comparing the experiments is the post-edge feature becoming attenuated which indicates the high number of coordinated water molecules when the number of acetonitrile molecules dominates.[70]

RIXS measurements were also conducted on water-acetonitrile mixtures.[71] When the excitation energy is 534 eV (corresponding to the first pre-edge feature in the O K NEXAFS spectrum of liquid water), a sharp peak appears at ~ 526 eV (Figure 2.17) and corresponds to the oxygen lone-pair molecular orbitals. At higher excitation energies than 534 eV, this peak splits in two features, d_1 and d_2 (Figure 2.17). For example, Tokushima[72,73] considered that this stems from the occurrence of two distinct hydrogen-bonded motifs: one tetrahedral and one strongly distorted hydrogen bond configuration. Therefore, the spectrum can be decomposed in two individual spectra characteristic to each motif that differentiate through an energy offset caused by the core hole effects that depend on the specific hydrogen bond coordination. Alternatively, Fuchs[74] suggested that the additional feature is an indication for an ultrafast dissociation process consisting in the creation of proton transfer dynamics with the immediate effect of OH^- species formation. In this case, the spectrum can be decomposed in two spectra: one characteristic to non-dissociated water molecules and one to OH^- species. The validity of these two theoretical approaches of Tokushima and Fuchs can be verified experimentally. Thus, when the water concentration is decreased, the hydrogen bond coordination decreases and the $1b_2$, $3a_1$ and $1b_1$ features (Figure 2.17) do not change significantly, while the d_2 peak becomes less intense, indicating that the spectral decomposition approaches do not correlate with the measurements.

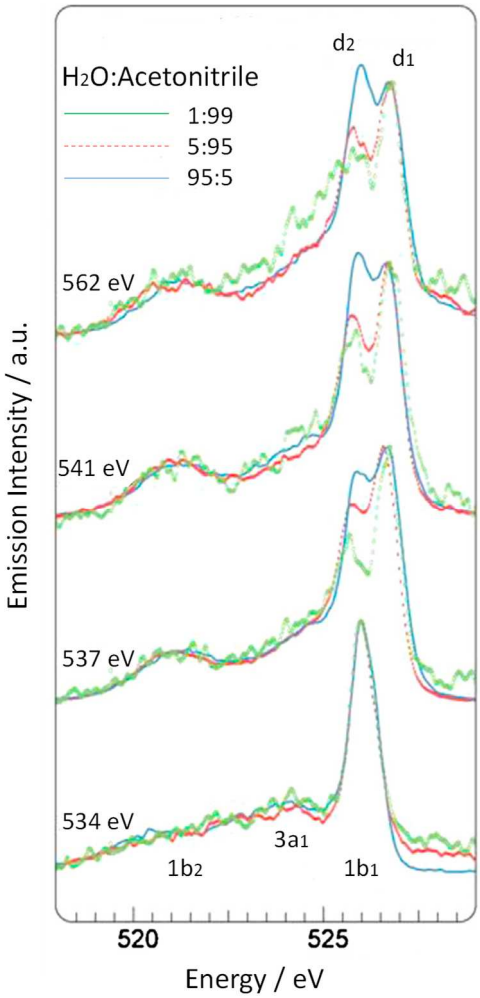


Figure 2.17. O K RIXS spectra of water-acetonitrile binary mixtures of volume ratios 1:99 (green dotted line), 5:95 (red dotted line), 95:5 (blue line) at the excitation energies indicated on the left side. Figure adapted from reference [71].

To explain the deviation from ideality of dimethylsulfoxide (DMSO)-water mixtures, RIXS spectra at the O K-edge reveal the occurrence of local electronic changes triggered by intermolecular interactions.[75] When using excitation energies of 532.4, 533.0, 533.6 eV (Figure 2.18) that are close to the first resonance of DMSO (yet below the excitation resonances for water), the features arising in the RIXS spectra are specific solely to DMSO. In the region of ~ 533 eV, the excited electron presents a localised character and is able to fill back the core hole, generating an elastic peak. The features 1 and 2 (Figure 2.18) start to shift even when water is at its lowest concentration in the binary mixture. A possible explanation is the S–O bond from the DMSO molecule becoming weaker as the water concentration increases, due to strong hydrogen bonds formed with water

molecules. The intensity ratio variation between the two main spectral features 3 and 4 (Figure 2.18) was assigned to the cross section variation for the scattering over orbitals LUMO and LUMO+1. Conclusions arising from this study reveal that the hydrogen bond network becomes disrupted even in water-rich DMSO solutions, and at 75 vol % DMSO the interaction between molecules occurs as clustering, thus lowering significantly the freezing temperature of the binary mixture.

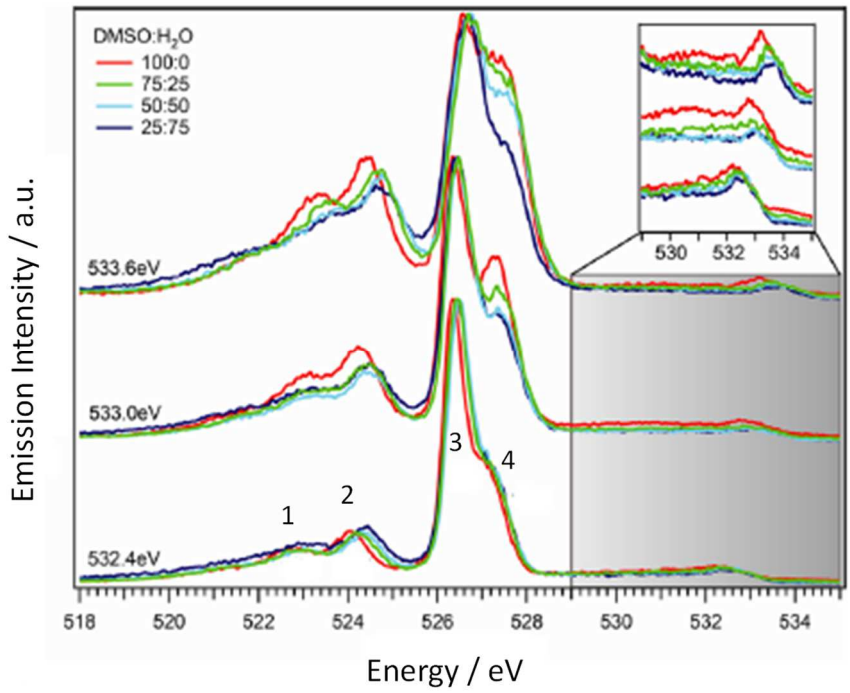


Figure 2.18. O K RIXS spectra of DMSO-water binary mixtures of volume ratios 100:0 (red), 75:25 (green), 50:50 (light blue), 25:75 (dark blue) at excitation energies 532.4 (bottom), 533.0 (middle) and 533.6 eV (top) with the elastic peaks magnified in an inset for better visibility. Figure adapted from reference [75].

Comparing the RIXS and X-ray emission spectra of DMSO with the ones of acetone, the elastic scattering peaks at 532 eV are less prominent for DMSO (Figure 2.19), which suggests that the excited electron is less localised here than in acetone.[76] The resemblance of DMSO and acetone O K-edge spectra indicates that the local interactions at the oxygen moiety are similar in both cases. As the highest energy peak corresponds to an electron decay from HOMO to the core hole, the double intensity of this peak in the acetone

spectrum suggests that HOMO has more oxygen p-character here than in the DMSO molecule. Particularly for acetone and DMSO, the HOMO is an oxygen lone-pair orbital which is, therefore, less pronounced for DMSO.

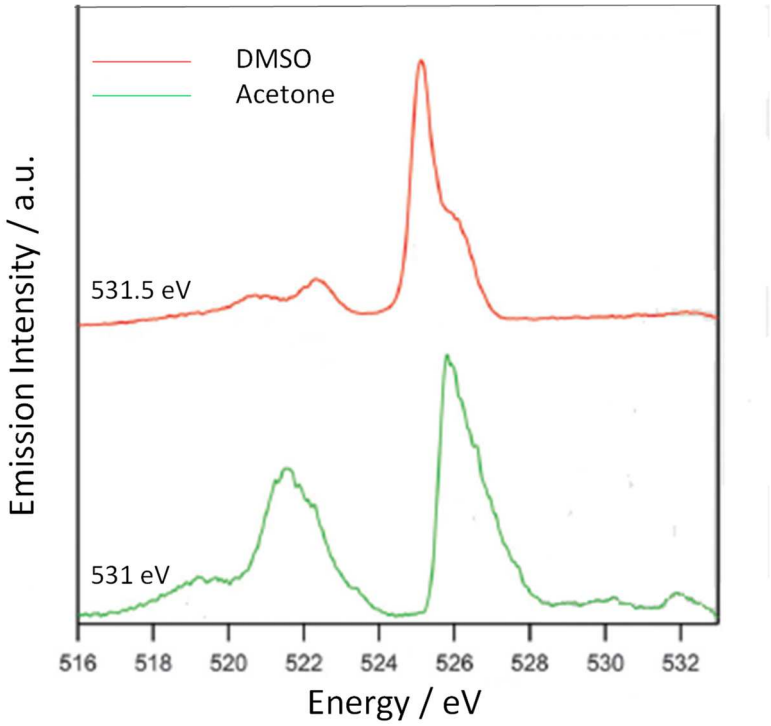


Figure 2.19. O K RIXS spectra of DMSO (red) and acetone (green) at the excitation energies indicated on the left side. Figure adapted from reference [76].

Another example of RIXS studies in solution highlights differences induced by dimerisation of hemin in water as opposed to DMSO, where hemin is encountered as monomer.[77] The absence of loss features in the spectra suggests the delocalised character of valence orbitals at the nitrogen equivalent moieties. At the same time, the narrower peaks for hemin in NaOH aqueous solution compared to hemin in DMSO can be explained by the increased degree of orbital localisation caused by solute-solute interactions. The more intense elastic peaks in the case of hemin in NaOH aqueous solution (dimer form) indicate a higher cross-section of the N 2p→1s transition and a preservation of the density of states with p character.

A different applicability of the liquid microjet is to determine the orbital symmetry of molecules in liquid phase where, unlike solid phase, molecules are randomly oriented. A RIXS study on acetonitrile[78] allowed preferential excitation of molecules of specific orientation relative to the incident polarisation vector. Here, the advantage of using the microjet is that it replaces membranes that could introduce artifacts[79] or create an additional undesired effect of molecular orientation.

References

[1] J. Hulliger, *Angew. Chem. Int. Ed.*, 1994, **33**, 143.
[2] K. Chadwick, R. J. Davey, G. Dent, R. G. Pritchard, C. A. Hunter, D. Musumeci, *Cryst. Growth Des.*, 2009, **9**, 1990.
[3] R. J. Davey, J. Garside, *From Molecules to Crystallisers. An Introduction to Crystallisation*, Oxford University Press, New York, 2001.
[4] J. W. Mullins, *Crystallisation*, 4th edition, Elsevier Butterworth-Heinemann, Oxford, 2001.
[5] J. Morcillo, E. Gallego, F. Peral, *J. Mol. Struct.*, 1987, **157**, 353.
[6] F. Peral, E. Gallego, *J. Mol. Struct.*, 1997, **415**, 187.
[7] F. Peral, E. Gallego, *Spectrochim. Acta, Part A*, 2000, **56**, 747.
[8] F. Peral, E. Gallego, *J. Mol. Struct.*, 1992, **274**, 105.
[9] F. Peral, E. Gallego, *J. Mol. Struct.*, 1994, **326**, 59.
[10] R. C. Denney, R. Sinclair, *Visible and Ultraviolet Spectroscopy*, John Wiley & Sons, London, 1987.
[11] R. A. Sullivan, R. J. Davey, G. Sadiq, G. Dent, K. R. Back, J. H. ter Horst, D. Toroz, R. B. Hammond, *Cryst. Growth Des.*, 2014, **14**, 2689.
[12] R. J. Davey, G. Dent, R. K. Mughal, S. Parveen, *Cryst. Growth Des.*, 2006, **6**, 1788.
[13] S. A. Kulkarni, E. S. McGarrity, H. Meekes, J. H. ter Horst, *Chem. Commun.*, 2012, **48**, 4983.
[14] R. J. Davey, S. L. M. Schroeder, J. H. ter Horst, *Angew. Chem. Int. Ed.* 2013, **52**, 2166.
[15] A. Spitaleri, C. A. Hunter, J. F. McCabe, M. J. Packer, S. L. Cockroft, *Cryst. Eng. Comm.*, 2004, **6**, 489.
[16] A. Saito, K. Igarashi, M. Azuma, H. Ooshima, *J. Chem. Eng. Jpn.*, 2002, **35**, 1133.
[17] R. C. Burton, E. S. Ferrari, R. J. Davey, J. L. Finney, D. T. Bowron, *J. Phys. Chem. B*, 2009, **113**, 5967.
[18] J. L. Finney, A. K. Soper, *Chem. Soc. Rev.*, 1994, **23**, 1.
[19] C. A. Koh, R. P. Wisbey, X. P. Wu, R. E. Westacott, A. K. Soper, *J. Chem. Phys.*, 2000, **113**, 6390.
[20] E. F. Aziz, N. Ottosson, S. Eisebitt, W. Eberhardt, B. Jagoda-Cwiklik, R. Vacha, P. Jungwirth, B. Winter, *J. Phys. Chem. B*, 2008, **112**, 12567.
[21] E. F. Aziz, *Local Electronic and Geometric Structure of Ions, Molecules, and Polymers in Liquid Phase*, PhD thesis, Berlin, 2006.
[22] M. J. Thomasson, C. R. Seabourne, B. M. Sattelle, G. A. Hembury, J. S. Stevens, A. J. Scott, E. F. Aziz, S. L. M. Schroeder, *Faraday Discuss.*, 2015, **179**, 269.
[23] H. G. Alison, R. J. Davey, J. Garside, M. J. Quayle, G. J. T. Tiddy, D. T. Clarke, G. R. Jones, *Phys. Chem. Chem. Phys.*, 2003, **5**, 4998.
[24] M. Faubel, S. Schlemmer, J. P. Toennies, *Z. Phys. D*: *At. Mol. Clusters*, 1988, **10**, 269.
[25] K. R. Wilson, B. S. Rude, J. Smith, C. Cappa, D. T. Co, R. D. Schaller, M. Larsson, T. Catalano, R. J. Saykally, *Rev. Sci. Instrum.*, 2004, **75**, 725.
[26] K. M. Lange, R. Könnecke, S. Ghadimi, R. Golnak, M. A. Soldatov, K. F. Hodeck, A. Soldatov, E. F. Aziz, *Chem. Phys.*, 2010, **377**, 1.
[27] L. Gontrani, R. Caminiti, L. Beneivenni, C. Sadun, *Chem. Phys. Lett.*, 1999, **301**, 131.
[28] B. M. Messer, C. D. Cappa, J. D. Smith, K. R. Wilson, M. K. Gilles, R. C. Cohen, R. J. Saykally, *J. Phys. Chem. B*, 2005, **109**, 5375.
[29] L. Stryer, *Biochemistry*, W. H. Freeman and Co., New York, 1975.

[30] M. L. Gordon, G. Cooper, C. Morin, T. Araki, C. C. Turci, K. Kaznatcheev, A. P. Hitchcock, *J. Phys. Chem. A*, 2003, **107**, 6144.
[31] Y. Zubavichus, M. Zharnikov, A. Schaporenko, M. J. Grunze, *Electron Spectrosc. Relat. Phenom.*, 2004, **134**, 25.
[32] B. M. Messer, C. D. Cappa, J. D. Smith, W. S. Drisdell, C. P. Schwartz, R. C. Cohen, R. J. Saykally, *J. Phys. Chem. B*, 2005, **109**, 21640.
[33] D. N. Kelly, C. P. Schwartz, J. S. Uejio, A. M. Duffin, A. H. England, R. J. Saykally, *J. Chem. Phys.*, 2010, **133**, 101103.
[34] H. Sigel, N. Corfu, *Eur. J. Biochem.*, 1996, **240**, 508.
[35] K. Scheller, F. Hofstetter, P. Mitchell, B. Prijs, H. Sigel, *J. Am. Chem. Soc.*, 1981, **103**, 247.
[36] M. Ukai, A. Yokoya, K. Fujii, Y. Saitoh, *Radiat. Phys. Chem.*, 2008, **77**, 1265.
[37] Y. Zubavichus, A. Shaporenko, V. Korolkov, M. Grunze, M. Zharnikov, *J. Phys. Chem. B*, 2008, **112**, 13711.
[38] O. Plekan, V. Feyer, R. Richter, M. Coreno, M. de Simone, K. Prince, A. Trofimov, E. Gromov, I. Zaytseva, J. Schirmer, *Chem. Phys.*, 2008, **347**, 360.
[39] J. S. Uejio, C. P. Schwartz, A. M. Duffin, W. S. Drisdell, R. C. Cohen, R. J. Saykally, *Proc. Natl. Acad. Sci. U.S.A.*, 2008, **105**, 6809.
[40] M. Ide, Y. Maeda, H. Kitano, *J. Phys. Chem. B*, 1997, **101**, 7022.
[41] E. F. Aziz, N. Ottosson, S. Bonhommeau, N. Bergmann, W. Eberhardt, M. Chergui, *Phys. Rev. Lett.*, 2009, **102**, 68103.
[42] K. Kaznacheyev, A. Osanna, C. Jacobsen, O. Plashkevych, O. Vahtras, H. Agren, *J. Phys. Chem. A*, 2002, **106**, 3153.
[43] R. Golnak, K. Atak, E. Suljoti, K. F. Hodeck, K. M. Lange, M. A. Soldatov, N. Engel, E. F. Aziz, *Phys. Chem. Chem. Phys.*, 2013, **15**, 8046.
[44] Y. Horikawa, T. Tokushima, Y. Harada, O. Takahashi, A. Chainani, Y. Senba, H. Ohashi, A. Hiraya, S. Shin, *Phys. Chem. Chem. Phys.*, 2009, **11**, 8676.
[45] J. Gråsjö, E. Andersson, J. Forsberg, L. Duda, E. Henke, W. Pokapanich, O. Bjorneholm, J. Andersson, A. Pietzsch, F. Hennies, J. E. Rubensson, *J. Phys. Chem. B*, 2009, **113**, 16002.
[46] K. Atak, S. I. Bokarev, M. Gotz, R. Golnak, K. M. Lange, N. Engel, M. Dantz, E. Suljoti, O. Kühn, E. F. Aziz, *J. Phys. Chem. B*, 2013, **117**, 12613.
[47] C. P. Schwartz, J. S. Uejio, A. M. Duffin, A. H. England, D. N. Kelly, D. Prendergast, R. J. Saykally, *Proc. Natl. Acad. Sci. USA*, 2010, **107**, 14008.
[48] N. Engel, S. I. Bokarev, E. Suljoti, R. Garcia-Diez, K. M. Lange, K. Atak, R. Golnak, A. Kothe, M. Dantz, O. Kühn, E. F. Aziz, *J. Phys. Chem. B*, 2014, **118**, 1555.
[49] Y. Tamenori, O. Takahashi, K. Yamashita, T. Yamaguchi, K. Okada, *J. Chem. Phys.*, 2009, **131**, 174311.
[50] K. R. Wilson, M. Cavalleri, B. S. Rude, R. D. Schaller, A. Nilsson, L. G. M. Pettersson, N. Goldman, T. Catalano, J. D. Bozek, R. J. J. Saykally, *Physics-Condens. Matter*, 2002, **14**, L221.
[51] K. R. Wilson, R. D. Schaller, D. T. Co, R. J. Saykally, B. S. Rude, T. Catalano, J. D. Bozek, *J. Chem. Phys.*, 2002, **117**, 7738.
[52] M. Cavalleri, H. Ogasawara, L. G. M. Pettersson, A. Nilsson, *Chem. Phys. Lett.*, 2002, **364**, 363.
[53] H. Bluhm, D. F. Ogletree, C. S. Fadley, Z. Hussain, N. Salmeron, *J. Phys.-Condens. Matter*, 2002, **14**, L227.
[54] P. Parent, C. Laffon, C. Mangeney, F. Bournel, M. Tronc, *J. Chem. Phys.*, 2002, **117**, 10842.
[55] U. Bergmann, P. Wernet, P. Glatzel, M. Cavalleri, L. G. M. Pettersson, A. Nilsson,

S. P. Cramer, *Phys. Rev. B*, 2002, **6609**, U30.
[56] O. Björneholm, F. Federmann, S. Kakar, T. Möller, *J. Chem. Phys.*, 1999, **111**, 546.
[57] D. T. Bowron, M. H. Krisch, A. C. Barnes, J. L. Finney, A. Kaprolat, M. Lorenzen, *Phys. Rev. B*, 2000, **62**, R9223.
[58] S. Myneni, Y. Luo, L. Å. Näslund, M. Cavalleri, L. Ojamäe, H. Ogasawara, A. Pelmenschikov, P. Wernet, P. Väterlein, C. Heske, Z. Hussain, L. G. M. Pettersson, A. Nilsson, *J. Phys.-Condens. Matter*, 2002, **14**, L213.
[59] P. Wernet, D. Nordlund, U. Bergmann, M. Cavalleri, M. Odelius, H. Ogasawara, L. Å. Näslund, T. K. Hirsch, L. Ojamäe, P. Glatzel, L. G. M. Pettersson, A. Nilsson, *Science*, 2004, **304**, 995.
[60] G. A. Jeffrey, *An Introduction to Hydrogen Bonding*, Oxford University Press, Oxford, 1997.
[61] J. M. Hermida-Ramon, M. A. Rios, *J. Phys. Chem. A*, 1998, **102**, 2594.
[62] D. J. Frurip, L. A. Curtiss, M. Blander, *J. Phys. Chem.*, 1978, **82**, 2555.
[63] E. Knözinger, R. Wittenbeck, *J. Mol. Spectrosc.*, 1984, **105**, 314.
[64] X.-S. Xu, *J. Struct. Chem.*, 2008, **27**, 316.
[65] I. H. Suzuki, N. Saito, *Chem. Phys.*, 2000, **253**, 351.
[66] I. H. Suzuki, N. Saito, *Int. J. Mass Spectrom.*, 2000, **198**, 165.
[67] M. C. Nelson, J. Murakami, S. L. Anderson, D. M. Hanson, *J. Chem. Phys.*, 1987, **86**, 4442.
[68] P. Ball, *Chem. Rev.* 2008, **108**, 74.
[69] E. W. Lang, H. D. Lüdemann, *Angew. Chem.*, 1982, **94**, 351; *Angew. Chem. Int. Ed.* 1982, **21**, 315.
[70] K. M. Lange, A. Kothe, E. F. Aziz, *Phys. Chem. Chem. Phys.*, 2012, **14**, 5331.
[71] K. M. Lange, E. Suljoti, E. F. Aziz, *J. Electron Spectrosc. Relat. Phenom.*, 2013, **188,** 101.
[72] T. Tokushima, Y. Harada, O. Takahashi, Y. Senba, H. Ohashi, L. G. M. Pettersson, A. Nilsson, S. Shin, *Chem. Phys. Lett.*, 2008, **460**, 387.
[73] T. Tokushima, Y. Harada, Y. Horikawa, O. Takahashi, Y. Senba, H. Ohashi, L. G. M. Pettersson, A. Nilsson, S. Shin, *J. Electron Spectrosc. Relat. Phenom.*, 2010, **177**, 192.
[74] O. Fuchs, M. Zharnikov, L. Weinhardt, M. Blum, M. Weigand, Y. Zubavichus, M. Bar, F. Maier, J. D. Denlinger, C. K. Heske, M. Grunze, E. Umbach, *Phys. Rev. Lett.*, 2008, **100**, 027801.
[75] N. Engel, K. Atak, K. M. Lange, M. Gotz, M. Soldatov, R. Golnak, E. Suljoti, J.-E. Rubensson, E. F. Aziz, *J. Phys. Chem. Lett.*, 2012, **3**, 3697.
[76] K. M. Lange, E. F. Aziz, *Chem. Soc. Rev.*, 2013, **42**, 6840.
[77] R. Golnak, J. Xiao, K. Atak, J. S. Stevens, A. Gainar, S. L. M. Schroeder, E. F. Aziz, *Phys. Chem. Chem. Phys.*, 2015, **17**, 29000.
[78] B. Dierker, E. Suljoti, K. Atak, K. M. Lange, N. Engel, R. Golnak, M. Dantz, K. Hodeck, M. Khan, N. Kosugi, E. F. Aziz, *New J. Phys.*, 2013, **15**, 093025.
[79] T. Tokushima, Y. Horikawa, H. Arai, Y. Harada, O. Takahashi, L. G. M. Pettersson, A. Nilsson, S. Shin, *J. Chem. Phys.*, 2012, **136**, 04451.

CHAPTER 3

METHODOLOGY

3.1. X-ray absorption spectroscopy

3.1.1. General aspects

Synchrotron radiation is generated by charged particles in accelerated motion, travelling at the speed of light owing to a perturbation field that interacts with their corresponding electric field. In synchrotron radiation sources (*i.e.* storage rings periodically re-filled with electrons that are originally accelerated in the linear accelerator and the booster), the electrons move along a circular path (Figure 3.1) with relativistic energies as they pass through bending magnets to produce X-rays.[1] Along the storage ring, the radio-frequency generators have the role of feeding energy to the circulating electron bunches to compensate for their energy losses during the X-ray emission. The resulting radiation is then directed towards beamlines surrounding the storage ring in the experimental hall.

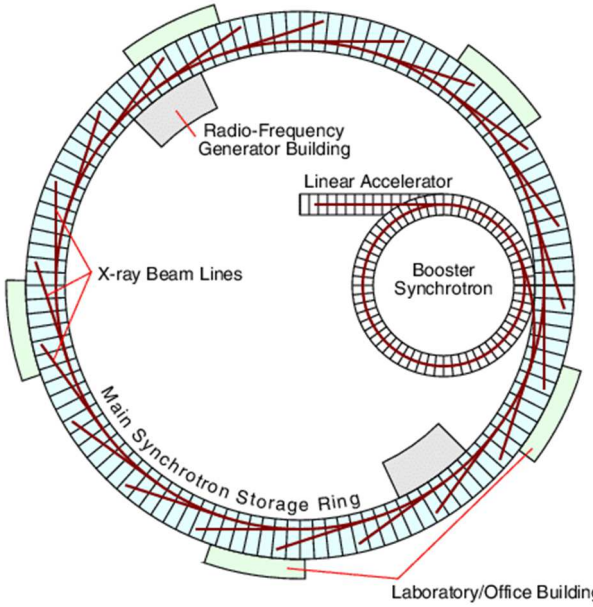


Figure 3.1. Schematic of a synchrotron light source. Figure reproduced from reference [2].

X-rays, discovered by Röntgen in 1895, can lead to electron promotions from the atomic core level. The electrons ejected out in the continuum possess certain energies that equal the difference in energy between the incoming beam, and the electron in the ground state; this phenomenon is called the photoelectric effect, for which Einstein later received a Nobel Prize in 1921.

X-ray absorption spectroscopy (XAS) involves the use of X-rays with the effect of promoting electrons from inner core levels to unoccupied levels.[3] The ionisation potential (IP) is a positive quantity of energy, corresponding to the energy required for an electron from an inner core shell to be ejected out in the continuum.

The XAS spectrum can be split in two main regions: near-edge X-ray absorption fine structure (NEXAFS) and extended X-ray absorption fine structure (EXAFS), represented in Figure 3.2.

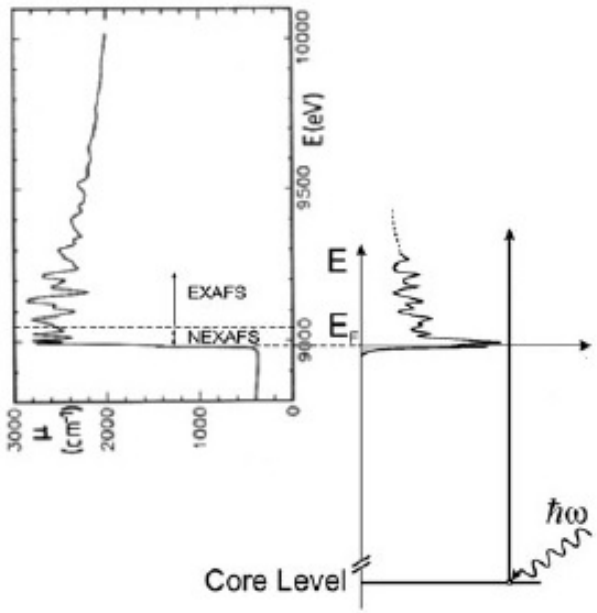


Figure 3.2. Schematic representation of an XAS spectrum where the absorption coefficient, μ, is plotted as a function of energy (the dotted line delimits the NEXAFS and EXAFS regions), and the energy diagram of the sample irradiation process at the core level with photons of energy $\hbar\omega$ ($\hbar$ – reduced Planck constant, ω – angular frequency), where E_F represents the Fermi energy level. Figure reproduced from reference [3].

EXAFS extends from circa 40-50 eV beyond the IP up to a few hundred eV, while NEXAFS refers more to the region around the IP. NEXAFS supplies information on the electronic structure, local structure/environment and chemical state, being complementary to EXAFS, which provides details on the coordination number and oxidation state. Additionally, the terms NEXAFS and X-ray absorption near-edge structure (XANES) are interchangeable, although NEXAFS usually refers to soft X-rays applied to organic elements, whereas XANES refers to the use of hard X-rays for elements with a higher atomic number (generally inorganic systems).

Measurements of C, N and O K-edge spectra have started from the early 1970s. The initial experiments were implemented on single-crystals providing information on the structure of small organic molecules immobilised on the surface of polymer films.[4]

NEXAFS measurements are sensitive tools for investigating electronic and structural properties, with large applications to gas-phase atoms or molecules, species adsorbed on the surface, solid state or solutions.[4] It was also demonstrated that NEXAFS provides information on the existence of hydrogen-bond interactions, the first measurements being applied to liquid water and ice,[5-7] where two different types of –OH groups exist in the hydrogen-bonded water molecules: hydrogen-donor and non-hydrogen-donor –OH groups.

3.1.2. NEXAFS technique

The NEXAFS technique uses X-rays to irradiate the sample. When a photon of a known energy hits an atom, its energy is transferred to the atomic system.[8] Figure 3.3 illustrates the cascade of processes that occur upon irradiation. If the energy of the photon is higher than the binding energy of an inner core-shell electron, the electron is ejected out into the continuum, and a core hole is created (photoelectron emission), which is subsequently filled by another electron originating from a higher energy level through a relaxation process followed by one of the possibilities below:

i. X-ray fluorescence, when the energy resulted from relaxation is converted under the form of an emitted photon
ii. Auger effect, when the resulted energy is used for the ejection of another electron (Auger electron)

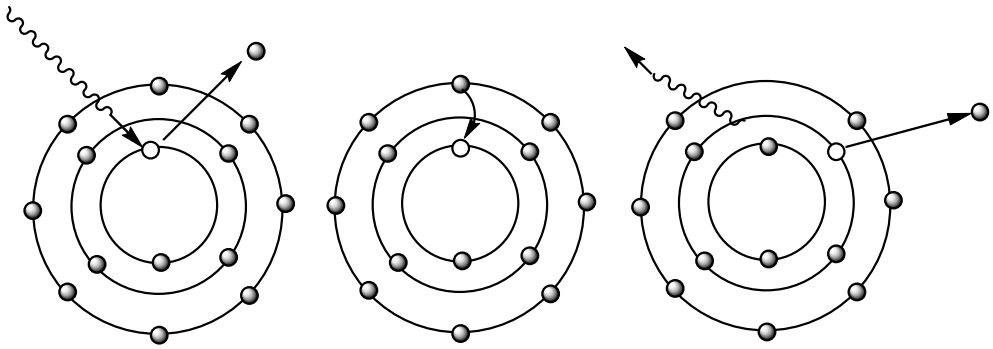


Figure 3.3. X-ray photoelectron emission (left), relaxation (middle), fluorescence/emission of an Auger electron (right).

3.1.3. The absorption coefficient

Beer's law can be expressed as:

$$I = I_0 \times e^{-\mu \times d} \qquad \textit{Equation 3.1}$$

This relationship represents an exponential dependence of the X-ray intensity, *I*, on the sample thickness, *d*, starting from the position where the incident X-ray intensity, I_0, hits the sample (Figure 3.4).

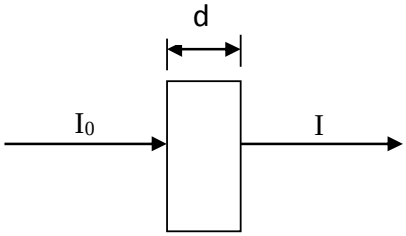


Figure 3.4. Schematic representation of a sample of thickness *d*, irradiated with an incident X-ray intensity, I_0; after the absorption process, the outcoming radiation has a reduced intensity, *I*.

The coefficient that accompanies the thickness, *d*, in Equation 3.2 is called the absorption coefficient or absorptivity, μ, and in NEXAFS, it is plotted against the photon energies.

The absorptivity is a decreasing function of energy,[9] and can be approximated as a value that depends on the sample density, ρ, the atomic number, *Z*, atomic mass, *A*, and the X-ray energy, *E*:

$$\mu \approx \frac{\rho Z^4}{AE^3}$$ *Equation 3.2*

As μ is proportional to Z^4, there are visible differences of several orders of magnitude between organic (*e.g.* C, N, O, S), and inorganic elements, especially metals (*e.g.* Na, Fe, Pb), as shown in Figure 3.5.

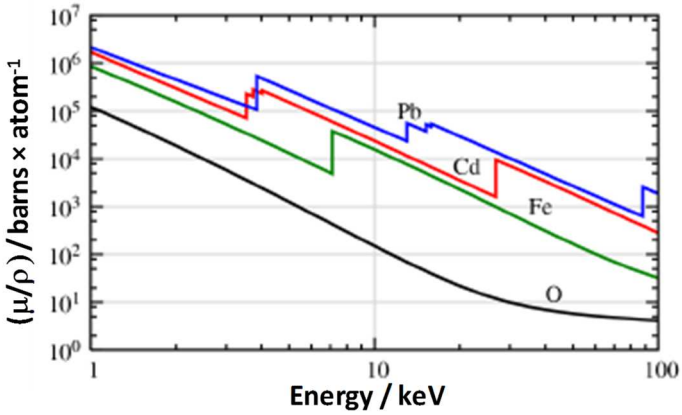


Figure 3.5. Absorption cross-section (μ/ρ) for several elements of X-ray energy over the range 1 to 100 keV. The sharp edges are assigned to core-level binding energies of the atoms.

3.1.4. Koopman's theorem

Koopman's theorem is a well-known theoretical approach which assumes that the IP of a specific electron, annotated as *i*, equals the negative value of its orbital energy, E_i.

At the same time, the IP derives from the energy difference between the excited species, E^* (obtained by promotion of the electron *i*), and the ground-state species, E_0.[10] Therefore, these relationships can be written as:

$$IP = E^* - E_0 = -E_i$$ *Equation 3.3*

The IP is a positive energetic quantity due the ionisation process, which requires energy from outside for the electron to be promoted to an unoccupied state. Following Equation 3.3, the orbital energy calculated for the electron *i* is negative for bound states. Additionally, the orbital energy, E_i, represents the sum of kinetic energy, potential energy within the nuclear field and the interaction energy of electron *i* with the other electrons.[4]

In NEXAFS, a core hole is generated upon electron promotion from a ground state (*i.e.* initial state) to an unoccupied state, determining the other electronic orbitals to become

contracted, and therefore will have higher absolute energies with respect to the vacuum level, which represents a final-state effect. When resonant inelastic X-ray spectroscopy (RIXS) is involved, the core hole subsequently decays when it is filled back with a valence electron during a relaxation process that could also introduce final-state effects on the molecular orbital energies.

3.1.5. Origin of molecular orbitals

Using a simplistic general approach available for elements from the first and second row of the periodic table, the atoms in the structure of a molecule may form bonds that involve 1s (*i.e.* hydrogen, helium), or 2s and 2p (*i.e.* second row elements) electrons. Based on that, the resulting molecular orbitals are described as combinations of the atomic orbitals that form them.

Taking a diatomic molecule as an example and choosing the z axis as being the internuclear axis, the σ molecular orbitals (bonding and anti-bonding) will be formed from atomic 1s, 2s, or $2p_z$ orbitals, while the π orbitals (bonding and anti-bonding) result from mixing of $2p_x$ or $2p_y$ orbitals, as shown in Figure 3.6.

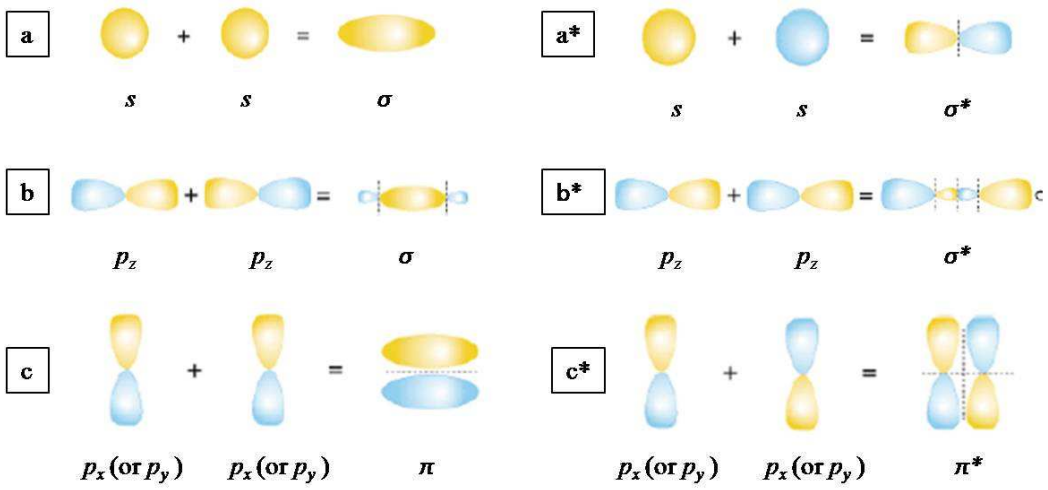


Figure 3.6. Diagram representing the formation of bonding σ and anti-bonding σ* molecular orbitals from individual atomic s orbitals (a and a*) or individual atomic p_z orbitals (b and b*), and formation of bonding π and anti-bonding π* molecular orbitals from p_x (or p_y) individual atomic orbitals (c and c*); the yellow colour represents the positive sign of the wave function characterising the orbital, while the blue colour is the negative sign of the wave function characterising the orbital. Figure adapted from reference [11].

In the simple case of diatomic molecules, the σ orbitals possess a symmetry around the internuclear axis, whereas the π orbitals are symmetric with respect to a plane containing the internuclear axis.

3.1.6. Resonances in NEXAFS

Resonances in the K NEXAFS spectra reflect electron transitions from a 1s core-level state to states with energetic values below and 40-50 eV beyond the IP.

Taking a diatomic molecule as a relatively simple example, the molecular orbitals can be classified as occupied (filled with electrons) or unoccupied (vacant). σ* orbitals can be found at higher energies than the IP. π* orbitals are present only in molecules with at least one double or triple bond; this aspect is usually reflected in their characteristic NEXAFS spectra as π* resonances are below the IP, as shown in Figure 3.7. Features just below the IP look sharp and low in intensity; they correspond to the Rydberg type of transitions and are solutions of the Schrödinger equation.

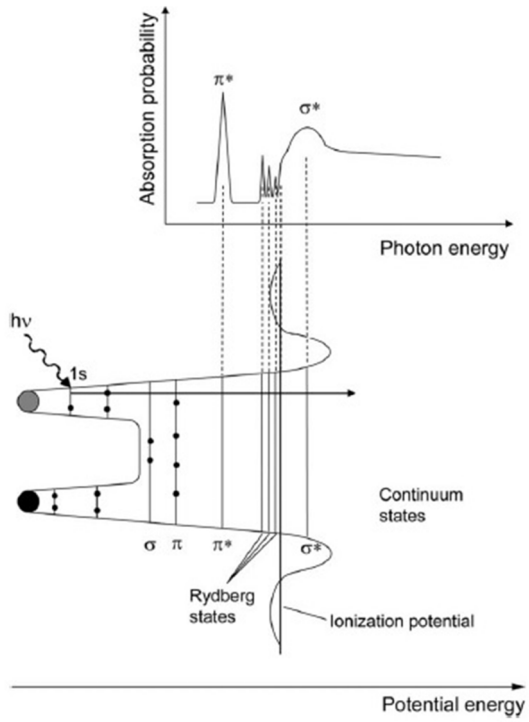


Figure 3.7. Distribution of resonances (marked with dotted lines) in a NEXAFS spectrum for a diatomic molecule, where transitions occur (upon irradiation with photons of energy hν) from 1s core orbitals to unoccupied molecular orbitals; from left to right: π* features correspond to electronic transitions from 1s to π* orbitals, σ* features to transitions from 1s to σ* orbitals, and the sharp step-like small features between π* and σ* represent the Rydberg transitions. The ionisation potential separates the π* transitions from σ* transitions, and continuum states correspond to energies higher than σ* transitions. Figure reproduced from reference [12].

Electrons promoted to π* unoccupied orbitals decay more easily than σ* electrons, due to a smaller energy gap with respect to the core levels. Generally, the width of a resonance depends directly on the probability of an excited electron to decay, therefore π* features appear in the spectra as sharper than the σ* ones.

An interesting aspect is the energy position of the π* and σ* transitions for simple molecules with single, double or even triple bonds (*e.g.* C=C, C=O, C=N, C≡N, N≡N, $H_2C=O$, $F_2C=O$). The approach is restricted specifically to simple molecular systems, due to the fact that here resonances are assumed to involve localised bonds. A K-edge transition involves electron promotions from 1s orbitals to unoccupied molecular orbitals

localised around the bonds, and therefore transitions to the same molecular orbital can be seen from the 1s orbital of each atom participating in the bonding. For example, in the C=O molecule, the π* transition can be noticed in both the C and O K-edge spectra.

Starting with the π* resonances, Table 3.1 contains the calculated relative positions of the π* transitions with respect to the IP. The molecules are grouped upon the increasing order of the total atomic number, Z_{tot} (sum of all individual atomic numbers of the atoms participating to the formation of the bond within the molecule). Inside each of these groups, there are molecules with different bond orders *i.e.* triple and double bonds.[4]

As a general observation, the energy difference, δ, is almost constant inside each group and does not depend on the atom in the periodic table, hybridisation, or bond length.[4] However, this difference increases with Z_{tot} (*e.g.* for C=C, δ is around −6 eV, while for C=O, it is around −8 eV).

Table 3.1. Calculated positions (δ) of the π* resonance energies (*E*) relative to the ionisation potential (IP) for simple molecules containing localised (non-conjugated) double or triple bonds; marked with an asterisk are the atoms for which assignments were made relative to their edges.[4]

Molecule	Z_{tot}	1s ionisation potential IP / eV	π* resonance position *E* / eV	$\delta = E - IP$ / eV
HC≡CH	12	291.1	285.9	−5.2
$H_2C=CH_2$	12	290.8	284.7	−6.1
$F_2C=CF_2$	12	296.5	290.1	−6.4
$HC^*{\equiv}N$	13	293.4	286.4	−7.0
$HC{\equiv}N^*$	13	406.8	399.7	−7.1
N≡N	14	409.9	401.0	−8.9
$C^*{\equiv}O$	14	296.2	287.3	−8.9
$C{\equiv}O^*$	14	542.5	534.1	−8.4
$H_2C^*=O$	14	294.5	286.0	−8.5
$H_2C=O^*$	14	539.4	530.8	−8.6
$F_2C^*=O$	14	299.6	290.9	−8.7
$F_2C=O^*$	14	540.8	532.7	−8.1

Unlike π* resonances, σ* resonances show a more complex energy dependence. In Table 3.2, similar to Table 3.1, the σ* resonance position (*E*), the ionisation potential (IP), the difference between these values (δ), and the bond length are presented for molecules with localised bonds. Similarly, the molecules are grouped upon the increasing order of Z_{tot}.

There is strong evidence that δ decreases when bond length increases within a group of same Z_{tot}, which is equivalent to the fact that in the spectrum, σ* resonances lie closer to the IP as the bond length increases; this concept is also known as 'bond length with a ruler'[13-16] and is an immediate consequence of the σ orbitals lying along the internuclear axis. Although this topic was debated and several controversies were generated in the past,[17-22] there is enough evidence of empirical background, as well as theoretical support[22,23] for its interpretation as multiple scattering resonances dependent on the bond length. σ* shape resonances can be understood as a back and forth scattering of the photoelectron wave between two atoms. Therefore, assuming the theoretical model of a one-dimensional square well potential, it was shown that the position of the σ* resonance relative to the IP (δ) depends on the square of the inverse of the internuclear distance (*R*), as described by Equation 3.4:

$$\delta = \frac{h^2}{8m}\frac{n^2}{R^2}, \qquad \textit{Equation 3.4}$$

where *h* is Planck's constant, *m* the electron mass, *n* is the characteristic quantum number of the energy level in or above the well, *R* the internuclear distance between the absorbing atom and its neighbouring scatterer.[4] In fact, the term value δ can be approximated as a linear dependence on *R* over a narrow bond length range (1.1–1.5 Å), for which measurements are normally made in organic systems.[13,23]

Another observation when analysing the data is that δ decreases with Z_{tot} more rapidly than in the case of π* resonances.[4] However, these trends in energy positions for both σ* and π* relative to IP are restricted to simple molecules where the bonds do not extend over the whole molecule, but are centred around the atoms directly involved in the bond formation. Therefore, more complex systems where bonds are delocalised, or where bond-bond interactions are present (*e.g.* conjugated bonds in aromatic systems), do not fall into this category and may not follow such clear trends.

Table 3.2. Calculated positions (δ) of σ* resonances relative to the ionisation potential (IP) for simple molecules containing localised (non-conjugated) bonds; marked with an asterisk are the atoms for which assignments were made relative to their edges.[4]

Molecule	Z_{tot}	1s ionisation potential IP / eV	σ* resonance position E / eV	$\delta = E - IP$ / eV	Bond length / Å
HC≡CH	12	291.1	310.0	18.9	1.203
$H_2C=CH_2$	12	290.8	301.0	10.2	1.337
$H_3C–CH_3$	12	290.7	291.2	0.5	1.533
$HC^*≡N$	13	293.4	307.9	14.5	1.158
$HC≡N^*$	13	406.8	420.8	14.0	1.158
$H_3C^*–NH_2$	13	291.6	291.5	−0.1	1.465
$H_3C–N^*H_2$	13	405.1	404.6	−0.5	1.465
$F_3C–N^*O$	13	410.5	407.5	−3.0	1.546
N≡N	14	409.9	418.9	9.0	1.095
$C^*≡O$	14	296.2	304.0	7.8	1.128
$C≡O^*$	14	542.5	550.0	7.5	1.128
$H_2C^*=O$	14	294.5	300.5	6.0	1.209
$H_2C=O^*$	14	539.4	544.0	4.6	1.209
$B^*–F_3$	14	202.8	205.1	2.3	1.313
$B–F^*_3$	14	694.8	699.0	4.2	1.313

3.1.7. Detection methods

The most common electron-based detection methods employed for soft X-ray measurements are partial electron yield (PEY) and total electron yield (TEY), as the X-ray transmission is normally preferred for hard X-rays (> 5-10 keV).[24]

In the PEY, only electrons with a kinetic energy that is over a threshold limit (Auger electrons) are detected.[4,25] This method is very surface sensitive because the mean free path of an Auger electron is around a few tenths of 1 Å. The electrons with a kinetic energy smaller than the threshold limit are stopped from entering the detector with the help of a retarding grid of negative potential.

Alternatively, the TEY involves detection of all the generated electrons, independent of their kinetic energy, and gives bulk sensitive information because all the photoemitted

electrons including the ones that may have suffered energy loss after inelastic scattering[25] are taken into account. Here, electrons with a low energy (below ~ 20 eV) dominate and form an '*inelastic tail'*. Only a part of these represent inelastically scattered Auger electrons from the adsorbate (~ 1 %) and are involved in the generated NEXAFS signal of the spectrum.[15]

When irradiating the adsorbate with X-rays, the photons penetrate the substrate to a depth given by the inverse of the absorptivity, φ. This is also called the linear X-ray absorption coefficient and depends on the atomic volume density (ρ_v) of the adsorbate, and on the absorption cross section (σ).

$$\varphi = \rho_v \times \sigma \qquad \textit{Equation 3.5}$$

Generally, the inelastic mean free path of the electron (also called the electron scattering length) is a function of kinetic energy.[26,27] Upon photon absorption of energies comprised in the range of 250 eV to 600 eV, the mean free paths of the primary photoelectrons and subsequently generated Auger electrons are smaller than 10 Å. These primary electrons form a cascade during the inelastic scattering until they reach the surface, yet only a fraction of them possess an energy that is high enough to cross the surface potential barrier. Therefore, the TEY signal will be thus limited to the optimum electron escape depth needed for this fraction of electrons.

Calculations for the electron yield are generally numerical approximations due to the cascade effects and electron scattering that occur upon irradiation. The aim of finding a relationship between the TEY (or PEY) as functions of photon energy, X-ray incidence angle and electron kinetic energy is difficult to achieve and a comprehensive theory would seem ambitious to develop; instead of complex assumptions, an approximate approach for calculating these yields is generally used. The theory built around calculating the TEY as effective as possible consists in presenting the X-ray which interacts under an incidence angle α with the exposed sample. The absorption coefficient φ(α) is a measure of this X-ray–sample interaction and is expressed by the relationship:

$$\varphi(\alpha) = \frac{\varphi}{\sin \alpha} \qquad \textit{Equation 3.6}$$

At an angle α < 90°, the linear absorptivity φ(α) is greater than the φ, which is the linear absorption at a normal incidence angle.

In the fluorescence yield (FY) method, the photons emitted upon sample irradiation are counted with the help of a detector called a photodiode. FY is bulk sensitive, providing information on a deep region of the sample, as opposed to PEY and TEY techniques, which are more surface sensitive because of the small escape depth of the photogenerated electrons.[28] As it is a photon-in/photon-out technique, FY is able to monitor surface and bulk reactions over a large range of surface temperatures.[29]

3.1.8. Detectors

In the case of electron yield detection, the X-ray beam resulting from the monochromator is collimated with the help of mobile blades and meets a grid which was *in situ* coated with a metal that has no absorption edge in the energy range of interest. A typical example is represented by gold, which may be sometimes manufactured in such a way that it is covered with a monolayer of carbon and oxygen adsorbates that are transparent within the analysis K-edge range. In the case of TEY, electrons coming off the grid are attracted towards the channeltron by a small potential difference of ~ 50 V. A floating battery box renders a positive potential in the range of 2-3 kV at the other end of the channeltron. Voltage is then converted to frequency and fed into a computer scaler. The signal from the sample is then detected by a cylindrical mirror analyser. For the PEY, the simple partial yield detector usually consists of two high transmission metal grids for retardation and a double channel plate for electron multiplication. The first grid is maintained at the ground voltage and the second at a negative retarding voltage, which can be in the range of −230 to −180 eV for carbon, −340 to −290 eV for nitrogen and −480 to −430 eV for oxygen. The phosphorus covered screen is a component that helps for the alignment of the X-ray beam in the chamber and of the sample in the beam. For soft X-ray spectroscopies, the FY detectors are represented by semiconductor diodes (*e.g.* GaAsP) that are sensitive to incoming photons above a threshold energy value, for which their quantum yield increases with photon energy.

3.2. X-ray emission spectroscopy

X-ray emission spectroscopy (XES) was first implemented a few decades ago in the area of solid-state physical chemistry, but underwent only recently a remarkable evolution.[30-36]

XES is a spectroscopy complementary to XAS with much greater information insight, as apart from the photoexcitation of a core-level electron, it takes into account decay of an upper energy level electron to the resulted core hole, with the emission of a photon. The simplest case is the non-resonant X-ray emission spectroscopy, when the excitation energy is sufficiently high to remove the core level electron to the continuum. Alternatively, when the core level electron is excited to an energy level close around the IP (as in the case of NEXAFS), the emission spectrum depends on the energy of the incident photon. This type of technique is called resonant X-ray emission spectroscopy (RXES) as it resonates with one of the X-ray transitions encountered in the absorption spectra.

A particular type of RXES is RIXS in that it provides bulk-sensitive details due to its photon-in/photon-out characteristic.[37] The intensity of the signal is usually much weaker than that of an XAS process because the photon emission presents a reduced efficiency. Among the particular features of RIXS, the element and orbital specificity are the most important as it can differentiate between inequivalent chemical bonding sites. As depicted in Figure 3.8, in most of the cases RIXS involves excitation of a 1s core electron to an unoccupied orbital, thus creating a core hole, which is subsequently filled by an electron from an occupied valence orbital.[38]

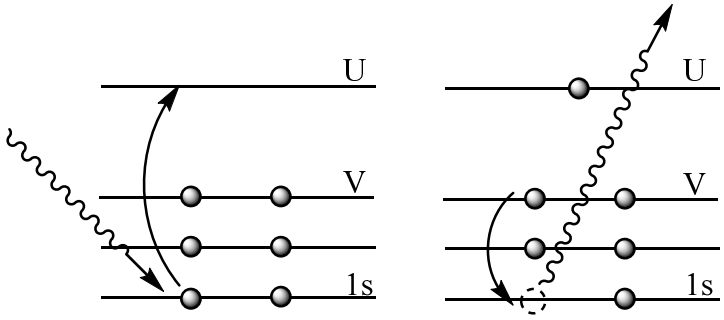


Figure 3.8. Excitation of a core-level 1s electron to an unoccupied orbital (U) after irradiation (left), followed by a decay of an electron from a valence orbital (V) to the 1s orbital, with the emission of a photon (right).

RIXS is bulk sensitive because the penetration depth of the hard X-ray incident photons is of about several microns, while for soft X-ray incident photons this is 0.1 microns at maximum. Another characteristic of RIXS is that it requires comparably smaller sample volumes for analysis (*e.g.* surfaces, thin films) than other techniques that involve neutron sources. In spite of these considerable advantages, RIXS is not as widely used as NEXAFS or X-ray photoelectron spectroscopy (XPS) because it involves a substantial incident photon flux to obtain high-resolution spectra. Another main practical limitation is the energy resolution, which is determined by the availability of photon sources of high enough intensity.

Depending on the changes in the emitted photon's energy and momentum, the scattering mechanisms can be direct or indirect.[39,40] In the case of direct RIXS, the electron is excited from the core level to an empty valence orbital, generating a core hole, which is then filled by a decaying electron from a different valence orbital. For this direct process to occur, both of these transitions involved need to be allowed. Alternatively, the indirect RIXS is more complex. Here, the core level electron is promoted to an unoccupied orbital above the IP, where it decays to fill back the core hole. In both cases, a quasi-elastic line with energy losses due to vibrations is generated in the spectrum.

3.3. Liquid microjet technique

3.3.1. General aspects

Standard use of X-ray core-level spectroscopies involves solid surfaces with the aim to obtain various details on the local environment, electronic structure or local geometry. In the case of molecules naturally found in solution-state (*e.g.* proteins, peptides), their local environment may be altered if these are in a solid, non-dissolved state; thus, the liquid microjet X-ray spectroscopy becomes suitable here. The uniqueness of this revolutionary technique presents the advantage that the potentially induced X-ray beam damage and heating of the analysed sample can be avoided due to its continuous renewal. Moreover, the absence of a membrane in the design of the experimental setup avoids several potential issues presented below:

i. the membrane is normally transparent for the beam energy range used so that the materials this is made of do not interact with the X-rays; nevertheless, measuring techniques that require a high incoming influx will face photon loss due to generally longer data acquisition times. For example, if the beam intensity is reduced in dilute solutions, the measurements may be compromised.

ii. hydrophilic or hydrophobic interactions between sample and membrane may reflect changes in the resulting spectra.

iii. oxidation of the Si_3N_4 membrane over time may change the sample consistency, and induce an apparent molecular orientation in the measured spectra.

3.3.2. Experimental setups

With respect to soft X-ray spectroscopies for liquids, the high vapour pressure in detriment to ultra-high vacuum challenging conditions determined a slow evolution in commissioning compatible experimental setups. The first attempt was an enclosed static liquid cell designed at Uppsala University, where liquid water was investigated behind a thin Si_3N_4 membrane.[41] In order to reach better energy resolutions, a slitless spectrometer equipped with a cylindrical varied line spacing (VLS) grating and an illuminated charge-coupled device (CCD) detector was developed for measurements of liquid water running through a flow cell.[42,43] Several years later, the group from Uppsala University adapted the static cell to a flow cell[44] to investigate the electronic structure of various functional groups in aqueous solutions.

In 1997, Faubel developed the first fast-flowing apparatus for photoelectron spectroscopy suitable for volatile liquids using the liquid microjet technique under vacuum conditions.[45] An updated version of this setup is currently available at beamline U125 from BESSY (Figure 3.9), and consists of a main (experimental) chamber where the microjet coming out of the nozzle is irradiated perpendicularly with X-rays. The resulting photoelectrons traverse a skimmer into the electron analyser (detection) chamber. The pressure in the experimental chamber (~ 10^{-5} mbar) is maintained by a turbomolecular pump and a liquid nitrogen trap, while the pressure in the detection chamber (~ 10^{-9} mbar) is maintained by two turbomolecular pumps.[46] The pressure difference is delivered by a differential pump (skimmer) that is placed between the liquid jet chamber and the beamline chamber.

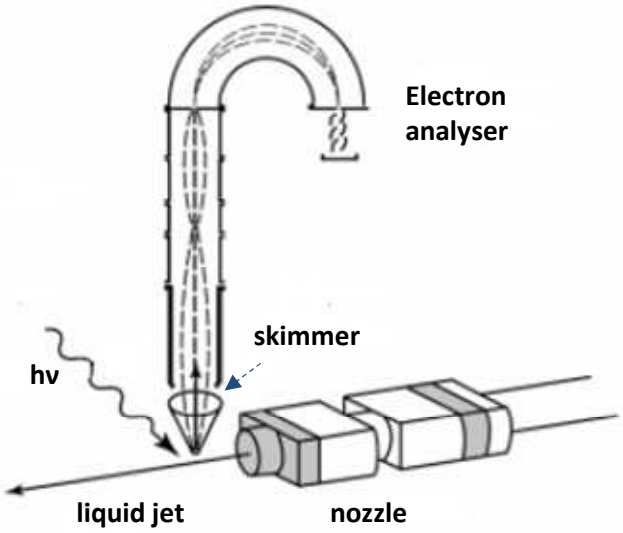


Figure 3.9. Experimental setup for photoelectron spectroscopy (XPS) using a liquid microjet at the third-generation synchrotron radiation facility from BESSY II, comprising a skimmer through which photoelectrons pass from the liquid jet chamber towards the detection chamber. Figure adapted from reference [46].

The results presented in this thesis were carried out at the U41-PGM beamline from the electron storage ring station BESSY II, Helmholtz Zentrum Berlin, using the LiXEdrom endstation.[47-50] The strong features of LiXEdrom is that the liquid jet coming out of the nozzle is extremely small (~ 20 μm) and owing to the fast speed of ~ 80 m/s the pressure in the vacuum chamber can be maintained without the need of a membrane.

The spectrometer possesses a GaAsP diode that records the total fluorescence yield, while the partial fluorescence yield is recorded by a CCD fluorescence screen and microchannel plate (MCP). By rotating the holder, one can choose between different blazed gratings on a range from 20 to 1000 eV (Grating #1: 20-50 eV with a line density 400 l/mm; Grating #2: 50-175 eV with a line density 600 l/mm; Grating #3: 175-400 eV with a line density 1200 l/mm; Grating #4: 400-1000 eV with a line density 1200 lines/mm), as shown in Figure 3.10.

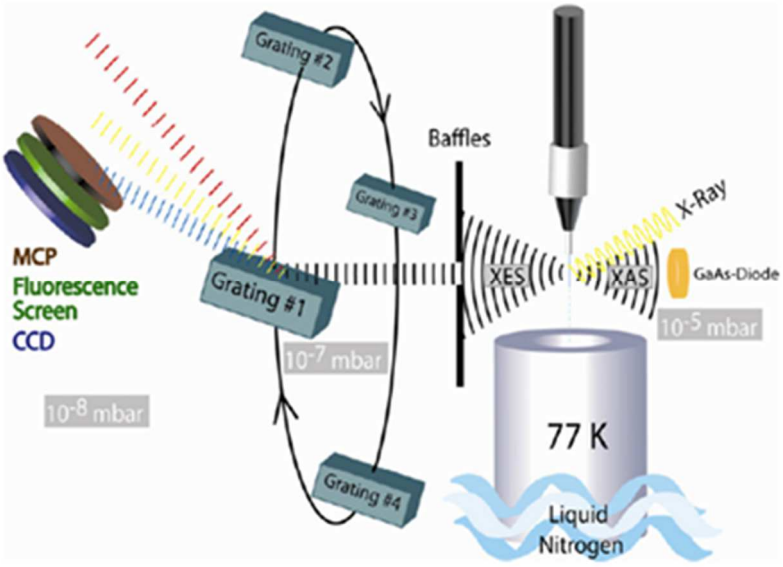


Figure 3.10. Schematic arrangement of the LiXEdrom endstation with the liquid jet delivered through a nozzle and irradiated by the X-ray beam; XAS (NEXAFS) and XES (RIXS) measurements can be recorded simultaneously within the laminar region of the jet, after which droplets start to form and subsequently freeze, being trapped inside of a liquid nitrogen (low-temperature) container. Figure reproduced from reference [50].

3.3.3. Experimental parameters

Solvents that were previously used for the liquid jet are water, acetone, acetonitrile, methanol, ethanol, isopropanol and dimethylsulfoxyde. The maximum volume of the stock solution that can be used is 1000 mL, on a jet flow rate in the range of 0.4-1 mL/min. The solution for analysis is loaded into the liquid jet reservoir by a peristaltic pump (Figure 3.11) and then injected through the microjet nozzle into the NEXAFS chamber, by a high-performance liquid chromatography (HPLC) pump. The diameter of the glass capillary nozzle can vary from 15 to 25 μm. Impurities need to be filtered away to avoid blocking the micron size nozzle, therefore microfilters are installed on the pathway of the microjet before exiting the capillary. The flow in the microjet is laminar for the first 3-5 mm from the nozzle followed by droplets prone to freeze after 20 mm.

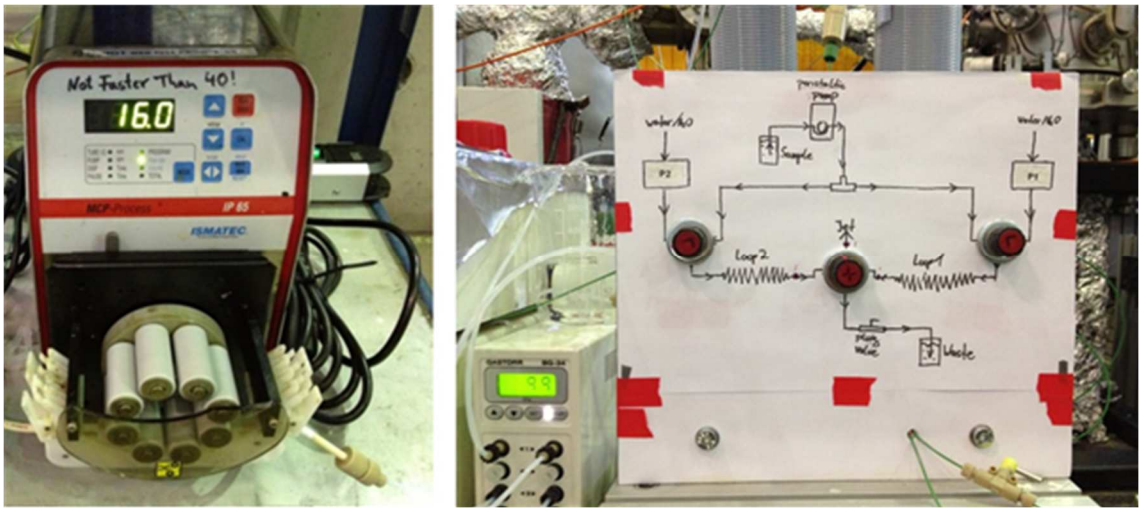


Figure 3.11. Peristaltic pump used for pumping the stock solution in the experimental setup from U41-PGM at BESSY (left), and schematic representation of the solution circuit before reaching the experimental chamber (right), where the sample is pushed by the peristaltic pump into the loops with additional help from pumped water/pure solvent.

3.4. UV-Vis technique

3.4.1. Beer-Lambert law

In the UV-Vis technique, when a certain radiation wavelength matches the energy gap between an occupied orbital (core type, σ or π molecular type) and an unoccupied molecular orbital (σ* or π* type), an electronic transition occurs. The UV-Vis spectrum results from plotting the degree of radiation absorbance through the sample as a function of wavelength. The absorbance is directly proportional to the number of absorbing molecules of the sample (in the case of solutions, to the molar concentration), as illustrated by the Beer-Lambert law:

$$A = \varepsilon \times c \times l, \qquad \textit{Equation 3.7}$$

where A is the absorption, ε is the molar absorptivity, c is the molarity, and l is the path length through the sample.

3.4.2. Electronic transitions in the UV-Vis theory

In Figure 3.12 an energetic diagram of the orbitals is shown. Here, the molecular orbitals are a mixture of the contributing atomic orbitals and for every bonding orbital that is formed (σ, π), there is a corresponding energetically symmetric anti-bonding orbital (σ*, π*). Thus, the lowest energy occupied orbitals are, in general, σ orbitals, which implies that σ* orbitals are the highest in energy. The π orbitals have a higher energy than the σ, while the π* orbitals possess a lower energy than the σ*. The unshared electrons of the non-bonding atomic orbitals (n) lie usually in the middle of the diagram; their energy is higher than the bonding orbitals as from an energetic point of view, these electrons do not contribute to the formation of a bond.

Figure 3.12 intuitively suggests that a transition from a non-bonding orbital to a π* orbital requires usually less energy than a transition from a π to a π* orbital. The transition energy is inversely proportional to the wavelength of the absorbed light:

$$E = \frac{hc}{\lambda}, \qquad \textit{Equation 3.8}$$

where *h* is Planck's constant, *c* is the speed of light, and λ is the wavelength.

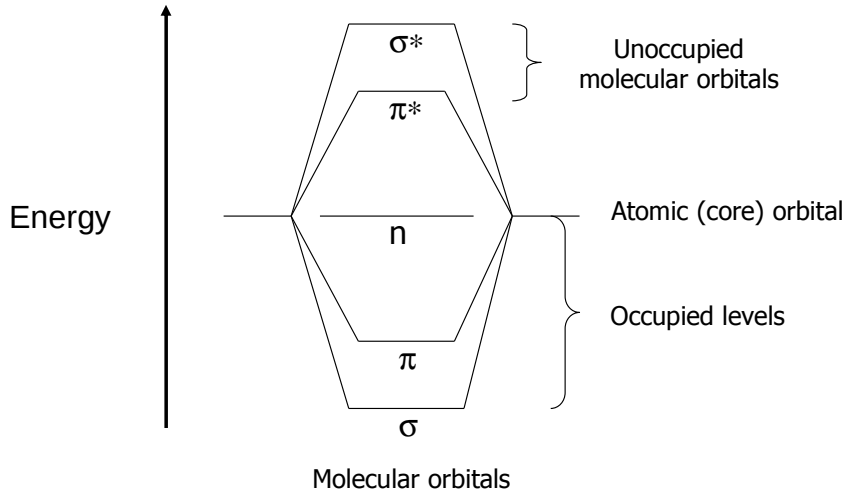


Figure 3.12. Energetic diagram of molecular orbitals showing the bonding (σ, π), non-bonding (n) and anti-bonding orbitals (σ*, π*).

The immediate implication is that n→π* transitions are generally encountered at longer wavelengths than the π→π* transitions in a UV-Vis absorption spectrum; this observation

will be used further in Chapter 7 to outline differences between the spectra for PABA aqueous solutions of high, intermediate and low pH.

3.5. X-ray powder diffraction

3.5.1. Diffraction principles

Knowledge on how atoms are arranged within a crystal structure is essential to understand their physicochemical properties. X-ray powder diffraction (XRPD) derives from single crystal X-ray diffraction and generally provides valuable information on the structure, and atomic spacing involving bond angles, unit cell parameters and bond lengths. This analysis technique is suitable for systems where crystallites are orientated in many directions, and characterised by various interatomic plane distances.

In the diffraction experiment, an X-ray of wavelength comparable to the interatomic spacing hits the sample while a detector records the intensities of the diffracted resulting waves. The diffraction pattern constitutes a spectrum based on atomic periodicities in the sample.[51] Long repeat (large interatomic) distances cause diffraction at small angles, whereas short repeat (small interatomic) distances determine diffraction at wide angles.

The purity of a compound can be reflected by its diffraction pattern; therefore, sharp and well-defined peaks are attributed to defect-free, pure samples, while broadened, distorted or suppressed peaks are addressed to presence of impurities, defects, atomic dislocations or internal strains. For samples without a long-range order (amorphous), the diffraction pattern generally present very broad, diffuse features.[52]

3.5.2. Bragg's law

Figure 3.13 presents the diffraction process occurring when an incident radiation beam hits the surface of a sample.

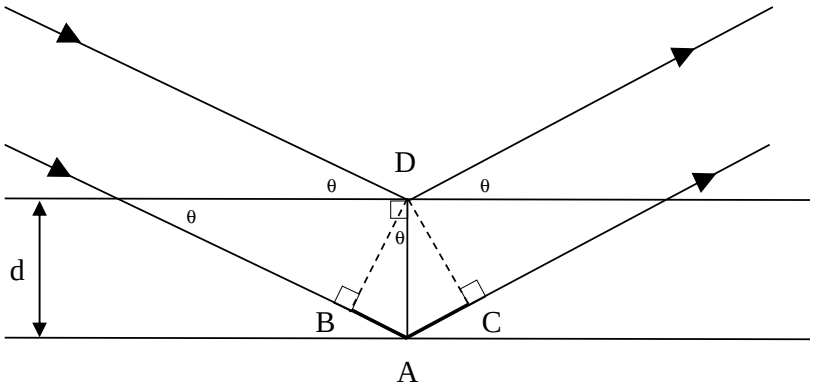


Figure 3.13. Geometry of wave scattering from two planes separated by distance *d*, forming a diffraction angle θ; drawing dashed lines as normal to each wave, the difference in path length for the wave scattered from the top plane and the one scattered from the bottom plane is marked with dark lines, and would represent the sum between the distances AB and AC.

From the trigonometry of the ABD triangle:

$$AB = d \times sin\theta \qquad \textit{Equation 3.9}$$

As ABD and ACD are equivalent triangles, AB = AC.

The immediate consequence is that the path length can be written as:

$$AB + AC = 2 \times d \times sin\theta \qquad \textit{Equation 3.10}$$

The constructive wave interference occurs when the path length equals an integer multiple, *n*, of the radiation wavelength, λ, as in Equation 3.11:

$$n \times \lambda = 2 \times d \times sin\theta, \qquad \textit{Equation 3.11}$$

which is also known as Bragg's law.

For a plane (*hkl*), Equation 3.11 becomes general as:

$$n \times \lambda = 2 \times d_{hkl} \times sin\theta, \qquad \textit{Equation 3.12}$$

where *h*, *k* and *l* are known as Miller indices.

A diffraction pattern of a sample is composed of distinct peaks, each corresponding to an interplanar distance d_{hkl}. In the particular case of cubic crystals, of lattice parameter a_0, the d_{hkl} distance can be written as:

$$d_{hkl} = \frac{a_0}{\sqrt{h^2+k^2+l^2}}, \qquad \textit{Equation 3.13}$$

From Bragg's law (Equation 3.12), the measured angle $2\theta_{hkl}$ that is characteristic to a (*hkl*) diffraction peak becomes:

$$2\theta_{hkl} = 2 \times \arcsin\left(\frac{n \times \lambda \times \sqrt{h^2 + k^2 + l^2}}{2 \times a_0}\right) \qquad \textit{Equation 3.14}$$

3.6. Density functional theory

In the recent years, density functional theory (DFT) has been adopted by quantum chemists as it is widely used for the simulation of potential energy surfaces in molecules.

The main concern of the theoretical aspects lies in the calculation of the ground state energy of a collection of *n* electrons. The energy of this system represents the solution of the non-relativistic Schrödinger equation for a multi-electron atom, which can be rendered by the Born-Oppenheimer time-independent approximation[53] contained in Equation 3.15:

$$\hat{\mathrm{H}}\Psi(r_1, r_2, \ldots, r_n) = E\Psi(r_1, r_2, \ldots, r_n), \qquad \textit{Equation 3.15}$$

where $\hat{H}$ is the Hamiltonian operator, Ψ the state wave function, which depends on *n* coordinates specific to each electron: r_1, r_2, …, r_n, and *E* is the system energy.

DFT differs from other wave function based methods that approximately solve the Schrödinger equation (*e.g.* Hartree-Fock, perturbation theory) by using the electron density concept, which is three-dimensional regardless of the number of electrons, *n*, in the system.[54] The immediate consequence is the possibility of applying the method to large systems comprising even hundreds of atoms, and thus has become one of the most widely used electronic structure approaches. [53,55,56]

Another essential approximation in DFT is represented by the local density approximation (LDA), which has the capability of reproducing in satisfactory manner experimental results, and is even improved compared to the Hartree-Fock theory, when calculating the strength of the molecular bonds. Additionally, lattice constants can be obtained with an accuracy of ~ 1% in the LDA,[57] but has its limitations when considering more rigorous cases such as ferromagnetic materials.

3.7. Time-dependent density functional theory

Time-dependent density functional theory (TD-DFT) investigates the dynamics of systems where time-dependent potentials (*e.g.* electric or magnetic field) have an impact over the excitation energies, photoabsorption spectra and frequency-related properties, unlike DFT which is used to find the ground state of a quantum system.

Zangwill and Soven[58] were the first to extrapolate the DFT approach to the time domain. A significant progress was made in the last 20 years when several physics aspects, such as vibrational and solvent effects, were considered,[59] which improves the accuracy of calculating vertical transitions. An example of an approach that takes into account the solvent shifts on the absorption spectra is the polarisable continuum model (PCM),[60] composed of two parts; the slow part represents molecular orientations that do not have enough time to re-adjust, while the fast part consists of solvent electronic degrees of freedom with re-adjusting capability. In general, the calculated results for excitation energies tend to be approximated with fairly good standard errors (a few tenths of an eV), yet depend on the types of excitation and systems considered.[61] Apart from its applications in spectroscopy, TD-DFT can also be used in real-time dynamics in non-perturbative fields or in determining the ground-state exchange-correlation energy from a response function.

In a TD-DFT method, the direct target to solve is the time-dependent non-relativistic Schrödinger equation for the polyelectronic wave function Ψ:

$$\hat{\mathrm{H}}(t)\Psi(r_1, r_2, \dots, r_n, t) = i\frac{\partial\Psi(r_1, r_2, \dots, r_n, t)}{\partial t}, \qquad \textit{Equation 3.16}$$

where $\hat{H}$ is the Hamiltonian operator, Ψ the state wave function, which depends on n coordinates specific to each electron: r_1, r_2, …, r_n, and time t (i is the imaginary unit).

TD-DFT has yet its own downsides under the consideration of charge-transfer estimations, two- or multi-electron excitations or conduction polymer scale-up due to its limited accuracy of the functionals employed.[62-64]

3.8. Polarisable continuum model

Originally formulated about 35 years ago by Miertus,[65] this model is based on generating overlapping spheres for each atom of the solute molecule, which form a cavity inside of a uniform dielectric continuum that represents the solvent.

In a PCM-solvated system, where the electrostatic, and dispersion-repulsion interactions are considered, both direct and indirect effects can be disseminated. Direct effects are generated by solvent-induced modifications on the electronic charge distribution of the solute, whereas indirect effects are reflected by changes in bond lengths, angles, or torsional angles.

In the PCM method, the free energy of solvation can be represented by a sum of three different terms, *i.e.* of electrostatic, dispersion-repulsion and cavitation type:

$$G_{solvation} = G_{electrostatic} + G_{dispersion-repulsion} + G_{cavitation} \qquad \textit{Equation 3.17}$$

The model system reproduces the case of a very dilute solution composed of a single solute molecule immersed into an infinite solvent 'reservoir'. Each atom within the solute is assigned a van der Waals radius, and then the cavitation energy is computed using a van der Waals surface resulted from the interlocking spheres. The dispersion-repulsion energy is calculated with a solvent-accessible surface with radii augmented by the solvent, and for the electrostatic energy term the initial cavity surface is increased by a scaling factor considering that the first solvation layer does not have the same dielectric properties as the rest of the bulk of the continuum.[66]

The cavity formed at the interface solute-solvent continuum should reproduce the molecular shape as much as possible, otherwise they may lead to deformations in the charge distribution after solvent polarisation that can have an impact on the accuracy of the results.

3.9. Full-multiple scattering

Full-multiple scattering (FMS) calculations are widely used to interpret NEXAFS spectra with the help of the FEFF[67] software tool developed by John Rehr and collaborators. In

these simulations, the absorption coefficient, μ, is plotted as a function of energy and then compared with the experimental results.

After an incident photon hits an atom, the outcome may be either fluorescence or Auger effect. The photoelectron that is emitted is thus scattered at the neighbouring atoms within the molecule, namely it is multiply scattered.

FEFF calculates self-consistent orbitals to obtain the scattering potentials usually using a Hartree-Fock approach and can reproduce well NEXAFS spectra; however, in some cases, the simulated spectra show shifts of a few eV compared to the experimental results. Using trimming parameters, those deviations can be corrected to a certain extent (*e.g.* shrinking the cluster size by reducing the atomic coordinates and interatomic distances[68]) and adjusted by rigid-shifting the energy axis where needed.

References

[1] S. Calvin, *XAFS for Everyone*, Taylor & Francis Group, Boca Raton, 2013.
[2] http://pd.chem.ucl.ac.uk/pdnn/inst2/work.htm
[3] H. Wende, *Rep. Progr. Phys*., 2004, **67**, 2105.
[4] J. Stohr, *NEXAFS Spectroscopy*, Springer: Heidelberg, 2003.
[5] P. Wernet, D. Nordlund, U. Bergmann, M. Cavalleri, M. Odelius, H.Ogasawara, L. A. Näslund, T. K. Hirsch, L. Ojamäe, P. Glatzel, L. G. M. Pettersson, A. Nilsson, *Science*, 2004, **304**, 995.
[6] M. Cavalleri, H. Ogasawara, L. G. M. Pettersson, A. Nilsson, *Chem. Phys. Lett.*, 2002, **364**, 363.
[7] S. Myneni, Y. Luo, L. Å. Näslund, M. Cavalleri, L. Ojamäe, H. Ogasawara, A. Pelmenschikov, P. Wernet, P. Väterlein, C. Heske, Z. Hussain, L. G. M. Pettersson, A. Nilsson, *J. Phys.: Condens. Matter*, 2002, **14**, L213.
[8] A. Bianconi, *Appl. Surf. Sci.*, 1980, **6**, 392.
[9] M. Newville, *Fundamentals of XAFS*, University of Chicago, Chicago, 2004, http://xafs.org/Tutorials?action=AttachFile&do=get&target=Newville_xas_fundamentals.pdf.
[10] G. G. Hall, *Proc. R. Soc. Lond. A*, 1949, **198**, 1.
[11] P. Atkins, J. de Paula, *Physical Chemistry*, Oxford University Press, 8th edition, Oxford, 2006.
[12] G. Hahner, *Chem. Soc. Rev*., 2006, **35**, 1244.
[13] F. Sette, J. Stöhr, A. P. Hitchcock, *J. Chem. Phys.*, 1984**, 81**, 4906.
[14] J. Stöhr, F. Sette, A. L. Johnson, *Phys. Rev. Lett.*, 1984, **53**, 1684.
[15] J. S. Stevens, A. Gainar, E. Suljoti, J. Xiao, R. Golnak, E. F. Aziz, S. L. M. Schroeder, *Chem. Eur. J.*, 2015, **21**, 7256.
[16] A. P. Hitchcock, C. E. Brion, *J. Phys. B*, 1981**, 14**, 4399.
[17] D. Arvanitis, N. Haack, G. Ceballos, H. Wende, K. Baberschke, A. L. Ankudinov, J. J. Rehr, *J. Electron Spectrosc. Relat. Phenom.*, 2000, **113**, 57.
[18] M. N. Piancastelli, D. W. Lindle, T. A. Ferrett, D. A. Shirley, *J. Chem. Phys.*, 1987, **87**, 3255.
[19] M. N. Piancastelli, D. W. Lindle, T. A. Ferrett, D. A. Shirley, *J. Chem. Phys.*, 1987, **86**, 2765.
[20] B. Kempgens, H. M. Köppe, A. Kivimäki, M. Neeb, K. Maier, U. Hergenhahn, A. M. Bradshaw, *Phys. Rev. Lett*., 1997, **79**, 35.
[21] B. Kempgens, H. M. Köppe, A. Kivimäki, M. Neeb, K. Maier, U. Hergenhahn, A. M. Bradshaw, *Surf. Sci. Lett.*, 1999, **425**, L376.
[22] N. Haack, G. Ceballos, H. Wende, K. Baberschke, D. Arvanitis, A. L. Ankudinov, J. J. Rehr, *Phys. Rev. Lett.*, 2000, **84**, 614.
[23] V. L. Shneerson, D. K. Saldin, W. T. Tysoe, *Surf. Sci.*, 1997, **375**, 340.
[24] Y. Zubavichus, M. Zharnikov, A. Schaporenko, M. Grunze, *J. Electr. Spectr. Rel. Phen*., 2004, **134**, 25.
[25] K. Ihm, T-H. Kang, D. H. Lee, S.-Y. Park, K.-J. Kim, B. Kim, J. H. Yang, C. Y. Park, *Surf. Sc.*, 2006, **600**, 3729.
[26] I. Lindau, W. E. Spicer, *J. Electron Spectrosc.*, 1974, **3**, 409.
[27] C. J. Powell, *Surf. Sci.*, 1974, **44**, 29.
[28] N. Masahito, K. Takuya, K. Retsuo, I. Takeshi, T. Kikuo, M. Takashi, *Jpn. J. Appl. Phys.*, 2012, **51**, 01AB02.
[29] J. G. Chen, B. Frühberger, M. L. Colaianni, *J. Vac. Sci. Technol. A*, 1996, **14**, 1668.

[30] L.-C. Duda, J. Nordgren, G. Dräger, S. Bocharov, T. Kirchner, *J. Electron Spectrosc. Relat. Phenom.*, 2000, **27**, 110.
[31] G. Ghiringhelli, N. B. Brookes, E. Annese, H. Berger, C. Dallera, M. Grioni, L. Perfetti, A. Tagliaferri, L. Braicovich, *Phys. Rev. Lett.*, 2004, **92**, 117406.
[32] L. Braicovich, L. J. P. Ament, V. Bisogni, F. Forte, C. Aruta, G. Balestrino, N. B. Brookes, G. M. De Luca, P. G. Medaglia, F. Miletto Granozio, M. Radovic, M. Salluzzo, J. van den Brink, and G. Ghiringhelli, *Phys. Rev. Lett.*, 2009, **102**, 167401.
[33] T. Tokushima, Y. Harada, H. Ohashi, Y. Senba, S. Shin, *Rev. Sci. Instrum.*, 2006, **77**, 063107.
[34] J. Nordgren, J. H. Guo, *J. Electron Spectrosc. Relat. Phenom.*, 2000, **110**, 1.
[35] G. Ghiringhelli, A. Piazzalunga, C. Dallera, G. Trezzi, L. Braicovich, T. Schmitt, V.N. Strocov, R. Betemps, L. Patthey, X. Wang and M. Grioni, *Rev. Sci. Instrum.*, 2006, **77**, 113108.
[36] O. Fuchs, L. Weinhardt, M. Blum, M. Weigand, E. Umbach, M. Bär, C. Heske, J. Denlinger, Y.-D. Chuang, W. McKinney, Z. Hussain, E. Gullikson, M. Jones, P. Batson, B. Nelles, R. Follath, *Rev. Sci. Instrum.*, 2009, **80**, 063103.
[37] A. Kotani, S. Shin, *Rev. Mod. Phys.*, 2001, **73**, 203.
[38] L. J. P. Ament, M. van Veenendaal, T. P. Devereaux, J. P. Hill, J. van den Brink, *Rev. Mod. Phys.*, 2011, **83**, 705.
[39] J. van den Brink, M. van Veenendaal, *J. Phys. Chem. Solids*, 2005, **66**, 2145.
[40] M. van Veenendaal, *Phys. Rev. Lett.*, 2006, **96**, 117404.
[41] J.-H. Guo, Y. Luo, A. Augustsson, J.-E. Rubensson, C. Såthe, H. Ågren, H. Siegbahn, J. Nordgren, *Phys. Rev. Lett.*, 2002, **89**, 137402.
[42] T. Tokushima, Y. Harada, O. Takahashi, Y. Senba, H. Ohashi, L. G. M. Pettersson, A. Nilsson, S. Shin, *Chem. Phys. Lett.*, 2008, **460**, 387.
[43] T. Tokushima, Y. Harada, Y. Horikawa, O. Takahashi, Y. Senba, H. Ohashi, L. G. M. Pettersson, A. Nilsson, S. Shin, *J. Electron Spectrosc. Relat. Phenom.*, 2010, **177**, 192.
[44] J. Forsberg, L. C. Duda, A. Olsson, T. Schmitt, J. Andersson, J. Nordgren, J. Hedberg, C. Leygraf, T. Aastrup, D. Wallinder, J. H. Guo, *Rev. Sci. Instrum.*, 2007, **78**, 083110.
[45] M. Faubel, B.Steiner, J. P. Toennies, *J. Chem. Phys.*, 1997, **106**, 9013.
[46] B. Winter, M. Faubel, *Chem. Rev.*, 2006, **106**, 1176.
[47] K. Atak, S. I. Bokarev, M. Gotz, R. Golnak, K. M. Lange, N. Engel, M. Dantz, E. Suljoti, Oliver Kühn, E. F. Aziz. *J. Phys. Chem. B*, 2013, **117**, 12613.
[48] R. Golnak, K. Atak, E. Suljoti, K. F. Hodeck, K. M. Lange, M. A. Soldatov, N. Engel, E. F. Aziz, *Phys. Chem. Chem. Phys.*, 2013, **15**, 8046.
[49] K. Atak, N. Engel, K. M. Lange, R. Golnak, M. Gotz, M. Soldatov, J.-E. Rubensson, N. Kosugi, E. F. Aziz, *Chem. Phys. Chem.*, 2012, **13**, 3106.
[50] K. M. Lange, R. Könnecke, S. Ghadimi, R. Golnak, M. A. Soldatov, K. F. Hodeck, A. Soldatov, E. F. Aziz, *Chem. Phys.*, 2010, **377**, 1.
[51] C. Suryanarayana, M. G. Norton, *X-ray Diffraction: A Practical Approach*, Springer Science+Business Media, New York, 1998.
[52] C. Hammond, *The Basics of Chrystallography and Diffraction*, 4th edition, Oxford University Press, Oxford, 2015.
[53] R. G. Parr, W. Yang, *Density-Functional Theory of Atoms and Molecules*, OUP, Oxford, 1989.
[54] B. Li, *Density-Functional Theory and Quantum Chemistry Studies on "dry" and "wet" NaCl(001)*, PhD thesis, Technical University of Berlin, Berlin, 2009.
[55] R. G. Parr, *Ann. Rev. Phys. Chem.*, 1983, **34**, 631.
[56] T. Ziegler, *Chem. Rev.*, 1991, **91**, 651.

[57] V. L. Moruzzi, J. F. Janak, A. R. Williams, *Calculated Electronic Properties of Metals*, Pergamon, New York, 1978.
[58] A. Zangwill, P. Soven, *Phys. Rev. A*, 1980, **21**, 1561.
[59] M. E. Casida, *J. Mol. Struct.*, 2009, **914**, 3.
[60] M. Cossi, V. Barone, *J. Chem. Phys.*, 2000, **112**, 2427.
[61] M. A. L. Marques, N. T. Maitra, F. M. S. Nogueira, E. K. U. Gross, A. Rubio, *Fundamentals of Time-Dependent Density Functional Theory*, Springer, Berlin, 2012.
[62] A. Dreuw, J. L. Weisman, M. Head-Gordon, *J. Chem. Phys.*, 2003, **119**, 2943.
[63] N. T. Maitra, F. Zhang, F. J. Cave, K. Burke, *J. Chem. Phys.*, 2004, **120**, 5932.
[64] M. Dierksen, S. Grimme, *J. Chem. Phys.*, 2004, **120**, 8425.
[65] S. Miertus, E. Scrocco, J. Tomasi, *Chem. Phys.*, 1981, **55**, 117.
[66] V. Barone, M. Cossi, J. Tomassi, *J. Chem. Phys.*, 1997, **107**, 3210.
[67] D. Cabaret, P. Sainctavit, P. Ildefonse, A.-M. Flank, *Am. Mineral.*, 1998, **83**, 300.
[68] E. Bosman, J. Thieme, *J. Phys. Conf. Ser.*, 2009, **186**, 012004.

CHAPTER 4

SOLID-STATE NEXAFS OF SACCHARIDES

4.1. Introduction

Saccharides, also known as carbohydrates and sugars, are an important class of natural products and are well known as foods and components of ingestible formulated products.[1-6] They have a propensity to present various non-crystalline and crystalline forms, and interest in intra- and intermolecular bonding in their solid state structures goes back to the earliest days of organic chemistry.[7] Saccharides can exist as monomers or, covalently bonded, as dimers, oligomers or polymers, in which molecular monomers are linked by glycosidic bonds. Well known monosaccharides are glucose (dextrose, also known as '*blood sugar*'), fructose (or '*fruit sugar*'), and ribose, an essential component of ribonucleic acid (RNA). Among disaccharides, well known representatives are sucrose ('*table/beet/cane sugar*', or simply '*sugar*') and lactose ('*milk sugar*'), which is commonly used in pharmaceutical products as an excipient. The most abundant polysaccharide occurring in nature is cellulose,[8] which confers support, resistance and rigidity to plants. In terms of classification by the number of carbons, saccharides with five carbon atoms in their molecules are called pentoses (*e.g*. ribose, xylose, lyxose), while six-membered ring saccharides are called hexoses (*e.g*. glucose, fructose, galactose).

Identifying mono-, di-, oligo- and polysaccharides spectroscopically is often important for understanding the properties of formulated products.[9] X-ray core level spectroscopies are structurally incisive methods that are increasingly used for the analysis of interfacial species in complex products, devices and biomaterials.[9-14] The most commonly used variant, X-ray photoelectron spectroscopy (XPS), provides information on the chemical state of atoms through sensitivity to the local electron density variations caused by changes in chemical bonding, which lead to chemical shifts of core level binding energies, and interpretation can be complemented or enhanced with the results of density functional theory (DFT) calculations.[15] Careful C 1s core level binding energy measurements by

XPS of fructose, xylose, glucose, galactose, maltose, α-lactose, β-lactose, and cellulose have previously revealed that XPS is sensitive enough to distinguish mono-, di- and polysaccharides.[16]

Near-edge X-ray Absorption Fine Structure (NEXAFS, also called X-ray Absorption Near-edge Structure, XANES)[17] provides more incisive information than XPS about local chemical and electronic structure in condensed matter. Determination of bond lengths, static and dynamic disorder as well as coordination numbers and coordination geometry around the X-ray absorbing atoms in organic compounds can sometimes also be achieved by extended X-ray absorption fine-structure analysis (EXAFS).[17,18] However, C K EXAFS analysis is often not possible because of reciprocal space limitations caused by (i) spectral overlap of absorption edges from other elements (for the present case of saccharides the O K-edge ~ 250 eV above the C K-edge) and (ii) instrumental limitations such as source transmission variations due to contamination on optics, abrupt variations in photon flux or non-linearities in I_0 monitoring. However, NEXAFS offers the possibility to obtain information about bond lengths in organic molecules through determination of σ* shape resonance energies relative to the ionisation potential of the core level from which excitation takes place. σ* shape resonances arise because photoexcited core electrons can be temporarily trapped by a centrifugal potential barrier near the absorption edge of the molecule, so electrons traverse the molecule multiple times before escaping from the molecular core by tunnelling.[17,19-23] The use of shape resonance energy shifts for determining bond lengths has been referred to as the 'bond length with a ruler' method.[17,20-22,24] The method relies on the determination of the core level binding energies (*i.e.*, variations in the core level ionisation potentials, IPs) from the absorption spectra and the energetic position of the σ* shape resonances relative to the IP.[25] We have recently reported evidence that this method can also be applied to C–N bonds of organic molecules in solution and in the organic solid state.[22] The likely reason for this transferability of a method originally developed for gas phase molecules is that intermolecular interactions in the organic condensed phase are much weaker than the internal covalent bonds in the molecules. While the exact physical origins of σ* shape resonances, and thus the validity of their use for bond-length determination, has been controversially discussed in the past,[25-30] there is considerable empirical evidence, as well as theoretical support[30,31] for their interpretation as multiple scattering resonances dependent on the bond length.[23]

Full deconvolution and interpretation of all features in NEXAFS spectra can be a complex task, but our previous success in detecting differences between saccharides by core level binding energy analysis in XPS[16] led us to hypothesise that the structurally more incisive NEXAFS spectroscopy may also be sensitive to differences in structure and bonding in saccharide systems. Note that unlike diffraction methods, core level spectroscopies do not require long-range order and can therefore also be applied to non-crystalline forms. A deeper understanding of the core level spectra of the crystalline forms would therefore establish a basis for the characterisation of local interactions in materials without long-range order, such as melts, solutions, coatings or nanocrystalline forms. Literature references describing applications of NEXAFS to saccharides have been rather sporadic[32-34] and generally did not relate the spectra to specific structural details.

In this chapter C K NEXAFS data of a series of monosaccharides (fructose, xylose, glucose, galactose), disaccharides (maltose monohydrate, α-lactose monohydrate, anhydrous β-lactose) and one polysaccharide (cellulose), will be discussed with the aim to provide a database of spectral fingerprints and establish what level of structural detail is evident in these spectra. In particular, it will be shown by reference to the crystal structures that the σ* shape resonance analysis of the NEXAFS provides access to rather small variations in intramolecular C–OH bond lengths in the molecules.

4.2. Experimental section

4.2.1. Materials

Polycrystalline powder samples were obtained from Sigma-Aldrich, UK. Fructose, xylose, galactose, maltose monohydrate and cellulose has ≥ 99% purity. Glucose was 99.5% pure. α-lactose monohydrate and β-lactose had been provided by Sanofi-Aventis (Alnwick, UK) and were the same samples examined in our previous XPS study.[14] All samples were used as provided. Thick layers of powder were spread onto double-sided copper tape (VWR, US) with a clean thin spatula, ensuring complete covering of the tape surface and then pressed on mechanically to obtain a compact layer. Figure 4.1 summarises the chemical structures of the analysed saccharides. The pentose (xylose) in the first row has the chemical formula $C_5H_{10}O_5$, the hexoses (fructose, glucose, galactose)

in the second row have the general chemical formula $C_6H_{12}O_6$, the disaccharides (maltose, α-lactose, β-lactose) in the third row have the general formula $C_{12}H_{22}O_{11}$. The polysaccharide (cellulose) in the fourth row is a polymer of β-glucose, having the formula $[C_6H_{10}O_5]_n$.

Generally, monosaccharides crystallise in the pyranose form, *i.e.*, as 6-membered rings; fructose can crystallise in either a pyranose or furanose (5-membered ring) form; the pyranose form of fructose is usually obtained by industrial crystallisation processes, whereas the furanose form occurs in natural products. Xylose is naturally found in pyranose form. For the disaccharides, maltose, α-lactose, and β-lactose are formed of two monosaccharide units in pyranose form. The polysaccharide, cellulose, is formed of linked glucose units in the pyranose form.

4.2.2. X-ray powder diffraction (XRPD)

A Rigaku Miniflex XRPD instrument was used to determine the polymorphic form of each saccharide powder, taking diffraction patterns over a 2θ range from 5° to 40°, at a rate of 1°/min with a 0.02° step. The X-ray source operated with 30 kV voltage and 15 mA current. The experimental patterns were compared with calculated crystal structure patterns based on published crystal structures and are included in the Appendix B of this thesis (Figures B1-B8). Results indicated that the forms were β-fructose, α-xylose, α-glucose, α-galactose, β-maltose, α-lactose monohydrate, anhydrous β-lactose. The positions of the broad features in the pattern of amorphous cellulose (Figure B8 from the Appendix B) indicated that residual local order was determined by nuclei of the I_α polymorph of cellulose.

4.2.3. NEXAFS

NEXAFS spectra were measured at the National Institute of Standards and Technology U7A beamline of the National Synchrotron Light Source (NSLS) at Brookhaven National Laboratory (BNL) in New York, USA. Partial electron yield (PEY) C K-edge spectra were obtained using a channeltron electron multiplier with the sample positioned at the magic angle (54.7°) relative to the incident beam, with a photon flux of approximately

5×10^{10} photons/s. An entrance grid bias of –150 V was used for the C K-edge in the case of galactose and –50 V for the rest of saccharides.

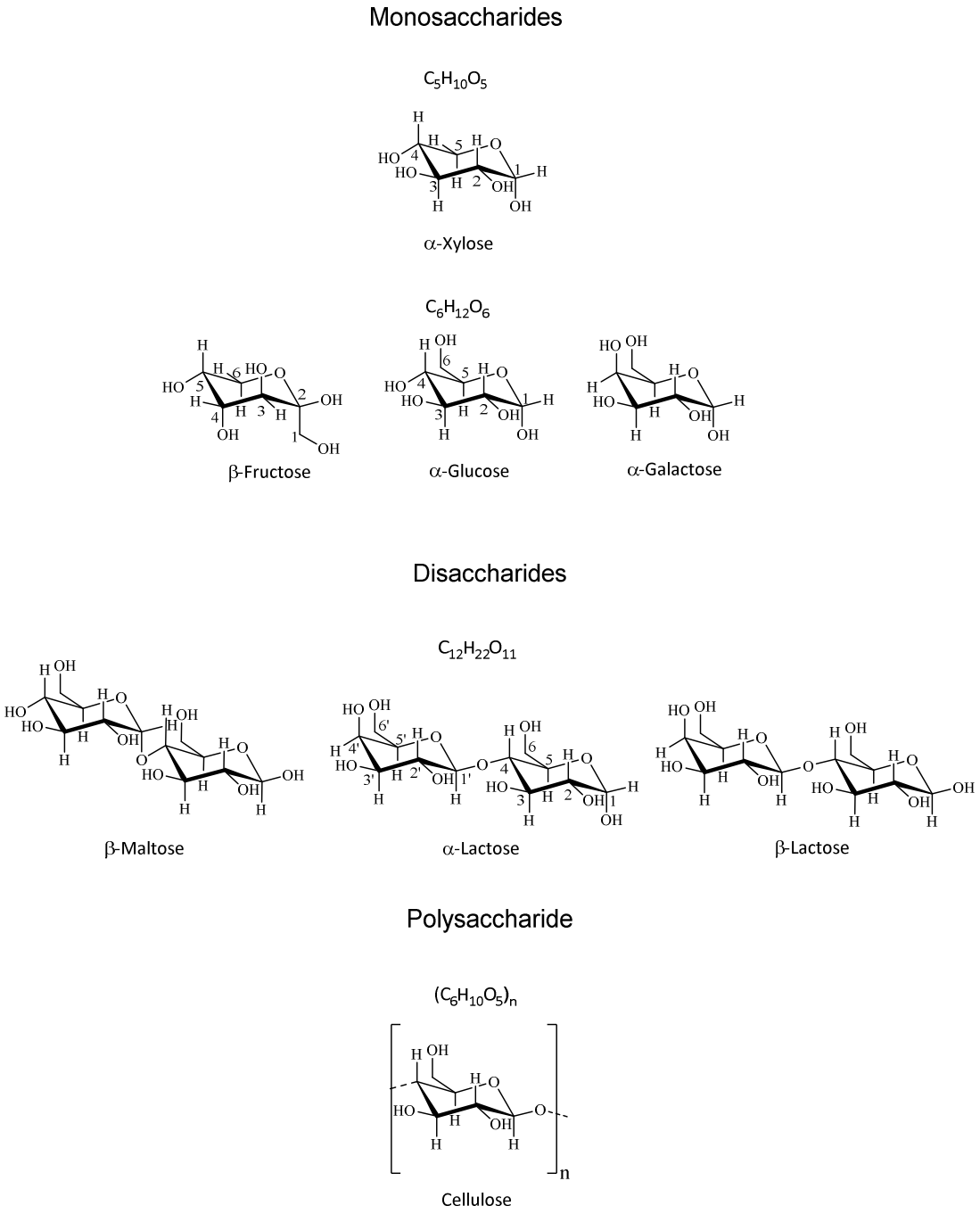


Figure 4.1. The saccharides analysed by NEXAFS. The IUPAC numbering of the carbon atoms is included for xylose, fructose, glucose and α-lactose. The other mono- and disaccharides and the polysaccharide adopt analogous numbering.

At –50 V grid bias the detected electron-yield signal is essentially the total Auger yield, *i.e.*, the probing depth of the electron-yield signal is determined by the escape depth of all C KLL Auger electrons, including the whole background of inelastically scattered electrons on the low energy side of the Auger emission lines,[35-38] which probes deeper than the elastic Auger or photoemission peaks. A previous experimental determination of the C K-edge probing depth in polymers yielded a value of ~ 2 nm,[39] while experiments with signal attenuation in Cr metal indicate for photons at the C K-edge that the maximum probing depth is approximately 3.5 nm.[40] These values are equivalent to probing a few molecular layers deep into the saccharide materials and should hence confer bulk sensitivity. In line with this, our previous XPS analysis of the same saccharide materials through C 1s emission with an Al K_α source, which is associated with signal attenuation lengths only slightly higher (~ 3 nm),[41] yielded chemical compositions consistent with those expected from the saccharide stoichiometries. Moreover, for molecular materials surface effects on internal bond lengths should be negligible, as intermolecular interactions in organic molecular solids are weak relative to the strength of the internal bonds in the molecules. The strongest intermolecular interaction in the saccharide bulk is hydrogen bonding. It is therefore expected that the C–O bond lengths derived from the bulk crystal structures used for the interpretation of our data are good models.

A toroidal spherical grating monochromator with 600 lines mm^{-1} grating was used to acquire the C K-edge data, yielding an energy resolution of approximately 0.1 eV. After collection, the spectra were normalised by the simultaneously recorded drain current from an *in situ* gold-coated, 90% transmission grid (I_0) which is situated in the X-ray beam path, to eliminate the incident beam intensity fluctuations and absorption features in the beamline optics. The monochromator energy scale calibration was performed through the 285.1 eV π^* graphite absorption by a carbon mesh located further upstream of I_0 (I_0_up). All spectra were processed using standard pre- and post-edge normalisation methods. The peak-fitting capability of Athena software[42,43] was employed to determine the energetic positions of all NEXAFS transitions at the C K-edge, and the spectrum was modelled with Gaussians for the peaks and two *arctan* step functions for the edge steps arising at the IPs.[17] The positions of the IPs were determined assuming two constraints: the expected intensity ratio for OC–OH and C–OH moieties, and their energy difference, which was known from our previous XPS study to be 1.35 eV higher for the OC–OH hemiacetal/hemiketal groups.[16] The first IP was set as the experimental edge step

inflection point, *i.e.*, the zero-crossing of the second derivative of the experimental spectrum. For investigating the 'bond length with a ruler' correlation, the standard deviation for the C–OH bond lengths in the reported crystal structures was ≤ 0.007 Å.[44-49] For the curve-fitted C K-edge spectra, by fitting with different models (variation of peak positions, functions and the number of included peak functions) and using different initial values for the peak parameters it was found that the error in energy positions for comparably good final fits was ≤ 0.02 eV.

4.2.4. Sample contamination and X-ray degradation

Slightly varying levels of adventitious surface contamination were previously observed in the XPS data for the same set of saccharides.[16] In the NEXAFS spectra this contamination is evident as weak pre-edge peaks in the C K-edge spectra. Similar to the constant peak position of the hydrocarbon contaminant in XPS C 1s results,[16] these pre-edge peaks were in a narrow energy range, from 285.3 to 285.4 eV. We carefully checked by examination of the signal from the incident beam monitor (a gold coated mesh) whether varying levels of surface contamination on it might have led to normalisation errors that could account for these pre-edge features. However, no variations in the mesh signal were evident that could cause the observed signals.

In the previous XPS analysis of saccharides, it was also noted that the intensity of the peaks from the C–OH and O–C–O hemiacetal/hemiketal moieties in the C 1s spectra decreased as a function of X-ray exposure[16] due to radiative damage, with the strongest effect on ribose and xylose.[16] In the NEXAFS measurements, no significant changes in peak positions and intensities, nor formation of new peaks were observed as a function of exposure to the X-ray beam, even after six scans of the C K-edge.

4.2.5. FEFF8 full-multiple scattering calculations

FEFF8.2[50] was used to simulate the σ* shape resonances in the C K NEXAFS for a cluster cut from the crystal structure of glucose (CCDC reference: GLUCSA03).[51] The XANES, FMS and SCF cards were used – FMS for full-multiple scattering X-ray absorption near-edge structure (XANES) calculation; SCF to enable self-consistent field iterations. To identify the shape resonance contributions due to scattering near the potential of the

C–OH single bonds we also ran calculations were run in which the C–OH bond length in the hemiacetal/hemiketal position was increased and decreased by 0.1 Å relative to the equilibrium value in the crystal structure (Figure 4.2 and Figures B9-B14 from Appendix B).

4.3. Results and Discussion

The C K NEXAFS spectra of the saccharides are presented in Figures 4.2 and 4.3 together with the results of the fitting analysis involving Gaussian and edge step functions. For a deeper analysis of these spectra we need to recall that saccharide molecules present only single C–C, C–H and C–O bonds (Figure 4.1), so it is expected that the NEXAFS spectra are dominated by transitions to unoccupied, anti-bonding σ* states. The most prominent features are σ* shape resonances, which arise from multiple scattering of the photoelectron wave along the internuclear axis between the absorbing atom and its neighbour; thus the σ* shape resonance energies depend sensitively on the associated intramolecular bond lengths.

In a previous study of the C K-edge spectra of a dimannoside derivative adsorbed on Au, the prominent σ* resonance at ~ 289 eV was attributed to the C–OH moieties of the saccharide, and the feature at ~ 291 eV to the OC–OH moiety.[34] In line with this assignment, the intensity of peak 1 is reduced for xylose (Figure 4.2) as there is one less contributing carbon (*i.e.* one less C–OH bond in the structure) so the relative intensities of the peaks arising from the two moieties are different compared to the other monosaccharides. The small pre-edge peak at ~ 285 eV in the spectra can be attributed to adventitious contamination with unsaturated hydrocarbons containing C=C bonds,[34] while the second pre-edge peak at ~ 288 eV, observed as a shoulder, is associated with C–H transitions.[33] The broader features at higher photon energies around 294 eV may arise from C–C bonds.

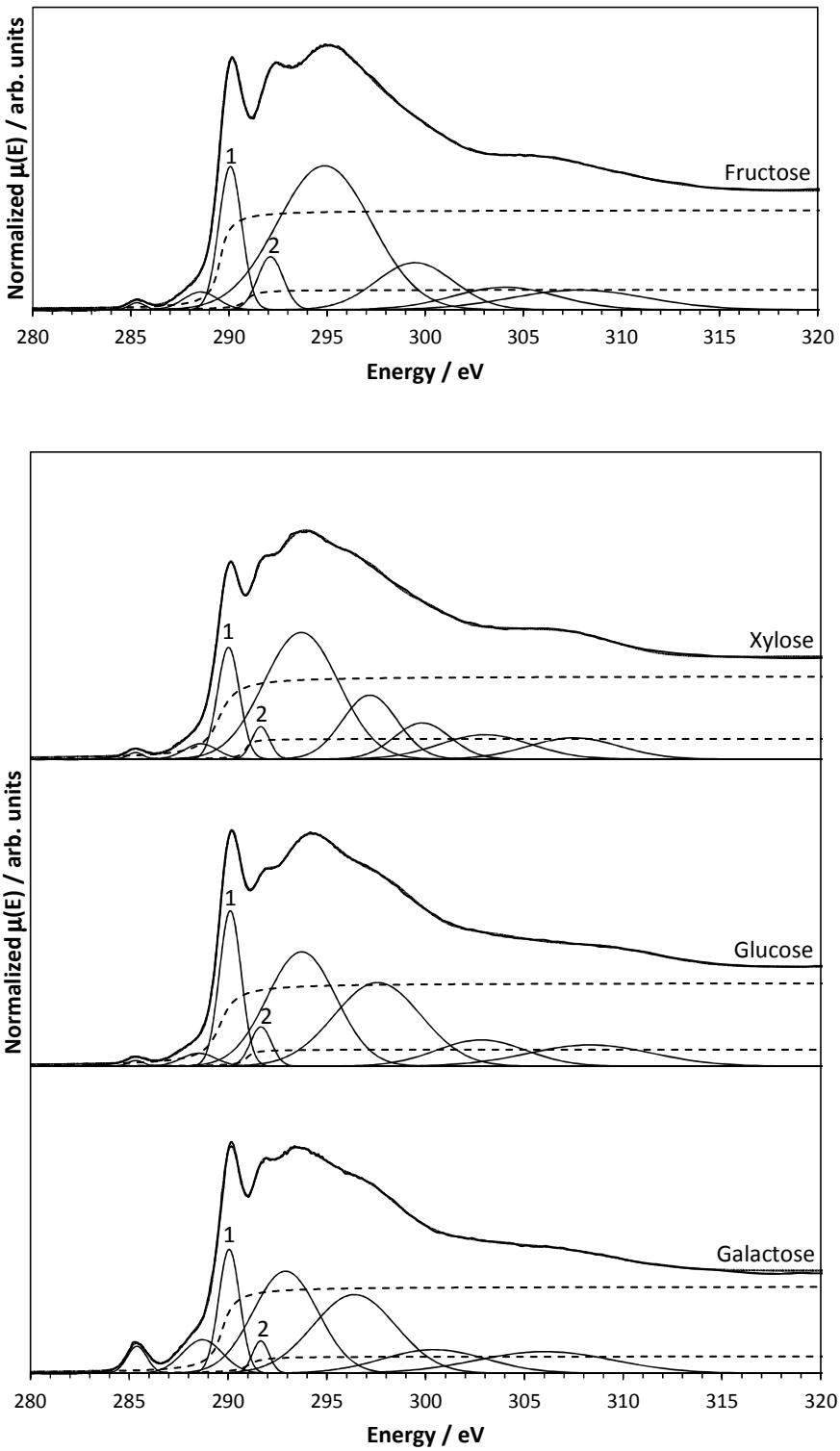


Figure 4.2. Experimental C K NEXAFS spectrum of fructose (top plot) and the C K NEXAFS spectra (bottom plot) for the monosaccharides xylose (top), glucose (middle), and galactose (bottom). Each spectrum is accompanied by best fits obtained to determine peak positions, with the $\sigma^*_{\underline{C}\text{–OH}}$ and $\sigma^*_{O\underline{C}\text{–OH}}$ peaks marked with numbers 1 and 2, respectively. The IP edge steps are marked with dashed lines.

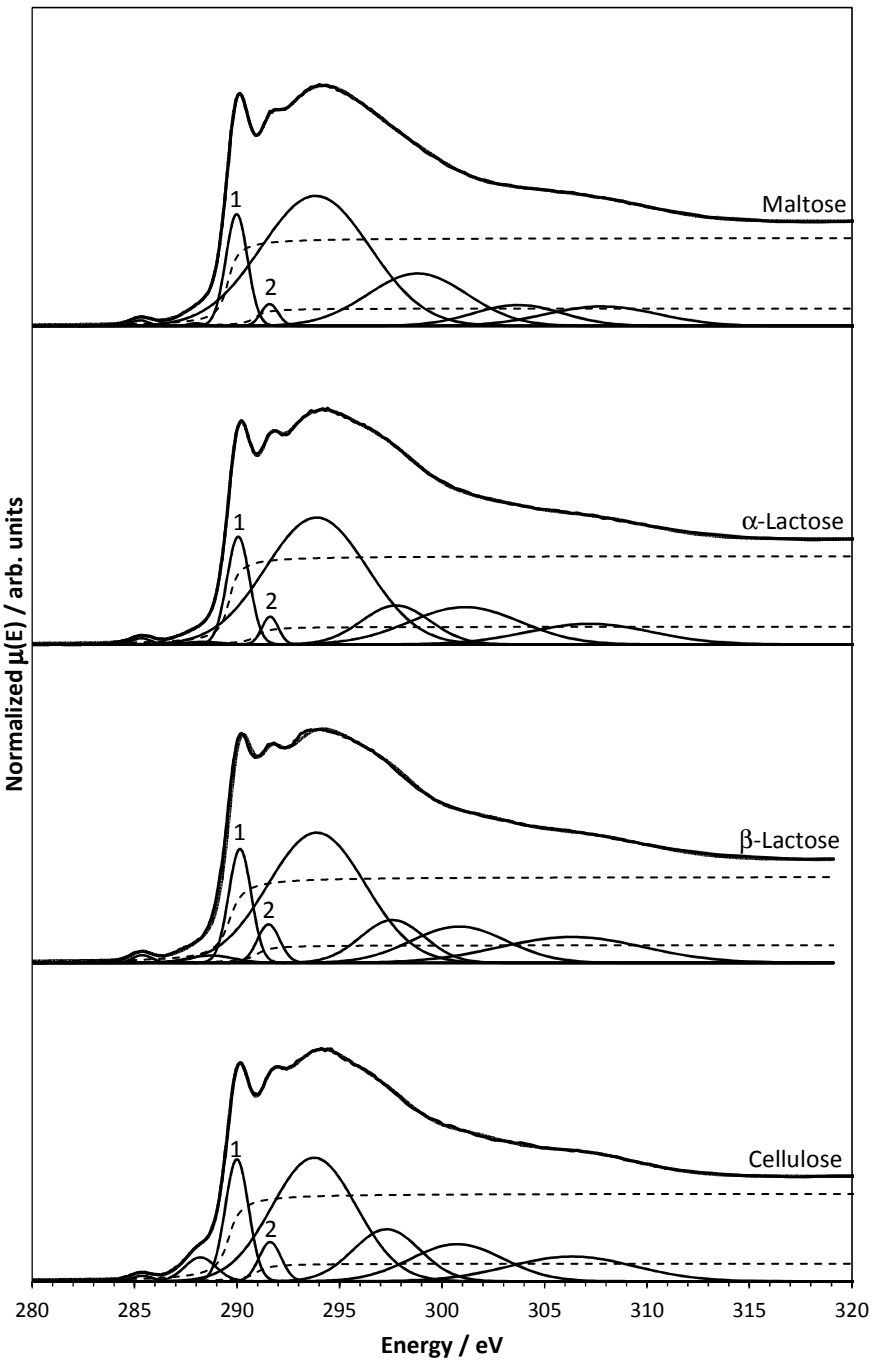


Figure 4.3. Experimental C K NEXAFS spectra for the di- and polysaccharides studied: maltose (top), α-lactose (second from top), β-lactose (third from top), and cellulose (bottom), alongside fits with the $\sigma^*_{\underline{C}\text{–OH}}$ and $\sigma^*_{O\underline{C}\text{–OH}}$ peaks marked with numbers 1 and 2, respectively. The IP edge steps are marked with dashed lines.

To examine how the σ* transitions associated with C–OH bonds vary in energetic position as a function of bond length, FEFF8 calculations were performed on each inequivalent C

atom in the crystal structure of glucose, first with the equilibrium C–OH bond length in the crystal structure and then for the same bond shortened and elongated by 0.1 Å. One set of resulting spectra is presented in Figure 4.4 and the others in Figures B9-B13 from the Appendix B. The strong shifts of the shape resonances originally in the region around 290–292 eV are evident, with a shorter bond length resulting in a blue shift (higher photon energy) by approximately +4.7 eV whereas the longer bond length results in a red shift (lower photon energy) by about –4.2 eV.

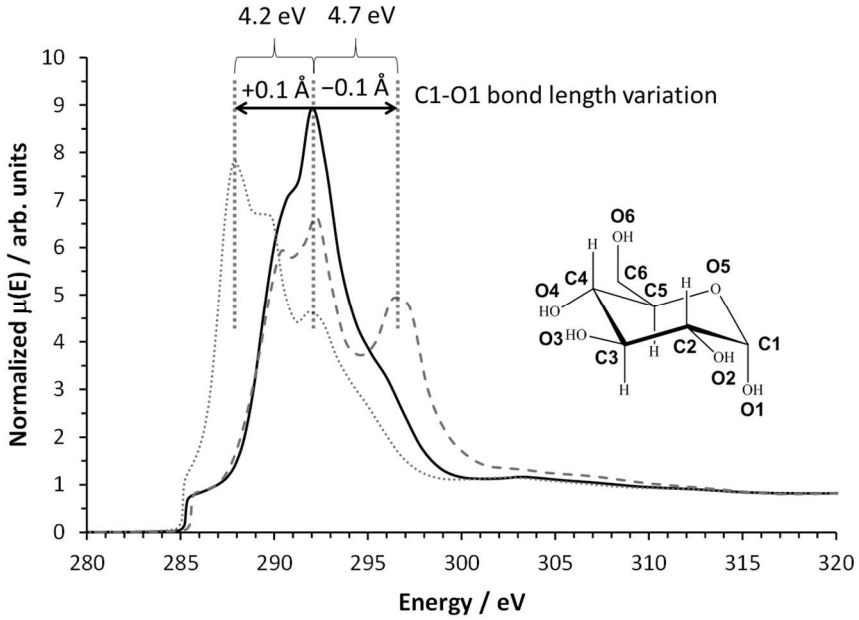


Figure 4.4. FEFF8 simulations for C1 absorbing atom in: an unaltered crystal structure of glucose (black line), an altered crystal structure of glucose with C1–O1 bond shortened by 0.1 Å (dashed grey line), and an altered crystal structure of glucose with C1–O1 bond elongated by 0.1 Å (dotted grey line).

As mentioned above, during peak fitting of the spectra in Figures 4.2 and 4.3, a 1.35 eV difference in energy between the IPs arising from C–OH and OC–OH moieties was assumed, as this is (with a margin of about ±0.05 eV) the value of the C 1s binding energy separations observed by XPS.[16] The intensity ratios between the two edge step functions representing the IPs were fixed to be in line with the stoichiometries of the saccharides.[16] Note that in the extended X-ray absorption fine structure (EXAFS) region at higher photon energies, where the spectra are determined by scattering of the photoelectron wave

by neighbouring atoms, the fitted Gaussian functions have no physical meaning, but modelling of the post-edge background was required to ensure accurate fitting results in the spectral regions of the electronic transitions of interest, *i.e.* around and below the IPs. The parameters derived with the fitting results (shown in Figures 4.2 and 4.3) are summarised in Table 4.1. It can be seen that the $\sigma^*_{\underline{C}\text{–OH}}$ transitions are in a narrow range of energies, from 290.0 eV to 290.1 eV. The $\sigma^*_{O\underline{C}\text{–OH}}$ transitions are found between 291.5 eV and 292.1 eV.

Table 4.1. IPs and centroid energies of the σ* shape resonances determined by curve fitting; the term values δ are calculated as the difference between each σ* centroid energy and its corresponding IP.

Saccharide	$IP_{\underline{C}\text{–OH}}$ / eV	$\sigma^*_{\underline{C}\text{–OH}}$ / eV	$\delta_{\underline{C}\text{–O}}$ ($\sigma^*_{\underline{C}\text{–OH}} - IP_{\underline{C}\text{–OH}}$) / eV	$IP_{O\underline{C}\text{–OH}}$ / eV	$\sigma^*_{O\underline{C}\text{–OH}}$ / eV	$\delta_{O\text{–}\underline{C}\text{–O}}$ ($\sigma^*_{O\underline{C}\text{–OH}} - IP_{O\underline{C}\text{–OH}}$) / eV
β-Fructose	289.55	290.07	0.52	290.90	292.11	1.21
α-Xylose	289.55	290.02	0.47	290.90	291.65	0.75
α-Glucose	289.60	290.11	0.51	290.95	291.66	0.71
α-Galactose	289.70	290.06	0.36	291.05	291.65	0.60
β-Maltose monohydrate	289.55	289.98	0.43	290.90	291.58	0.68
α-Lactose monohydrate	289.62	290.06	0.44	290.97	291.61	0.65
β-Lactose	289.63	290.14	0.51	290.98	291.53	0.55
Cellulose	289.60	289.99	0.39	290.95	291.61	0.66

We can now turn to examining how the crystallographically determined C–OH bond lengths correlate with the shifts in the σ* shape resonances and IPs. This 'bond-length-with-a-ruler' analysis[17,20-22] consists of plotting the term value δ, which is the energy difference between the σ* shape resonance energy and its corresponding IP (Table 4.1), as a function of the length of the bond from which it arises. Because of the significant IP difference between the C–OH and OC–OH moieties in the saccharides, their shape resonances are distinct and their bond lengths can be determined separately. There are of course several C–OH moieties with slightly different bond lengths in each saccharide, so their average bond length was used for the analysis (Table 4.2), as available from their published crystal structures.[44-49,52,53] The OC–OH hemiacetal or hemiketal groups are associated with a shorter C–OH bond length (Table 4.2) than for the other C–OH

groups,[44-49,52,53] and averaging of two values has to be used in the cases of the disaccharides and for cellulose, which is composed of repetitive interconnected motifs of two distinct cyclic glucose units bound by a glycosidic bond (*i.e.* cellulose can also be regarded as a polymer of a disaccharide). For the monosaccharides no averaging was required as there are no inequivalent OC–OH groups in the crystal structures. It is also pointed out that contributions from the rather short endocyclic C–O bonds of the hemiacetal and hemiketal groups were omitted from bond length averaging as they are located in the energy region beyond 293 eV, where their contribution cannot be distinguished against the strong background arising from the C–C bonds. The other endocyclic C–O bond is always similar in length to the C–OH bonds and was therefore included in their average.

Table 4.2. C–OH bond lengths determined by crystal structure analysis, where (1) represents the C–OH bond length from the hemiacetal or hemiketal group (as in Figure 4.1), and (av) represents the average values of the other the C–OH bond lengths (including one endocyclic C–O bond – see main text).

Saccharide	CSD identifier	C–OH bond length at OC–OH (1) / Å	average length of remaining C–OH (av)/ Å
β-Fructose	FRUCTO11[44]	1.412	1.423
α-Xylose	XYLOSE[47]	1.393	1.424
α-Glucose	GLUCSA03[45]	1.391	1.428
α-Galactose	ADGALA01[46]	1.409	1.437
β-Maltose monohydrate	MALTOS11[48]	1.402	1.418
α-Lactose monohydrate	LACTOS03[49]	1.390	1.425
β-Lactose	BLACTO02[53]	1.381	1.423
Cellulose	PADTUL[52]	1.403	1.427

The resulting plot of the δ values as a function of the bond lengths is shown in Figure 4.5. It is evident that apart from two outliers (the OC–OH groups in β-lactose and fructose; see discussion below), there is a good linear correlation despite the small range (~ 0.05 Å) of bond length variations covered (Figure 4.5). More precisely, for a change in the C–OH bond length by 0.047 Å, a term value shift of 0.393 eV occurs (which is equivalent to a shift of ~ 0.8 eV for a 0.1 Å variation). Assuming that the term values can

be determined with an accuracy of approximately 0.05 eV, this suggests a sensitivity of the method to bond length variations slightly better than ~ 0.01 Å, making the precision of the technique comparable to accuracies achievable by standard X-ray crystal structure determination. The trend observed in these experimentally derived figures is in good agreement with the C–OH shape resonance shifts predicted by the FEFF8 calculations (*vide supra*), which indicated shifts of approximately 4.5 eV per 0.1 Å variation in bond length. However, the smaller variation of term values in the experimental data may originate from the strain in the cyclic saccharide molecules that impacts on the bond angles and bond-bond interactions between adjacent bonds.[54]

Turning to the outliers in the plot in Figure 4.5, for fructose the C–OH bond at the OC–OH group is considerably longer (1.412 Å) by comparison with the other saccharides (range 1.389-1.409 Å). The associated term value is also higher than for the other saccharides. Both deviations are likely due to the fact that the OC–OH group in fructose represents a hemiketal rather than a hemiacetal moiety. As a result, there is an additional C–C (rather than C–H) bond at the OC–OH moiety.[55] Interestingly, this result suggests that the NEXAFS analysis of C–OH σ* shape resonances and IPs provides an avenue to distinguishing hemiketals from hemiacetals (*i.e.* ketoses from aldoses). However, more systematic work should be carried out to confirm whether this observation can be generalised.

β-lactose, the other outlier in Figure 4.5, has a significantly shorter C–OH bond at the OC–OH group than the other saccharides. It is distinguished from the other hemiacetals by a different configuration at the anomeric position, in which the OH group is positioned *cis* (rather than *trans*) to the $–CH_2OH$ group in the molecule. Moreover, unlike the other two disaccharides (β-maltose and α-lactose), the two monosaccharide units in the β-lactose structure are rotated relative to each other around the glycosidic bond. These local structure differences may perhaps combine to cause weak electronic structure variations and hence, *e.g.*, slight differences in core hole relaxation or local screening differences that may affect the term value δ. A deeper analysis of other saccharide structures would be required to confirm or refute these points.

In any case, the outliers appear to underline an important constraint of the bond-length-with-a-ruler methodology, namely that variations in the term value δ can only be compared for bonds in which the two atoms are in a similar chemical state, so additional

dependencies on functional group (as in, aldehyde vs. ketone) identity and (in condensed matter) secondary interactions with other functional groups due to conformational differences, configurational isomerism and/or functional group interactions in the crystal structure may influence the results.

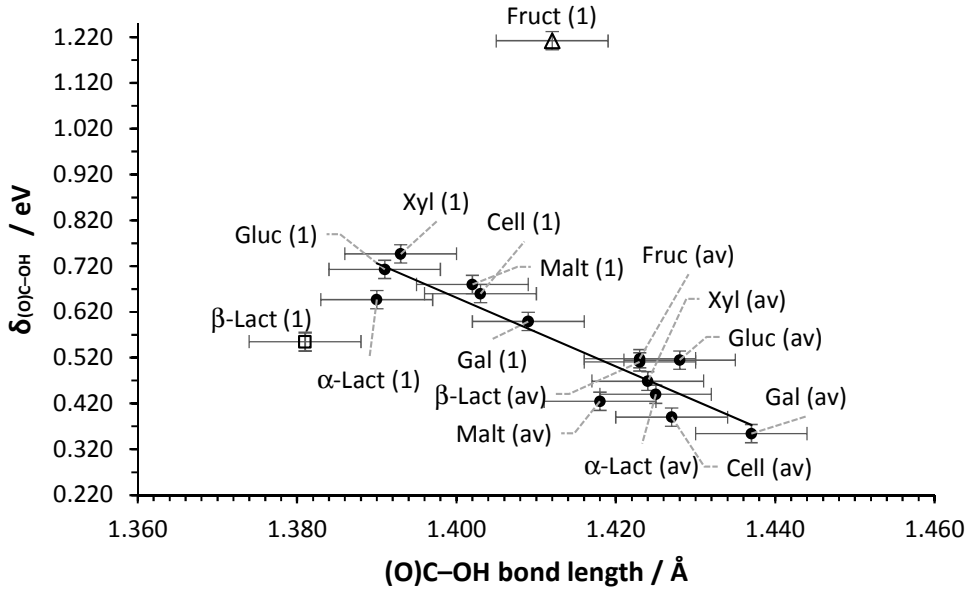


Figure 4.5. Linear correlation between term value δ and (O)C–OH bond length for the crystal structures of saccharides, where Fruct = fructose, Xyl = xylose, Gluc = glucose, Gal = galactose, Malt = maltose, α-Lact = α-lactose, β-Lact = β-lactose, Cell = cellulose; annotation (1) refers to the OC–OH bond length at the hemiacetal/hemiketal position (as in Figure 4.1), and (av) refers to the average value of the remaining C–OH bond lengths; fructose and β-lactose, outliers away from the fitted line (see discussion in the text), are represented with an empty triangle and an empty square, respectively.

The possibility to additionally determine term values in the O K NEXAFS spectra was investigated, but the determination and interpretation of these values is more complex than for the C K-edge due to the additional influence of intermolecular hydrogen bonding, so that no firm conclusions could be drawn. Overall however, the C K-edge results that has been obtained indicate that for hemiacetals with α configuration at the anomeric centre the small relative energy variations between IPs and shape resonance features in the C K NEXAFS are, within experimental error, consistent with the variations of the C–O bond lengths known from single crystal structure analysis. Our data lead to an

estimate of a sensitivity to bond length variations on the order of ~ 0.01 Å, which is in very good agreement with the previously reported sensitivity for gas phase molecules.[21] These results suggest that careful use of solid state NEXAFS can provide very incisive structure information for molecular materials, and should perhaps be considered more for the characterisation of local bonding in non-crystalline systems, not only dissolved species as in a recent study,[22] but also colloidal systems as well as melts and similar amorphous phases.

4.4. Conclusions

Analysis of the C K NEXAFS spectra of solid monosaccharides, disaccharides and a polysaccharide indicates that the precise determination of IPs and σ* shape resonance energies is sensitive enough to distinguish between different saccharides through the minute variations in C–O bond lengths in these systems. The analysis of the experimental data as well as FEFF8 calculations show that this sensitivity arises because bond length variations are associated with term value shifts, permitting determination of bond length variations of ~ 0.01 Å. These observations suggest that NEXAFS can be used to obtain valuable bond length information for other molecular materials of similar complexity. Very importantly, NEXAFS is known to be applicable to non-crystalline systems (*e.g.* species in solutions, melts, amorphous phases and adsorbed species), opening up an avenue towards structural information that can be used to confront computationally derived models for systems without long-range order. Moreover, the C K-edges of fructose and β-lactose do not fit into the trend set up by the other saccharides, suggesting that NEXAFS may also be sensitive enough to distinguish hemiketals from hemiacetals, as well as α and β forms of the otherwise identical saccharides, but more systematic experimental work should examine these hypotheses.

References

[1] R. C. Rowe, P. J. Sheskey, S. C. Owen, *Handbook of Pharmaceutical Excipients*, Pharmaceutical Press and American Pharmacists Association, London and Washington, D.C., 2005.
[2] V. Holsinger, *Fundamentals of Dairy Chemistry*, 3rd ed., Springer, Berlin, Heidelberg, New York, 1999.
[3] D. C. Johnson, W. R. Lacourse, *Carbohydrate Analysis*, Elsevier, Amsterdam, 1995.
[4] P. Collins, R. Ferrier, *Monosaccharides*, John Wiley & Sons, Chichester, 1995.
[5] M. F. Chaplin, *Carbohydrate Analysis - A Practical Approach*, Oxford University Press, Oxford, 1994.
[6] T. K. Lindhorst, *Essentials of Carbohydrate Chemistry and Biochemistry*, Wiley-VCH, Weinheim, 2007.
[7] G. M. Brown, H. A. Levy, *Science*, 1965, **147**, 1038.
[8] J. R. Colvin, *The Biochemistry of Plants*, Academic Press, New York, 1980.
[9] I. Murrieta-Pazos, C. Gaiani, L. Galet, R. Calvet, B. Cuq, J. Scher, *J. Food Eng.* 2012, **112**, 1.
[10] E. Kaivosoja, S. Virtanen, R. Rautemaa, R. Lappalainen, Y. Konttinen, *Eur. Cell. Mater.*, 2012, **24**, 60.
[11] F. Benetti, N. Marchettini, A. Atrei, *Appl. Surf. Sci.*, 2011, **257**, 2142.
[12] P. M. Dietrich, T. Horlacher, P. L. Girard-Lauriault, Th. Gross, A. Lippitz, H. Min, Th. Wirth, R. Castelli, P. Seeberger, W. E. Unger, *J. Carbohydr. Chem.*, 2011, **30**, 361.
[13] P. M. Dietrich, T. Horlacher, Th. Gross, Th. Wirth, R. Castelli, A. G. Shard, M. Alexander, P. H. Seeberger, W. E. Unger, *Surf. Interface Anal.*, 2010, **42**, 1188.
[14] E. A. Willneff, B. A. Ormsby, J. S. Stevens, C. Jaye, D. A. Fischer, S. L. M. Schroeder, *Surf. Interface Anal.*, 2014, **46**, 776.
[15] J. S. Stevens, C. R. Seabourne, C. Jaye, D. A. Fischer, A. J. Scott, S. L. M. Schroeder, *J. Phys. Chem. B*, 2014, **118**, 12121.
[16] J. S. Stevens, S. L. M. Schroeder, *Surf. Interface Anal.*, 2009, **41**, 453.
[17] J. Stöhr, *NEXAFS Spectroscopy*, Springer, Berlin, Heidelberg, 1992.
[18] B. X. Yang, J. Kirz, T. K. Sham, *Phys. Rev. A*, 1987, **36**, 4298.
[19] *Atomic, Molecular and Optical Physics*, National Academic Press (National Research Council), Washington, D.C., 1986.
[20] F. Sette, J. Stöhr, A. P. Hitchcock, *J. Chem. Phys.*, 1984, **81**, 4906.
[21] J. Stöhr, F. Sette, A. L. Johnson, *Phys. Rev. Lett.*, 1984, **53**, 1684.
[22] J. S. Stevens, A. Gainar, E. Suljoti, J. Xiao, R. Golnak, E. F. Aziz, S. L. M. Schroeder, *Chem. Eur. J.*, 2015, **21**, 7256.
[23] M. N. Piancastelli, *J. Electron Spectrosc. Relat. Phenom.*, 1999, **100**, 167.
[24] A. P. Hitchcock, C. E. Brion, *J. Phys. B*, 1981, **14**, 4399.
[25] D. Arvanitis, N. Haack, G. Ceballos, H. Wende, K. Baberschke, A. L. Ankudinov, J. J. Rehr, *J. Electron Spectrosc. Relat. Phenom.*, 2000, **113**, 57.
[26] M. N. Piancastelli, D. W. Lindle, T. A. Ferrett, D. A. Shirley, *J. Chem. Phys.*, 1987, **87**, 3255.
[27] M. N. Piancastelli, D. W. Lindle, T. A. Ferrett, D. A. Shirley, *J. Chem. Phys.*, 1987, **86**, 2765.
[28] B. Kempgens, H. M. Köppe, A. Kivimäki, M. Neeb, K. Maier, U. Hergenhahn, A. M. Bradshaw, *Phys. Rev. Lett.*, 1997, **79**, 35.
[29] B. Kempgens, H. M. Köppe, A. Kivimäki, M. Neeb, K. Maier, U. Hergenhahn, A. M. Bradshaw, *Surf. Sci. Lett.*, 1999, **425**, L376.

[30] N. Haack, G. Ceballos, H. Wende, K. Baberschke, D. Arvanitis, A. L. Ankudinov, J. J. Rehr, *Phys. Rev. Lett.*, 2000, **84**, 614.
[31] V. L. Shneerson, D. K. Saldin, W. T. Tysoe, *Surf. Sci.*, 1997, **375**, 340.
[32] A. G. Hardie, J. J. Dynes, L. M. Kozak, P. M. Huang, *J. Mol. Catal. A*, 2009, **308**, 114.
[33] D. Solomon, J. Lehmann, J. Kinyangi, B. Liang, K. Heymann, L. Dathe, K. Hanley, S. Wirick, C. Jacobsen, *Soil Sci. Soc. Am. J.*, 2009, **73**, 1817.
[34] P. M. Dietrich, T. Horlacher, P. L. Girard-Lauriault, T. Gross, A. Lippitz, H. Min, T. Wirth, R. Castelli, P. H. Seeberger, W. E. S. Unger, *Langmuir*, 2011, **27**, 4808.
[35] S. L. M. Schroeder, G. D. Moggridge, R. M. Ormerod, T. Rayment, R. M. Lambert, *Surf. Sci. Lett.*, 1995, **324**, L371.
[36] S. L. M. Schroeder, *Solid State Commun.* 1996, **98**, 405.
[37] S. L. M. Schroeder, *J. Phys. IV France*, 1997, **7 C2**, 153.
[38] S. L. M. Schroeder, G. D. Moggridge, R. M. Lambert, T. Rayment, *Spectroscopy for Surface Science (Advances in Spectroscopy, Vol.26)*, Wiley & Sons, New York, Chichester, 1998.
[39] K. E. Sohn, M. D. Dimitriou, J. Genzer, D. A. Fischer, C. J. Hawker, E. J. Kramer, *Langmuir*, 2009, **25**, 6341.
[40] B. H. Frazer, B. Gilbert, B. R. Sonderegger, G. De Stasio, *Surf. Sci.* 2003, **537**, 161.
[41] S. Tanuma, C. J. Powell, D. R. Penn, *Surf. Interface Anal.*, 1994, **21**, 165.
[42] B. Ravel, M. Newville, *J. Synchrotron Rad.*, 2005, **12**, 537.
[43] B. Ravel, M. Newville, *Phys. Scr.*, 2005, **T115**, 1007.
[44] J. A. Kanters, G. Roelofsen, B. P. Alblas, I. Meinders, *Acta Cryst., Sect. B*, 1977, **33**, 665.
[45] A. Mostad, *Acta Chem. Scand.*, 1994, **48**, 276.
[46] J. Ohanessian, H. Gillier-Pandraud, *Acta Cryst., Sect. B*, 1976, **32**, 2810.
[47] A. Hordvik, *Acta Chem. Scand.*, 1971, **25**, 217.
[48] M. E. Gress, G. A. Jeffrey, *Acta Cryst., Sect. B*, 1977, **33**, 2490.
[49] J. H. Noordik, P. T. Beurskens, P. Bennema, R. A. Visser, R. O. Gould, *Z. Kristallogr.*, 1984, **168**, 59.
[50] A. L. Ankudinov, S. D. Conradson, J. M. de Leon, J. J. Rehr, *Phys. Rev. B*, 1998, **57**, 7518.
[51] G. M. Battle, G. M Ferrence, F. H. Allen, *J. Appl. Cryst.*, 2010, **43**, 1208.
[52] Y. Nishiyama, J. Sugiyama, H. Chanzy, P. Langan, *J. Am. Chem. Soc.*, 2003, **125**, 14300.
[53] S. Garnier, S. Petit, G. Coquerel, *J. Therm. Anal. Calorim.*, 2002, **68**, 489.
[54] J. Stöhr, *NEXAFS Spectroscopy*, Springer: Heidelberg, 2003.
[55] G. A. Jeffrey, *Carbohydr. Res.*, 1973**, 28**, 233.

CHAPTER 5

LIQUID-STATE NEXAFS/RIXS OF PABA IN WATER AT HIGH/LOW pH AND IN METHANOL SOLUTION

5.1. Introduction

The ability to observe chemical and structural changes in organic molecules in solution is of central importance for progress in chemical and life sciences. For example, the activity of biological systems is strongly influenced by pH variations causing changes in chemical speciation,[1] while the structural outcomes of self-assembly and crystallisation processes define physical and chemical properties of products.[2,3] In particular, details of the structural evolution of nucleating solutions are of contemporary interest[3,4] because of their far-reaching impact on both structural and morphological outcomes, the tailoring of product properties, and the design of materials.

A crucial pre-requisite for further elucidation of molecular structure in such processes is having the experimental capability to incisively study conformational structure, bonding and electronic properties of organic molecular solutes *in situ*.[3,5] Core-level spectroscopies are emerging as suitable probes for this task, as they are highly sensitive to local electronic structure and bonding. Even X-ray photoelectron spectroscopy (XPS),[6] one of the most facile core-level spectroscopies in the laboratory, permits the characterisation of protonation and hydrogen bonding in both solid-state and solution acid-base systems through determination of chemical shifts in core-level binding energies.[7,8]

Near-edge X-ray absorption fine structure (NEXAFS), also known as X-ray absorption near-edge structure (XANES), is chemically and structurally much more incisive,[9,10] as it involves excitation of core electrons to unoccupied valence orbitals that are readily interpreted by use of molecular orbital (MO) calculations. In addition, σ* shape resonances in the NEXAFS are very sensitive to bond length variations, thus providing incisive molecular structure parameters, even for non-crystalline systems.[11] Resonant inelastic X-ray scattering (RIXS) probes occupied valence orbitals when the correlation of the valence electrons is weak, by monitoring the transitions of electrons from occupied

valence orbitals into core-hole states.[12] The combination of NEXAFS and RIXS thus provides a comprehensive picture of the local electronic structure and is ideally placed for investigations of local environment and bonding of molecular solutes. Recent development of the liquid microjet technique[13,14] allows windowless *in situ* probing of solute species in solution and, importantly for organics, alleviates potential problems caused by radiation damage. In addition, the ultra-fast nature of the electronic transitions ($\sim 10^{-15}$ s)[9] means that every detected event represents a snapshot of a molecule rather an average over vibrational or conformational changes, thereby permitting population analysis when transient species co-exist.[8,15]

This chapter will present how NEXAFS and RIXS microjet studies can be combined with density functional theory (DFT) calculations to elucidate the pH- and solvent-dependent changes in the local chemical and electronic environment of *para*-aminobenzoic acid (PABA, Figure 5.1). PABA is an intermediate for folic acid synthesis and has pharmaceutical applications.[16] Depending on pH and solvent it forms neutral, zwitterionic, cationic and anionic species in solution.[17,18] Direct observations of these chemically and electronically distinct species are possible by monitoring the nitrogen 1s core-level excitation (nitrogen K-edge), thereby providing electronic structure signatures for aqueous solutions with varying pH as well as for alcoholic solutions. This chapter will also present a correlation of the electronic structure and bond length changes with the effects of ionisation at Brønsted donor and acceptor moieties, and a comparison with non-ionic species in the solid-state and DFT molecular orbitals.

5.2. Experimental section

5.2.1. Solid-state NEXAFS

The solid-state PABA sample (β-PABA[19]) was formed through an aqueous slurry of the commercial form (>99%, Sigma-Aldrich, UK) at 5°C. Solid-state NEXAFS measurements were performed at the U7A beamline of the National Synchrotron Light Source (NSLS) at Brookhaven National Laboratory, NY. Partial electron yield (PEY) spectra for the N K-edge were collected via a channeltron electron multiplier with the sample at the magic angle (54.7°) relative to the incident beam. An entrance grid bias of −150 V was used for PEY collection and a monochromator with a 600 l/mm grating,

which provided energy resolution of ~ 0.15 eV. After collection, the spectra were normalised by the simultaneously recorded drain current from an *in situ* gold-coated, 90% transmission grid (I_0) placed in the incident X-ray beam to eliminate the effect of incident beam intensity fluctuations and beamline optics absorption features, and the monochromator energy was calibrated using the 400.6 eV first π^* feature of a titanium nitride grid.

Peak fitting and normalisation was performed using the Athena software,[20] with *arctan* steps for the IPs (edge steps) and Gaussian functions for the peaks.[9,21,22] For investigation of the relationship of C–N bond length with the NEXAFS term value δ_{C-N} (σ^*_{C-N} − IP), standard deviation for the C–N bond lengths from single crystal X-ray diffraction (XRD) is ≤ 0.003 Å [19,23,24] and good correspondence is observed for the fitted IP energy shifts with the XPS N 1s core binding energies (≤ 0.15 eV standard deviation).[21] The correlation between the N NEXAFS and C–N bond length reveals that a 0.1 eV change in the term value δ_{C-N} corresponds to a 0.0025 Å alteration in C–N bond length.

5.2.2. Solution-state *in situ* NEXAFS

Solution-state N K-edge spectra were recorded with the LiXEdrom endstation[14] at the U41-PGM beamline of the BESSY II synchrotron at Helmholtz Zentrum Berlin (HZB), by using the liquid microjet technique. The 0.5 M acidic (HCl) and basic (NaOH) aqueous solutions and 0.5 M methanol solution of PABA were prepared and filtered at ambient pressure to remove impurities or undissolved crystals. Partial fluorescence yield (PFY) data were recorded in scanning mode by a grating with line density 1200 lines/mm and radius 7.5 m dispersing the emitted photon energy from the sample,[14] and subsequently a detector consisting of a charge-coupled device (CCD), fluorescence screen and microchannel plate (MCP) stack collecting the amplified signal. The sample, grating and detector were arranged in Rowland circle geometry for accurate focussing. Although total fluorescence yield –TFY– can be recorded with a GaAsP photodiode mounted in the vicinity of the liquid jet, there were problems with the signal recording and solute could crystallise on the photodiode, interfering with the signal). Use of the microjet ensures fresh sample is probed by the X-ray beam, vastly minimising the potential for any X-ray induced damage. An 18 μm diameter glass nozzle with 0.6 mL/min flow rate was used, and the measurement performed around 2 mm from the nozzle within the laminar part of

the jet flow (droplets start to form after 3-5 mm, with the resulting frozen residues collected by a liquid nitrogen trap) with 10^{-5} mbar pressure in the main chamber. Beamline energy calibration was performed with N_2 gas (TEY XAS 1s→2p π transition) and the resolution was ≤ 0.1 eV.

5.2.3. Solution-state *in situ* RIXS

Microjet N RIXS spectra were recorded at BESSY II[14,25] by using the same setup and solutions as for the NEXAFS. N RIXS data were recorded at multiple excitation energies corresponding to NEXAFS resonances. In line with this, similar spectral features at constant emission energy with varying excitation energy were observed, and only changes in relative intensity occurred across the NEXAFS transition thresholds (excitation energies of 404.5, 410, 408 eV for the anionic, non-ionic and cationic species, respectively, were used for further analysis).

5.2.4. DFT calculations

Non-ionic, anionic, and cationic PABA monomers were optimised with the B3LYP functional and 6-31G* basis set in Gaussian 09[26] to obtain molecular orbitals[27] for the ground state. This permitted identification of electronic structure changes originating from chemical (initial-state) variations.[9] Strictly speaking, NEXAFS interpretation should also include the influence of relaxation effects due to the presence of the core hole, which were expected to affect the final state of the observed electronic transitions. We have recently included such effects in CASTEP calculations for crystalline PABA[21] and found that the inclusion of final-state effects did not alter the MO interpretation of the NEXAFS substantially. The reason for this insensitivity lies in a combination of the strong localised character of core-level excitations and the weakness of intermolecular interactions (with hydrogen bonding dominant) of organic molecules relative to internal covalent bonding and protonation effects. The ground-state calculations used for the interpretation of the NEXAFS are therefore expected to somewhat overestimate the absolute energies associated with the π* transitions, but less so the relative energies and sequence of the unoccupied π* states.

5.3. Results and discussion

5.3.1. Expected speciation in solution and in the solid state

The solubility of PABA in aqueous solution is very low (0.03 M) and dissolution leads to a complex mixture at the isoelectric point (pH 3.69), comprising non-ionic and zwitterionic neutral species[17] alongside small percentages of cationic and anionic species (Figure 5.1).[28] Based on the pK_a values,[28] upon increasing the pH above 4.87 the carboxylate anionic species dominates, while decreasing the pH below 2.50 results in solutions containing almost entirely the cationic species (Figure 5.1). These ionic species have greatly increased solubility (> 0.5 M) and allow microjet NEXAFS and RIXS measurements to be recorded with good signal-to-noise ratio. In the solid state, PABA exists as neutral, non-zwitterionic molecules bonded through a series of intermolecular hydrogen bonds and π-π interactions,[19,23,29] providing a spectrum of the non-ionic form for comparison with the ionic species expected at high and low pH. For comparison we have also examined alcoholic solutions, in which PABA is expected to primarily exist in the non-ionic neutral form (*i.e.* non-zwitterionic).[17]

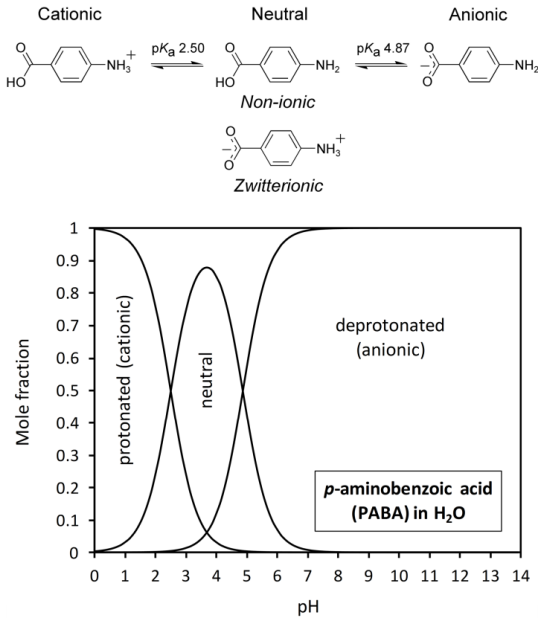


Figure 5.1. Equilibria between the cationic, neutral and anionic PABA species in water (top) and speciation diagram as a function of pH (bottom).

5.3.2. NEXAFS

Comparison of the N K NEXAFS spectra for non-ionic, solid-state PABA and *in situ* measurements at high/low pH in aqueous solution (Figure 5.2) reveals marked differences between the electronic structure of the species at pH 11 and pH 1, as well as variations between the non-ionic and pH 11 species. Two pre-edge resonances (energetically below the ionisation potential, IP), are evident for the non-ionic form and at pH 11, which arise from 1s→π* transitions due to conjugation of the nitrogen lone pair with the aromatic ring, and these are followed by the broader 1s→σ* resonances that are more susceptible to variations in geometric structure.[9]

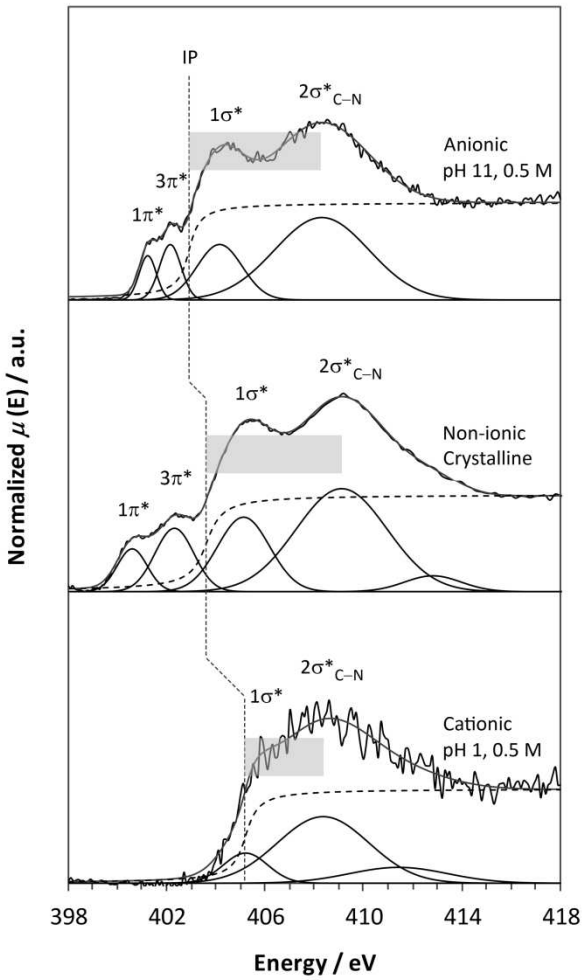


Figure 5.2. N K NEXAFS spectra for anionic (top), non-ionic solid-state β- (middle), and cationic PABA (bottom), together with the fits to the data, IPs marked with vertical dashed lines, and the magnitude of the term value δ_{C-N} indicated by shaded boxes.

The IP for the species present at pH 1 is very different from that observed at pH 11, with a positive energy shift of 2.3 eV (Figure 5.2, Table 5.1). This increased IP for the pH 1 solution species reflects acquisition of a positive charge on nitrogen for the cationic species (decreased electron density, orbital contraction). A slight decrease in IP occurs at pH 11 compared to the non-ionic form (Table 5.1), with conjugation of the carboxyl and amine groups through the aromatic ring leading to some orbital screening and redistribution of electron density occurring with change to the anionic species (Figure 5.1). Due to the absence of the lone electron pair on the nitrogen atom with NH_3^+, the cationic nitrogen will not be able to participate in the π molecular orbitals with the aromatic ring, therefore no pre-edge π^* peaks show in the pH 1 N K-edge (Figure 5.2), further signifying transformation to the cationic species. Interestingly, the π^* resonances become closer to one another in the anionic species at pH 11 compared to the non-ionic form, leading to a smaller $\Delta\pi^*$ value for the anionic species (Figure 5.2, Table 5.1).

Table 5.1. Experimental N K NEXAFS resonance and IP energies from Figure 5.2 along with the term value $\delta_{C-N.}$

Nitrogen [eV]	Non-ionic β-PABA	Anionic pH 11	Cationic pH 1	Non-ionic methanol
IP	403.60	402.90	405.20	403.47
$1\pi^*$	400.58	401.22	/	400.68
$3\pi^*$	402.30	402.14	/	401.98
$1\sigma^*_{N-H}$	405.11	404.15	405.20	403.66
$2\sigma^*_{C-N}$	409.09	408.30	408.36	408.66
$\Delta\pi^*$	1.72	0.92	/	1.30
δ_{C-N} (σ^*_{C-N} – IP)	5.49	5.40	3.16	5.19

DFT calculations[†] of the unoccupied MOs with nitrogen contributions allow us to pinpoint the electronic transitions leading to pre-edge π^* nitrogen intensity in the anionic and non-ionic NEXAFS and to interpret the magnitude of the calculated energy shifts. Figure 5.3 shows the lowest unoccupied MOs (LUMOs) for non-ionic and anionic PABA,

[†] Acknowledged work of Dr Joanna S. Stevens

revealing the calculated gap between the two lowest energy nitrogen π* orbitals (Δπ*) is over twice as large for non-ionic PABA as for the anionic species. In addition, the π* energies for anionic PABA are raised above that for PABA, with the nitrogen 1π* becoming the LUMO+1 rather than the LUMO (Figure 5.3).

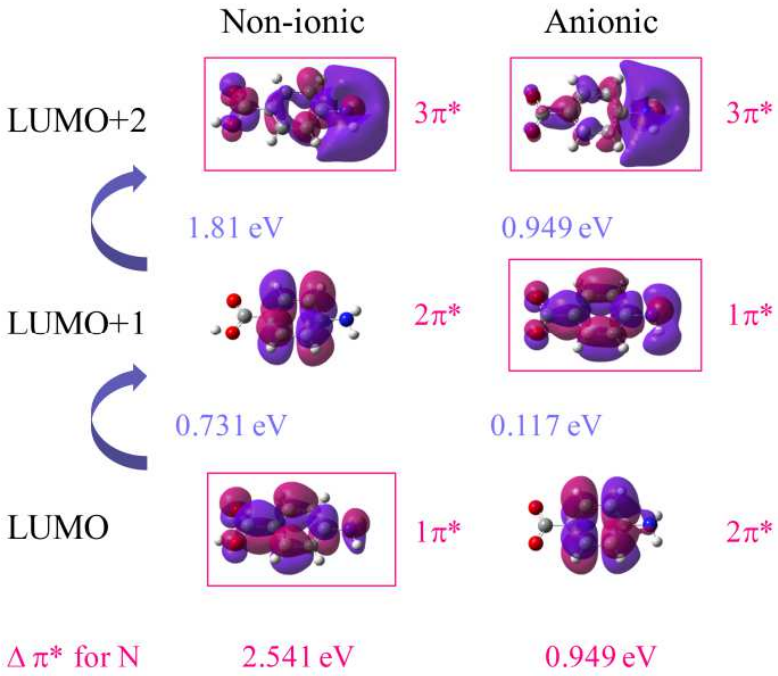


Figure 5.3. Lowest unoccupied molecular orbitals for the neutral non-ionic and anionic monomers of PABA. The π* orbitals are labelled 1π*, 2π*, 3π* from the lowest energy relative to non-ionic PABA and those with weight on the nitrogen absorber are marked with boxes; transitions from the N 1s core level to these marked orbitals are visible in the N K NEXAFS.

The experimental trend in the IP (Figure 5.2) for the different PABA species is also reproduced by the calculations (Figure 5.4), with a decrease in core-level N 1s energy for cationic species (thus increased IP) and slight increase for anionic relative to non-ionic PABA (decreased IP). The nitrogen π* LUMOs for the anionic species mirror the change in the core-level energy, with an increase in energy relative to non-ionic PABA (Figure 5.4). The energy gap between the 1π* and 3π* MOs also narrows for the anionic species (Figures 5.3 and 5.4); this decrease in the gap between the nitrogen π* resonances (Δπ*) for the anionic species parallels that observed in the experimental NEXAFS (Figure 5.2), explaining it by a significant increase in the 1s→1π* energy and small decrease in the

1s→3π* energy relative to non-ionic PABA occurring with formation of the carboxylate form (Table 5.2). Comparison of calculations for the zwitterionic form (Table 5.2) illustrates that its presence would significantly raise the IP, albeit perhaps not quite as much as for the cationic species, alongside the loss of resolved pre-edge π* peaks through the positive charge on the nitrogen. While the magnitude of the calculated energy shifts may not be accurate due to the absence of core-hole relaxation and screening from intermolecular interactions, the direction and trends observed correlate well with the experimental data, indicating that variations in ground-state (initial-state) properties allow the interpretation of electronic structure probed through core-level excitation, tending to dominate over final-state effects.

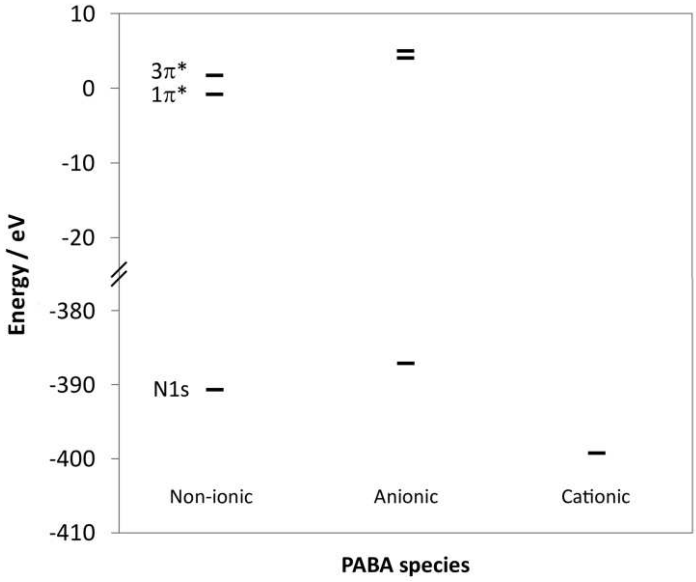


Figure 5.4. Calculated orbital energies for PABA showing the trends in nitrogen π* (top) and core 1s (bottom) orbital energies for non-ionic, anionic and cationic species.

Table 5.2. Calculated orbital energies and shifts for different PABA species.

For N / eV	Non-ionic	Anionic	Cationic	Zwitterionic
N 1s	–390.66	–387.10	–399.24	–396.02
N 1s→1π*	389.86	391.16	/	/
N 1s→3π*	392.40	392.11	/	/
Δπ* (1π*→3π*)	2.54	0.95	/	/

5.3.3. C–N bond lengths‡

The experimental N K NEXAFS spectra (Figure 5.2) also show variation in the position of the broader, post-edge σ* resonances relative to the IP for the non-ionic, anionic, and cationic species. For a bond between two atoms X and Y the term value $\delta_{X-Y,}$ the difference between the energy of its σ^*_{X-Y} orbital and the IP of either X or Y, provides a highly sensitive measure of the bond length.[9,30] As can be seen in Figure 5.2, and from the term values (δ_{C-N}) in Table 5.1, the separation between the energy of σ^*_{C-N} resonance and the IP is of a similar magnitude for anionic and non-ionic PABA. In contrast, the energy of σ^*_{C-N} and the IP are much closer for the cationic species, which reflects an increase in IP and decrease in the energy of σ^*_{C-N} resonance following protonation of the amino group, and results in a substantially smaller δ_{C-N} value (Figure 5.2). As the C–N bond lengths are known from X-ray diffraction (XRD) crystal structures for the solid state (β non-ionic form[19] and cationic HCl salt,[24] as well as the additional non-ionic α form[23] previously investigated,[21] Table 5.3), these can be used to obtain a calibration plot for bond length determination in the solutions. Plotting the term value δ_{C-N} against the C–N bond distances in the corresponding crystal structures indicates a linear dependence (Figure 5.5, black data points). The correlation between XRD-derived bond lengths and NEXAFS-derived term values is very high (Figure 5.5), which suggests sensitivity to bond length variations of less than 0.005 Å (Table 5.3). The term values for the solution species are similar to those found in their corresponding solid-state structures, thus suggesting that determination of the term values δ_{C-N} measures their bond length reliably (Table 5.3, Figure 5.5 grey data points) in both the solid state and in solution.

‡ Partly acknowledged work of Dr Joanna S. Stevens on the bond length determination; synthesis of cationic HCl salt, and peak fitting represents my own work

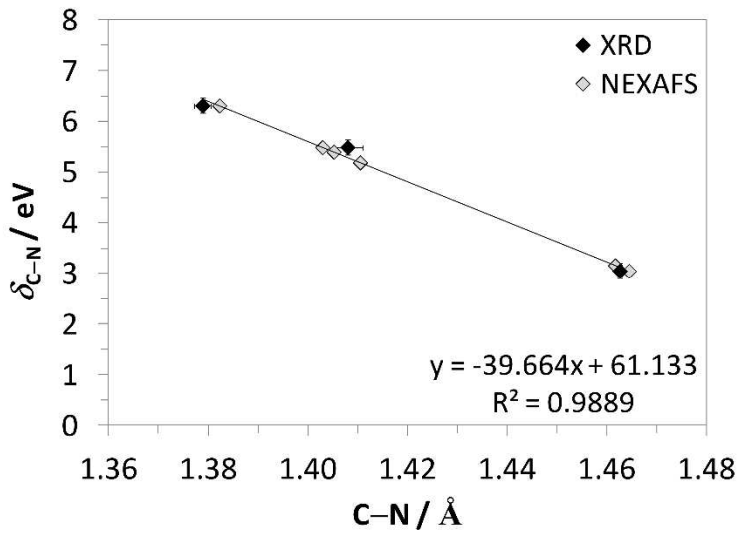


Figure 5.5. Correlation between the NEXAFS term value (δ_{C-N}) and C–N bond lengths from XRD[19,23,24] (black), and comparison with C–N bond lengths obtained from NEXAFS (grey) for the non-ionic and ionic PABA species (Table 5.3).

The observed impact on the C–N bond length is of the order of less than tenths of an Ångstrom, but appears to be within the detection limit. Certainly the expected trend, overall longest C–N bonds for protonated nitrogen in the cationic species and shorter bonds for the neutral as well as the non-ionic species, is evident for both the solid and the liquid phase. This indicates strong potential for a generalisation of the σ* shape resonance analysis to other organic molecular solutes, and even 'measuring' bond lengths in solution.

Table 5.3. Experimental C–N bond lengths from XRD and those derived from NEXAFS

Nitrogen	δ_{C-N} /eV	XRD C−N /Å	NEXAFS C−N /Å
Non-ionic α-PABA	6.31[a]	1.379(2)[b]	1.3822
Non-ionic β-PABA	5.49	1.408(3)[c]	1.4029
Anionic pH 11	5.40	[1.410(2)[d]]	1.4052
Non-ionic methanol	5.19	/	1.4105
Cationic pH 1	3.16	/	1.4617
Cationic HCl salt	3.05	1.4626(5)[e]	1.4644

[a] Ref. 21
[b] Ref. 23
[c] Ref. 19
[d] C–N bond length from the hydrated Na salt crystal structure[31] shown for comparison
[e] Ref. 24

5.3.4. Methanol species

Comparison with the N K NEXAFS for PABA obtained in methanol (Figure 5.6) reveals a shape in agreement with that of non-protonated nitrogen species (Figure 5.2). The 1π* resonance (with its previously demonstrated sensitivity to the chemical species present) is consistent with those observed for the non-ionic form rather than anionic species (Table 5.1). If the molecules in the 0.5 M methanol solution were primarily zwitterionic (Figure 5.1), a spectrum resembling that of the cationic species (Figure 5.2) would be expected, with no pre-edge π* resonances and an increased IP resulting from the positive charge on nitrogen. The fact that this is not observed provides direct evidence that (at least the majority of) PABA molecules in methanol exist in the neutral, uncharged (non-ionic) form. The smaller δ_{C-N} for the methanol species relative to the solid state (Table 5.1) suggests that the C–N bond is slightly longer in the methanol solution, and applying the correlation above gives a predicted C–N bond length of 1.4044 Å (Table 5.3).

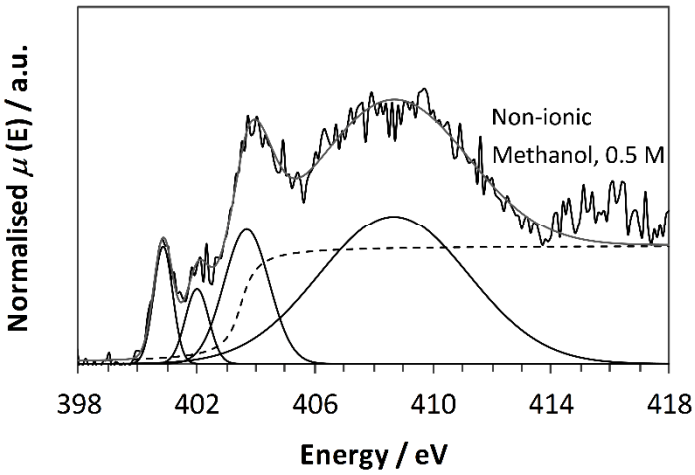


Figure 5.6. N K NEXAFS of PABA in 0.5 M methanol.

5.3.5. RIXS[§]

While NEXAFS probes electron transitions from the core level to unoccupied orbitals, RIXS involves transitions from occupied valence orbitals to the core hole after excitation.[9] For nitrogen, this follows the 2p→1s transitions. The N RIXS spectra for the non-ionic and anionic species (methanol and pH 11 solutions, respectively) initially look

[§] Acknowledged work of Dr Joanna S. Stevens

similar, whereas that for pH 1 is clearly distinguished (Figure 5.7). Following protonation of nitrogen for the cationic species at low pH, the highest energy peaks arising from the occupied 2p π valence orbitals→1s are absent due to the use of the nitrogen lone pair in forming an additional N–H bond and not part of the π MOs.

Closer inspection of the N RIXS for anionic and non-ionic species reveals a significant shift to lower energy for the anionic species (Figure 5.8, left). This is particularly noticeable towards the higher-energy region of the RIXS data. As the highest energy (inelastic) peak in the RIXS arises from the decay of electrons from the nitrogen highest occupied MO (HOMO) to the core N 1s level, this indicates a lower energy for this MO for the anionic species.

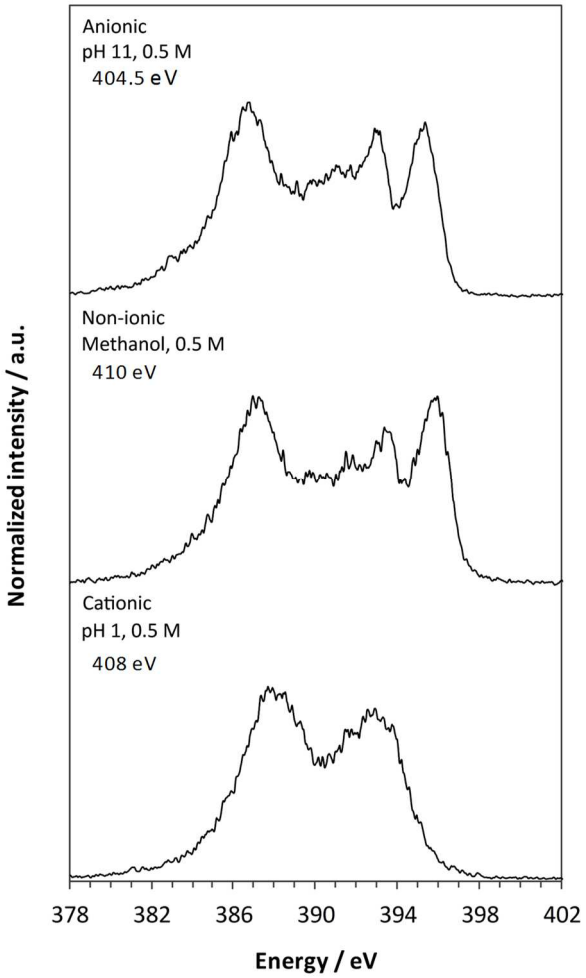


Figure 5.7. N K RIXS spectra for anionic (top), non-ionic (middle), and cationic (bottom) species in solution at excitation energies 404.5, 410 and 408 eV respectively.

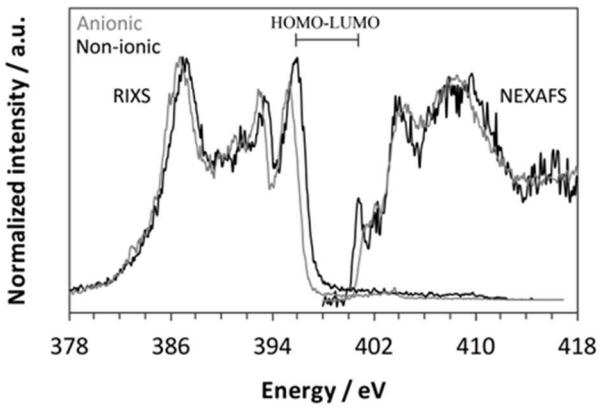


Figure 5.8. N K RIXS and K NEXAFS showing the shift in energy between anionic (grey) and non-ionic (black) solution species and the HOMO↔LUMO gap.

Comparison of the relative energies for the HOMOs with nitrogen contributions for non-ionic and anionic species reveals that the HOMO→1s leads to the first RIXS peak for the non-ionic form (Figure 5.9). In contrast, this MO is lowered in energy for the anionic species, becoming the HOMO–3 (Figure 5.9), thereby elucidating the shift to lower energy seen experimentally (Figure 5.8). Further insight can be obtained through combination of the RIXS and NEXAFS data (Figure 5.8), providing a model of the local density of valence states. The energy difference between the highest-energy RIXS peak and lowest-energy NEXAFS peak is a probe of the gap between the HOMO and LUMO per atomic chemical state (Figure 5.8); an alternative method for the band gap is to take the difference of the intersections between the first peak slopes and the backgrounds (tending to underestimate).[32] For nitrogen in the non-ionic form, this represents the HOMO↔LUMO gap, whereas for the anionic species, this is for HOMO–3↔LUMO+1 (Figure 5.9). Experimentally, this leads to energy gaps of 5.06 and 5.89 eV for the non-ionic and anionic PABA species respectively, which compares favourably with the predicted values of 5.01 and 5.86 eV.

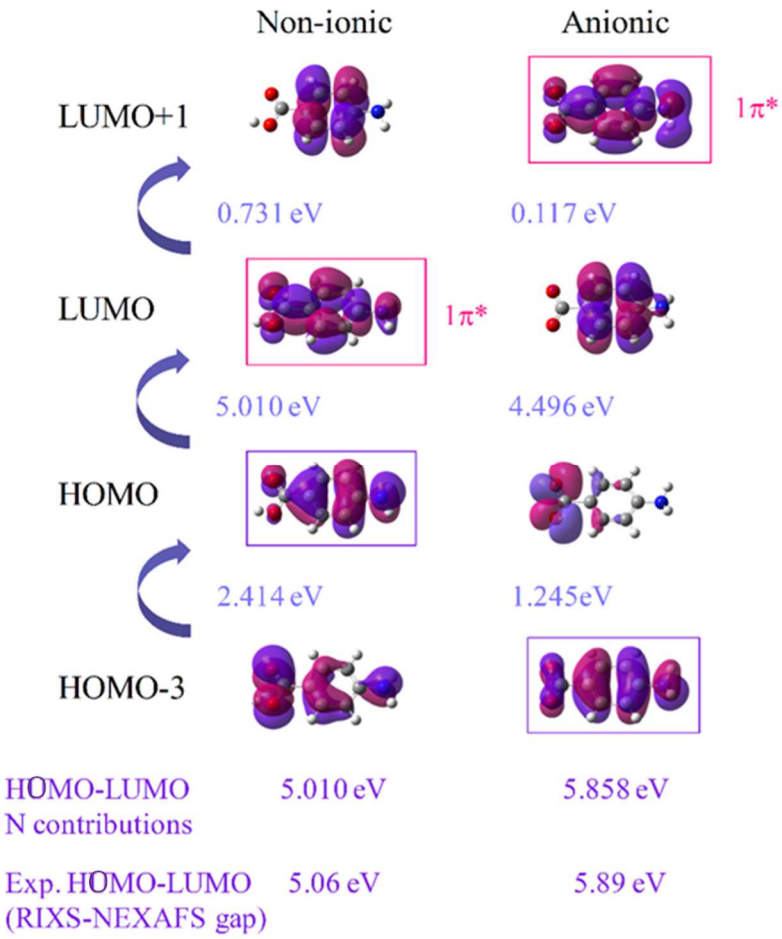


Figure 5.9. Highest occupied and lowest unoccupied molecular orbitals involved in the first RIXS (purple outline) and NEXAFS (pink outline) resonances, respectively, with the experimental and calculated nitrogen HOMO↔LUMO gap for neutral and anionic PABA.

5.4. Conclusions

Directly monitoring the core-level transitions of the amino group by N NEXAFS and RIXS as a function of pH successfully characterises the chemical and electronic state of PABA species in aqueous solution. Formation of the cationic species by protonation of the amino group at low pH leads to a significant shift in IP to higher energy, along with an absence of π* resonances in NEXAFS and of π valence peaks in RIXS. Although the amino group is not protonated in the anionic and non-ionic species, differences are observed in both NEXAFS and RIXS. There is a slight shift to low energy for the IP and

a widening of the HOMO↔LUMO gap for the anionic species, in agreement with predictions of DFT calculations. In methanol, the NEXAFS resembles that of non-ionic PABA, with no indications of the presence of the zwitterionic form. Structural as well as chemical and electronic changes impact the NEXAFS spectra, with variation of the C–N bond length influencing the energy of the σ^*_{C-N} resonance relative to the IP, thereby providing access to bond length determination of solutes in solution by comparison with NEXAFS data of PABA species in the solid state.

References

[1] M. N. Chatterjea, R. Shinde, *Textbook of Medical Biochemistry*, 8th edition, Jaypee Brothers Medical Publishers (P) Ltd, New Delhi, 2012; J. M. Berg, J. L. Tymoczko, L. Stryer, *Biochemistry*, 5th edition, Freeman, San Francisco, 2002.
[2] J. T. Carstensen, *Pharmaceutical Principles of Solid Dosage Forms*, Technomic Publishing Co., Lancaster, PA, 1993.
[3] R. J. Davey, S. L. M. Schroeder, J. H. ter Horst, *Angew. Chem., Int. Ed.*, 2013, **52,** 2166.
[4] R. J. Davey, K. Allen, N. Blagden, W. I. Cross, H. F. Lieberman, M. J. Quayle, S. Righini, L. Seton, G. J. T. Tiddy, *Cryst Eng Comm.*, 2002, **4**, 257; I. Weissbuch, M. Lahav, L. Leiserowitz, *Cryst. Growth Des.*, 2003, **3**, 125; D. Gebauer, H. Cölfen, *Nano Today*, 2011, **6**, 564; J. Anwar, D. Zahn, *Angew. Chem., Int. Ed.*, 2013, **50**, 1996; D. Gebauer, M. Kellermeier, J. D. Gale, L. Bergstrom, H. Colfen, *Chem. Soc. Rev.*, 2014, **43**, 2348.
[5] N. Pienack, W. Bensch, *Angew. Chem., Int. Ed.*, 2011, **50**, 2014.
[6] D. Briggs, M. P. Seah, P. M. A. Sherwood, in *Practical Surface Analysis, Volume 1: Auger and X-ray Photoelectron Spectroscopy*, 2nd edition, John Wiley & Sons, Chichester, 1990.
[7] J. S. Stevens, S. J. Byard, C. C. Seaton, G. Sadiq, R. J. Davey, S. L. M. Schroeder, *Phys. Chem. Chem. Phys.*, 2014, **16**, 1150; J. S. Stevens, S. J. Byard, C. C. Seaton, G. Sadiq, R. J. Davey, S. L. M. Schroeder, *Angew. Chem., Int. Ed.*, 2011, **50**, 9916; D. Nolting, E. F. Aziz, N. Ottosson, M. Faubel, I. V. Hertel, B. Winter, *J. Am. Chem. Soc.*, 2007, **129**, 14068.
[8] D. Nolting, N. Ottosson, M. Faubel, I. V. Hertel, B. Winter, *J. Am. Chem. Soc.*, 2008, **130**, 8150.
[9] J. Stöhr, *NEXAFS Spectroscopy*, 2nd edition, Springer-Verlag, Berlin, 1992.
[10] G. Hahner, *Chem. Soc. Rev.*, 2006, **35**, 1244; J. Lehmann, D. Solomon, J. Brandes, H. Fleckenstein, C. Jacobson, J. Thieme, *Biophysico-Chemical Processes Involving Natural Nonliving Organic Matter in Environmental Systems*, John Wiley & Sons, Inc., 2009; P. Leinweber, J. Kruse, F. L. Walley, A. Gillespie, K.-U. Eckhardt, R. I. R. Blyth, T. Regier, *J. Synchrotron Radiat.*, 2007, **14**, 500; B. M. Messer, C. D. Cappa, J. D. Smith, K. R. Wilson, M. K. Gilles, R. C. Cohen, R. J. Saykally, *J. Phys. Chem. B*, 2005, **109**, 5375; E. Otero, S. G. Urquhart, *J. Phys. Chem. A*, 2006, **110**, 12121.
[11] D. Arvanitis, N. Haack, G. Ceballos, H. Wende, K. Baberschke, A. L. Ankudinov, J. J. Rehr, *J. Electron Spectrosc. Relat. Phenom.*, 2000, **113**, 57; N. Haack, G. Ceballos, H. Wende, K. Baberschke, D. Arvanitis, A. L. Ankudinov, J. J. Rehr, *Phys. Rev. Lett.*, 2000, **84**, 614.
[12] A. Nilsson, L. G. M. Pettersson, *Surf. Sci. Rep.*, 2004, **55**, 49.
[13] M. Faubel, B. Steiner, J. P. Toennies, *J. Chem. Phys.*, 1997, **106**, 9013; K. M. Lange, R. Könnecke, M. Soldatov, R. Golnak, J.-E. Rubensson, A. Soldatov, E. F. Aziz, *Angew. Chem., Int. Ed.*, 2011, **50**, 10621.
[14] K. M. Lange, R. Könnecke, S. Ghadimi, R. Golnak, M. A. Soldatov, K. F. Hodeck, A. Soldatov, E. F. Aziz, *Chem. Phys.*, 2010, **377**, 1.
[15] E. Ito, H. Oji, T. Araki, K. Oichi, H. Ishii, Y. Ouchi, T. Ohta, N. Kosugi, Y. Maruyama, T. Naito, T. Inabe, K. Seki, *J. Am. Chem. Soc.*, 1997, **119**, 6336.
[16] A. F. Wagner, K. Folkers, *Vitamins and Coenzymes*, Interscience Publishers, New York, 1964; E. Staunton, W. R. Todd, H. S. Mason, J. T. V. Bruggen, *Textbook of Biochemistry*, The MacMillan Company, New York, 1967.
[17] B. van de Graaf, A. J. Hoefnagel, B. M. Wepster, *J. Org. Chem.*, 1981, **46**, 653.

[18] R. A. Sullivan, R. J. Davey, G. Sadiq, G. Dent, K. R. Back, J. H. ter Horst, D. Toroz, R. B. Hammond, *Cryst. Growth Des.*, 2014, **14**, 2689.
[19] *Gaussian 09, Revision B.01*, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, et al., 2009.
[20] F. Sette, J. Stöhr, A. P. Hitchcock, *J. Chem. Phys.*, 1984, **81**, 4906.
[21] T. F. Lai, R. E. Marsh, *Acta Cryst.*, 1967, **22**, 885.
[22] M. Colapietro, A. Domenicano, *Acta Cryst.*, 1980, **B36**, 354.
[23] B. Ravel, *ATHENA User Guide*,
http://cars9.uchicago.edu/~ravel/software/doc/Athena/html/athena.pdf, 2009.
[24] M. L. Deviney, R. C. Anderson, W. A. Felsing, *J. Am. Chem. Soc.*, 1957, **79**, 2371.
[25] J. S. Stevens, C. R. Seabourne, C. Jaye, D. A. Fischer, A. J. Scott, S. L. M. Schroeder, *J. Phys. Chem. B*, 2014, **118**, 12121.
[26] G. Salas-Ciminigo, *C. R. Acad. Sci., Ser. C (Chim)*, 1968, **267**, 1304.
[27] R. G. Wilks, University of Saskatchewan (Saskatoon), 2009.
[28] B. Ravel, M. Newville, *J. Synchrotron Radiat.*, 2005, **12**, 537; M. Newville, *J. Synchrotron Radiat.*, 2001, **8**, 322.
[29] M. A. Soldatov, K. M. Lange, M. D. Gotz, N. Engel, R. Golnak, A. Kothe, E. F. Aziz, *Chem. Phys. Lett.*, 2012, **546**, 164; K. M. Lange, E. Suljoti, E. F. Aziz, *J. Electron Spectrosc. Relat. Phenom.*, 2013, **188**, 101.
[30] *GaussView, Version 5*, Roy Dennington, Todd Keith, and John Millam ed., R. Dennington, T. Keith, and J. Millam, Semichem Inc., Shawnee Mission, 2009.
[31] S. Athimoolam, S. Natarajan, *Acta Cryst., Sect. C: Cryst. Struct. Commun.*, 2007, **63**, o514.
[32] S. Gracin, A. Fischer, *Acta Cryst., Sect. E: Struct. Rep. Online*, 2005, **61**, o1242.

CHAPTER 6

NEXAFS OF NEUTRAL PABA SPECIES IN WATER AND IN ALCOHOL SOLUTIONS

6.1. Introduction

Understanding the molecular basis for changes in solute and solution structure as a function of solvent composition and the presence of other components such as salts, surfactants, acids and bases is a pre-requisite for generating a deep understanding of speciation and pre-nucleation processes in solution.[1-3] In chapter 5 it was shown how detailed information about the properties of *para*-aminobenzoic acid (PABA) solutions at high and low pH could be obtained from N K near-edge X-ray absorption fine structure (NEXAFS) and resonant inelastic X-ray scattering (RIXS) spectroscopies of solution microjets,[4] providing structural and chemical state information on the ionic forms of PABA. The spectrum of neutral PABA species at intermediate pH could at the time not be obtained due to the low concentrations in near-neutral aqueous solutions. However, it would be interesting to obtain the spectrum of neutral PABA, as it can exist in either non-ionic (uncharged) or zwitterionic form.[5] The proportion of each, and the level of their self-association, is known to vary between solvents,[6] but there are not many techniques that are at the same time structurally incisive like NEXAFS and allow the determination of the balance between non-ionic and zwitterionic neutral organic species in solution.

N K NEXAFS is expected to be extremely sensitive to the balance between the two species, as the previous N K-edge study highlighted the strong effect of protonation at the amino group of PABA.[7] The N 2p lone pair of the non-protonated primary amino group is electronically conjugated to the π system of the aromatic ring, representing an sp^2-like electronic configuration at the N centre. As a result, N 1s electrons can be excited into the unoccupied π^* states that arise from conjugation to the aromatic system. In contrast, protonation leads to the formation of an sp^3-like configuration at the N centre, resulting in the loss of the electronic transitions to π^* states. This probably singles out NEXAFS as one of the spectroscopic techniques most sensitive to the protonation state of the amino

group, and examination of the N K-edge spectrum of the neutral species should be able to determine speciation unequivocally.

In the following, the speciation in aqueous solutions will be investigated at pH values that are characteristic for mixtures of cationic, neutral and anionic species of PABA. Analysis of the electronic signature for PABA in these spectra will provide an indication of whether the nature of interactions alters as the concentration is increased towards the saturation limit, and with variation of the solvent.

6.2. Experimental section

6.2.1. Materials

The PABA crystalline powder (≥ 99%) was purchased from Sigma-Aldrich (UK). For the preparation of solutions, laboratory distilled and deionised water, methanol (≥ 99.9%) from Carl Roth (Germany), ethanol (≥ 99.8%) and 2-propanol (≥ 99.8%) from Sigma-Aldrich (Germany) were used as solvents. pH of the solution was adjusted by dropwise addition of 1.2 M aqueous solution of NaOH or of HCl.

6.2.2. *In situ* solution-state NEXAFS

The N K-edge spectra were recorded with the LiXEdrom endstation[8-11] at the U41-PGM beamline from BESSYII at Helmholtz Zentrum Berlin (HZB), by employing the liquid microjet technique, thus allowing continuously fresh liquid sample to be probed by the X-ray beam and avoiding potential radiation damage. The 0.05 M (pH 2.20), 0.05 M (pH 2.50), 0.03 M (pH 3.69), 0.05 M (pH 4.85) aqueous solutions, 0.30 M, 0.60 M, 0.75 M ethanol solutions, and 0.30 M 2-propanol solution were prepared and filtered at ambient pressure to remove impurities or undissolved crystals. The experimental setup allows spectra to be recorded in partial fluorescence yield (PFY) scanning mode by a grating with line density 1200 lines/mm,[11] while the amplified signal is collected by a detector consisting of a charged-couple device (CCD), fluorescence screen and microchannel plate (MCP) stack.

The measurements were performed using a 23 μm diameter glass nozzle with 0.7 mL/min flow rate within the laminar region (first 2 mm away from nozzle), after which droplets

start to form. The temperature varies along the jet from 26 °Cat ca. 1 mm distance from nozzle to lower values, when the resulting frozen residues are collected by a liquid nitrogen trap. Energy calibration was performed using N_2 gas (against the total electron yield for X-ray absorption spectroscopy on the 1s→2p π transition) with a resolution of ≤ 0.1 eV. Peak fitting and normalisation were performed using the Athena software,[12] with *arctan* steps for the IPs (edge steps) and Gaussian functions for the peaks.[12-14]

Total fluorescence yield (TFY) spectra were recorded simultaneously with a GaAsP photodiode mounted in the vicinity of the liquid jet; nevertheless, likely due to solvent condensation or solute crystallisation on the diode surface (thus interfering with the signal), these were of poor quality and not used for further analysis.

The concentration range that is accessible near saturation while still into the metastable regime was limited because of the formation of solids around the microjet nozzle, which led to jet instability and/or nozzle blockage. In the case of water, also ice can accumulate inside the main chamber, build up, grow and subsequently obstruct the nozzle.

6.2.3. Density functional theory (DFT) calculations

Gaussian 09[15] software was used for DFT[16] calculations on a geometry optimised non-ionic PABA monomer with the B3LYP functional and 6-31G* basis set to obtain the molecular orbital (MO)[17] energies for the ground state.

6.2.4. Prediction of speciation as a function of pH

The expected concentration of ionic and neutral species can be calculated as a function of pH from the acidity constants of PABA, K_{a1} and K_{a2}. Experimental techniques for the determination of acidity constants have recently been reviewed.[18] For the present study values reported in the previous literature were used.[19-21]

For a concentration of 1 mol/L, the sum of the partial concentrations of species (cationic, anionic, neutral) is equal to unity:

$$[cationic] + [neutral] + [anionic] = 1, \qquad \textit{Equation 6.1}$$

where the [cationic] is the partial molar concentration for cationic PABA species, [neutral] for the neutral PABA species, and [anionic] for the anionic PABA species.

The dissociation constants K_{a1} and K_{a2} are responsible for equilibria in solution[22,23] and can be expressed as:

$$K_{a1} = \frac{[neutral]\times[H^+]}{[cationic]}, \text{ and } K_{a2} = \frac{[anionic]\times[H^+]}{[neutral]} \qquad \textit{Equation 6.2}$$

Therefore, the neutral species concentration can be re-written as:

$$[neutral] = \frac{K_{a1}\times[cationic]}{[H^+]} \qquad \textit{Equation 6.3}$$

Likewise, the anionic species concentration can be expressed as:

$$[anionic] = \frac{K_{a2}\times[neutral]}{[H^+]}, \qquad \textit{Equation 6.4}$$

which, based on Equation 6.3, is equivalent to:

$$[anionic] = \frac{K_{a1}\times K_{a2}\times[cationic]}{[H^+]^2} \qquad \textit{Equation 6.5}$$

Knowing that:

$$pH = -\log[H^+],\ \ pK_a = -\log K_a\ , \qquad \textit{Equation 6.6}$$

Equations 6.3 and 6.5 become:

$$[neutral] = \frac{10^{-pK_{a1}}\times[cationic]}{10^{-pH}}, [anionic] = \frac{10^{-pK_{a1}-pK_{a2}}\times[cationic]}{10^{-2pH}} \qquad \textit{Equation 6.7}$$

Thus, going back to Equation 6.1,

$$[cationic] + \frac{10^{-pK_{a1}}\times[cationic]}{10^{-pH}} + \frac{10^{-pK_{a1}-pK_{a2}}\times[cationic]}{10^{-2pH}} = 1, \qquad \textit{Equation 6.8}$$

from which the partial concentration for the cationic species can be expressed as a function of pH:

$$[cationic] = \frac{1}{1+10^{pH-pK_{a1}}+10^{2pH-pK_{a1}-pK_{a2}}} \qquad \textit{Equation 6.9}$$

The partial concentrations of neutral and anionic species can then be written as:

$$[neutral] = \frac{10^{pH-pK_{a1}}}{1+10^{pH-pK_{a1}}+10^{2pH-pK_{a1}-pK_{a2}}}, \text{ and} \qquad \textit{Equation 6.10}$$

$$[anionic] = \frac{10^{2pH-pK_{a1}-pK_{a2}}}{1+10^{pH-pK_{a1}}+10^{2pH-pK_{a1}-pK_{a2}}}$$ *Equation 6.11*

The partial concentration of a species in the solution of known concentration 1 mol/L is equivalent to the mole fraction for that particular species, as shown for PABA in Figure 6.1.

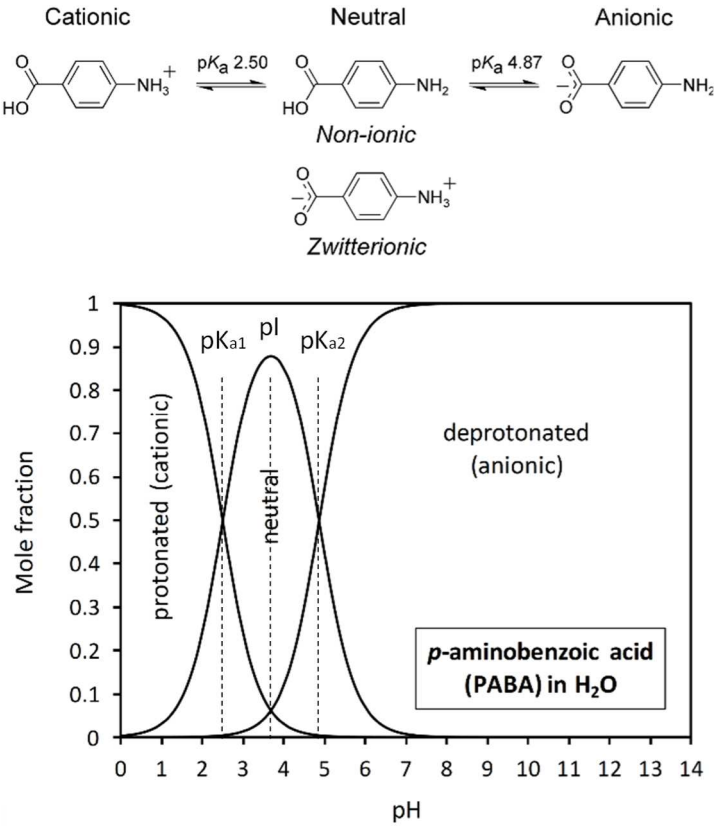


Figure 6.1. Equilibria between cationic, neutral and anionic PABA species in water (top) and speciation composition in mole fractions of PABA aqueous solutions as a function of pH (bottom); dashed lines mark the pK_a values at pH 2.50 and 4.87, and the isoelectric point pI 3.69.

6.3. Results and discussion

6.3.1. Speciation distribution diagram of PABA in aqueous solutions

In water, the neutral PABA molecule presents amphoteric character as it can either become cationic by accepting a proton at the $-NH_2$ group at low pH values (acidic environment), or anionic by donating a proton from the –COOH group at high pH values

(basic environment). Figure 6.1 shows the dependence of mole fraction (partial concentrations) as a function of pH below pH ~ 1 and above pH ~ 11 the speciation is almost entirely cationic and anionic, respectively. Of particular interest is the point where the concentrations of cationic and neutral species are equal, which corresponds to pK_{a1}. Likewise, concentrations of neutral and anionic species are equal at pK_{a2}. The concentration of the neutral species is at a maximum at the pH value corresponding to the isoelectric point (pI, half of the sum of pK_{a1} and pK_{a2}). Figure 6.1 highlights that in aqueous solutions of pH values near the isoelectric point a mixture of species is expected, consisting of mostly neutral but also a few % of cationic and anionic PABA (see also Table 6.1).

Table 6.1. Mole fractions (contributions) of PABA species in water at extreme and intermediate pH values.

pH	Cationic	Neutral	Anionic
1.00	0.9694	0.0306	/
2.20	0.6662	0.3331	0.0007
pK_{a1} (2.50)	0.4989	0.4989	0.0021
pI (3.69)	0.0576	0.8848	0.0576
pK_{a2} (4.87)	0.0021	0.4989	0.4989
11.00	/	/	1

6.3.2. Solution-state NEXAFS for PABA in water at intermediate pH values

The nature of the neutral present in aqueous solutions as a function of pH was examined through direct comparison and peak-fitting of spectra. In the absence of significant interactions between the solute molecules the NEXAFS spectrum of PABA mixtures is expected to be a linear combination of the spectra of the individual solute molecules. For example, the NEXAFS spectrum of an aqueous PABA solution at pH = 2.50 should be composed of spectra *s* from cationic, neutral and anionic components (Figure 6.1) as follows:

$$TS = (0.4989 \times s_{cat}) + (0.4989 \times s_{neut}) + (0.0021 \times s_{an}), \qquad \textit{Equation 6.12}$$

where TS represents the total (original) spectrum of the solution at pH = 2.50 and s_{cat}, s_{neut}, s_{an} are the spectra for cationic, neutral and anionic PABA, respectively.

As shown in Figure 5.2 in Chapter 5, the N K-edge spectrum of cationic species in solution differs strongly from that of the anionic species and of the non-ionic form in the solid state, as the pre-edge peaks arising from N 1s→π* transitions to 1π* and 3π* are absent due to protonation, which forms an N–H bond with the lone pair electrons at the nitrogen centre.

From Figure 6.1 and Table 6.1, the mole fraction of the cationic species (and consequently the contribution of the individual cationic spectrum to the total spectrum) should decrease when the pH is raised from 2.20 to 3.69 (pI). With progressive reduction of the cation concentration with increasing pH, the pre-edge peaks 1π* and 3π* become more prominent, and the 1σ* feature grows in intensity as the proportion of the neutral species increases (Figure 6.2).

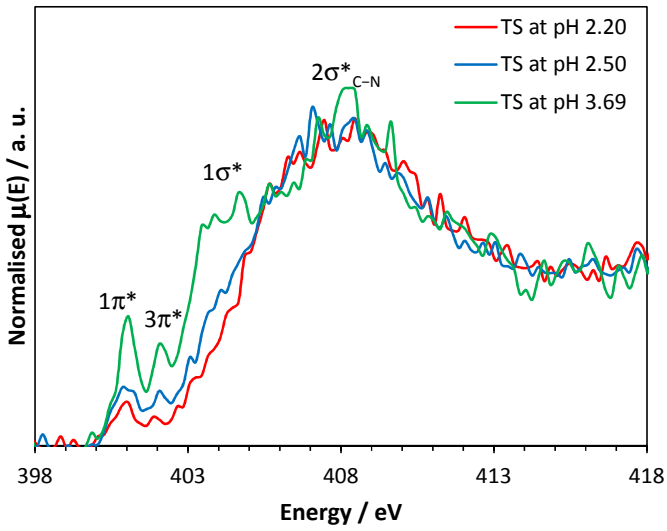


Figure 6.2. N K NEXAFS total spectra (TS) of PABA aqueous solutions at pH 2.20 (red), 2.50 (blue) and 3.69 (green).

When compared to the methanol solution, the spectrum of PABA at pI = 3.69 ('normal' aqueous solution) presents a similar overall shape, yet the intensities of the 1π* and 3π* peaks are reduced (Figure 6.3). This observation supports the rationale accounting for PABA being > 99% in its non-ionic form in methanol (Chapter 5), while previous work indicated that only ~ 89.5% in aqueous solution at pI should be in non-ionic form, while the remainder should be zwitterionic.[5] The zwitterionic form is not expected to present

π* resonances and its presence therefore appears to be in line with the slightly reduced intensities of the π* resonances relative to the methanolic solution.

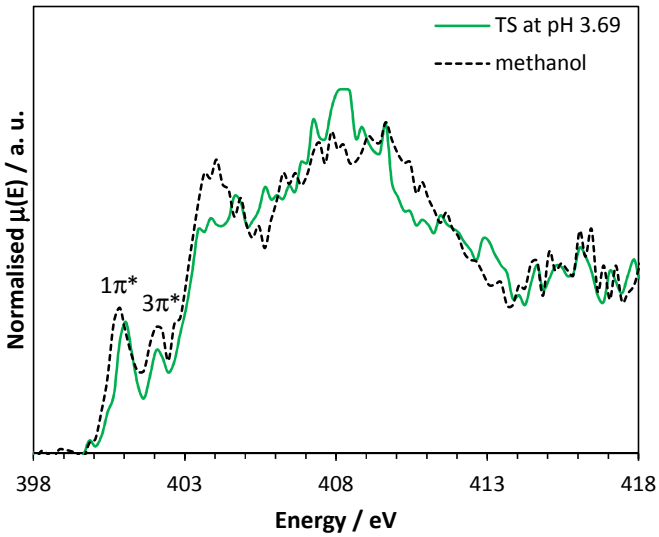


Figure 6.3. N K NEXAFS of PABA in water at pI 3.69 (green) and in methanol (dashed black line).

Raising the pH of the aqueous solutions above the isoelectric point to pH = 4.87, the pre-edge peaks arising from transitions to 1π* and 3π* become slightly broader, perhaps indicating the additional presence of some anionic species (Figure 6.4).

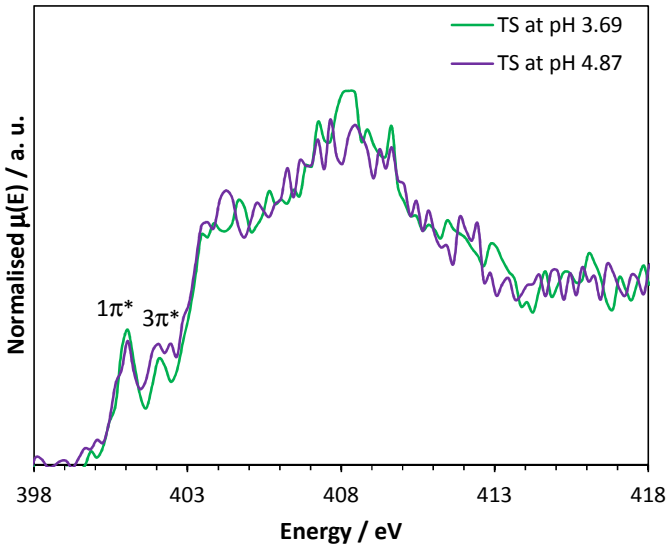


Figure 6.4. N K NEXAFS for PABA aqueous solutions at pH = 3.69 (green) and pH = 4.87 (purple).

Having recorded the spectra of cationic and anionic PABA (Figure 5.2, Chapter 5) it is now possible to obtain the spectra for neutral PABA species in solution at different pH by subtracting the corresponding contributions from s_{cat} and s_{an} from the original spectra:

i. For pH = 2.50:

$$s_{neut} = TS - (0.4989 \times s_{cat}) - (0.0021 \times s_{an}), \qquad \textit{Equation 6.13}$$

ii. For pH = 4.87:

$$s_{neut} = TS - (0.4989 \times s_{an}) - (0.0021 \times s_{cat}), \qquad \textit{Equation 6.14}$$

The predicted speciation distribution serves more as an orientational theoretical guide rather than an exact appreciation of the individual contributions from the PABA species. However, the relatively high level of noise resulted from the spectra substraction may alter the absolute energy positions and intensities of the shape resonances, although the overall shape resembling the one of the original spectrum of PABA in water at pI = 3.69 suggests the dominance of the non-ionic over the zwitterionic species.

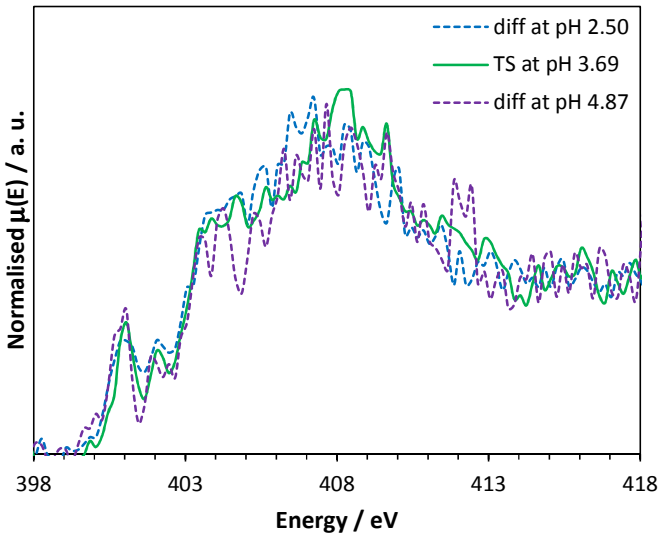


Figure 6.5. N K NEXAFS difference spectra of aqueous PABA at pH = 2.50 (blue dashed line), pH = 4.87 (purple dashed line), and the original spectrum at pH = pI = 3.69 (green line).

If the original spectra at pH = 2.50 (where the composition is half cationic and half neutral) and pH = 4.87 (half anionic and half neutral) are peak fitted keeping the same

ionisation potentials (IPs), resonance positions and full widths at half maximum (FWHM) as for the cationic and anionic standard spectra, the presence of the neutral species in each of these two mixtures alongside the cationic or anionic becomes evident (Table 6.2, Figure 6.6 and 6.7).

Table 6.2. Experimental N K NEXAFS resonances and IP energies for PABA in water at pH = 2.50 and pH = 4.87.

Nitrogen [eV]	pH = 2.50		pH = 4.87	
	Cationic	Neutral	Anionic	Neutral
IP	405.20	403.20	402.90	403.20
$1\pi^*$	/	400.95	401.15	400.95
$3\pi^*$	/	402.02	402.25	402.02
$1\sigma^*_{N-H}$	405.20	403.56	404.15	403.56
$2\sigma^*_{C-N}$	408.36	407.55	408.30	407.90

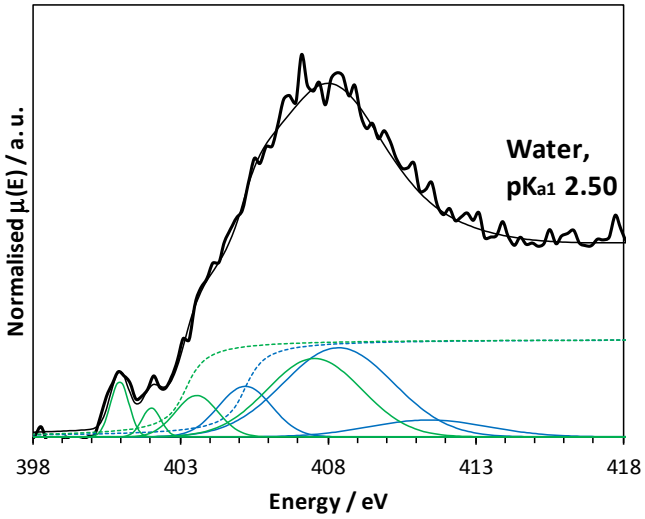


Figure 6.6. N K NEXAFS for aqueous PABA at pH = 2.50, together with the fits to the data and IPs marked with dashed lines. Contributions from cationic species are marked in blue, while the ones from neutral species in green.

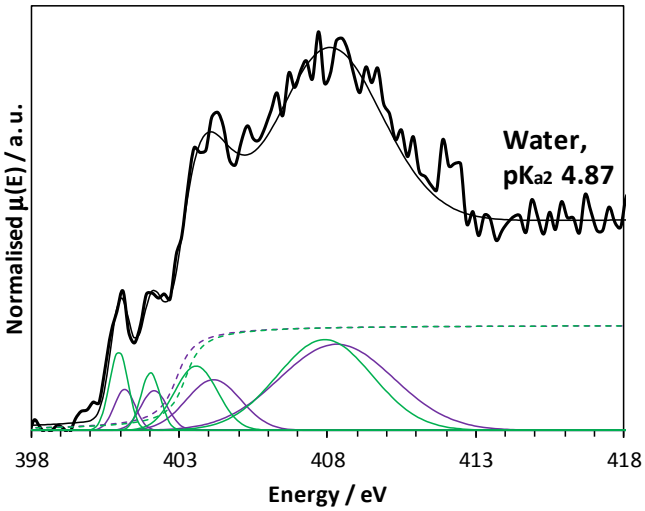


Figure 6.7. N K NEXAFS for aqueous PABA at pH = 4.87, together with the fits to the data, and IPs marked with dashed lines. Contributions from anionic species are marked in purple, while the ones from neutral species in green.

Specifically referring to the second σ* resonance (*i.e.* $2\sigma^*_{C-N}$) from the contributions of the neutral species indicated in Table 6.2, the distinct energy values 407.55 eV and 407.90 eV (for pH = 2.50 and pH = 4.87, respectively) may possibly be explained through the electrostatic influence of the charged component species (*i.e.* cationic and anionic, respectively) creating an electric field that perturbs the regular electronic structure of the other component from the mixture, neutral PABA.

6.3.3. Solution-state NEXAFS for PABA in alcohols as a function of concentration

Having established the solvent and pH sensitivity of NEXAFS in Chapter 5 and in the previous sub-section the possibility of detecting solute-solute and solute-solvent interactions as a function of concentration up to the saturation limit, and perhaps into the pre-nucleation supersaturated regime was examined. The influence of the type of alcohol on the NEXAFS spectrum was examined as well.

An initial comparison of the NEXAFS spectra of PABA solutions at different concentrations below saturation (0.83 M) in ethanol shows the close resemblance of the spectra (Figure 6.8). To support this observation, a systematic peak-fitting analysis

indicated a quasi-constant $\Delta\pi^*$ energy separation of the first two pre-edge peaks, $1\pi^*$ and $3\pi^*$ (Figure 6.9, Table 6.3).

As described in Chapters 4 and 5, the position of the broader, post-edge σ^* shape resonances with respect to the IP, named as the term value (δ_{C-N}), depends linearly on the bond length.[24-26] Therefore, the calibration curve presented in Figure 5.5 (Chapter 5) can serve as a tool to determine C–N bond lengths (BL_{C-N}) for PABA in solution, using the following linear dependence:

$$\delta_{C-N} = -39.664 \times BL_{C-N} + 61.133 \qquad \textit{Equation 6.15}$$

The results for PABA molecules in undersaturated solutions (0.30 M and 0.60 M) are quite close, with a slight shortening of the C–N bond intimated near-saturation (0.75 M), getting somewhat closer to that of the solid-state (Table 6.3).

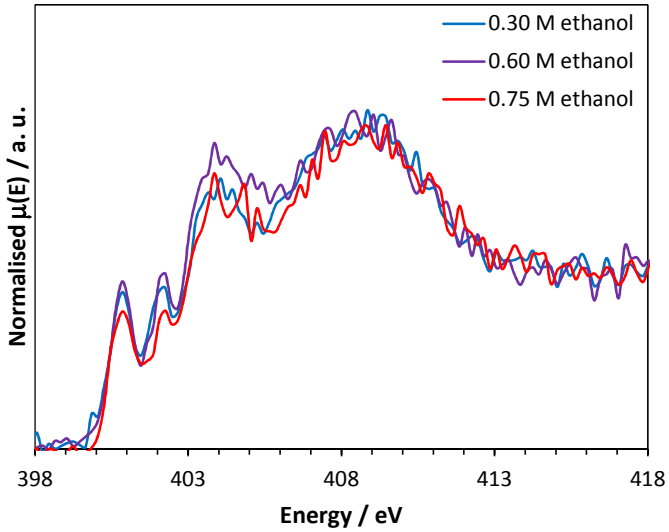


Figure 6.8. N K NEXAFS spectra of PABA in ethanol at various concentrations: 0.30 M (blue), 0.60 M (purple), and 0.75 M (red).

Table 6.3. Peak positions, and IPs at the N K-edge from Figure 6.9, along with the determined C–N bond lengths.

Nitrogen [eV]	0.30 M PABA ethanol	0.60 M PABA ethanol	0.75 M PABA ethanol	Non-ionic β-PABA
IP	403.20	403.20	403.20	403.60
$1\pi^*$	400.81	400.84	400.82	400.58
$3\pi^*$	402.04	402.06	402.10	402.30
$1\sigma^*_{N-H}$	403.55	403.60	403.62	405.11
$2\sigma^*_{C-N}$	408.50	408.55	408.68	409.09
$\Delta\pi^*$	1.23	1.22	1.28	1.72
δ_{C-N}	5.30	5.35	5.48	5.49
C–N	1.4076	1.4064	1.4031	1.4029

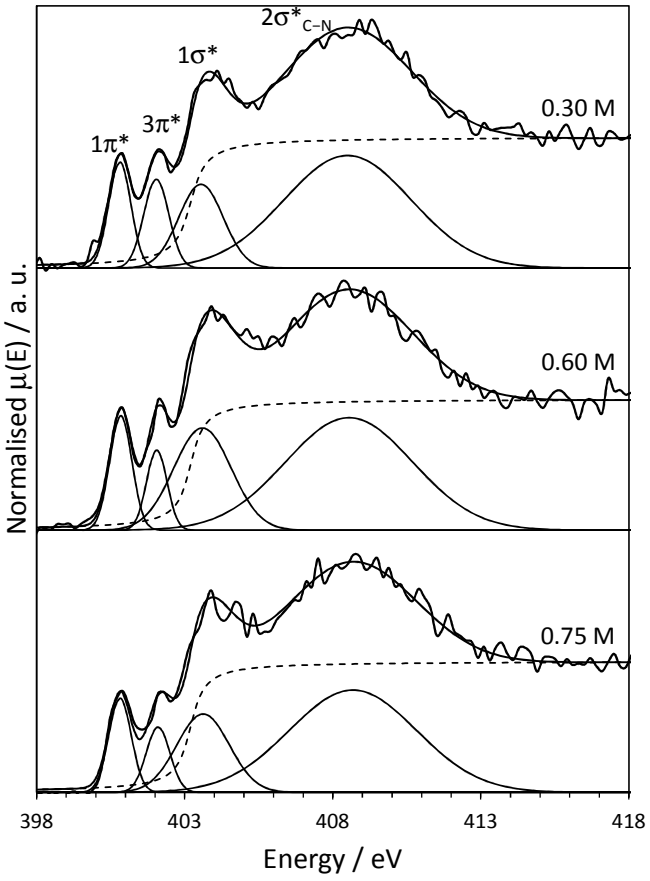


Figure 6.9. N K NEXAFS spectra for 0.30 M PABA in ethanol (top), 0.60 M PABA in ethanol (middle), 0.75 M PABA in ethanol (bottom), together with the fits to the data; IPs are marked with dashed lines.

Investigating the influence of the nature of the alcohol solvent on the type of intermolecular interactions in solution, no significant energy shifts are noticed for the overall shape of the spectra, apart from extra intensity around 401.5 eV for the 2-propanol solution (Figure 6.10). This feature appears to be too sharp for an additional resonance, so it is likely to be an artefact. Similar to PABA in methanol (Chapter 5), (at least the majority of) PABA in ethanol and 2-propanol is found in the neutral, non-ionic form[5] as the IP (~ 403.2 eV) and 1π* (~ 400.8 eV) resonance (that characterise the type of chemical species – Chapter 5) are closer to the non-ionic form rather than anionic species (Table 5.1 from Chapter 5). If the molecules were primarily zwitterionic, then the spectrum would resemble that of the cationic species (Figure 5.2 from Chapter 5), for which pre-edge π* resonances are absent and the IP is significantly different (405.2 eV).

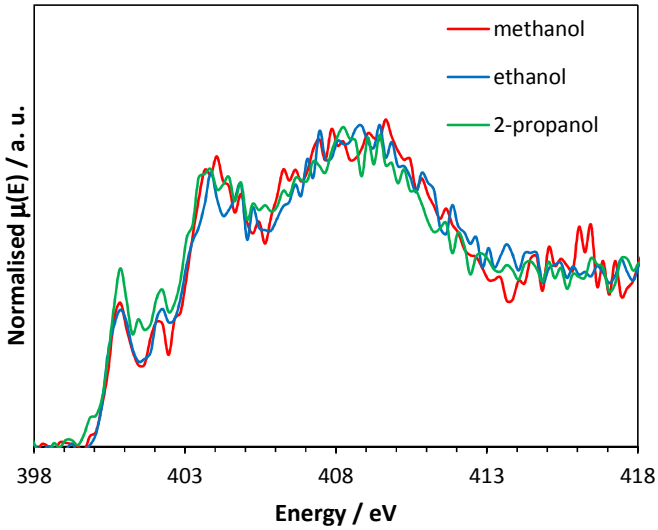


Figure 6.10. N K NEXAFS spectra of PABA in methanol (red), ethanol (blue), 2-propanol (green).

Focussing on the pre-edge spectral region, a direct comparison for solid-state PABA polymorphs (*i.e.* α and β) with PABA in alcohol solution (*e.g.* ethanol) reveals that the position of the N 1s→1π* peak is somewhat shifted to lower energy for α-PABA (Figure 6.11), reflecting variations in the local environment and type of –COOH interactions at the opposite end of the ring. As the spectrum of PABA in alcohol resembles that of the β polymorph, this suggests a similarity of the –COOH hydrogen bonding motifs for the

two systems, *i.e.* indication of solvation without PABA dimer formation in alcohols. Furthermore, a shift by ~ 0.2 eV of the 1π* peak to higher energy for PABA in water at pH = pI compared to PABA in ethanol indicates that in an aqueous environment the local electronic structure and molecular interactions are different from the situation in the alcohols. The position of the 1π* peak for PABA in water is closer to its equivalent in the spectrum of solid β-PABA than α-PABA (which contains hydrogen-bonded carboxylic acid dimers), further supporting the view that hydrogen-bonded carboxylic acid dimers are absent.

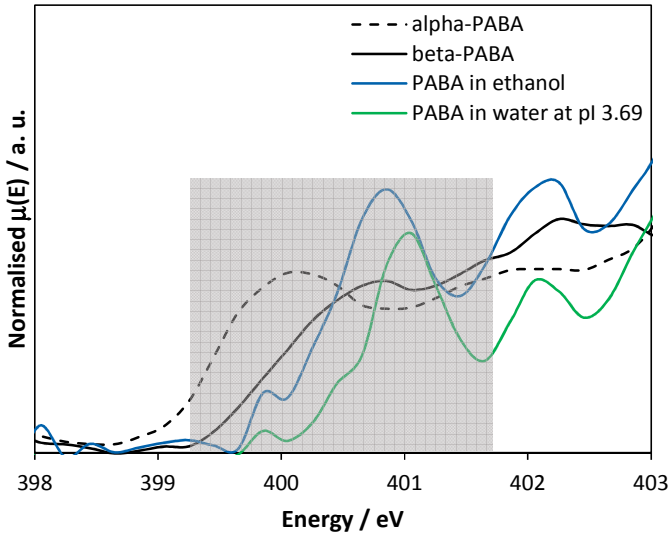


Figure 6.11. Pre-edge region of the N K NEXAFS spectra for α-PABA (black dashed line), β-PABA (black line), PABA in ethanol (blue line), and PABA in water at pI (green line) with the 1π* peak of interest highlighted by a grey area.

To correlate the 1π* and 3π* experimental resonances to the lowest unoccupied MOs (LUMOs) involved in the electronic transitions, DFT calculations for a ground-state non-ionic PABA monomer were carried out (Figure 6.12) using the same methodology as in Chapter 5. The LUMO+2 (= 3π*) orbital is evident through density of states (DOS) localised at the $–NH_2$ moiety, while for the LUMO orbital (= 1π*) the DOS is more spread out over the rest of the molecule. This indicates that the energy shift of the LUMO orbital, which is reflected in the position of the 1π* peak, would also depend on local

interactions/electronic changes occurring at the other end of the molecule such as the amino moiety, as shown in Figure 6.11.

As in ethanol solutions hydrogen bonds between the $–NH_2$ moiety of PABA and solvent/solute molecule broadly exist as N–H$\cdots$O and O–H$\cdots$N (possibly inducing more pronounced variations in the electron density at this moiety), the subtle shifts of the 3π* peak with concentration (Table 6.3) may therefore be influenced by variations in these interactions.

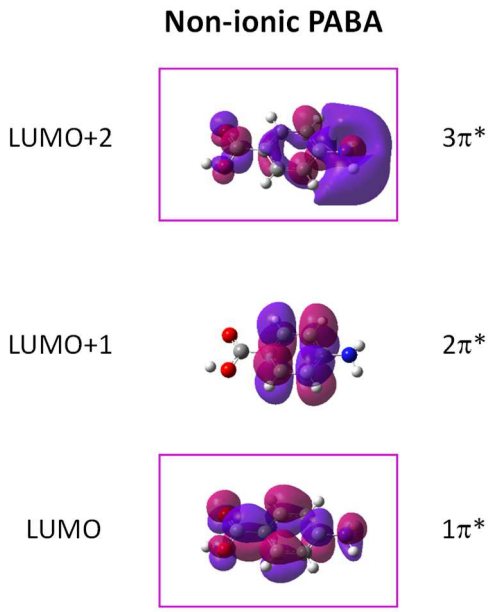


Figure 6.12. Unoccupied molecular orbitals for the non-ionic PABA monomer. The π* orbitals are labelled LUMO (1π*), LUMO+1 (2π*), LUMO+2 (3π*); transitions from the N 1s core level to these marked orbitals are visible in the N K-edge spectra.

6.4. Conclusions

Employing the liquid microjet technique at the N K-edge for aqueous solutions across pH values near the isoelectric point, and for different alcohol solutions provides insight into the dominant interactions characterising chemical speciation and for detecting electronic changes caused by local intermolecular interactions.

When raising the pH up to the isoelectric point pI the spectra show an intensity increase for the $1\pi^*$, $3\pi^*$ and $1\sigma^*$ peaks, which indicates that the neutral species' contribution grows progressively in the mixtures at the expense of cationic species. The spectrum at a pH value above the pI reveals slight broadening of the $1\pi^*$ and $3\pi^*$ peaks likely due to an increase in the contribution of anions. Linear combination analysis reveals that the spectrum of the neutral species at pH values with mixed speciation is virtually identical to the spectrum at pH = pI where the neutral form is dominant. Deconvolution by curve fitting using the standard spectra of the ionic species further confirms this.

Comparisons between ethanol solutions of varying concentrations up to near-saturation, and between alcohol solutions do not reveal significant differences, indicating that the local environment and interactions do not change significantly. However, a shift of the $1\pi^*$ peak is observed in the aqueous solution when compared to the alcohol solutions. DFT calculations support the experimental shifts of the $1\pi^*$ peak by means of the nature of interactions at the opposite end of the $-NH_2$ moiety, suggesting absence of dimers in alcohols and in water.

References

[1] M. N. Chatterjea, R. Shinde, *Textbook of Medical Biochemistry*, 8th edition, Jaypee Brothers Medical Publishers (P) Ltd, New Delhi, 2012; J. M. Berg, J. L. Tymoczko, L. Stryer, *Biochemistry*, 5th edition, Freeman, San Francisco, 2002.
[2] J. T. Carstensen, *Pharmaceutical Principles of Solid Dosage Forms*, Technomic Publishing Co., Lancaster, PA, 1993.
[3] R. J. Davey, S. L. M. Schroeder, J. H. ter Horst, *Angew. Chem., Int. Ed.*, 2013, **52,** 2166.
[4] B. Winter, M. Faubel, *Chem. Rev.*, 2006, **106**, 1176.
[5] B. van de Graaf, A. J. Hoefnagel, B. M. Wepster, *J. Org. Chem.*, 1981, **46**, 653.
[6] R. A. Sullivan, R. J. Davey, G. Sadiq, G. Dent, K. R. Back, J. H. ter Horst, D. Toroz, R. B. Hammond, *Cryst. Growth Des.*, 2014, **14**, 2689.
[7] J. S. Stevens, A. Gainar, E. Suljoti, J. Xiao, R. Golnak, E. F. Aziz, S. L. M. Schroeder, Chem. – Eur. J., 2015, **21**, 7256.
[8] K. Atak, S. I. Bokarev, M. Gotz, R. Golnak, K. M. Lange, N. Engel, M. Dantz, E. Suljoti, Oliver Kühn, E. F. Aziz. *J. Phys. Chem. B*, 2013, **117**, 12613.
[9] R. Golnak, K. Atak, E. Suljoti, K. F. Hodeck, K. M. Lange, M. A. Soldatov, N. Engel, E. F. Aziz, 2013, **15**, 8046.
[10] K. Atak, N. Engel, K. M. Lange, R. Golnak, M. Gotz, M. Soldatov, J.-E. Rubensson, N. Kosugi, E. F. Aziz, *Chem. Phys. Chem.*, 2012, **13**, 3106.
[11] K. M. Lange, R. Könnecke, S. Ghadimi, R. Golnak, M. A. Soldatov, K. F. Hodeck, A. Soldatov, E. F. Aziz, *Chem. Phys.*, 2010, **377**, 1.
[12] B. Ravel, *ATHENA User Guide*,
http://cars9.uchicago.edu/~ravel/software/doc/Athena/html/athena.pdf, 2009.
[13] J. Stöhr, *NEXAFS Spectroscopy*, 2nd edition, Springer-Verlag, Berlin, 1992.
[14] J. S. Stevens, C. R. Seabourne, C. Jaye, D. A. Fischer, A.J.Scott, S. L. M. Schroeder, *J. Phys. Chem. B*, 2014, **118**, 12121.
[15] *Gaussian 09, Revision B.01*, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, et al., 2009.
[16] B. Li, *Density-Functional Theory and Quantum Chemistry Studies on "dry" and "wet" NaCl(001)*, PhD thesis, Technical University of Berlin, Berlin, 2009.
[17] *GaussView, Version 5*, Roy Dennington, Todd Keith, and John Millam ed., R. Dennington, T. Keith, and J. Millam, Semichem Inc., Shawnee Mission, 2009.
[18] J. Reijenga, A. van Hoof, A. van Loon, B. Teunissen, *Anal. Chem. Insights*, 2013, **8**, 53.
[19] J. Jover, R. Bosque, J. Sales, *Comb. Sci.*, 2008, **27**, 563.
[20] R. D. Lide, *CRC Handbook of Chemistry and Physics*, 90th edition, CRC Press/Taylor and Francis, Boca Raton, 2010.
[21] B. Ghalami-Choobar, H. Dezhampanah, P. Nikparsa, A. Ghiami-Shomami, *J. Chem. Eng. Data*, 2012, **57**, 2802.
[22] L. Pauling, *The Nature of the Chemical Bond and the Structure of Molecules and Crystals; an Introduction to Modern Structural Chemistry*, 3rd edition, Cornell University Press, New York, 1960.
[23] P. Atkins, J. de Paula, *Physical Chemistry*, Oxford University Press, 8th edition, Oxford, 2006.
[24] F. Sette, J. Stöhr, A. P. Hitchcock, *J. Chem. Phys.*, 1984**, 81**, 4906.
[25] J. Stöhr, F. Sette, A. L. Johnson, *Phys. Rev. Lett.*, 1984, **53**, 1684.
[26] A. P. Hitchcock, C. E. Brion, *J. Phys. B*, 1981**, 14**, 4399.

CHAPTER 7

UV-VIS OF PABA IN WATER AND ALCOHOLS

7.1. Introduction

Fundamental understanding of the physical properties of organic materials is central to advancing the development and engineering of new materials and products.[1-4] Investigating the chemical and structural properties of organic solutes can provide details of their physicochemical properties determined by solvation and self-association. *Para*-aminobenzoic acid (PABA) has therapeutic effects[5] and sunscreen properties[6,7] due to its capability to absorb UV radiation determined by the electronic conjugation of the *para* orientated amino and carboxylic acid groups through the aromatic ring.[8,9]

The results in chapters 5 and 6 have shown how core-level spectroscopies such as near-edge X-ray absorption fine structure (NEXAFS) and resonant inelastic X-ray scattering (RIXS) give detailed insight into the electronic state of PABA species in solution and in the solid state, including the anionic and cationic forms in water.[10] As illustrated in Figure 7.1, NEXAFS involves promotion of electrons from 1s core orbitals to unoccupied molecular orbital (MO) levels (*e.g.* lowest unoccupied MO, LUMO), whilst RIXS probes transitions from occupied MOs (*e.g.* highest occupied MO, HOMO) to 1s core orbitals. UV-Vis spectroscopy is quite similar to these core-level techniques in that it probes electronic transitions, but from occupied MOs such as σ, n (non-participant) or π type orbitals to σ^* or π^* unoccupied MOs.[11,12,13] The combined measurement NEXAFS, RIXS and UV-Vis in conjunction with electronic structure calculations (particularly density functional theory, DFT, and time-dependent density functional theory, TD-DFT) therefore provides an opportunity to arrive at a deeper understanding of the electronic properties in the valence region, which can be difficult to elucidate based on UV-Vis spectroscopy alone. This chapter will reflect the effect of pH in PABA aqueous solutions and will show how laboratory UV-Vis and synchrotron NEXAFS-RIXS spectroscopies complement each other in providing access to information about the speciation of PABA, and opening up a methodology to interpret the laboratory UV-Vis results more deeply

and quantitatively. A solvent influence on the UV-Vis spectra of aqueous, methanolic and ethanolic PABA solutions will be demonstrated to stem from polarity differences between water and the two alcohols. Moreover, the analysis of the data will highlight that solute-solute aggregation in these systems is negligible.

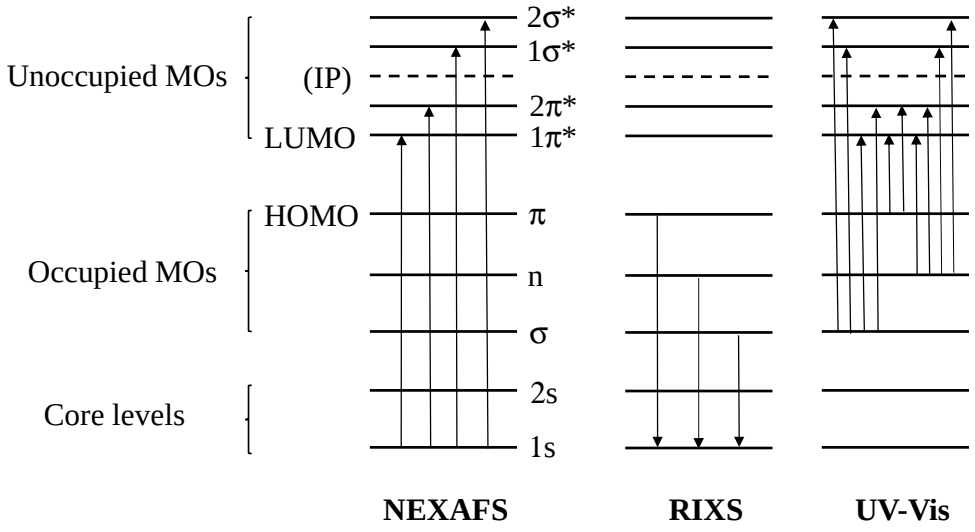


Figure 7.1. Possible electronic transitions (marked with vertical arrows) between orbital levels (represented by horizontal lines) in NEXAFS (left), RIXS (middle), and UV-Vis (right). The ionisation potential (IP) is marked with a horizontal dashed line.

7.2. Experimental section

7.2.1. Materials

PABA crystalline powder (99% purity) was purchased from Sigma-Aldrich (UK). For the preparation of solutions, laboratory distilled and deionised water, methanol (≥ 99.9%) from Fisher Scientific (UK), and ethanol (≥ 99.8%) from Fisher Scientific (UK) were used.

7.2.2. Transmission UV-Vis

For the aqueous systems, a saturated stock solution of 0.03 M PABA in water was prepared, and successively diluted to reach the detection range of the spectrophotometer. The high and low pH solutions were obtained by dropwise addition of 1.2 M aqueous

solution of NaOH (for high pH) or 1.2 M aqueous solution of HCl (for low pH), and measured using a benchtop S220 SevenCompact™ Mettler-Toledo pH-meter. The alcohol solutions (methanol and ethanol) for analysis were obtained by successive dilutions from 0.03 M stock solutions.

The UV-Vis spectra were measured with a 1 cm path length quartz cuvette inserted into a Beckman DU 520 spectrophotometer. The wavelength range was from 210 to 320 nm, with an accuracy of ±1 nm and a repeatability of ±0.1 nm. At wavelengths shorter than 210 nm the absorption of the solutions became too high to obtain reliable spectra, and these data were excluded from the plots.

7.2.3. Time-Dependent (TD) DFT calculations

Non-ionic, anionic and cationic PABA monomers were ground-state optimised with the B3LYP functional and 6-31G* basis set in Gaussian09[14] to obtain the DFT MOs.[15] Using these optimised monomers in a water continuum, TD-DFT calculations were then employed to generate the simulated UV-Vis spectra and the MOs involved in the transitions.

7.2.4. Solution-state *in situ* NEXAFS/RIXS

Solution-state N K-edge spectra were recorded with the LiXEdrom endstation[16] at the U41-PGM beamline of the BESSY II synchrotron at Helmholtz Zentrum Berlin (HZB), by using the liquid microjet technique. Partial fluorescence yield (PFY) data were recorded in scanning mode by a grating with line density 1200 lines/mm and radius 7.5 m dispersing the emitted photon energy from the sample,[16] and subsequently a detector consisting of a charge-coupled device (CCD), fluorescence screen and microchannel plate (MCP) stack collecting the amplified signal. The sample, grating and detector were arranged in Rowland circle geometry for accurate focussing. Although total fluorescence yield (TFY) can be recorded with a GaAsP photodiode mounted in the vicinity of the liquid jet, there were problems with the signal recording and solute could crystallise on the photodiode, interfering with the signal. N K RIXS data were recorded at multiple excitation energies corresponding to NEXAFS resonances. In line with this, similar spectral features at constant emission energy with varying excitation energy were

observed, and only changes in relative intensity occurred across the NEXAFS transition thresholds (excitation energies of 404.5, 410, 408 eV for the anionic, non-ionic and cationic species, respectively, were used for further analysis). An 18 μm diameter glass nozzle with 0.6 mL/min flow rate was used, and the measurement performed around 2 mm from the nozzle within the laminar part of the jet flow (droplets start to form after 3-5 mm, with the resulting frozen residues collected by a liquid nitrogen trap) with 10^{-5} mbar pressure in the main chamber. Beamline energy calibration was performed with N_2 gas (TEY XAS 1s→2p π* transition) and the resolution was ~ 0.1 eV.

7.3. Results and discussion

7.3.1. Effect of pH on the UV-Vis spectra of aqueous PABA solutions

This section presents changes in the UV-Vis absorption spectra induced by protonation/deprotonation of PABA in water, as well as evidence of almost absent solute self-associations.

At the isoelectric point, pH ~ 3.7, PABA is present mostly in its neutral, non-ionic form, with a small percentage of zwitterionic form,[17] which has no significant impact on the absorption spectra. One main peak is observed in the UV-Vis range at ~ 275 nm, and another peak at ~ 220 nm (Figure 7.2). The TD-DFT calculation of the UV-Vis spectrum of a non-ionic PABA monomer in a water continuum generates a spectrum (Figure 7.3) that is similar in shape to that observed experimentally, just overall shifted to shorter wavelengths (Figure 7.2).

An important aspect, probably responsible for the differences between the experimental and theoretical spectra, is that in the polarisable continuum model (PCM) calculations involving a self-consistent reaction field (SCRF) the solvent polarity creates an electric field that induces a surface charge density on the solute molecule so that electrostatic forces arise. The calculation does not fully take into account interactions such as van der Waals, hydrogen bonding or π–π stacking that are often encountered in solution.[18,19] Table 7.1 shows that the primary transition from HOMO to LUMO gives rise to the main absorption peak at ~ 267 nm from Figure 7.3.

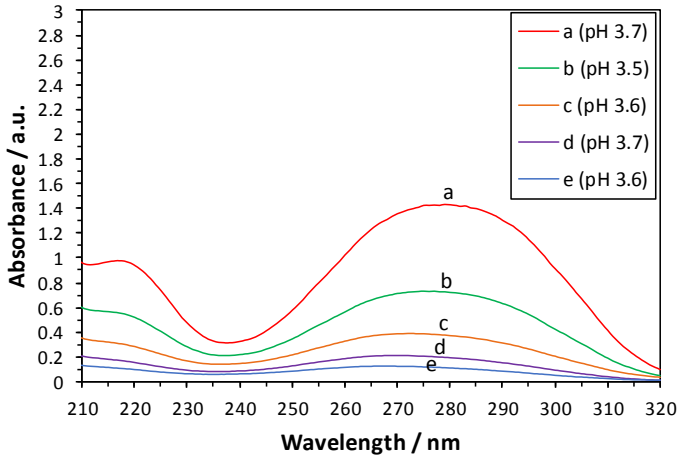


Figure 7.2. UV-Vis absorption spectra of neutral PABA in water at intermediate pH for dilute solutions: 6.39×10^{-5} M (red a), 3.19×10^{-5} M (green b), 1.60×10^{-5} M (orange c), 7.99×10^{-6} M (purple d), 3.99×10^{-6} M (blue e).

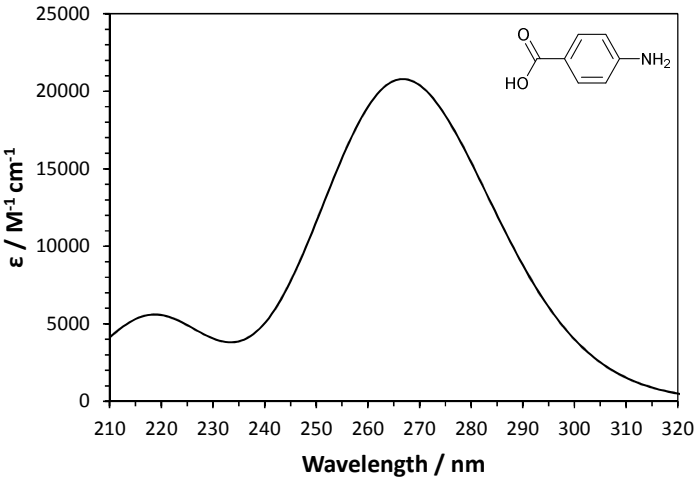


Figure 7.3. TD-DFT calculated spectrum of non-ionic PABA monomer in water.

Table 7.1. Comparison between main absorption peak wavelengths encountered in the experimental UV-Vis and TD-DFT simulated spectra of PABA species in water, along with their corresponding primary transitions.

PABA species	Wavelength (experimental) / nm	Wavelength (simulated) / nm	Primary transition
Non-ionic	~ 275	266.96	HOMO→LUMO
Anionic	~ 265	228.26	HOMO–1→LUMO
Cationic	~ 225	218.95	HOMO–1→LUMO

At high pH (at which PABA is almost entirely anionic – Figure 5.1, Chapter 5), there is a main peak in the UV-Vis spectrum at ~ 265 nm (Figure 7.4) and a rising absorption band below 210 nm. TD-DFT calculations for a monomeric PABA anion in a water continuum return a similar shape of the overall spectrum (Figure 7.5), yet the main absorption peak is shifted to a shorter wavelength at ~ 228 nm. This peak is attributed to the HOMO–1→LUMO primary transition as shown in Table 7.1.

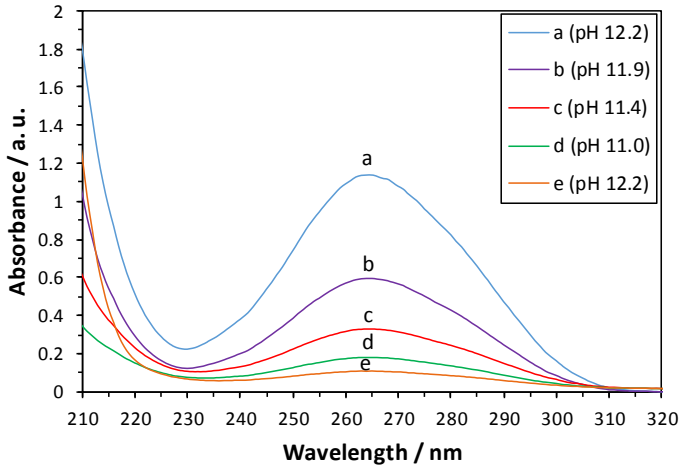


Figure 7.4. UV-Vis absorption spectra of anionic PABA in water at high pH for dilute solutions: 6.10×10^{-5} M (blue a), 3.05×10^{-5} M (purple b), 1.53×10^{-5} M (red c), 7.63×10^{-6} M (green d), 3.77×10^{-6} M (orange e).

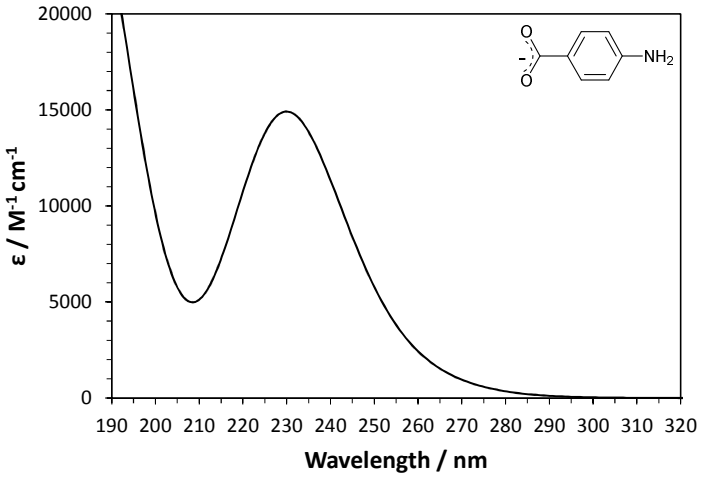


Figure 7.5. TD-DFT simulated spectrum of anionic PABA monomer in water.

At low pH values, PABA is almost entirely present as the cationic species (Figure 5.1, Chapter 5). The UV-Vis spectra presents one main peak at ~ 225 nm, and another weak absorption centred at ~ 275 nm (Figure 7.6). The TD-DFT calculated spectrum for monomeric cation in a water continuum is presented in Figure 7.7, where the main peak is slightly shifted to a shorter wavelength (~ 219 nm) than in the experimental UV-Vis spectrum (Figure 7.6). This primary transition peak is associated with the HOMO–1→LUMO transition (Table 7.1).

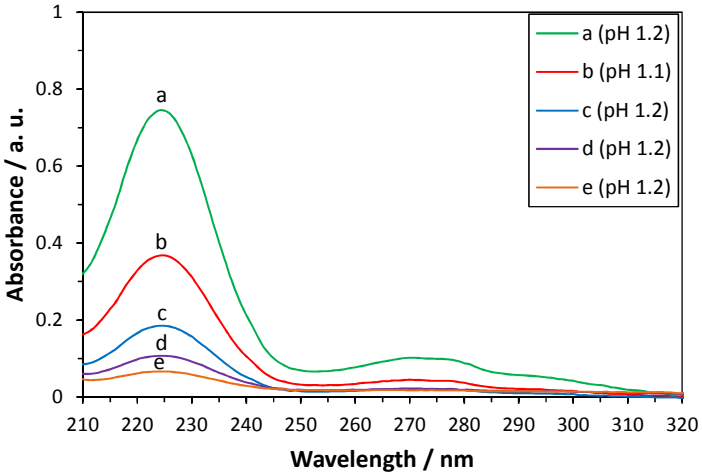


Figure 7.6. UV-Vis absorption spectra of cationic PABA in water at low pH for dilute solutions: 5.39×10^{-5} M (green a), 2.57×10^{-5} M (red b), 1.25×10^{-5} M (blue c), 6.12×10^{-6} M (purple d), 2.95×10^{-6} M (orange e).

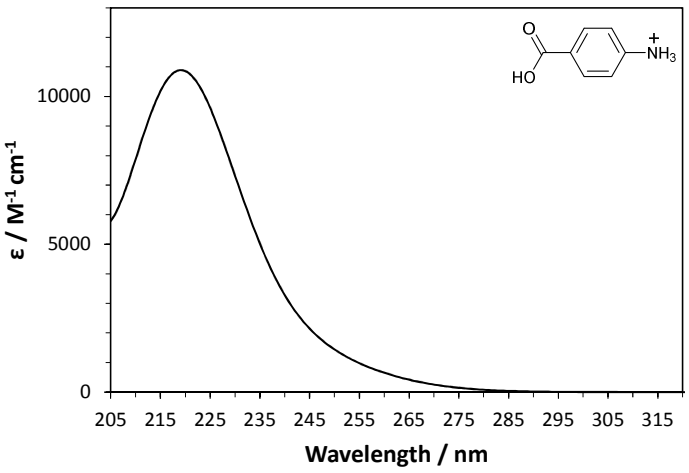


Figure 7.7. TD-DFT simulated spectrum of cationic PABA monomer in water.

In the absence of concentration-dependent self-association or strong variations in solvation structure the absorbance depends linearly on the concentration:[20]

$$A = \varepsilon c l, \qquad \textit{Equation 7.1}$$

where A is the absorption, ε is the molar absorptivity, c is the molarity, and l is the absorption path length through the sample. Plotting the absorbance values at the main peaks (~ 275 nm for neutral PABA, ~ 265 nm for anionic PABA, and ~ 225 nm for cationic PABA) as a function of concentration, linear dependencies are obtained in each case across the range of concentrations studied (Figure 7.8), which implies that speciation changes in solvation and self-associations are weak (*i.e.* monomeric forms are widely encountered) in aqueous systems.

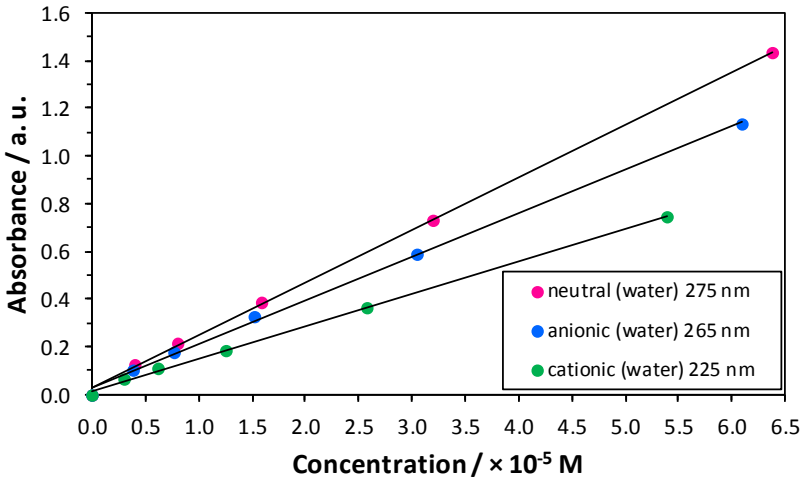


Figure 7.8. Linear dependence of main peak absorption values as a function of concentration for neutral (pink dots), anionic (blue dots), cationic PABA aqueous solutions (green dots).

7.3.2. Correlation between trends in UV-Vis, NEXAFS and RIXS for PABA in aqueous solutions

The main differences observed in the UV-Vis spectra of cationic, neutral and anionic PABA in aqueous solution are the wavelength shifts of the main absorption peaks in the order cationic < anionic < neutral. At low pH (cationic PABA) the absorption maximum is at ~ 225 nm, while the low intensity peak at ~ 275 nm (observable in the intermediate pH, neutral PABA spectrum) is highly attenuated. This aspect may be interpreted in terms of the forbidden n→π* transition assigned to the suppressed peak at ~ 275 nm, due to the effect of protonation at the amino moiety at low pH,[21] which engages the 'n' lone pair electrons of the nitrogen centre in a covalent bond (Chapter 5) rather than allowing them to be promoted to π* unoccupied levels.[10] In the case of high pH solutions (anionic PABA), the main absorption peak occurs at ~ 265 nm (compared to ~ 275 nm in the neutral PABA spectrum) due to the negatively charged carboxylate group that is less susceptible to π-conjugation with the aromatic ring (π electrons are less delocalised) than the carboxylic group.[22,23]

The TD-DFT calculations reproduce the observed trend, with the main absorption shifted to shorter wavelengths compared to non-ionic PABA (cationic < anionic < non-ionic, Table 7.1). Within this trend, it is noticeable that the calculated value for the anionic

species is significantly lower than the experimentally observed value. It has been suggested previously that calculated charge transfer excitations can be less accurate for ionic π systems due to underestimated excitation energies of ionic components.[24] Perhaps it also plays a role here the electron withdrawing nature of the carboxylate, and thus electrons are more consolidated around this group (*vide supra*).

The MOs generated from the TD-DFT calculations for the non-ionic, anionic and cationic monomers in water (Figure 7.9) reveal that the primary transition in neutral PABA (~ 267 nm, 4.64 eV, Figure 7.3) stems from a transition from orbital A (HOMO) to the 1π* (LUMO). The shift of the main absorption peak to shorter wavelength (*i.e.* larger energy gap) in anionic PABA (Figure 7.5) is due to widening energy separation between A and 1π* (~ 228 nm, 5.43 eV), with A lowering in energy and becoming HOMO–1. Similarly, the noted shift to lower wavelength for cationic PABA (~ 219 nm, 5.66 eV) can be explained by the primary transition between A, which is now the HOMO–1, and 1π*.

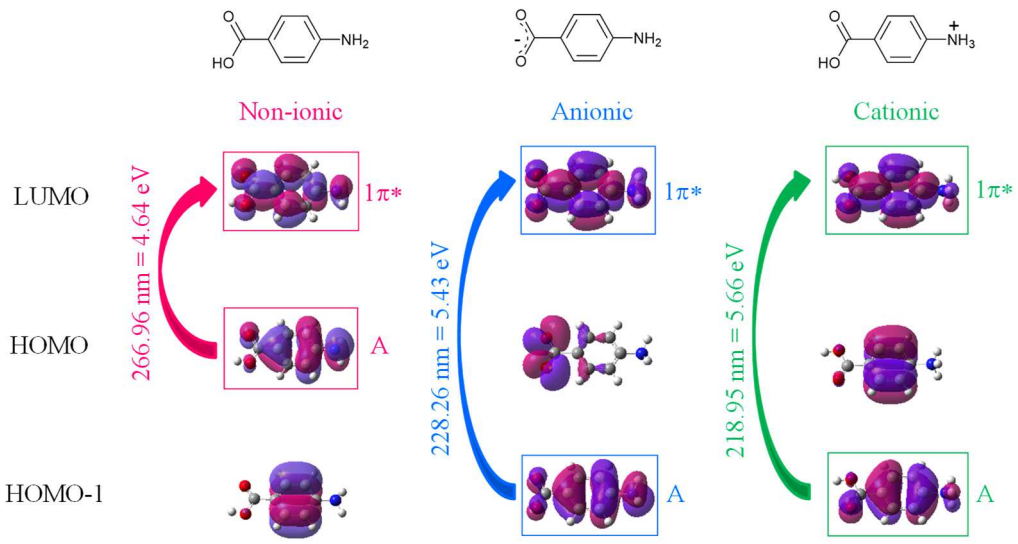


Figure 7.9. Transitions (arrows) between MOs of interest labelled as A and 1π* (marked with boxes) in the TD-DFT calculations for non-ionic (left), anionic (middle) and cationic (right) PABA monomers in a water continuum. From comparison with the experimental UV-Vis spectra, these transitions are the ones allowed for the simulated main absorption peaks.

These observations can now be correlated with the MO information for these systems that has been previously obtained by RIXS and NEXAFS (Chapter 5).[10] Figure 7.10 illustrates that the experimentally determined HOMO↔LUMO gap for nitrogen (which for the cationic PABA is not applicable due to the protonation of nitrogen) is correlated with the calculated energy gap between orbitals A and 1π* from the DFT calculations for the non-ionic and anionic species. In the case of non-ionic PABA, orbital A corresponds to HOMO and orbital 1π* to the LUMO, whilst for the anionic species, orbital A lowers in energy to become HOMO–3 and orbital 1π* raises up in energy as LUMO+1.

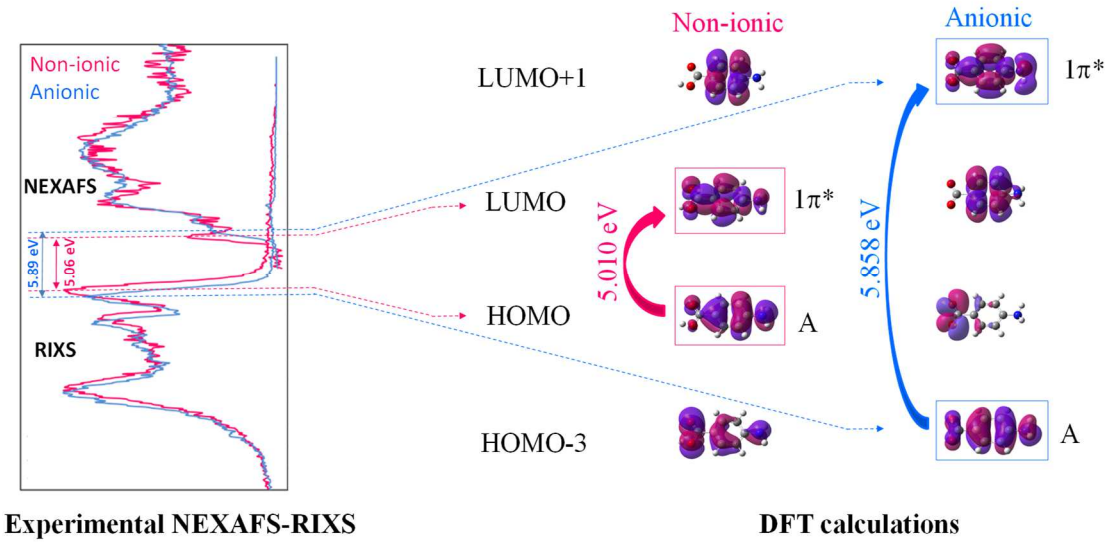


Figure 7.10. N K RIXS and NEXAFS showing the HOMO↔LUMO gap for non-ionic and anionic species, correlated with transitions (arrows) between DFT MOs of interest labelled as A and 1π*.

Although core hole relaxation effects are expected to affect the energetic position of features in the NEXAFS and RIXS data, the impact on taking the experimental RIXS-NEXAFS energy gap as a measure of the HOMO-LUMO is believed to be of the order of at most a few 100 meV; additionally, when such effects were included in CASTEP calculations for crystalline PABA[25] the inclusion of final-state effects did not change the MO interpretation significantly. For the DFT calculations, omitting relaxation effects may lead to a slight overestimation of transition energies, and thus of the RIXS-NEXAFS energy gap. The reason for the weak effects of relaxation are the strong localised core-

level excitations and the weak nature of the intermolecular interactions (*i.e.* hydrogen bonding) compared to intramolecular interactions (*i.e.* covalent bonds and protonation effects).

The differences noted in the A↔1π* gap presented in Table 7.2 between UV-Vis and RIXS-NEXAFS spectroscopies may be explained through the presence of a hole in the valence band in the UV-Vis that exerts a smaller effect over the final state than the core hole from RIXS-NEXAFS. The TD-DFT PCM calculations also involve a water continuum (thus PABA surface-polarised monomers with slightly different MO energy levels), yet this rather simplistic model is sufficient to picture the increase of the A↔1π* energy gap when moving from non-ionic to anionic PABA species.

Table 7.2. Values for the energy gap of interest A↔1π* reflected by the experimental UV-Vis, RIXS-NEXAFS, and by TD-DFT and DFT calculations.

Energy gap A↔1π*	Non-ionic PABA	Anionic PABA
UV-Vis (experimental)	~ 4.51 eV	~ 4.68 eV
TD-DFT (calculated)	4.64 eV	5.43 eV
RIXS-NEXAFS (experimental)	5.06 eV	5.89 eV
DFT (calculated)	5.01 eV	5.86 eV

7.3.3. Solvent effect in UV-Vis spectra of PABA solutions

This section aims to bring together the UV-Vis spectra of intermediate pH ~ 3.7 aqueous (where PABA is ~ 89.5% non-ionic), methanol and ethanol solutions (where PABA predominantly occurs as > 99% in its non-ionic form – Chapters 5 and 6).[13] Additionally, plots verifying the Beer-Lambert law[26,27] enhance the argument that PABA is mostly found as monomers in these polar solvents.

For PABA in methanol, one main peak shows up in the UV-Vis spectra at ~ 285 nm and another less intense peak at ~ 220 nm (Figure 7.11). The UV-Vis spectra of ethanol solutions present one main absorption peak at ~ 290 nm, and another peak at ~ 220 nm (Figure 7.12).

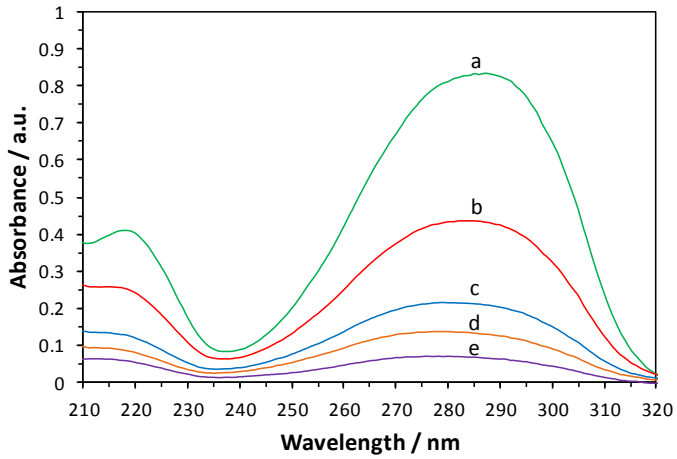


Figure 7.11. UV-Vis absorption spectra of PABA in methanol for dilute solutions: 3.19×10^{-5} M (green a), 1.60×10^{-5} M (red b), 7.99×10^{-6} M (blue c), 3.99×10^{-6} M (orange d), 2.00×10^{-6} M (purple e).

The UV-Vis spectra of ethanol solutions present one main absorption peak at ~ 290 nm, and another peak at ~ 220 nm (Figure 7.12).

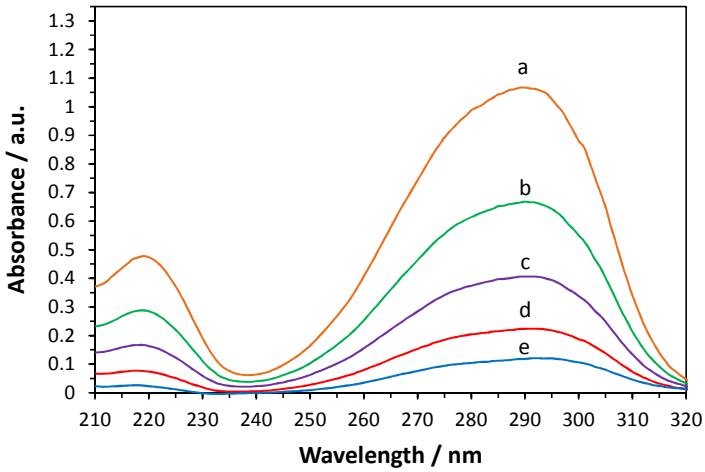


Figure 7.12. UV-Vis absorption spectra of PABA in ethanol for dilute solutions: 3.19×10^{-5} M (orange a), 1.60×10^{-5} M (green b), 7.99×10^{-6} M (purple c), 3.99×10^{-6} M (red d), 2.00×10^{-6} M (blue e).

The main absorption peaks for both methanol (~ 285 nm) and ethanol (~ 290 nm) are at slightly higher wavelength than for the neutral species in water (~ 275 nm). The increase

in solvent polarity from ethanol via methanol to water[28] is expected to cause a shift to shorter wavelengths known as a hypsochromic effect or negative solvatochromism[29] on the absorption peaks over the examined range 210-320 nm (Figure 7.13). In the UV-Vis theory, this characteristic is consistent with systems that possess substituents with non-bonding electron pairs ('n') on the aromatic ring, such as PABA presenting the lone pair electrons at the nitrogen atom; here, the solvent polarity lowers the energy of the non-bonding electrons and does not affect the final transition state π*.[22,23] The result is a larger energy gap (*i.e.* a shorter wavelength) of the n→π* absorption bands for the increasing polarity.

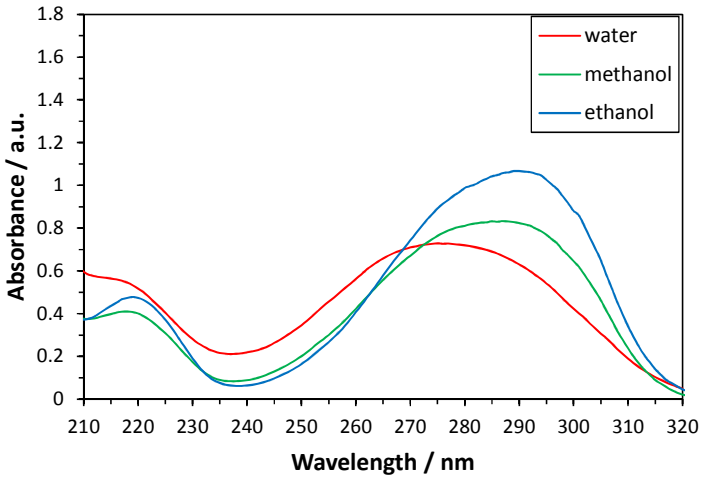


Figure 7.13. Hypsochromic effect observed for PABA in polar solvents: decreasing wavelength with increasing solvent polarity for ethanol (blue line), methanol (green line) and water (red line) for solutions of molarity 3.19×10^{-5} M.

The plot in Figure 7.14 summarises the absorption values of the main peaks (~ 285 nm for methanol, ~ 290 nm for ethanol) as a function of concentration. The linear dependence for the methanol solutions indicates fairly insignificant solute speciation changes due to solvation effects. For the ethanol system, a deviation from a linear dependence is observed, reflecting relatively stronger PABA self-association effects than for methanol or water. The explanation lies in the less polar character of ethanol compared to methanol and water, which determines PABA molecules to exhibit a higher degree of self-aggregation in ethanol.

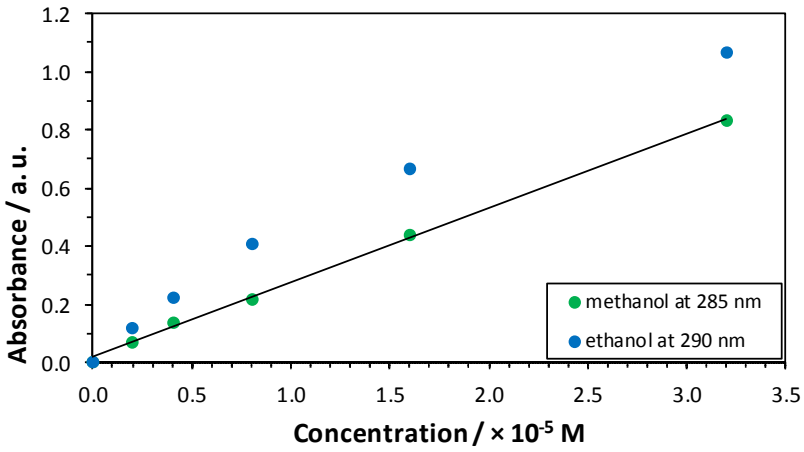


Figure 7.14. Dependence of absorbance as a function of concentration for PABA in methanol (green dots), and ethanol (blue dots).

7.4. Conclusions

Monitoring the electronic transitions of PABA in the UV range in aqueous solution at intermediate, high and low pH, and interpreting the results with TD-DFT calculations and the known NEXAFS and RIXS data provides a spectroscopic identification of PABA species that is complementary to the fine detail arising from the previously reported NEXAFS-RIXS combination (Chapter 5). The effect of pH can be translated in the UV-Vis by the shifts to shorter wavelengths in the main absorption peaks encountered in the spectra of high and low pH solutions with respect to the intermediate pH solutions, which are reflected by the TD-DFT simulations. These shifts are explained by A↔1π* energy gap alterations with the changes in electronic structure accompanying the pH variation; the DFT calculations associated with the experimental RIXS-NEXAFS identify well with the orbital gap observed in the TD-DFT simulations that reflects the main absorption peaks in the UV-Vis, thus bridging the synchrotron X-ray core-level spectroscopies NEXAFS and RIXS with the laboratory UV-Vis technique.

When solvent polarity is increased from ethanol to methanol, and then to water, a shift of the absorption peaks to shorter wavelengths (hypsochromic effect) in the experimental spectra is also noticed, arising from lowering of the energy of non-bonding electrons. When plotting the absorption values at the main peaks as a function of concentration, a linear dependence is obtained in the case of aqueous solutions of intermediate, high and low pH, and methanol solutions, with a major presence of PABA monomers in these systems (albeit less abundant in ethanol solutions).

References

[1] M. N. Chatterjea, R. Shinde, *Textbook of Medical Biochemistry*, 8th edition, Jaypee Brothers Medical Publishers (P) Ltd, New Delhi, 2012.
[2] J. M. Berg, J. L. Tymoczko, L. Stryer, *Biochemistry*, 5th edition, Freeman, San Francisco, 2002.
[3] J. T. Carstensen, *Pharmaceutical Principles of Solid Dosage Forms*, Technomic Publishing Co., Lancaster, PA, 1993.
[4] R. J. Davey, S. L. M. Schroeder, J. H. ter Horst, *Angew. Chem., Int. Ed.*, 2013, **52,** 2166.
[5] R. B. Failey, R. H. Childress, T. P. McConahay, *Am. J. Clin. Nutr.*, 1962, **10**, 158.
[6] D. R. Sambandan, D. Ratner, *J. Am. Acad. Dermatol.*, 2011, **64**, 748.
[7] C. Stenberg, T. Mellstrand, O. Larko, *Photodermatol*., 1987, **4**, 201.
[8] N. Serpone, A. Salinaro, A. V. Emeline, S. Horikoshi, H. Hidaka, J. Zhao, *Photochem. Photobiol. Sci.*, 2002, **1**, 970.
[9] H. Vainio, F. Bianchini, *Chemical and physical characteristics of sunscreen constituents*, IARC Handbooks of Cancer Prevention, Lyon, France, 2001.
[10] J. S. Stevens, A. Gainar, E. Suljoti, J. Xiao, R. Golnak, E. F. Aziz, S. L. M. Schroeder, *Chem. Eur. J.*, 2015, **21,** 7256.
[11] H. Wende, *Rep. Progr. Phys*., 2004, **67**, 2105
[12] L. J. P. Ament, M. van Veenendaal, T. P. Devereaux, J. P. Hill, J. van den Brink, *Rev. Mod. Phys.*, 2011, **83**, 705
[13] T. Owen, *Fundamentals of Modern UV-Visible Spectroscopy: A Primer*, Hewlett-Packard, Germany, 1996.
[14] *Gaussian 09, Revision B.01*, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, et al., 2009.
[15] *GaussView, Version 5*, R. Dennington, T. Keith, John Millam ed., R. Dennington, T. Keith, J. Millam, Semichem Inc., Shawnee Mission, KS, 2009.
[16] K. M. Lange, R. Könnecke, S. Ghadimi, R. Golnak, M. A. Soldatov, K. F. Hodeck, A. Soldatov, E. F. Aziz, *Chem. Phys.*, 2010, **377**, 1.
[17] B. van de Graaf, A. J. Hoefnagel, B. M. Wepster, *J. Org. Chem.*, 1981, **46**, 653.
[18] V. Barone, M. Cossi, J. Tomasi, *J. Chem. Phys.*, 1997, **107**, 3210.
[19] J. Tomasi, B. Mennucci, Roberto Cammi, *Chem. Rev*., 2005, **105**, 2999.
[20] P. Atkins, J. de Paula, *Physical Chemistry*, 10th edition, Oxford University Press, Oxford, 2014.
[21] F. Peral, E. Gallego, *Spectrochim. Acta Part A*, 2003, **59**, 1223.
[22] G. M. Lampman, D. L. Pavia, G. S. Kriz, J. R. Vyvyan, *Spectroscopy*, Brooks/Cole Cengage Learning, Belmont, 2001.
[23] D. Williams, I. Fleming, *Spectroscopic Methods in Organic Chemistry*, McGraw-Hill Education, Maidenhead, 2008.
[24] S. Grimme, M. Parac, *Chem. Phys. Chem.*, 2003, **3**, 292.
[25] J. S. Stevens, C. R. Seabourne, C. Jaye, D. A. Fischer, A. J. Scott, S. L. M. Schroeder, 2014, *J. Phys. Chem. B*, **118**, 12121.
[26] R. C. Denney, R. Sinclair, *Visible and Ultraviolet Spectroscopy*, John Wiley & Sons, London, 1987.
[27] H.-H. Perkampus, *UV-Vis Spectroscopy and Its Applications*, Springer-Verlag, Berlin, 1992.
[28] C. Reichardt, T. Welton, *Solvents and Solvent Effects in Organic Chemistry*, 3rd edition, Wiley-VCH Publishers, Weinheim, 2003.

[29] M. S. Masoud, A. E. Ali, M. A. Shaker, M. A. Ghani, *Spectrochim. Acta, Part A* 2004, **60**, 3155.

CHAPTER 8

CONCLUSIONS

The thesis presents the use of soft X-ray core-level (NEXAFS, RIXS) and UV-Vis spectroscopies applied to organic systems in both the solid and solution state, with the aim to understand their electronic structure changes determined by chemical structure, pH, concentration and nature of the solvent. Density functional theory (DFT), full-multiple scattering (FMS) and time-dependent density functional theory (TD-DFT) calculations enhance the value of the experimental data by providing an intuitive approach towards a careful interpretation of the electronic spectra.

Saccharides as solid state molecular materials were presented in Chapter 4 to outline the use of NEXAFS in detecting minute variations in their C–OH bond lengths, thus extending the applicability of the technique to complex organic molecules, including non-crystalline systems without long-range order such as amorphous phases. The sensitivity of the 'bond length with a ruler' approach is as good as that originally reported for gas-phase and adsorbed molecules. Some results indicate further analytical value in that the σ* shape resonance analysis may distinguish hemiketals from hemiacetals (*i.e.*, derived from ketoses and aldoses) as well as α from β forms of otherwise identical saccharides.

Turning now to the core-level spectroscopies of *para*-aminobenzoic acid (PABA) presented in Chapters 5 and 6, the 'bond length with a ruler' approach was this time extended to solution-state systems to reveal pH, solvent and concentration dependent spectral variations. The novelty represents the high precision in the determination of bond lengths for molecular species in solutions through comparison with NEXAFS data in the solid state, providing an avenue to examining computational predictions of solute properties and solute-solvent interactions. The pH changes induce local electronic structure and chemical state variations with formation of cationic and anionic species, compared to neutral (non-ionic) PABA. Thus, there are marked differences for the position of the ionisation potential (IP), along with an absence of π* resonances in NEXAFS and of π valence peaks in RIXS for the cationic species. A small shift of the IP to a lower energy and a wider HOMO↔LUMO gap is noticed for the anionic species and related to the DFT energy levels and MOs. Accounting now for the solvent influence on

the PABA molecule, the N K-edge spectra look similar in alcohol solutions (methanol, ethanol, 2-propanol), with no dramatic changes when concentration increases up to saturation. The species present in this case exist as neutral, non-ionic PABA, with no indication of zwitterionic form. The spectra of PABA in the alcohols resemble more that of β rather than the α-PABA, implying that the hydrogen bonding pattern is more similar to the former polymorph, which does not form –COOH dimers.[1,2] On the other hand, for the aqueous solutions the first peak (1π*) is higher in energy than for the alcohols, suggesting that the electronic structure and interactions are different in water. In addition, the sensitivity of 1π* to the interactions and local chemical state of the *para* –COOH group is reflected through energy shifts for the PABA polymorphs[3] and anionic PABA in water.[4] The second resonance peak (3π*) is more sensitive to the nature of interactions at the nitrogen moiety; for example, the additional hydrogen bonding involving the nitrogen in β-PABA increases the energy of 3π*.

Based on the NEXAFS-RIXS combination interpreted in Chapter 5 through DFT calculations, the UV-Vis spectra for PABA in aqueous solution (backed up with TD-DFT calculations) in Chapter 7 reflect main absorption peak shifts associated with changes in the electronic structure accompanying the pH variation. Additionally, a shift of the main absorption peak to shorter wavelengths (hypsochromic effect) with solvent polarity is observed. Studying the level of solute self-associations in solution through plots of absorption as a function of concentration, a fairly abundant presence of monomers in methanol and aqueous solutions is suggested, albeit this turns out to be reduced in ethanol.

With respect to future work, it would be of real interest to validate the characteristics of NEXAFS in differentiating between hemiketals and hemiacetals, or α and β forms from larger sets of saccharides. This step could have a significant impact in the food and pharmaceutical industries to identify distinct forms from mixtures of isomers.

On the other hand, for the PABA system a challenging aim is to obtain information from temperature-dependent core-level spectroscopic studies on the nucleation processes encountered in PABA solutions within the metastable zone, based on previously determined induction times for these systems.[5] In order to achieve this goal, the experimental setup at the LiXEdrom[6] endstation (where the *in situ* liquid microjet[7] experiments were conducted) could be improved through additional modifications that allow heating/cooling control of the stock solution to a set temperature for the metastable

zone/supersaturation level to be achieved. On the same note, three designs of small volume Conflat® cube cells were generated with the purpose to be integrated to the in-development I09 beamline at Diamond (Didcot, UK), each being compatible with a previous chamber developed for X-ray photoelectron spectroscopy (XPS) near-ambient pressures:

i. The prototype cell body for a windowed flow cell for NEXAFS fluorescence yield detection with liquid flow inlet/outlet and Si_3N_4 membrane windows was assembled and tested by running deionised water through the pipes with a peristaltic pump. The follow-up would be running PABA solutions of various concentrations to check whether the membrane can withstand the pressure.
ii. A microjet cell for windowless studies of liquids by electron yield and fluorescence yield detection of NEXAFS (similar to the microjet setup) is currently in the design phase and will be the technologically most advanced and ultimately the most versatile.
iii. An environmental NEXAFS cell was designed for studies of solid samples under controlled solvent vapour pressures by electron and fluorescence yield detection. This device is composed of a separate manipulator that can be fitted inside the cubic chamber. It will allow experiments on powders and single crystal facets under control of sample temperature (including cooling) and gas environment. The main use of this system will facilitate specific studies of surfaces. Polar angle control will allow studies of the single crystal surfaces as a function of X-ray polarization vector, and hence the determination of the orientation of adsorbed molecules.

Additionally, another set of beamtime experiments could be focussed on N RIXS for PABA in alcohols (methanol, ethanol, 2-propanol) and water at intermediate pH values to account for the determination of the HOMO↔LUMO gap through combination with the previously analysed N K NEXAFS. The very next step would be to identify the main absorption peak position in the UV-Vis spectra with this HOMO↔LUMO gap and compare the results with PABA aqueous solutions at low and high pH.

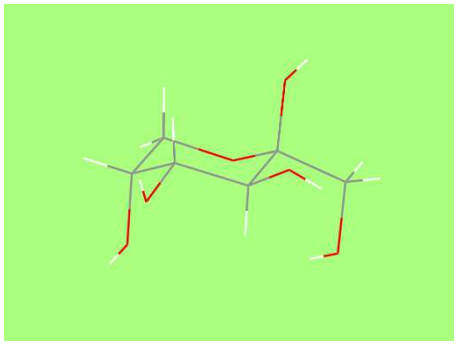
On a different note, comparative interpretation of systems that form dimers in non-polar solvents (*e.g.* benzoic acid in toluene) would serve as guidance to investigate how changes in the local hydrogen bonding pattern influences the electronic structure of the solute.

Following the same approach, other systems representing derivatives of benzoic acid can be studied by NEXAFS and complementary spectroscopy techniques such as RIXS and UV-Vis; it will be of potentially high interest to observe how substitution with various functional groups (*e.g.* $–NO_2$, –F, –Cl, –Br, –I, $–CONH_2$) at different positions on the benzene ring impacts on the local electronic structure and solute-solute/solute-solvent interactions. An alternative would be to change the relative positions of the substituents on the aromatic ring. These aspects would allow dissemination of the self-association mechanisms through systematic comparison of structurally related molecules.

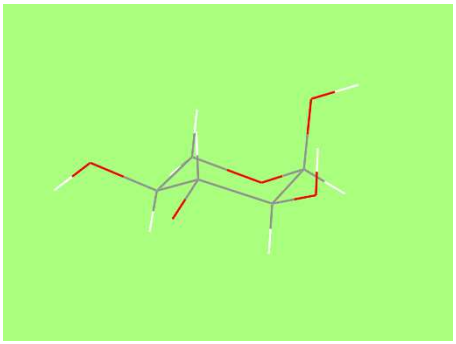
References

[1] T. F. Lai, R. E. Marsh, *Acta Cryst.*, 1967, **22**, 885.
[2] S. Gracin, A. Fischer, *Acta Cryst. Sect. E,* 2005, **61**, 1242.
[3] J. S. Stevens, C. R. Seabourne, C. Jaye, D. A. Fischer, A. J. Scott, S. L. M. Schroeder, *J. Phys. Chem. B*, 2014, **118**, 12121.
[4] J. S. Stevens, A. Gainar, E. Suljoti, J. Xiao, R. Golnak, E. F. Aziz, S. L. M. Schroeder, *Chem. Eur. J.*, 2015, **21**, 7256.
[5] R. A. Sullivan, R. J. Davey, G. Sadiq, G. Dent, K. R. Back, J. H. ter Horst, D. Toroz, R. B. Hammond, *Cryst. Growth Des.*, 2014, **14**, 2689.
[6] K. M. Lange, R. Könnecke, S. Ghadimi, R. Golnak, M. A. Soldatov, K. F. Hodeck, A. Soldatov, E. F. Aziz, *Chem. Phys.*, 2010, **377**, 1.
[7] K. R. Wilson, B. S. Rude, J. Smith, C. Cappa, D. T. Co, R. D. Schaller, M. Larsson, T. Catalano, R. J. Saykally, *Rev. Sci. Instrum.*, 2004, **75**, 725.

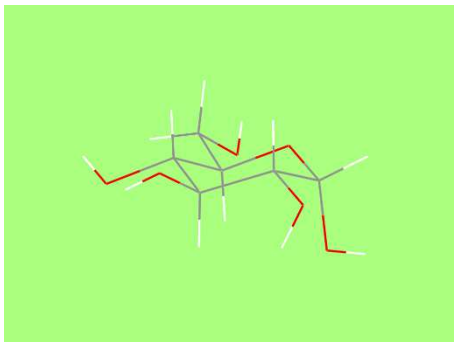
APPENDIX A. Crystal structures of saccharides, PABA polymorphs and PABA salts



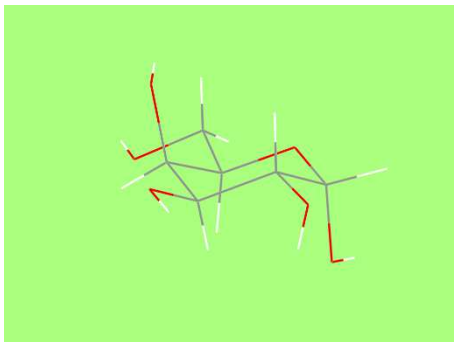
FRUCTO11 (β-Fructose)



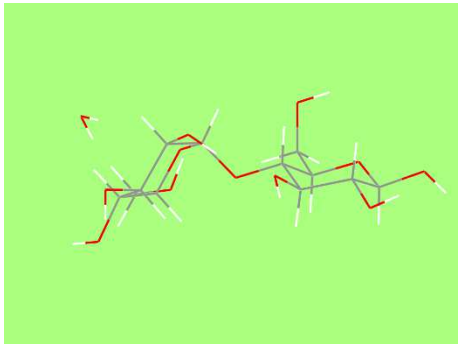
XYLOSE (α-Xylose)



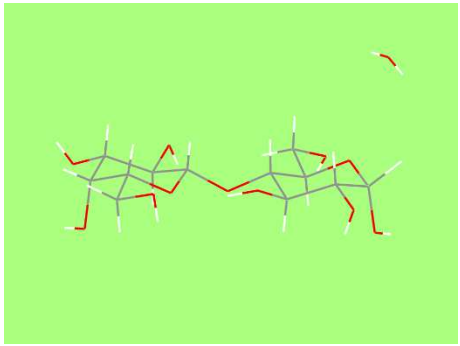
GLUCSA03 (α-Glucose)



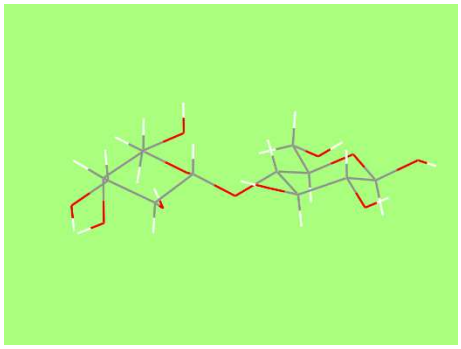
ADGALA01 (α-Galactose)



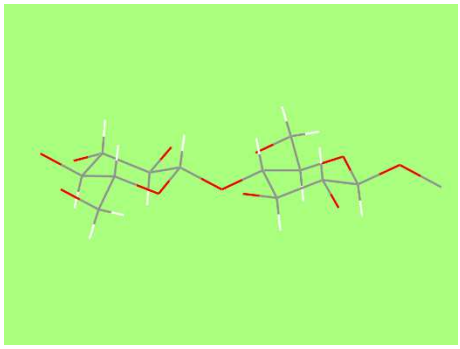
MALTOS11 (β-Maltose monohydrate)



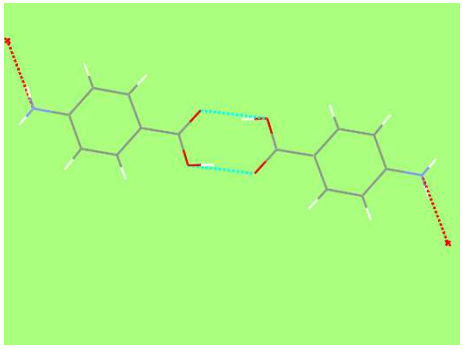
LACTOS03 (α-Lactose monohydrate)



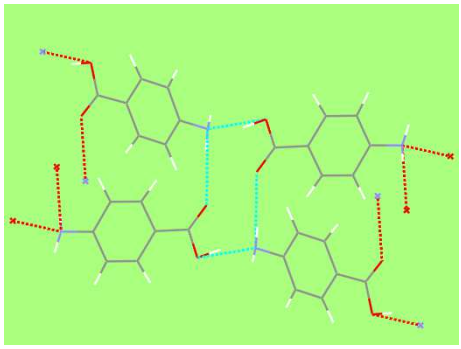
BLACTO02 (β-Lactose)



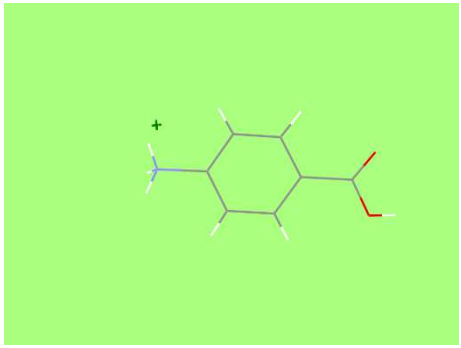
PADTUL (Cellulose)



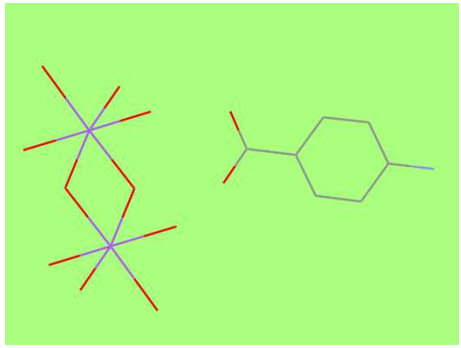
AMBNAC01 (non-ionic α-PABA)



AMBNAC04 (non-ionic β-PABA)



PAMBZA (cationic HCl salt of PABA)



NAABZH (anionic NaOH salt of PABA)

APPENDIX B. X-ray powder diffraction of saccharides and FEFF8 simulations on the crystal structure of glucose

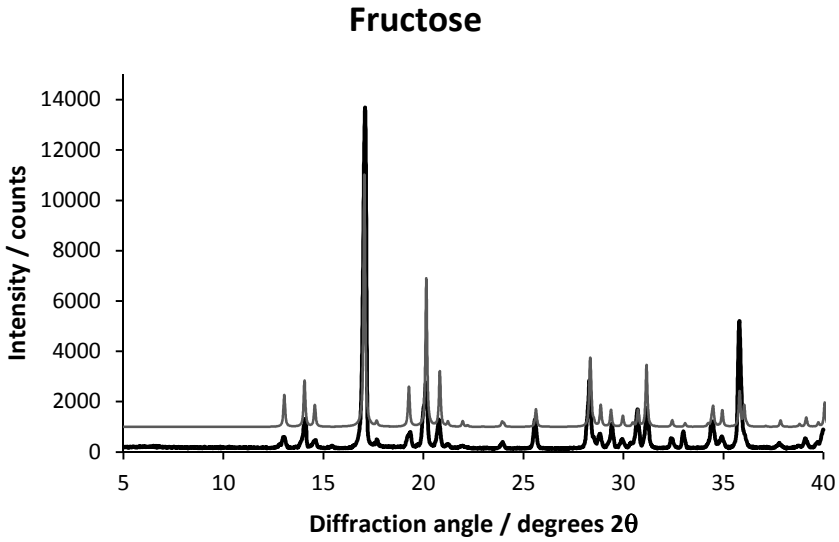


Figure B1. X-ray powder diffraction of experimental β-fructose (bold black), and calculated β-fructose (dark grey, Refcode FRUCTO11) patterns.

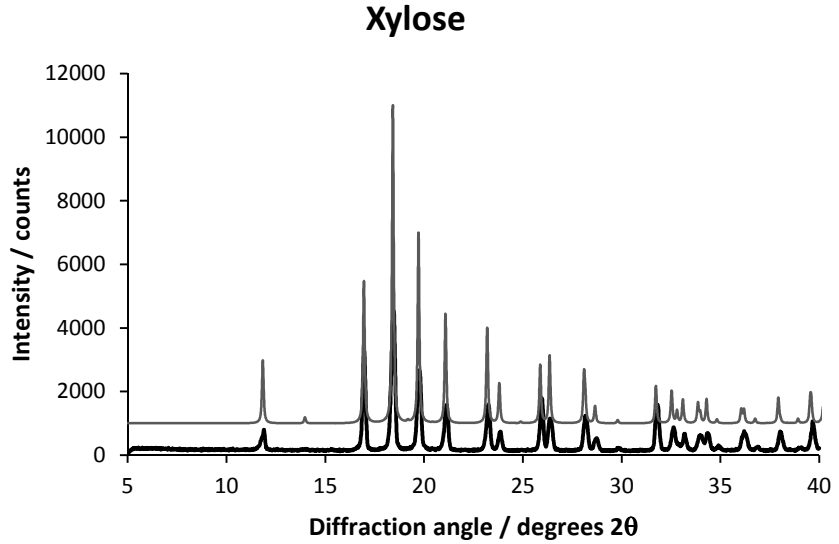


Figure B2. X-ray powder diffraction of experimental α-xylose (bold black), and calculated α-xylose (dark grey, Refcode XYLOSE) patterns.

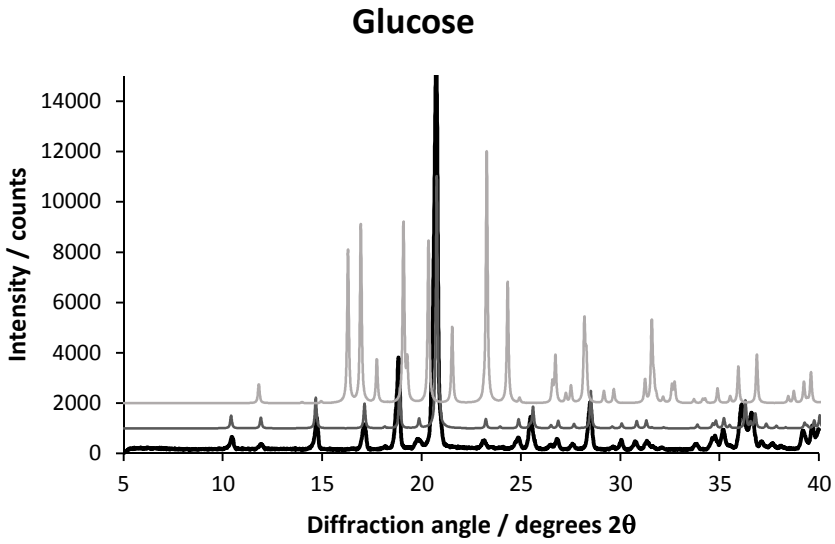


Figure B3. X-ray powder diffraction of experimental α-glucose (bold black), calculated α-glucose (dark grey, Refcode GLUCSA03), and calculated β-glucose (light grey, Refcode GLUCSE) patterns.

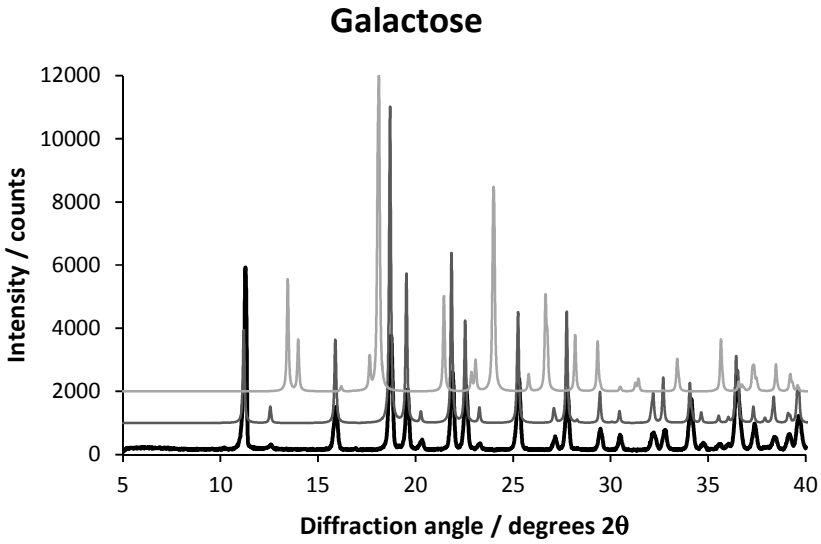


Figure B4. X-ray powder diffraction of experimental α-galactose (bold black), calculated α-galactose (dark grey, Refcode ADGALA01), and calculated β-galactose (light grey, Refcode BDGLOS01) patterns.

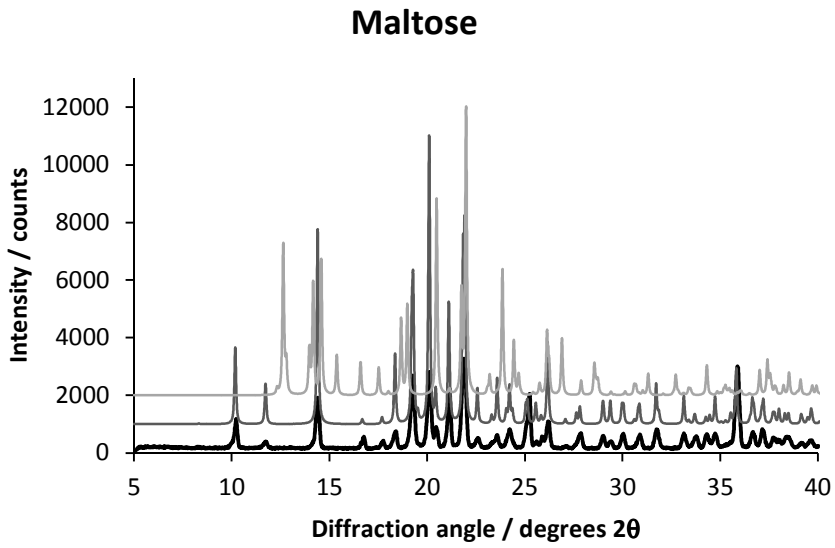


Figure B5. X-ray powder diffraction of experimental β-maltose monohydrate (bold black), calculated β-maltose monohydrate (dark grey, Refcode MALTOS11), and calculated α-maltose (light grey, Refcode MALTOT) patterns.

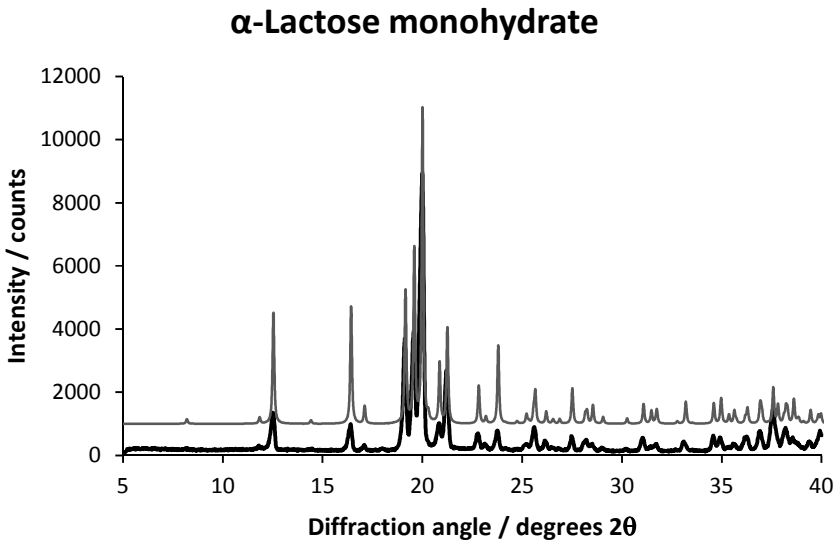


Figure B6. X-ray powder diffraction of experimental α-lactose monohydrate (bold black), and calculated α-lactose monohydrate (dark grey, Refcode LACTOS03) patterns.

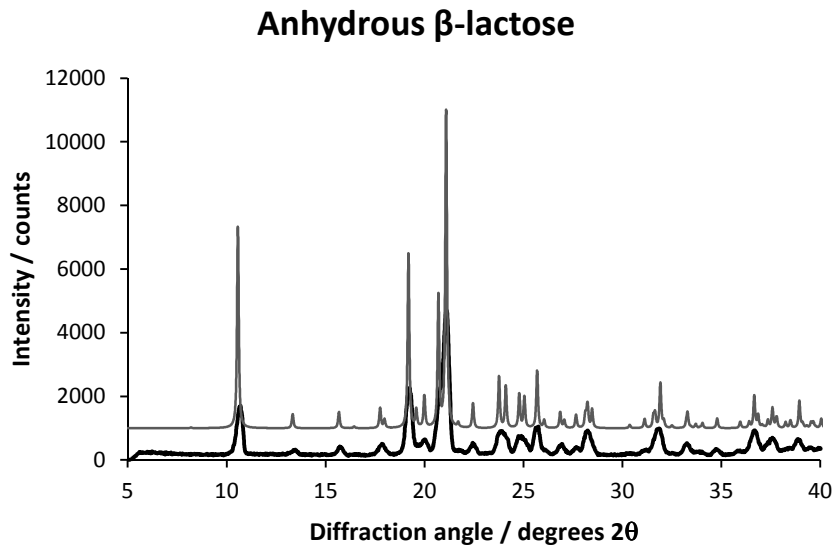


Figure B7. X-ray powder diffraction of experimental anhydrous β-lactose (bold black), and calculated anhydrous β-lactose (dark grey, Refcode BLACTO02) patterns.

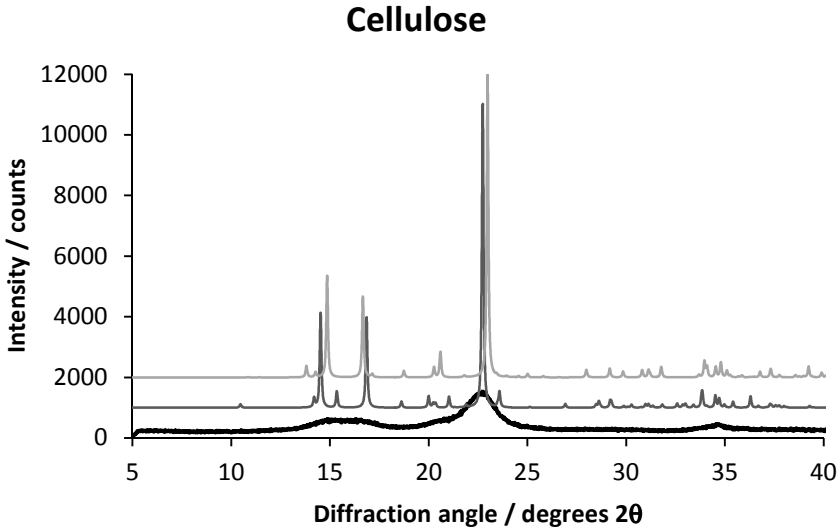


Figure B8. X-ray powder diffraction of experimental cellulose (bold black), calculated cellulose form I_α (dark grey, Refcode PADTUL), and calculated cellulose form I_β (light grey, Refcode JINROO01) patterns.

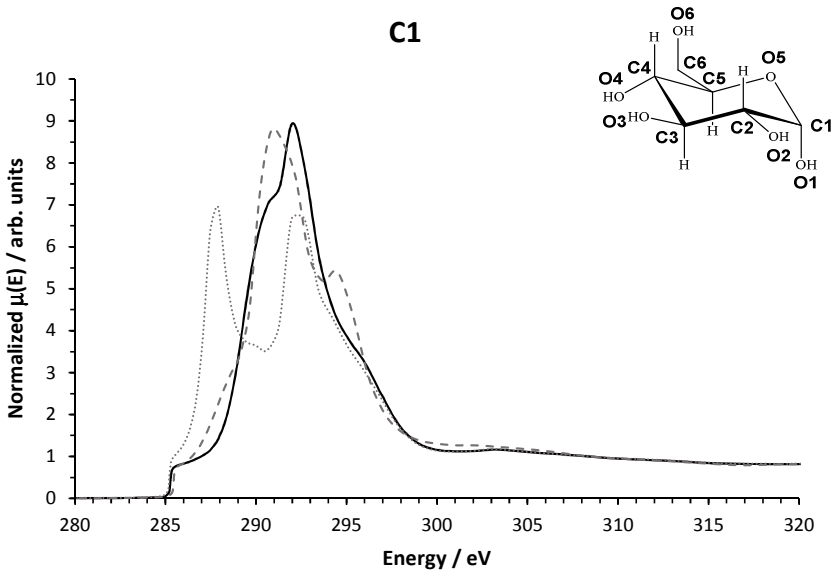


Figure B9. FEFF8 simulations for C1 absorbing atom in: an unaltered crystal structure of glucose (black line), an altered crystal structure of glucose with C1–O5 bond shortened by 0.1 Å (dashed grey line), and an altered crystal structure of glucose with C1–O5 bond elongated by 0.1 Å (dotted grey line); once C1–O5 bond altered, expected variations within the cycle on the C1–C2 and C5–O5 bond lengths are expected, albeit negligible.

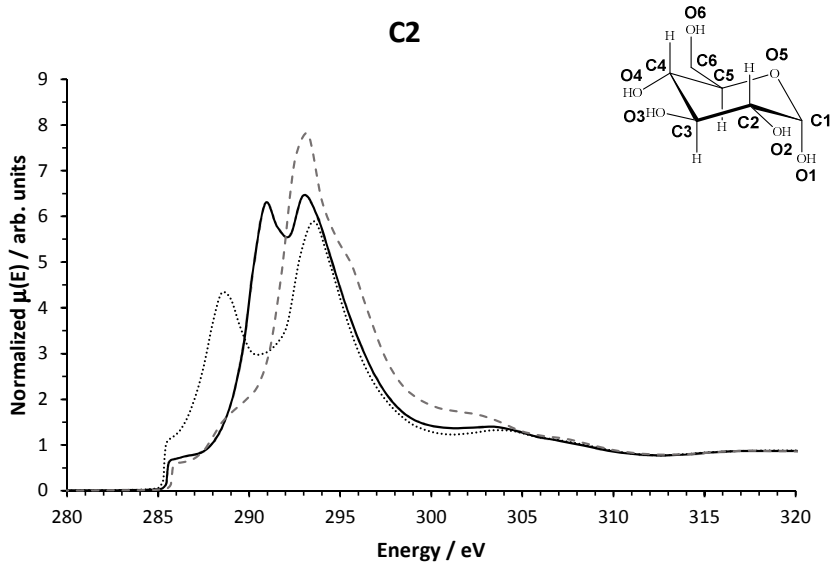


Figure B10. FEFF8 simulations for C2 absorbing atom in: an unaltered crystal structure of glucose (black line), an altered crystal structure of glucose with C2–O2 bond shortened by 0.1 Å (dashed grey line), and an altered crystal structure of glucose with C2–O2 bond elongated by 0.1 Å (dotted grey line).

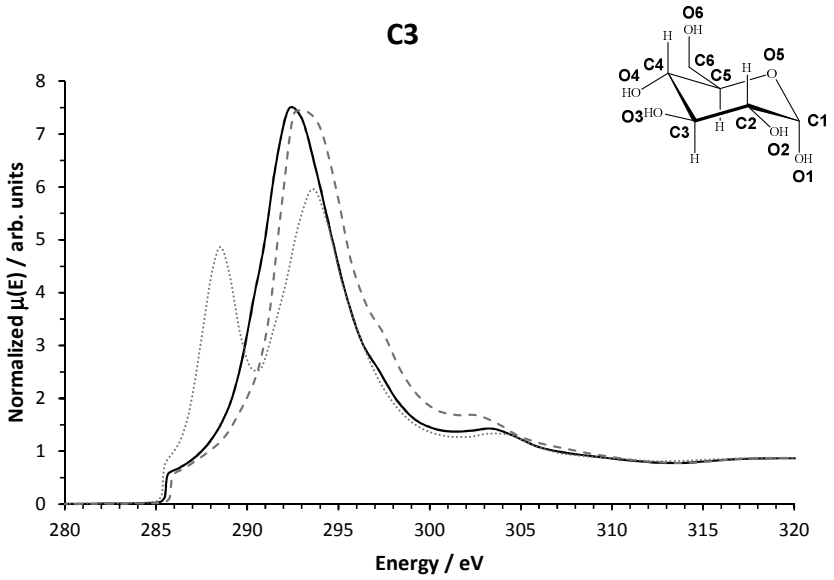


Figure B11. FEFF8 simulations for C3 absorbing atom in: an unaltered crystal structure of glucose (black line), an altered crystal structure of glucose with C3–O3 bond shortened by 0.1 Å (dashed grey line), and an altered crystal structure of glucose with C3–O3 bond elongated by 0.1 Å (dotted grey line).

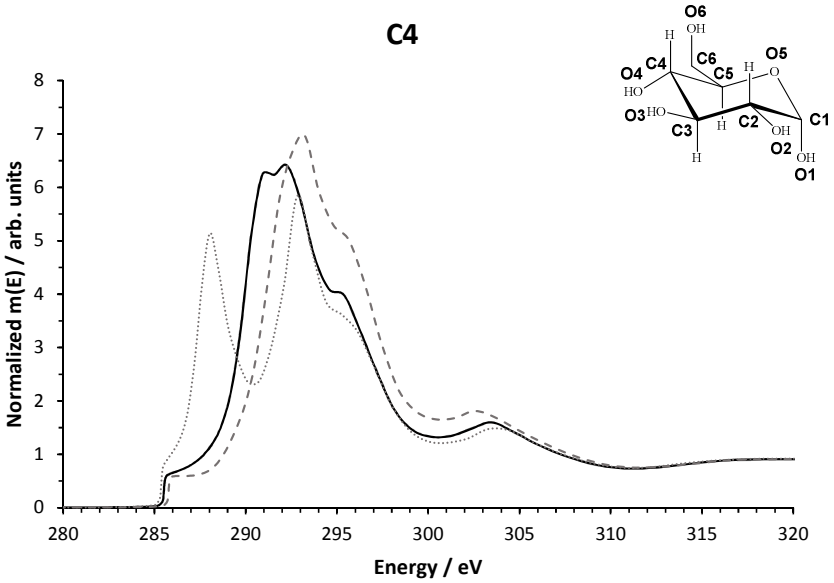


Figure B12. FEFF8 simulations for C4 absorbing atom in: an unaltered crystal structure of glucose (black line), an altered crystal structure of glucose with C4–O4 bond shortened by 0.1 Å (dashed grey line), and an altered crystal structure of glucose with C4–O4 bond elongated by 0.1 Å (dotted grey line).

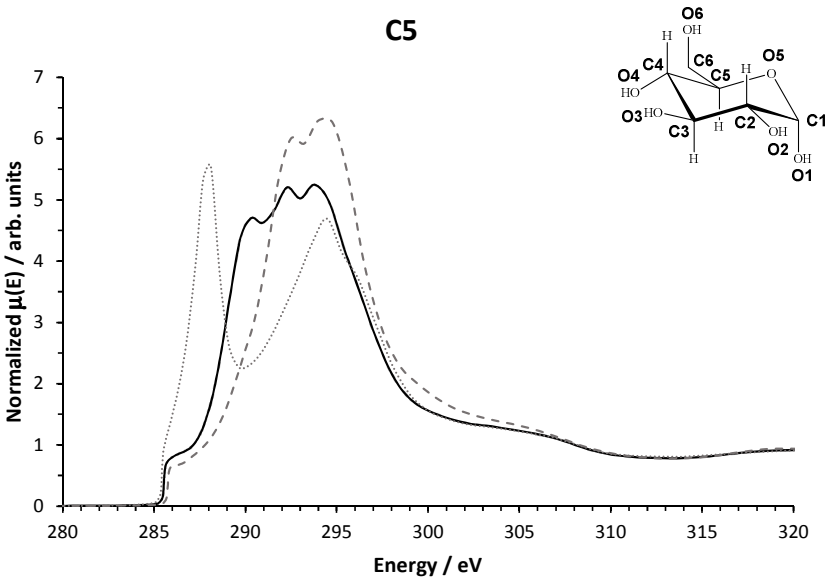


Figure B13. FEFF8 simulations for C5 absorbing atom in: an unaltered crystal structure of glucose (black line), an altered crystal structure of glucose with C5–O5 bond shortened by 0.1 Å (dashed grey line), and an altered crystal structure of glucose with C5–O5 bond elongated by 0.1 Å (dotted grey line); once C5–O5 bond altered, expected variations within the cycle on the C4–C5 and C1–O5 bond lengths are expected, albeit negligible.

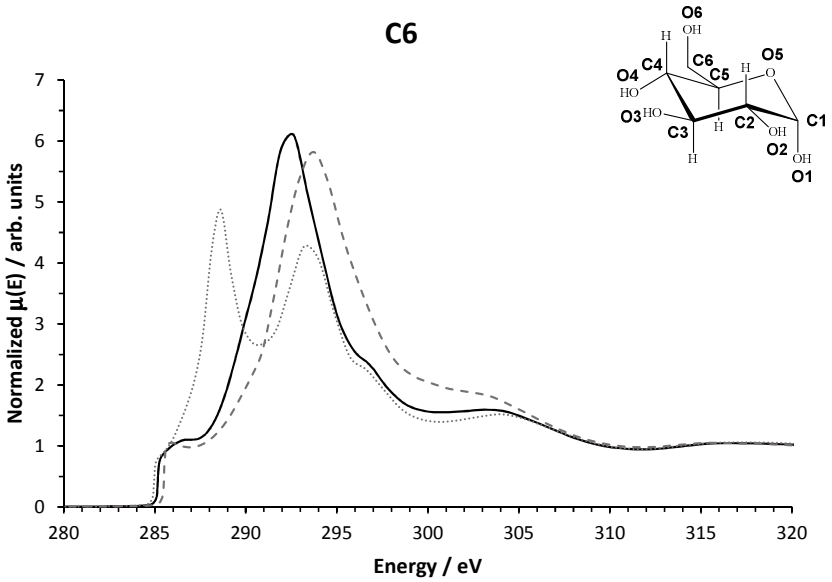


Figure B14. FEFF8 simulations for C6 absorbing atom in: an unaltered crystal structure of glucose (black line), an altered crystal structure of glucose with C6–O6 bond shortened by 0.1 Å (dashed grey line), and an altered crystal structure of glucose with C6–O6 bond elongated by 0.1 Å (dotted grey line).

APPENDIX C. Attended conferences, meetings, workshops

Conference participations

- Poster presentation at the Faraday Discussion *‚Nucleation - A Transition State to the Directed Assembly of Materials'*, Leeds Beckett University, 30th of March-1st of April 2015

- Poster presentation at the British Association for Crystal Growth (BACG2014), Leeds Beckett University, 13th-15th of July 2014
- Poster presentation at the Postgraduate Conference 2014, The University of Manchester, 4th of July 2014
- Oral talk on the current research pursued, during the Industrial Advisory Board meeting, The University of Manchester, January 2014
- Poster presentation at BESSY II User Meeting, Helmholtz Zentrum Berlin (HZB), Berlin, 5th-6th December 2013
- Poster presentation at the British Association of Crystal Growth (BACG2013), The University of Manchester, 16th-18th June 2013
- Poster presentation at the Postgraduate Conference 2013, The University of Manchester, 11th June 2013
- Poster presentation at the Surface Science of Biologically Important Interfaces (SSBII14) conference, The University of Manchester, 18th-19th October 2012
- Participation to I09 Workshop on Atomic and Electronic Structures of Surfaces and Interfaces, Diamond Light Source, Didcot, 9th-10th July 2012
- Poster presentation at the Postgraduate Conference 2012, The University of Manchester, 15th of June 2012
- Participation to Control and Prediction of the Organic Solid State (CPOSS) Conference *‚Crystal or Not, Where Do We Go From Here?'*, University of College London, 3rd of April 2012

Meetings

- Participations to project-related monthly Critical Mass meetings across The University of Manchester and University of Leeds, where presentations were given with

the aim to update project partners on the latest progress, thus having the opportunity to raise questions on the research progress

- Seminar on applications of X-ray absorption, emission and UV-Vis spectroscopies at The University of Manchester, February 2015
- Seminar on X-ray absorption techniques applied to crystalline sugars at The University of Manchester, February 2014
- Seminar on Near-Edge X-ray Absorption Fine Structure (NEXAFS) technique at The University of Manchester, September 2012

Professional development workshops attended within The University of Manchester

- ‚*Viva Survivor'*, 7th of October 2014
- ‚*Writing up Your Thesis'*, 25th of March 2014
- *'Procure to Pay – Requisitioner eTraining – R12'*, 16th of December 2013
- ‚*Planning the Final Year'*, 10th of December 2013
- ‚*Applying for Fellowships'*, 26th of November 2013
- ‚*Critical Reading'*, 14th of November 2013
- ‚*Science Communication Training'*, 24th of October 2013
- *'Oracle 12 Update'*, 2nd of October 2013
- ‚*Optimising Research Posters'*, 5th of June 2013
- ‚*Designing Effective PowerPoint Presentations'*, 12th of April 2013
- ‚*Academic Writing'*, 8th of November 2012
- ‚*Publishing Academic Papers'*, 20th of June 2012
- ‚*Lunchbite: Keeping up to date with RSS feeds'*, 19th of June 2012
- ‚*Searching the literature with Web of Knowledge and Scopus'*, 23rd of May 2012
- *'Beginners reference management with EndNote: Faculty of EPS'*, 18th of April 2012

APPENDIX D. Curriculum Vitae

Adrian Gainar

EDUCATION AND QUALIFICATIONS

The University of Manchester, PhD in Chemistry, 2012 – 2016

The project in collaboration with a large group of partners across several institutions, involved studies of nucleation/self-association mechanisms prior to crystallisation of para-aminobenzoic acid (PABA) in solutions through absorption and emission techniques, as well as applications in determining bond lengths in solid-state crystals and in solutions, which also constitutes one of the novelties. UV-Vis combined with such techniques and backed-up with calculations can contribute to a more complete general picture around the moieties of interest. The project required very good communication skills and analytical thinking reflected through periodic presentations/high quality reports and updates at monthly meetings and conferences.
I have participated to several experiments in Germany and USA where I organised the experiments, housekeeping, and storage of chemicals, solvents, materials and samples with great attention to detail that required working day or night shifts. I have also trained and supervised new users on several pieces of equipment and troubleshooted general technical issues.
I participated to several conferences and presented posters/ gave talks on my current research.

University of York, Master's degree in Chemistry, 2010 – 2012

The project involved synthesis of organic complex molecules and analytical characterisation of their properties. The difficulty consisted in obtaining the final compounds which proved to have a highly reduced solubility in all trivial organic solvents. Extensive separation using gas and liquid chromatography was also required. Mass spectrometry, extractions, separations, NMR, elemental analysis were employed to confirm compound identity and calorimetry, microscopy, X-ray diffraction to characterise their properties. A very good office/laboratory housekeeping was also required. The examiners were very satisfied with the results and acknowledged it was a difficult task to synthesise the target compounds.

University of Bucharest, BSc in Chemistry (with industrial placement), 2007 – 2010

The programme comprised valuable modules of all general areas of chemistry and related disciplines:

- Qualitative and quantitative analysis; Separation methods in biochemistry

❖ Molecular structure; Chemical kinetics and thermodynamics; Quantum biochemistry;
❖ Basics of organic chemistry; Natural products; Semi-synthesis of bioactive compounds; Colourants;
❖ General chemistry; Chemistry of elements; Bioinorganic chemistry;
❖ Elementary catalysis;
❖ General biochemistry; Proteins and nucleic acids;
❖ Mechanics and electromagnetism; Quantum physics;
❖ Mathematical analysis; Differential equations.

I was awarded the medal of Class of 2010 Leader (1st Class Honours) having the highest grade of 3 years of study, among approximately 60 students from the Faculty. The techniques involved were quantitative and qualitative analysis, titrations, pH determination, UV-Vis, chromatography, electrophoresis, extractions and separations. I also have experience in Biochemistry (industrial placement at the National Biochemistry Institute *‚Dr. I. Cantacuzino'* Bucharest), and have obtained the Scientific French Language Certificate and Psycho-Pedagogy Certificate.

‚*Grigore Moisil'* National College of Computer Science, Brasov, 2003 – 2007

I was awarded Class of 2007 Leader and Key of Success distinctions for scholar merits – the highest grade among approximately 180 students of the college. I also obtained perfect scores in Mathematics, English and Geography on my Baccalaureate examinations. I obtained the Certificate of Computer Science Knowledge with a maximum score, which included Computer Use and Project Presentation.

I participated to various Chemistry, Physics, Mathematics and Earth Science National Olympiads/Contests, and was awarded several prizes each year on scholar merits.

General School no 3, Rasnov, 1995 – 2003 finalised with National Examinations

Class of 2003 Leader and Key of Success distinctions awarded in 2002 on scholar merits – the highest grade among approximately 90 students. I participated to Chemistry and Mathematics National Olympiads and Contests, and was awarded excellence prizes for scholar merits.

AWARDS AND GRANTS

- Shortlisted for the Excellence EPSRC Doctoral Prize 2015/16 at The University of Manchester (2015)
- EPSRC Research Grant (2012-2015)
- ‚*Open Horizons'* Excellence Studentship (2010-2011)
- Medal of Class of 2010 Leader (1st Class Honours) for the highest grade of 3 years of study, among approximately 60 students from the Faculty (2010)
- ‚*Invent Your Future!'* Excellence Sctudentship (2008, 2009, 2010)
- University of Bucharest Merit-based Studentship (2007, 2008, 2009, 2010)

- Class of 2007 Leader and Key of Success distinctions on scholar merits – the highest grade among approximately 180 students (2007)
- Class of 2003 and Key of Success distinctions on scholar merits – the highest grade among approximately 90 students (2003)
- Various prizes (50+) at Chemistry, Physics, Mathematics, Earth Science National Contests (2002-2007)

PRACTICAL SKILLS

- Laboratory experience that conferred very good manual dexterity and ability to work following experimental procedures
- Very good organisational, planning and IT skills, attention to details; troubleshooting general technical issues; housekeeping comprising chemicals, solvent, materials, samples storage
- Data collection and analysis
- Liquid chromatography, mass spectrometry, titrations, extractions, pH determination, elemental analysis, UV-Vis, NMR, IR, absorption/emission spectroscopies, calorimetry, microscopy, X-ray diffraction
- Experienced in synthesis of complex organic molecules
- Gaussian and FEFF8 calculations for various molecular species (technique validation)

LANGUAGE SKILLS

- Fluent English (written and spoken)
- Fluent Romanian (written and spoken)
- French (intermediate in writing and speaking)
- German (beginner in writing and speaking)

APPENDIX E. Peer-reviewed publications

THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

NEXAFS Sensitivity to Bond Lengths in Complex Molecular Materials: A Study of Crystalline Saccharides

Adrian Gainar,[†] Joanna S. Stevens,[†] Cherno Jaye,[‡] Daniel A. Fischer,[‡] and Sven L. M. Schroeder*[,†,§,∥]

[†]School of Chemical Engineering and Analytical Science, The University of Manchester, Oxford Road, Manchester M13 9PL, United Kingdom

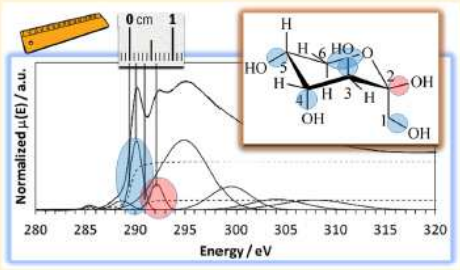
[‡]Material Measurement Laboratory, National Institute of Standards and Technology, Gaithersburg, Maryland 20899, United States

[§]School of Chemical and Process Engineering, University of Leeds, Leeds LS2 9JT, United Kingdom

[∥]DIAMOND Light Source Limited, Harwell Science and Innovation Campus, Chilton, Didcot OX11 0DE, United Kingdom

Supporting Information

ABSTRACT: Detailed analysis of the C K near-edge X-ray absorption fine structure (NEXAFS) spectra of a series of saccharides (fructose, xylose, glucose, galactose, maltose monohydrate, α-lactose monohydrate, anhydrous β-lactose, cellulose) indicates that the precise determination of IPs and σ^* shape resonance energies is sensitive enough to distinguish different crystalline saccharides through the variations in their average C–OH bond lengths. Experimental data as well as FEFF8 calculations confirm that bond length variations in the organic solid state of 10^{-2} Å can be experimentally detected, opening up the possibility to use NEXAFS for obtaining incisive structural information for molecular materials, including noncrystalline systems without long-range order such as dissolved species in solutions, colloids, melts, and similar amorphous phases. The observed bond length sensitivity is as good as that originally reported for gas-phase and adsorbed molecular species. NEXAFS-derived molecular structure data for the condensed phase may therefore be used to guide molecular modeling as well as to validate computationally derived structure models for such systems. Some results indicate further analytical value in that the σ^* shape resonance analysis may distinguish hemiketals from hemiacetals (i.e., derived from ketoses and aldoses) as well as α from β forms of otherwise identical saccharides.



INTRODUCTION

Saccharides, also known as carbohydrates and sugars, are an important class of natural products and are well-known as foods and components of ingestible formulated products.[1−6] They have a propensity to present various noncrystalline and crystalline forms, and interest in intra- and intermolecular bonding in their solid state structures goes back to the earliest days of organic chemistry.[7] Saccharides can exist as monomers or, covalently bonded, as dimers, oligomers, or polymers, in which molecular monomers are linked by glycosidic bonds. Well-known monosaccharides are glucose (dextrose, also known as *blood sugar*), fructose (or *fruit sugar*), and ribose, an essential component of ribonucleic acid (RNA). Among disaccharides, well-known representatives are sucrose (*table/beet/cane sugar*, or simply *sugar*) and lactose (*milk sugar*), which is commonly used in pharmaceutical products as an excipient. The most abundant polysaccharide occurring in nature is cellulose,[8] which confers support, resistance, and rigidity to plants. In terms of classification by the number of carbons, saccharides with five carbon atoms in their molecules are called pentoses (e.g., ribose, xylose, lyxose), while six-membered ring saccharides are called hexoses (e.g., glucose, fructose, galactose).

Identifying mono-, di-, oligo-, and polysaccharides spectroscopically is often important for understanding the properties of formulated products.[9] X-ray core level spectroscopies are structurally incisive methods that are increasingly used for the analysis of interfacial species in complex products, devices, and biomaterials.[9−14] The most commonly used variant, X-ray photoelectron spectroscopy (XPS), provides information on the chemical state of atoms through sensitivity to the local electron density variations caused by changes in chemical bonding, which lead to chemical shifts of core level binding energies, and interpretation can be complemented or enhanced with the results of density functional theory (DFT) calculations.[15] Careful C 1s core level binding energy measurements by XPS of fructose, xylose, glucose, galactose, maltose, α-lactose, β-lactose, and cellulose have previously

Received: July 23, 2015
Revised: October 12, 2015
Published: October 12, 2015

DOI: 10.1021/acs.jpcb.5b07159
J. Phys. Chem. B 2015, 119, 14373−14381

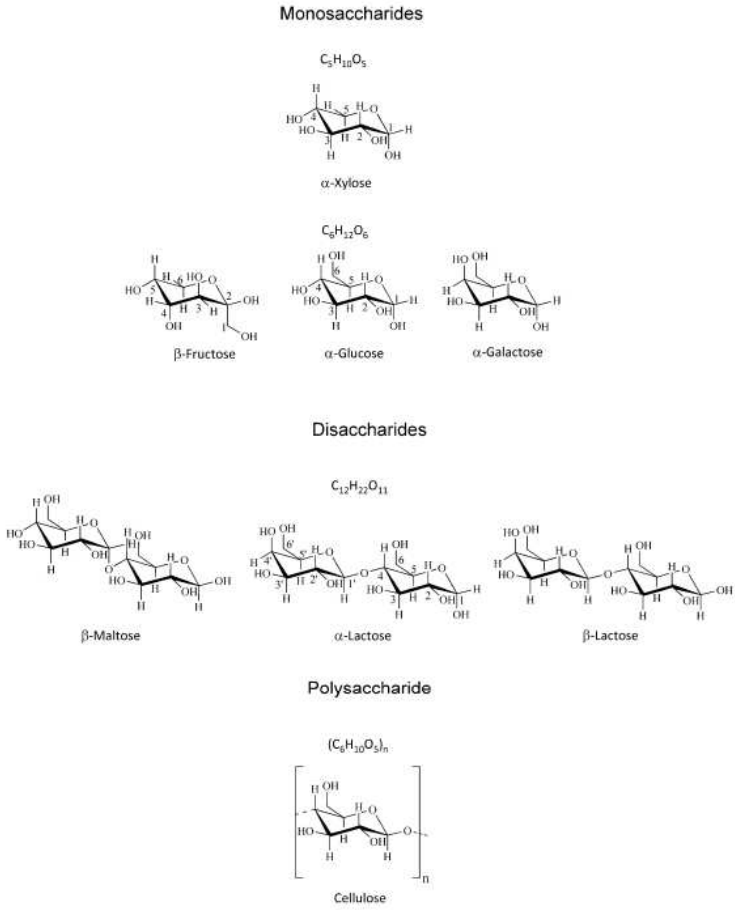


Figure 1. Saccharides analyzed by NEXAFS. IUPAC numbering of the carbon atoms is included for xylose, fructose, glucose, and α-lactose. Other mono- and disaccharides and the polysaccharide adopt analogous numbering.

revealed that XPS is sensitive enough to distinguish mono-, di-, and polysaccharides.[16]

Near-edge X-ray absorption fine structure (NEXAFS, also called X-ray absorption near-edge structure, XANES)[17] provides more incisive information than XPS about local chemical and electronic structure in condensed matter. Determination of bond lengths, static and dynamic disorder, as well as coordination numbers and coordination geometry around the X-ray absorbing atoms in organic compounds can sometimes also be achieved by extended X-ray absorption fine structure analysis (EXAFS).[17,18] However, C K-edge EXAFS analysis is often not possible because of reciprocal space limitations caused by (i) spectral overlap of absorption edges from other elements (for the present case of saccharides the O K-edge ~250 eV above the C K-edge) and (ii) instrumental limitations such as source transmission variations due to contamination on optics, abrupt variations in photon flux, or nonlinearities in I_0 monitoring. However, NEXAFS offers the possibility to obtain information about bond lengths in organic molecules through determination of σ^* shape resonance energies relative to the ionization potential of the core level from which excitation takes place. σ^* shape resonances arise because photoexcited core electrons can be temporarily trapped by a centrifugal potential barrier near the edge of the molecule, so electrons traverse the molecule multiple times before escaping from the molecular core by tunneling.[17,19−23] The use of shape resonance energy shifts for determining bond lengths has been referred to as the "bond length with a ruler" method.[17,20−22,24] The method relies on the determination of the core level binding energies (i.e., variations in the core level ionization potentials, IPs) from the absorption spectra and the energetic position of the σ^* shape resonances relative to the IP.[25] We recently reported evidence that this method can also be applied to C−N bonds of organic molecules in solution and in the organic solid state.[22] The likely reason for this transferability of a method originally developed for gas-phase

DOI: 10.1021/acs.jpcb.5b07159
J. Phys. Chem. B 2015, 119, 14373−14381

molecules is that intermolecular interactions in the organic condensed phase are much weaker than the internal covalent bonds in the molecules. While the exact physical origins of σ^* shape resonances, and thus the validity of their use for bond length determination, has been controversially discussed in the past,[25−30] there is considerable empirical evidence as well as theoretical support[30,31] for their interpretation as multiple scattering resonances dependent on the bond length.[23]

Full deconvolution and interpretation of all features in NEXAFS spectra can be a complex task, but our previous success in detecting differences between saccharides by core level binding energy analysis in XPS[16] led us to hypothesize that the structurally more incisive NEXAFS spectroscopy may also be sensitive to differences in structure and bonding in saccharide systems. Note that unlike diffraction methods, core level spectroscopies do not require long-range order and can therefore also be applied to noncrystalline forms. A deeper understanding of the core level spectra of the crystalline forms would therefore establish a basis for characterization of local interactions in materials without long-range order, such as melts, solutions, coatings, or nanocrystalline forms. Literature references describing applications of NEXAFS to saccharides have been rather sporadic[32−34] and generally did not relate the spectra to specific structural details.

In this paper we will discuss C K NEXAFS data of a series of monosaccharides (fructose, xylose, glucose, galactose), disaccharides (maltose monohydrate, α-lactose monohydrate, anhydrous β-lactose), and one polysaccharide (cellulose), with the aim to provide a database of spectral fingerprints and establish what level of structural detail is evident in these spectra. In particular, we show by reference to the crystal structures that the σ^* shape resonance analysis of the NEXAFS provides access to rather small variations in intramolecular C−OH bond lengths in the molecules.

EXPERIMENTAL SECTION

Polycrystalline powder samples were obtained from Sigma-Aldrich, UK. Fructose, xylose, galactose, maltose monohydrate, and cellulose have ≥99% purity. Glucose was 99.5% pure. α-Lactose monohydrate and β-lactose have been provided by Sanofi-Aventis (Alnwick, U.K.) and were the same samples examined in our previous XPS study.[14] All samples were used as provided. Thick layers of powder were spread onto sticky copper tape with a clean thin spatula, ensuring complete covering of the tape surface, and then pressed on mechanically to obtain a compact layer.

Figure 1 summarizes the chemical structures of the analyzed saccharides. The pentose (xylose) in the first row has the chemical formula $C_5H_{10}O_5$, the hexoses (fructose, glucose, galactose) in the second row have the general chemical formula $C_6H_{12}O_6$ and the disaccharides (maltose, α-lactose, β-lactose) in the third row have the general formula $C_{12}H_{22}O_{11}$. The polysaccharide (cellulose) in the fourth row is a polymer of β-glucose having the formula $[C_6H_{10}O_5]_n$.

Generally, monosaccharides crystallize in the pyranose form, i.e., as 6-membered rings; fructose can crystallize in either pyranose or furanose (5-membered ring) form; the pyranose form of fructose is usually obtained by industrial crystallization processes, whereas the furanose form occurs in natural products. Xylose is naturally found in pyranose form. For the disaccharides, maltose, α-lactose, and β-lactose are formed of two monosaccharide units in pyranose form. The polysaccharide, cellulose, is formed of linked glucose units in the pyranose form.

A Rigaku Miniflex XRPD instrument was used to determine the polymorphic form of each saccharide powder, taking diffraction patterns over a 2θ range from 5° to 40°, at a rate of 1°/min with a 0.02° step. The X-ray source (Cu K_α 1.5406 Å) operated with 30 kV voltage and 15 mA current. The experimental patterns were compared with calculated crystal structure patterns based on published crystal structures and are included in the Supporting Information (Figures S1−S8). Results indicated that the forms were β-fructose, α-xylose, α-glucose, α-galactose, β-maltose, α-lactose monohydrate, and anhydrous β-lactose. The positions of the broad features in the pattern of amorphous cellulose (Figure S8 in the Supporting Information) indicated that residual local order was determined by nuclei of the I_α polymorph of cellulose.

NEXAFS spectra were measured at the National Institute of Standards and Technology U7A beamline of the National Synchrotron Light Source (NSLS) at Brookhaven National Laboratory (BNL) in New York, United States. Partial electron yield (PEY) C K-edge spectra were obtained using a channeltron electron multiplier with the sample positioned at the magic angle (54.7°) relative to the incident beam, with a photon flux of approximately 5×10^{10} photons/s. An entrance grid bias of −150 V was used for the C K-edge in the case of galactose and −50 V for the rest of the saccharides.

At −50 V grid bias the detected electron-yield signal is essentially the total Auger yield, i.e., the probing depth of the electron-yield signal is determined by the escape depth of all C KLL Auger electrons, including the whole background of inelastically scattered electrons on the low-energy side of the Auger emission lines,[35−38] which probes deeper than the elastic Auger or photoemission peaks. A previous experimental determination of the C K-edge probing depth in polymers yielded a value of ∼2 nm,[39] while experiments with signal attenuation in Cr metal indicate for photons at the C K-edge that the maximum probing depth is approximately 3.5 nm.[40] These values are equivalent to probing a few molecular layers deep into the saccharide materials and should hence confer bulk sensitivity. In line with this, our previous XPS analysis of the same saccharide materials through C 1s emission with an Al Kα source, which is associated with signal attenuation lengths only slightly higher (∼3 nm),[41] yielded chemical compositions consistent with those expected from the saccharide stoichiometries. Moreover, for molecular materials surface effects on internal bond lengths should be negligible, as intermolecular interactions in organic molecular solids are weak relative to the strength of the internal bonds in the molecules. The strongest intermolecular interaction in the saccharide bulk is hydrogen bonding. It is therefore expected that the C−O bond lengths derived from the bulk crystal structures used for the interpretation of our data are good models.

A toroidal spherical grating monochromator with 600 lines mm^{-1} grating was used to acquire the C K-edge data, yielding an energy resolution of approximately 0.1 eV. After collection, the spectra were normalized by the simultaneously recorded drain current from an in situ gold-coated, 90% transmission grid (I_0) which is situated in the X-ray beam path, to eliminate the incident beam intensity fluctuations and absorption features in the beamline optics. The monochromator energy scale calibration was performed through the 285.1 eV π^* graphite absorption by a carbon mesh located further upstream of I_0 (I_0_up). All spectra were processed using standard pre- and

DOI: 10.1021/acs.jpcb.5b07159
J. Phys. Chem. B 2015, 119, 14373−14381

postedge normalization methods. The peak-fitting capability of Athena software[42,43] was employed to determine the energetic positions of all NEXAFS transitions at the C K-edge, and the spectrum was modeled with Gaussians for the peaks and two arctan step functions for the edge steps arising at the IPs.[17] The positions of the IPs were determined assuming two constraints: the expected intensity ratio for OC−OH and C−OH moieties and their energy difference, which was known from our previous XPS study to be 1.35 eV higher for the OC−OH hemiacetal/hemiketal groups.[16] The first IP was set as the experimental edge step inflection point, i.e., the zero crossing of the second derivative of the experimental spectrum. For investigating the "bond length with a ruler" correlation, the standard deviation for the C−OH bond lengths in the reported crystal structures was ≤0.007 Å.[44−49] For the curve-fitted C K-edge spectra, by fitting with different models (variation of peak positions, functions, and the number of included peak functions) and using different initial values for the peak parameters we found that the error in energy positions for comparably good final fits was ≤0.02 eV.

Slightly varying levels of adventitious surface contamination were previously observed in the XPS data for the same set of saccharides.[16] In the NEXAFS spectra this contamination is evident as weak pre-edge peaks in the C K-edge spectra. Similar to the constant peak position of the hydrocarbon contaminant in XPS C 1s results,[16] these pre-edge peaks were in a narrow energy range, from 285.3 to 285.4 eV. We carefully checked by examination of the signal from the incident beam monitor (a gold coated mesh) whether varying levels of surface contamination on it might have led to normalization errors that could account for these pre-edge features. However, no variations in the mesh signal were evident that could cause the observed signals.

In the previous XPS analysis of saccharides, it was also noted that the intensity of the peaks from the C−OH and O−C−O hemiacetal/hemiketal moieties in the C 1s spectra decreased as a function of X-ray exposure[16] due to radiative damage, with the strongest effect on ribose and xylose.[16] In the NEXAFS measurements, no significant changes in peak positions and intensities nor formation of new peaks were observed as a function of exposure to the X-ray beam, even after six scans of the C K-edge.

FEFF8.2[50] was used to simulate the σ^* shape resonances in the C K NEXAFS for a cluster cut from the crystal structure of glucose (CCDC reference: GLUCSA03).[51] The XANES, FMS, and SCF cards were used—FMS for full-multiple scattering X-ray absorption near-edge structure (XANES) calculation; SCF to enable self-consistent field iterations. To identify the shape resonance contributions due to scattering near the potential of the C−OH single bonds we also ran calculations in which the C−OH bond length in the hemiacetal/hemiketal position was increased and decreased by 0.1 Å relative to the equilibrium value in the crystal structure (see Figure 2 and Supporting Information).

■ RESULTS AND DISCUSSION

The C K NEXAFS spectra of the saccharides are presented in Figures 2 and 3 together with the results of the fitting analysis involving Gaussian and edge step functions. For a deeper analysis of these spectra we need to recall that saccharide molecules present only single C−C, C−H, and C−O bonds (Figure 1), so it is expected that the NEXAFS spectra are dominated by transitions to unoccupied, antibonding σ^* states.

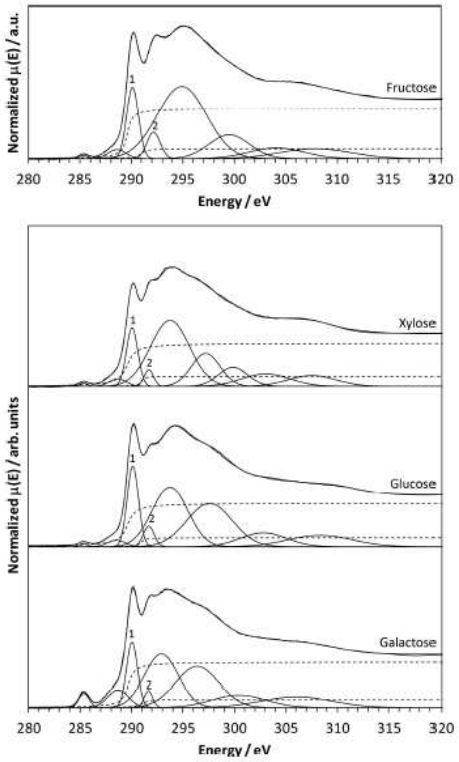


Figure 2. Experimental C K-edge spectrum of fructose (top plot) and C K-edge spectra (bottom plot) for the monosaccharides xylose (top), glucose (middle), and galactose (bottom). Each spectrum is accompanied by best fits obtained to determine peak positions, with the σ^*_{C-OH} and σ^*_{OC-OH} peaks marked by numbers 1 and 2, respectively. IP edge steps are marked with dashed lines.

The most prominent features are σ^* shape resonances, which arise from multiple scattering of the photoelectron wave along the internuclear axis between the absorbing atom and its neighbor; thus, the σ^* shape resonance energies depend sensitively on the associated intramolecular bond lengths.

In a previous study of the C K-edge spectra of a dimannoside derivative adsorbed on Au, the prominent σ^* resonance at ∼289 eV was attributed to the C−OH moieties of the saccharide and the feature at ∼291 eV to the OC−OH moiety.[34] In line with this assignment, the intensity of peak 1 is reduced for xylose (Figure 2) as there is one less contributing carbon (i.e., one less C−OH bond in the structure), so the relative intensities of the peaks arising from the two moieties are different compared to the other monosaccharides. The small pre-edge peak at ∼285 eV in the spectra can be attributed to adventitious contamination with unsaturated hydrocarbons containing C═C bonds,[34] while the second pre-edge peak at ∼288 eV, observed as a shoulder, is associated with C−H transitions.[33] The broader features at higher photon energies around 294 eV may arise from C−C bonds.

To examine how the σ^* transitions associated with C−OH bonds vary in energetic position as a function of bond length,

DOI: 10.1021/acs.jpcb.5b07159
J. Phys. Chem. B 2015, 119, 14373−14381

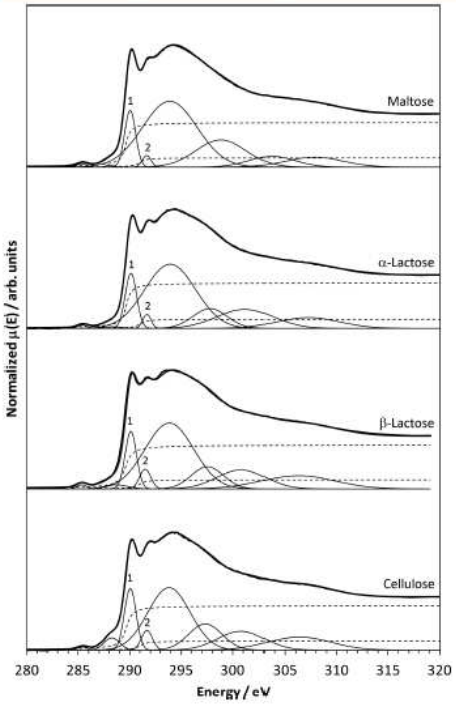


Figure 3. Experimental C K-edge spectra for the di- and polysaccharides studied, maltose (top), α-lactose (second from top), β-lactose (third from top), and cellulose (bottom), alongside fits with the $\sigma^*_{\underline{C}-OH}$ and $\sigma^*_{O\underline{C}-OH}$ peaks marked by numbers 1 and 2, respectively. IP edge steps are marked with dashed lines.

FEFF8 calculations were performed on each inequivalent C atom in the crystal structure of glucose, first with the equilibrium C−OH bond length in the crystal structure and then for the same bond shortened and elongated by 0.1 Å. One set of resulting spectra is presented in Figure 4 and the others in Figures S9−S13 of the Supporting Information. The strong shifts of the shape resonances originally in the region around 290−292 eV are evident, with a shorter bond length resulting in a blue shift (higher photon energy) by approximately +4.7 eV, whereas the longer bond length results in a red shift (lower photon energy) by about −4.2 eV.

As mentioned above, during peak fitting of the spectra in Figures 2 and 3, a 1.35 eV difference in energy between the IPs arising from C−OH and OC−OH moieties was assumed, as this is (with a margin of about ±0.05 eV) the value of the C 1s binding energy separations observed by XPS.[16] The intensity ratios between the two edge step functions representing the IPs were fixed to be in line with the stoichiometries of the saccharides.[16] Note that in the extended X-ray absorption fine structure (EXAFS) region at higher photon energies, where the spectra are determined by scattering of the photoelectron wave by neighboring atoms, the fitted Gaussian functions have no physical meaning, but modeling of the postedge background was required to ensure accurate fitting results in the spectral regions of the electronic transitions of interest, i.e., around and

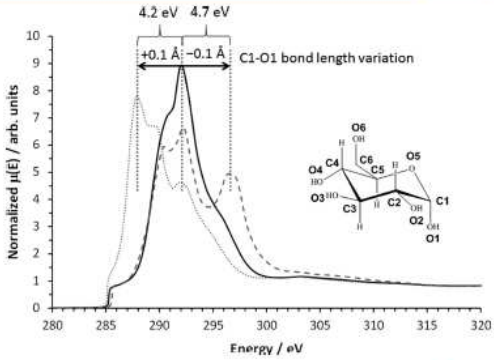


Figure 4. FEFF8 simulations for C1 absorbing atom in an unaltered crystal structure of glucose (black line), an altered crystal structure of glucose with the C1−O1 bond shortened by 0.1 Å (dashed gray line), and an altered crystal structure of glucose with the C1−O1 bond elongated by 0.1 Å (dotted gray line).

below the IPs. The parameters derived with the fitting results (shown in Figures 2 and 3) are summarized in Table 1. It can be seen that the $\sigma^*_{\underline{C}-OH}$ transitions are in a narrow range of energies, from 290.0 to 290.1 eV. The $\sigma^*_{O\underline{C}-OH}$ transitions are found between 291.5 and 292.1 eV.

We can now turn to examining how the crystallographically determined C−OH bond lengths correlate with the shifts in the σ^* shape resonances and IPs. This "bond length with a ruler" analysis[17,20−22] consists of plotting the term value δ, which is the energy difference between the σ^* shape resonance energy and its corresponding IP (Table 1), as a function of the length of the bond from which it arises. Because of the significant IP difference between the C−OH and the OC−OH moieties in the saccharides, their shape resonances are distinct, and their bond lengths can be determined separately. There are of course several C−OH moieties with slightly different bond lengths in each saccharide, so we used their average bond length for the analysis (Table 2), as available from their published crystal structures.[44−49,52,53] The OC−OH hemiacetal or hemiketal groups are associated with a shorter C−OH bond length (Table 2) than for the other C−OH groups,[44−49,52,53] and averaging of two values has to be used in the cases of the disaccharides and for cellulose, which is composed of repetitive interconnected motifs of two distinct cyclic glucose units bound by a glycosidic bond (i.e., cellulose can also be regarded as a polymer of a disaccharide). For the monosaccharides no averaging was required as there are no inequivalent OC−OH groups in the crystal structures. We also point out that contributions from the rather short endocyclic C−O bonds of the hemiacetal and hemiketal groups were omitted from bond length averaging as they are located in the energy region beyond 293 eV, where their contribution cannot be distinguished against the strong background arising from the C−C bonds. The other endocyclic C−O bond is always similar in length to the C−OH bonds and was therefore included in their average.

The resulting plot of the δ values as a function of the bond lengths is shown in Figure 5. It is evident that apart from two outliers (the OC−OH groups in β-lactose and fructose; see discussion below) there is a good linear correlation despite the small range (~0.05 Å) of bond length variations covered

DOI: 10.1021/acs.jpcb.5b07159
J. Phys. Chem. B 2015, 119, 14373−14381

Table 1. IPs and Centroid Energies of the σ^* Shape Resonances Determined by Curve Fitting[a]

saccharide	IP_{C-OH} (eV)	σ^*_{C-OH} (eV)	δ_{C-O} (σ^*_{C-OH} − IP_{C-OH}) (eV)	IP_{OC-OH} (eV)	σ^*_{OC-OH} (eV)	δ_{O-C-O} (σ^*_{OC-OH} − IP_{OC-OH}) (eV)
β-fructose	289.55	290.07	0.52	290.90	292.11	1.21
α-xylose	289.55	290.02	0.47	290.90	291.65	0.75
α-glucose	289.60	290.11	0.51	290.95	291.66	0.71
α-galactose	289.70	290.06	0.36	291.05	291.65	0.60
β-maltose monohydrate	289.55	289.98	0.43	290.90	291.58	0.68
α-lactose monohydrate	289.62	290.06	0.44	290.97	291.61	0.65
β-lactose	289.63	290.14	0.51	290.98	291.53	0.55
cellulose	289.60	289.99	0.39	290.95	291.61	0.66

[a]Term values δ are calculated as the difference between each σ^* centroid energy and its corresponding IP.

Table 2. C−OH Bond Lengths Determined by Crystal Structure Analysis[a]

saccharide	CSD identifier	C−OH bond length at OC−OH (1) (Å)	average length of remaining C−OH (av) (Å)
β-fructose	FRUCTO11[44]	1.412	1.423
α-xylose	XYLOSE[47]	1.393	1.424
α-glucose	GLUCSA03[45]	1.391	1.428
α-galactose	ADGALA01[46]	1.409	1.437
β-maltose monohydrate	MALTOS11[48]	1.402	1.418
α-lactose monohydrate	LACTOS03[49]	1.390	1.425
β-lactose	BLACTO02[53]	1.381	1.423
cellulose	PADTUL[52]	1.403	1.427

[a](1) represents the C−OH bond length from the hemiacetal or hemiketal group (as in Figure 1), and (av) represents the average values of the other the C−OH bond lengths (including one endocyclic C−O bond, see main text).

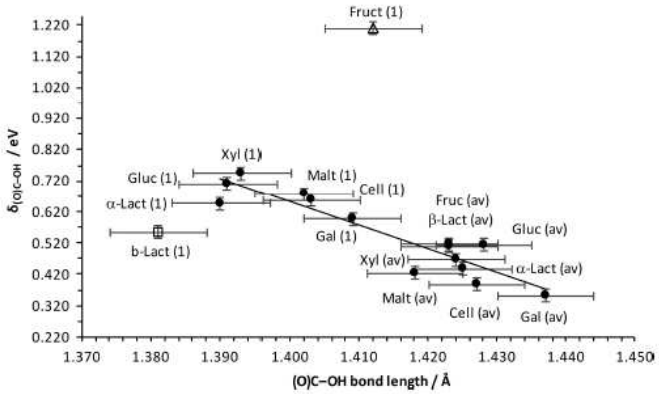


Figure 5. Linear correlation between term value δ and (O)C−OH bond length for the crystal structures of saccharides, where Fruct = fructose, Xyl = xylose, Gluc = glucose, Gal = galactose, Malt = maltose, α-Lact = α-lactose, β-Lact = β-lactose, Cell = cellulose; annotation (1) refers to the OC−OH bond length at the hemiacetal/hemiketal position (as in Figure 1), and (av) refers to the average value of the remaining C−OH bond lengths; fructose and β-lactose, outliers away from the fitted line (see discussion in the text), are represented with an empty triangle and an empty square, respectively.

(Figure 5). More precisely, for a change in the C−OH bond length by 0.047 Å, a term value shift of 0.393 eV occurs. Assuming that the term values can be determined with an accuracy of approximately 0.05 eV this suggests a sensitivity of the method to bond length variations slightly better than ~0.01 Å, making the precision of the technique comparable to accuracies achievable by standard X-ray crystal structure determination. We note that these experimentally derived figures are also in good agreement with the C−OH shape resonance shifts predicted by the FEFF8 calculations (vide supra), which similarly indicated shifts of approximately 4.5 eV per 0.1 Å variation in bond length.

Turning to the outliers in the plot in Figure 5, for fructose the C−OH bond at the OC−OH group is considerably longer (1.412 Å) by comparison with the other saccharides (range 1.389−1.409 Å). The associated term value is also higher than that for the other saccharides. Both deviations are likely due to the fact that the OC−OH group in fructose represents a hemiketal rather than a hemiacetal moiety. As a result there is an additional C−C (rather than C−H) bond at the OC−OH moiety.[54] Interestingly, this result suggests that the NEXAFS analysis of C−OH σ^* shape resonances and IPs provides an avenue to distinguishing hemiketals from hemiacetals (i.e., ketoses from aldoses). However, more systematic work should

DOI: 10.1021/acs.jpcb.5b07159
J. Phys. Chem. B 2015, 119, 14373−14381

be carried out to confirm whether this observation can be generalized.

β-Lactose, the other outlier in Figure 5, has a significantly shorter C−OH bond at the OC−OH group than the other saccharides. It is distinguished from the other hemiacetals by a different configuration at the anomeric position, in which the OH group is positioned cis (rather than trans) to the $-CH_2OH$ group in the molecule. Moreover, unlike the other two disaccharides (β-maltose and α-lactose), the two monosaccharide units in the β-lactose structure are rotated relative to each other around the glycosidic bond. These local structure differences may perhaps combine to cause weak electronic structure variations and hence, e.g., slight differences in core hole relaxation or local screening differences that may affect the term value δ. A deeper analysis of other saccharide structures would be required to confirm or refute these points.

In any case, the outliers appear to underline an important constraint of the bond length with a ruler methodology, namely, that variations in the term value δ can only be compared for bonds in which the two atoms are in a similar chemical state, so additional dependencies on functional group identity (as in, aldehyde vs ketone) and (in condensed matter) secondary interactions with other functional groups due to conformational differences, configurational isomerism, and/or functional group interactions in the crystal structure may influence the results.

We investigated the possibility to additionally determine term values in the O K NEXAFS spectra, but the determination and interpretation of these values is more complex than for the C K-edge due to the additional influence of intermolecular hydrogen bonding, so that no firm conclusions could be drawn. Overall, however, the C K-edge results we obtained indicate that for hemiacetals with α configuration at the anomeric center the small relative energy variations between IPs and shape resonance features in the C K NEXAFS are, within experimental error, consistent with the variations of the C−O bond lengths known from single-crystal structure analysis. Our data lead to an estimate of a sensitivity to bond length variations on the order of ∼0.01 Å, which is in very good agreement with the previously reported sensitivity for gas-phase molecules.[21] Moreover, FEFF8 analysis confirms that the observed experimental energy shifts of the shape resonances are consistent with theoretical expectations for saccharides. These results suggest that careful use of solid state NEXAFS can provide very incisive structure information for molecular materials and should perhaps be considered more for the characterization of local bonding in noncrystalline systems, not only dissolved species as in a recent study[22] but also colloidal systems as well as melts and similar amorphous phases.

CONCLUSIONS

Analysis of the C K NEXAFS spectra of solid monosaccharides, disaccharides, and a polysaccharide indicates that the precise determination of IPs and σ^* shape resonance energies is sensitive enough to distinguish between different saccharides through the minute variations in C−O bond lengths in these systems. The analysis of the experimental data as well as FEFF8 calculations show that this sensitivity arises because bond length variations of 10^{-1} Å are associated with term value shifts on the order of ±4.5 eV, permitting determination of bond length variations of ∼0.01 Å. These observations suggest that NEXAFS can be used to obtain valuable bond length information for other molecular materials of similar complexity. Very importantly, NEXAFS is known to be applicable to noncrystalline systems (e.g., species in solutions, melts, amorphous phases, and adsorbed species), opening up an avenue toward structural information that can be used to confront computationally derived models for systems without long-range order. Moreover, the C K-edges of fructose and β-lactose do not fit into the trend set up by the other saccharides, suggesting that NEXAFS may also be sensitive enough to distinguish hemiketals from hemiacetals as well as α and β forms of the otherwise identical saccharides, but more systematic experimental work should examine these hypotheses.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcb.5b07159.

X-ray powder diffraction patterns for each saccharide compared to calculated crystal structure patterns; results of FEFF8 calculations for each C atom in glucose with equilibrium, elongated and shortened C−O bond lengths (PDF)

Raw NEXAFS data files (ZIP)

Athena project file with fitted NEXAFS spectra (ZIP)

AUTHOR INFORMATION

Corresponding Author

*Phone/Fax: +44 (0)113 343 2401/2384. E-mail: s.l.m.schroeder@leeds.ac.uk.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

All data supporting this study are provided either in the results section of this paper or in the supplementary information accompanying it. We gratefully acknowledge support for A.G., J.S.S., and S.L.M.S. through an EPSRC Critical Mass Grant (EP/1013563/1). Use of the National Synchrotron Light Source, Brookhaven National Laboratory, was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. DE-AC02-98CH10886. Commercial names mentioned in this manuscript are for illustrative purposes and do not represent an endorsement by the National Institute of Standards and Technology. S.L.M.S. holds the Bragg Centenary Chair at the University of Leeds, which is supported by the Royal Academy of Engineering, Diamond Light Source Ltd and Infineum UK, Ltd.

REFERENCES

(1) Rowe, R. C.; Sheskey, P. J.; Owen, S. C. *Handbook of Pharmaceutical Excipients*; Pharmaceutical Press and American Pharmacists Association: London and Washington, D.C., 2005.

(2) Holsinger, V. *Fundamentals of Dairy Chemistry*, 3rd ed.; Springer: Berlin, Heidelberg, New York, 1999.

(3) Johnson, D. C.; Lacourse, W. R. *Carbohydrate Analysis*; Elsevier: Amsterdam, 1995.

(4) Collins, P.; Ferrier, R. *Monosaccharides*; John Wiley & Sons: Chichester, 1995.

(5) Chaplin, M. F. *Carbohydrate Analysis-A Practical Approach*; Oxford University Press: Oxford, 1994.

(6) Lindhorst, T. K. *Essentials of Carbohydrate Chemistry and Biochemistry*; Wiley-VCH: Weinheim, 2007.

(7) Brown, G. M.; Levy, H. A. Alpha-D-Glucose - Determination of Crystal and Molecular Structure by Neutron-Diffraction Analysis. *Science* **1965**, *147*, 1038.
(8) Colvin, J. R. *The Biochemistry of Plants*; Academic Press: New York, 1980.
(9) Murrieta-Pazos, I.; Gaiani, C.; Galet, L.; Calvet, R.; Cuq, B.; Scher, J. Food Powders: Surface and Form Characterization Revisited. *J. Food Eng.* **2012**, *112*, 1−21.
(10) Kaivosoja, E.; Virtanen, S.; Rautemaa, R.; Lappalainen, R.; Konttinen, Y. Spectroscopy in the Analysis of Bacterial and Eukaryotic Cell Footprints on Implant Surfaces. *Eur. Cell Mater.* **2012**, *24*, 60−73.
(11) Benetti, F.; Marchettini, N.; Atrei, A. ToF-SIMS and XPS Study of Ancient Papers. *Appl. Surf. Sci.* **2011**, *257*, 2142−2147.
(12) Dietrich, P. M.; Horlacher, T.; Girard-Lauriault, P. L.; Gross, Th. s.; Lippitz, A.; Min, H.; Wirth, Th.; Castelli, R.; Seeberger, P.; Unger, W. E. Multimethod Chemical Characterization of Carbohydrate-Functionalized Surfaces. *J. Carbohydr. Chem.* **2011**, *30*, 361−372.
(13) Dietrich, P. M.; Horlacher, T.; Gross, Th.; Wirth, Th.; Castelli, R.; Shard, A. G.; Alexander, M.; Seeberger, P. H.; Unger, W. E. Surface Analytical Characterization of Carbohydrate Microarrays. *Surf. Interface Anal.* **2010**, *42*, 1188−1192.
(14) Willneff, E. A.; Ormsby, B. A.; Stevens, J. S.; Jaye, Ch.; Fischer, D. A.; Schroeder, S. L. M. Conservation of Artists' Acrylic Emulsion Paints: XPS, NEXAFS and ATR FTIR Studies of Wet Cleaning Methods. *Surf. Interface Anal.* **2014**, *46*, 776−780.
(15) Stevens, J. S.; Seabourne, C. R.; Jaye, Ch.; Fischer, D. A.; Scott, A. J.; Schroeder, S. L. M. Incisive Probing of Intermolecular Interactions in Molecular Crystals: Core Level Spectroscopy Combined with Density Functional Theory. *J. Phys. Chem. B* **2014**, *118*, 12121−12129.
(16) Stevens, J. S.; Schroeder, S. L. M. XPS Analysis of Saccharides and Their X-ray Induced Degradation. *Surf. Interface Anal.* **2009**, *41*, 453−462.
(17) Stöhr, J. *NEXAFS Spectroscopy*; Springer Series in Surface Sciences 25; Springer: Berlin, Heidelberg, 1992.
(18) Yang, B. X.; Kirz, J.; Sham, T. K. Oxygen K-Edge Extended X-ray Absorption Fine Structure Studies of Molecules Containing Oxygen and Carbon Atoms. *Phys. Rev. A: At., Mol., Opt. Phys.* **1987**, *36*, 4298−4310.
(19) *Atomic, Molecular and Optical Physics*; National Academic Press (National Research Council): Washington, D.C., 1986.
(20) Sette, F.; Stöhr, J.; Hitchcock, A. P. Determination of Intramolecular Bond Lengths in Gas-Phase Molecules from K-Shell Shape Resonances. *J. Chem. Phys.* **1984**, *81*, 4906−4914.
(21) Stöhr, J.; Sette, F.; Johnson, A. L. Near-Edge X-ray Absorption Fine-Structure Studies of Chemisorbed Hydrocarbons: Bond Lengths with a Ruler. *Phys. Rev. Lett.* **1984**, *53*, 1684−1687.
(22) Stevens, J. S.; Gainar, A.; Suljoti, E.; Xiao, J.; Golnak, R.; Aziz, E. F.; Schroeder, S. L. M. Chemical Speciation and Bond Lengths of Organic Solutes by Core Level Spectroscopy: pH- and Solvent-Influence on p-Aminobenzoic Acid. *Chem. - Eur. J.* **2015**, *21*, 7256−7263.
(23) Piancastelli, M. N. The Neverending Story of Shape Resonances. *J. Electron Spectrosc. Relat. Phenom.* **1999**, *100*, 167−190.
(24) Hitchcock, A. P.; Brion, C. E. K-Shell Excitation of HF and F_2 Studied by Electron Energy-Loss Spectroscopy. *J. Phys. B: At. Mol. Phys.* **1981**, *14*, 4399−4413.
(25) Arvanitis, D.; Haack, N.; Ceballos, G.; Wende, H.; Baberschke, K.; Ankudinov, A. L.; Rehr, J. J. Shape Resonances of Oriented Molecules. *J. Electron Spectrosc. Relat. Phenom.* **2000**, *113*, 57−65.
(26) Piancastelli, M. N.; Lindle, D. W.; Ferrett, T. A.; Shirley, D. A. The Relationship Between Shape Resonances and Bond Lengths - Reply. *J. Chem. Phys.* **1987**, *87*, 3255−3256.
(27) Piancastelli, M. N.; Lindle, D. W.; Ferrett, T. A.; Shirley, D. A. The Relationship Between Shape Resonances and Bond Lengths. *J. Chem. Phys.* **1987**, *86*, 2765−2771.
(28) Kempgens, B.; Köppe, H. M.; Kivimäki, A.; Neeb, M.; Maier, K.; Hergenhahn, U.; Bradshaw, A. M. Reappraisal of the Existence of Shape Resonances in the Series C_2H_2, C_2H_4, and C_2H_6. *Phys. Rev. Lett.* **1997**, *79*, 35−38.
(29) Kempgens, B.; Köppe, H. M.; Kivimäki, A.; Neeb, M.; Maier, K.; Hergenhahn, U.; Bradshaw, A. M. On the Correct Identification of Shape Resonances in NEXAFS. *Surf. Sci.* **1999**, *425*, L376−L380.
(30) Haack, N.; Ceballos, G.; Wende, H.; Baberschke, K.; Arvanitis, D.; Ankudinov, A. L.; Rehr, J. J. Shape Resonances of Oriented Molecules: ab initio Theory and Experiment on Hydrocarbon Molecules. *Phys. Rev. Lett.* **2000**, *84*, 614−617.
(31) Shneerson, V. L.; Saldin, D. K.; Tysoe, W. T. On the Dependence with Bond Lengths of the Observed Energies of NEXAFS Resonances of Diatomic Molecules. *Surf. Sci.* **1997**, *375*, 340−352.
(32) Hardie, A. G.; Dynes, J. J.; Kozak, L. M.; Huang, P. M. The Role of Glucose in Abiotic Humification Pathways as Catalyzed by Birnessite. *J. Mol. Catal. A: Chem.* **2009**, *308*, 114−126.
(33) Solomon, D.; Lehmann, J.; Kinyangi, J.; Liang, B.; Heymann, K.; Dathe, L.; Hanley, K.; Wirick, S.; Jacobsen, C. Carbon (1s) NEXAFS Spectroscopy of Biogeochemically Relevant Reference Organic Compounds. *Soil Sci. Soc. Am. J.* **2009**, *73*, 1817−1830.
(34) Dietrich, P. M.; Horlacher, T.; Girard-Lauriault, P. L.; Gross, T.; Lippitz, A.; Min, H.; Wirth, T.; Castelli, R.; Seeberger, P. H.; Unger, W. E. S. Adlayers of Dimannoside Thiols on Gold: Surface Chemical Analysis. *Langmuir* **2011**, *27*, 4808−4815.
(35) Schroeder, S. L. M.; Moggridge, G. D.; Ormerod, R. M.; Rayment, T.; Lambert, R. M. What Determines the Probing Depth of Electron Yield XAS? *Surf. Sci.* **1995**, *324*, L371−L377.
(36) Schroeder, S. L. M. Towards a 'Universal Curve' for Total Electron-Yield XAS. *Solid State Commun.* **1996**, *98*, 405−409.
(37) Schroeder, S. L. M. Probing Depth of Total Electron-Yield XAFS: Monte-Carlo Simulations of Auger Electron Trajectories. *J. Phys. IV* **1997**, *7*, C2-153−C2-154.
(38) Schroeder, S. L. M.; Moggridge, G. D.; Lambert, R. M.; Rayment, T. Electron-Yield X-ray Absorption Spectroscopy in Gaseous Environments. In *Spectroscopy for Surface Science*; Advances in Spectroscopy; Clark, R. J. H., Hester, R. E., Eds.; Wiley & Sons: New York, Chichester, 1998; Vol. 26, pp 1−70.
(39) Sohn, K. E.; Dimitriou, M. D.; Genzer, J.; Fischer, D. A.; Hawker, C. J.; Kramer, E. J. Determination of the Electron Escape Depth for NEXAFS Spectroscopy. *Langmuir* **2009**, *25*, 6341−6348.
(40) Frazer, B. H.; Gilbert, B.; Sonderegger, B. R.; De Stasio, G. The probing depth of total electron yield in the sub-keV range: TEY-XAS and X-PEEM. *Surf. Sci.* **2003**, *537*, 161−167.
(41) Tanuma, S.; Powell, C. J.; Penn, D. R. Calculations of Electron Inelastic Mean Free Paths 0.5. Data for 14 Organic-Compounds Over the 50−2000 Ev Range. *Surf. Interface Anal.* **1994**, *21*, 165−176.
(42) Ravel, B.; Newville, M. ATHENA, ARTEMIS, HEPHAESTUS: Data Analysis for X-ray Absorption Spectroscopy Using IFEFFIT. *J. Synchrotron Radiat.* **2005**, *12*, 537−541.
(43) Ravel, B.; Newville, M. ATHENA and ARTEMIS: Interactive Graphical Data Analysis Using IFEFFIT. *Phys. Scr.* **2005**, *T115*, 1007−1010.
(44) Kanters, J. A.; Roelofsen, G.; Alblas, B. P.; Meinders, I. Crystal and Molecular-Structure of β-D-Fructose, with Emphasis on Anomeric Effect and Hydrogen-Bond Interactions. *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem.* **1977**, *33*, 665−672.
(45) Mostad, A. X-ray Crystallographic Study of alpha-Glucose at 140 K. *Acta Chem. Scand.* **1994**, *48*, 276−278.
(46) Ohanessian, J.; Gillier-Pandraud, H. Structure Cristalline de l'α-D-Galactose. *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem.* **1976**, *32*, 2810−2813.
(47) Hordvik, A. The Crystal and Molecular Structure of alpha-Xylose. *Acta Chem. Scand.* **1971**, *25*, 2175−2182.
(48) Gress, M. E.; Jeffrey, G. A. A Neutron Diffraction Refinement of Crystal Structure of β-Maltose Monohydrate. *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem.* **1977**, *33*, 2490−2495.
(49) Noordik, J. H.; Beurskens, P. T.; Bennema, P.; Visser, R. A.; Gould, R. O. Crystal-Structure, Polarity and Morphology of 4-O-β-D-Galactopyranosyl-α-D-Glucopyranose Monohydrate (α-Lactose Monohydrate): a Redetermination. *Z. Kristallogr.* **1984**, *168*, 59−65.

DOI: 10.1021/acs.jpcb.5b07159
J. Phys. Chem. B 2015, 119, 14373−14381

(50) Ankudinov, A. L.; Conradson, S. D.; de Leon, J. M.; Rehr, J. J. Relativistic XANES Calculations of Pu Hydrates. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1998**, *57*, 7518−7525.
(51) Battle, G. M.; Ferrence, G. M.; Allen, F. H. Applications of the Cambridge Structural Database in Chemical Education. *J. Appl. Crystallogr.* **2010**, *43*, 1208−1223.
(52) Nishiyama, Y.; Sugiyama, J.; Chanzy, H.; Langan, P. Crystal Structure and Hydrogen Bonding System in Cellulose I(alpha) from Synchrotron X-ray and Neutron Fiber Diffraction. *J. Am. Chem. Soc.* **2003**, *125*, 14300−14306.
(53) Garnier, S.; Petit, S.; Coquerel, G. Dehydration Mechanism and Crystallisation Behaviour of Lactose. *J. Therm. Anal. Calorim.* **2002**, *68*, 489−502.
(54) Jeffrey, G. A. Intramolecular Hydrogen-Bonding in Carbohydrate Crystal-Structures. *Carbohydr. Res.* **1973**, *28*, 233−241.

DOI: 10.1021/acs.jpcb.5b07159
J. Phys. Chem. B 2015, 119, 14373−14381

Supporting Information

NEXAFS Sensitivity to Bond Lengths in Complex Molecular Materials: A Study of Crystalline Saccharides

Adrian Gainar[1], *Joanna S. Stevens*[1], *Cherno Jaye*[2], *Daniel A. Fischer*[2] & *Sven L. M. Schroeder*[3,4*]

[1]School of Chemical Engineering and Analytical Science, The University of Manchester, Oxford Road, Manchester M13 9PL (UK)

[2]Material Measurement Laboratory, National Institute of Standards and Technology, Gaithersburg, Maryland 20899 (United States)

[3]School of Chemical and Process Engineering, University of Leeds, Leeds LS2 9JT (UK)

[4]DIAMOND Light Source Ltd., Harwell Science and Innovation Campus, Chilton, Didcot OX11 0DE (UK)

*Sven L. M. Schroeder: Tel. +44 (0)113 343 2401; Fax: +44 (0)113 343 2384

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

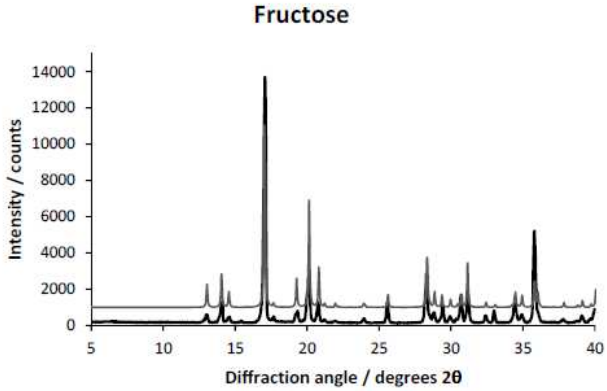


Figure S1. X-ray powder diffraction of experimental β-fructose (bold black), and calculated β-fructose (dark grey, Refcode FRUCTO11) patterns.

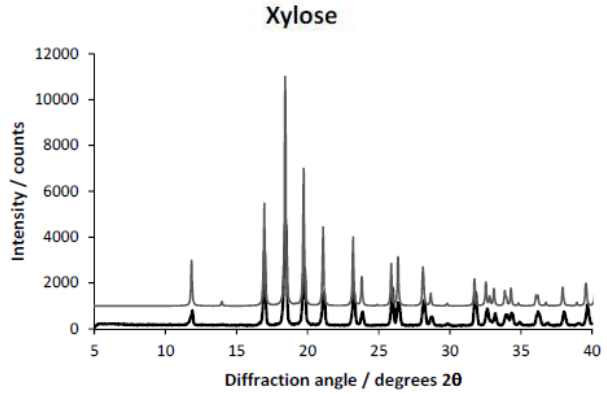


Figure S2. X-ray powder diffraction of experimental α-xylose (bold black), and calculated α-xylose (dark grey, Refcode XYLOSE) patterns.

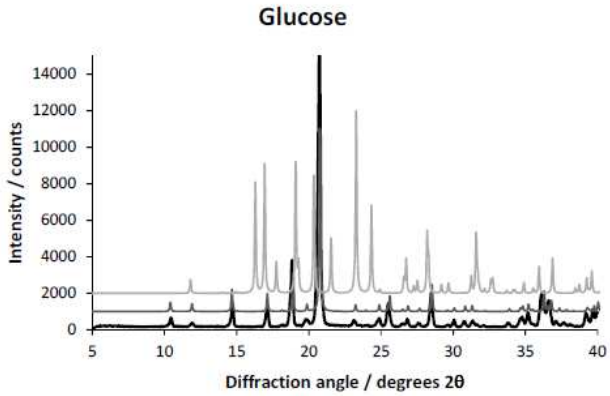


Figure S3. X-ray powder diffraction of experimental α-glucose (bold black), calculated α-glucose (dark grey, Refcode GLUCSA03), and calculated β-glucose (light grey, Refcode GLUCSE) patterns.

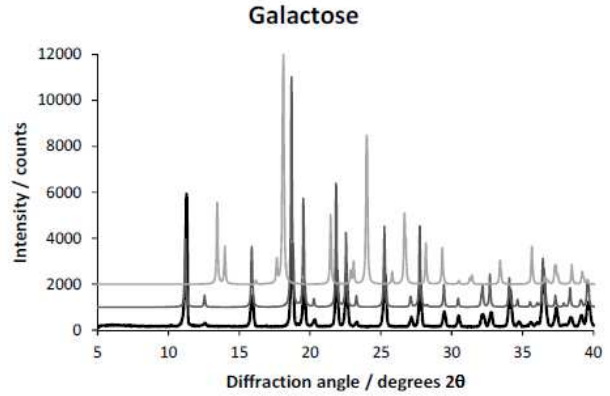


Figure S4. X-ray powder diffraction of experimental α-galactose (bold black), calculated α-galactose (dark grey, Refcode ADGALA01), and calculated β-galactose (light grey, Refcode BDGLOS01) patterns.

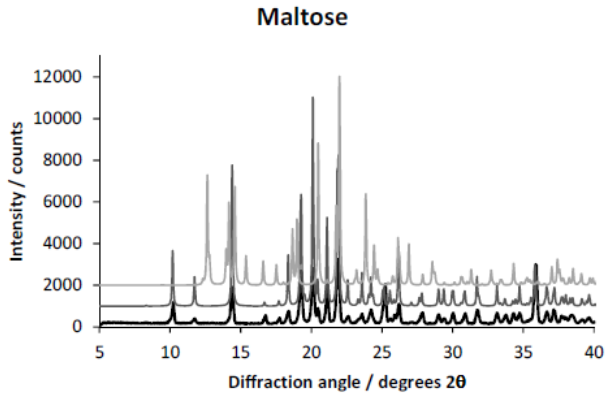


Figure S5. X-ray powder diffraction of experimental β-maltose monohydrate (bold black), calculated β-maltose monohydrate (dark grey, Refcode MALTOS11), and calculated α-maltose (light grey, Refcode MALTOT) patterns.

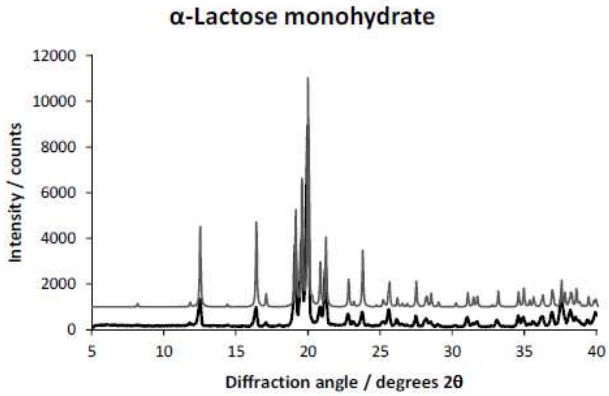


Figure S6. X-ray powder diffraction of experimental α-lactose monohydrate (bold black), and calculated α-lactose monohydrate (dark grey, Refcode LACTOS03) patterns.

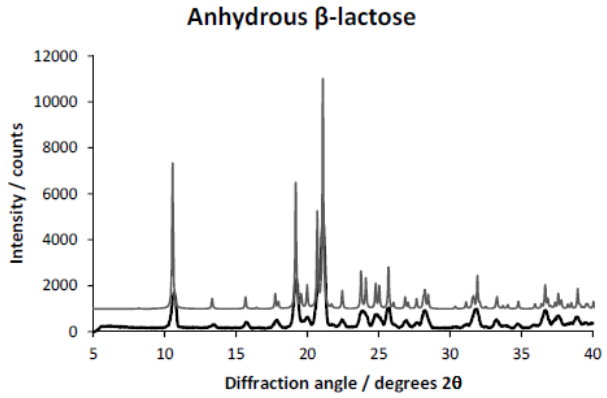


Figure S7. X-ray powder diffraction of experimental anhydrous β-lactose (bold black), and calculated anhydrous β-lactose (dark grey, Refcode BLACTO02) patterns.

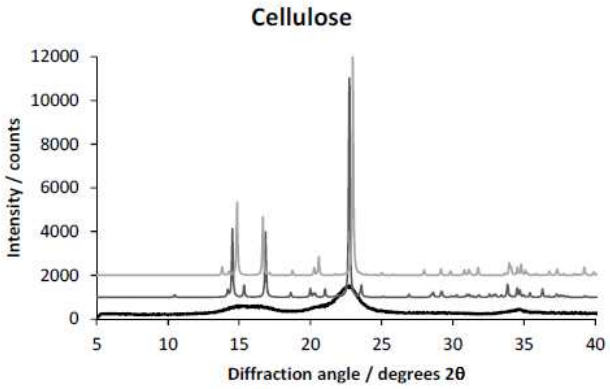


Figure S8. X-ray powder diffraction of experimental cellulose (bold black), calculated cellulose form I_α (dark grey, Refcode PADTUL), and calculated cellulose form I_β (light grey, Refcode JINROO01) patterns.

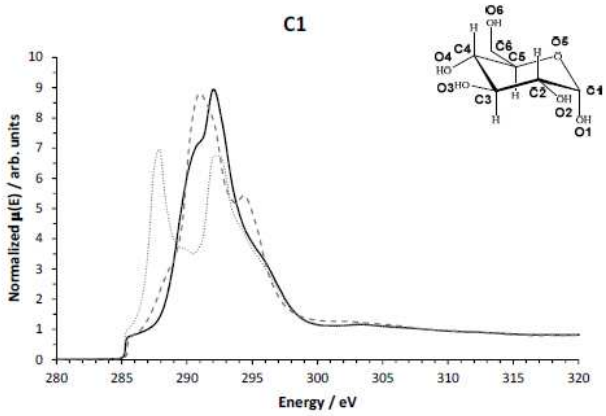


Figure S9. FEFF8 simulations for C1 absorbing atom in: an unaltered crystal structure of glucose (black line), an altered crystal structure of glucose with C1–O5 bond shortened by 0.1 Å (dashed grey line), and an altered crystal structure of glucose with C1–O5 bond elongated by 0.1 Å (dotted grey line); once C1–O5 bond altered, expected variations within the cycle on the C1–C2 and C5–O5 bond lengths are expected, albeit negligible.

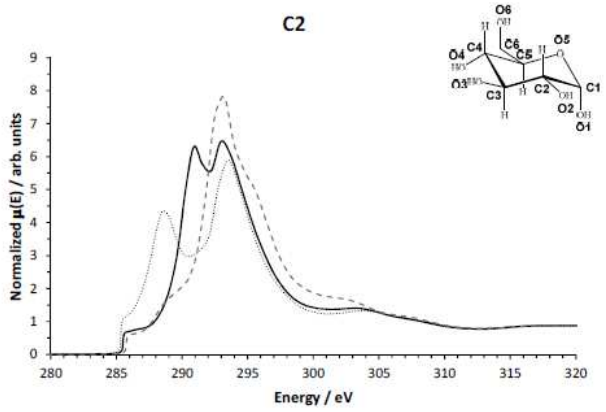


Figure S10. FEFF8 simulations for C2 absorbing atom in: an unaltered crystal structure of glucose (black line), an altered crystal structure of glucose with C2–O2 bond shortened by 0.1 Å (dashed grey line), and an altered crystal structure of glucose with C2–O2 bond elongated by 0.1 Å (dotted grey line).

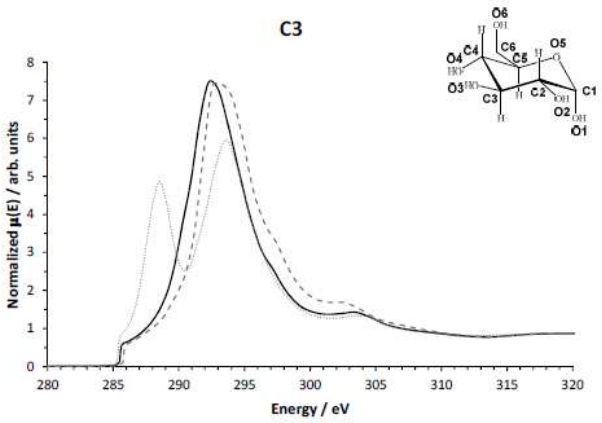


Figure S11. FEFF8 simulations for C3 absorbing atom in: an unaltered crystal structure of glucose (black line), an altered crystal structure of glucose with C3–O3 bond shortened by 0.1 Å (dashed grey line), and an altered crystal structure of glucose with C3–O3 bond elongated by 0.1 Å (dotted grey line).

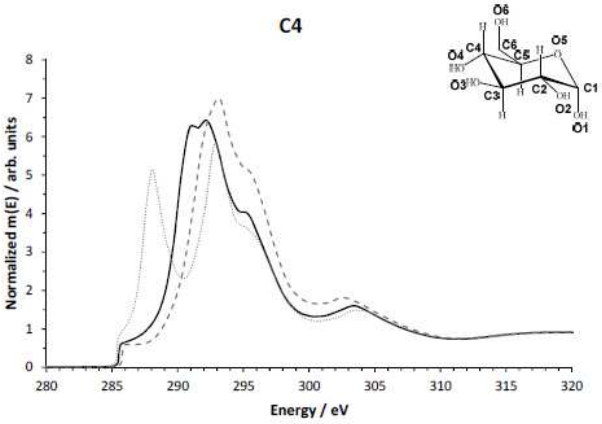


Figure S12. FEFF8 simulations for C4 absorbing atom in: an unaltered crystal structure of glucose (black line), an altered crystal structure of glucose with C4–O4 bond shortened by 0.1 Å (dashed grey line), and an altered crystal structure of glucose with C4–O4 bond elongated by 0.1 Å (dotted grey line).

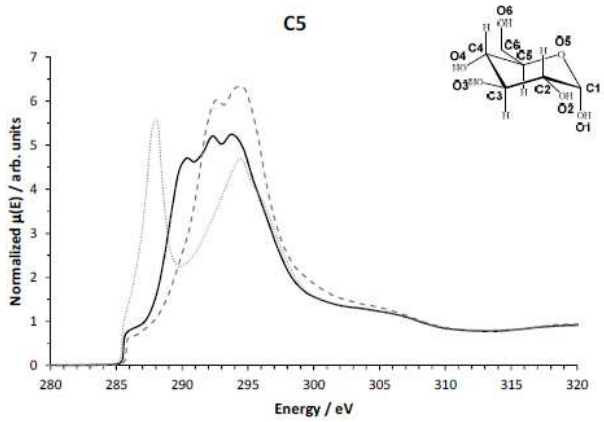


Figure S13. FEFF8 simulations for C5 absorbing atom in: an unaltered crystal structure of glucose (black line), an altered crystal structure of glucose with C5–O5 bond shortened by 0.1 Å (dashed grey line), and an altered crystal structure of glucose with C5–O5 bond elongated by 0.1 Å (dotted grey line); once C5–O5 bond altered, expected variations within the cycle on the C4–C5 and C1–O5 bond lengths are expected, albeit negligible.

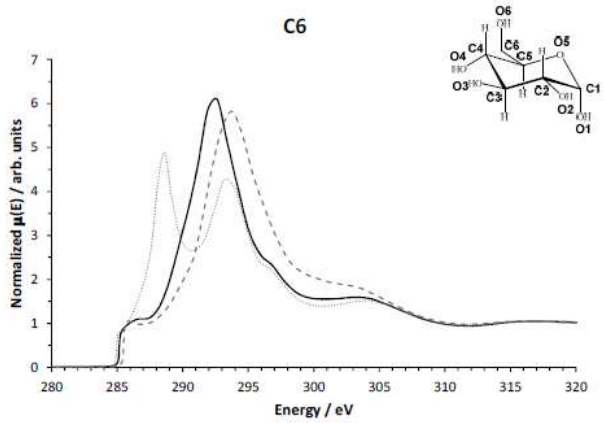


Figure S14. FEFF8 simulations for C6 absorbing atom in: an unaltered crystal structure of glucose (black line), an altered crystal structure of glucose with C6–O6 bond shortened by 0.1 Å (dashed grey line), and an altered crystal structure of glucose with C6–O6 bond elongated by 0.1 Å (dotted grey line).

ChemPubSoc Europe

DOI: 10.1002/chem.201405635

CHEMISTRY A European Journal Full Paper

Speciation

Chemical Speciation and Bond Lengths of Organic Solutes by Core-Level Spectroscopy: pH and Solvent Influence on *p*-Aminobenzoic Acid

Joanna S. Stevens,[a] Adrian Gainar,[a] Edlira Suljoti,[b] Jie Xiao,[b] Ronny Golnak,[b] Emad F. Aziz,[b, c] and Sven L. M. Schroeder*[a, d]

Abstract: Through X-ray absorption and emission spectroscopies, the chemical, electronic and structural properties of organic species in solution can be observed. Near-edge X-ray absorption fine structure (NEXAFS) and resonant inelastic X-ray scattering (RIXS) measurements at the nitrogen K-edge of *para*-aminobenzoic acid reveal both pH- and solvent-dependent variations in the ionisation potential (IP), $1s \rightarrow \pi^*$ resonances and HOMO–LUMO gap. These changes unequivocally identify the chemical species (neutral, cationic or anionic) present in solution. It is shown how this incisive chemical state sensitivity is further enhanced by the possibility of quantitative bond length determination, based on the analysis of chemical shifts in IPs and σ^* shape resonances in the NEXAFS spectra. This provides experimental access to detecting even minor variations in the molecular structure of solutes in solution, thereby providing an avenue to examining computational predictions of solute properties and solute–solvent interactions.

Introduction

The ability to observe chemical and structural changes in organic molecules in solution is of central importance for progress in chemical and life sciences. For example, the activity of biological systems is strongly influenced by pH variations causing changes in chemical speciation,[1] whereas the structural outcomes of self-assembly and crystallisation processes define physical and chemical properties of products.[2,3] In particular, details of the structural evolution of nucleating solutions are of contemporary interest[3,4] because of their far-reaching impact on both structural and morphological outcomes, the tailoring of product properties and the design of materials.

[a] Dr. J. S. Stevens,[+] A. Gainar,[+] Prof. S. L. M. Schroeder
School of Chemical Engineering and Analytical Science
School of Chemistry
The University of Manchester
Oxford Road, Manchester M13 9PL (UK)
E-mail: slm.schroeder@outlook.com

[b] Dr. E. Suljoti, Dr. J. Xiao, R. Golnak, Prof. E. F. Aziz
Joint Ultrafast Dynamics Lab in Solutions and at Interfaces (JULiq)
Institute of Methods for Material Development
Helmholtz-Zentrum Berlin für Materialien und Energie
Albert Einstein Strasse 15, 12489 Berlin (Germany)

[c] Prof. E. F. Aziz
Fachbereich Physik
Freie Universität Berlin
Arnimallee 14, 14195 Berlin (Germany)

[d] Prof. S. L. M. Schroeder
School of Chemical and Process Engineering
University of Leeds
Leeds LS2 9JT (UK)

[+] These authors contributed equally to this work.

A crucial prerequisite for further elucidation of molecular structure in such processes is having the experimental capability to incisively study conformational structure, bonding and electronic properties of organic molecular solutes in situ.[3,5] Core-level spectroscopies are emerging as suitable probes for this task, as they are highly sensitive to local electronic structure and bonding. Even X-ray photoelectron spectroscopy (XPS),[6] one of the most facile core-level spectroscopies in the laboratory, permits the characterisation of protonation and hydrogen bonding in both solid-state and solution acid–base systems through determination of chemical shifts in core-level binding energies.[7,8]

Near-edge X-ray absorption fine structure (NEXAFS), also known as X-ray absorption near-edge structure (XANES), is chemically and structurally much more incisive,[9,10] as it involves the excitation of core electrons to unoccupied valence orbitals that are readily interpreted by use of molecular orbital (MO) calculations. In addition, σ^* shape resonances in the NEXAFS are very sensitive to bond length variations, thus providing incisive molecular structure parameters, even for non-crystalline systems.[11] Resonant inelastic X-ray scattering (RIXS) probes occupied valence orbitals when the correlation of the valence electrons is weak, by monitoring the transitions of electrons from occupied valence orbitals into core-hole states.[12] The combination of NEXAFS and RIXS thus provides a comprehensive picture of the local electronic structure and is ideally placed for investigations of local environment and bonding of molecular solutes. Recent development of the liquid microjet technique[13,14] allows windowless in situ probing of solute species in solution and, importantly for organics, alleviates potential problems caused by radiation damage. In

addition, the ultra-fast nature of the electronic transitions ($\approx 10^{-15}$ s)[9] means that every detected event represents a snapshot of a molecule rather than an average over vibrational or conformational changes, thereby permitting population analysis when transient species coexist.[8,15]

Herein, we demonstrate how NEXAFS and RIXS microjet studies can be combined with density functional theory (DFT) calculations to elucidate the pH- and solvent-dependent changes in the local chemical and electronic environment of *para*-aminobenzoic acid (PABA, Figure 1). PABA is an intermediate for folic acid synthesis and has pharmaceutical applications.[16] Depending on pH and solvent it forms neutral, zwitterionic, cationic and anionic species in solution.[17,18] We will show how direct observation of these chemically and electronically distinct species is possible by monitoring the nitrogen 1s core-level excitation (nitrogen K-edge), thereby providing electronic structure signatures for aqueous solutions with varying pH as well as for alcoholic solutions. We will correlate electronic structure and bond length changes with the effects of ionisation at Brønsted donor and acceptor moieties, and compare the results with data for non-ionic species in the solid state and DFT molecular orbitals (MOs).

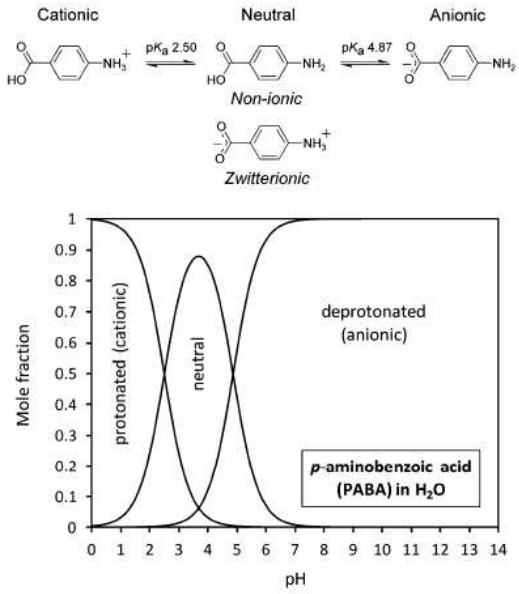


Figure 1. Equilibria between the cationic, neutral and anionic PABA species in water (top) and speciation diagram as a function of pH (bottom).

Results and Discussion

Expected speciation in solution and in the solid state

The solubility of PABA in aqueous solution is very low (0.03 M) and dissolution leads to a complex mixture at the isoelectric point (pH 3.69), which comprises non-ionic and zwitterionic neutral species[17] alongside small percentages of cationic and anionic species (Figure 1).[19] Based on the pK_a values,[19] upon increasing the pH above 4.87 the carboxylate anionic form dominates, whereas decreasing the pH below 2.50 results in solutions containing almost entirely the cationic form (Figure 1). These ionic forms have greatly increased solubility (> 0.5 M) and allow microjet NEXAFS and RIXS measurements to be recorded with good signal-to-noise ratio. In the solid state, PABA exists as neutral, non-zwitterionic molecules bonded through a series of intermolecular hydrogen bonds and π–π interactions,[20–22] thus providing a spectrum of the non-ionic form for comparison with the ionic species expected at high and low pH values. For comparison we have also examined alcoholic solutions, in which PABA is expected to exist primarily in the non-ionic neutral form (i.e., non-zwitterionic).[17]

NEXAFS

Comparison of the nitrogen K-edge NEXAFS spectra for non-ionic, solid-state PABA and in situ measurements at high/low pH in aqueous solution (Figure 2) reveals marked differences between the electronic structure of the species at pH 11 and pH 1, as well as variations between the non-ionic and pH 11 forms. Two pre-edge resonances (energetically below the ionisation potential, IP) are evident for the non-ionic form and at pH 11, which arise from 1s→π^* transitions due to conjugation of the nitrogen lone pair with the aromatic ring, and these are followed by the broader 1s→σ^* resonances that are more susceptible to variations in geometric structure.[9]

The IP for the species present at pH 1 is very different from that observed at pH 11, with a positive energy shift of 2.3 eV (Figure 2, Table 1). This increased IP for the pH 1 solution species reflects acquisition of a positive charge on nitrogen for the cationic form (decreased electron density, orbital contraction). A slight decrease in IP occurs at pH 11 compared to the non-ionic form (Table 1), with conjugation of the carboxyl and

Table 1. Experimental nitrogen K-edge NEXAFS resonance and IP energies from Figure 2 along with the term value δ_{C-N}.

Nitrogen [eV]	Non-ionic β-PABA	Anionic pH 11	Cationic pH 1	Non-ionic methanol
IP	403.60	402.90	405.20	403.47
$1\pi^*$	400.58	401.22	–	400.68
$3\pi^*$	402.30	402.14	–	401.98
$1\sigma^*_{N-H}$	405.11	404.15	405.20	403.66
$2\sigma^*_{C-N}$	409.09	408.30	408.36	408.66
$\Delta\pi^*$	1.72	0.92	–	1.30
δ_{C-N} (σ^*_{C-N}–IP)	5.49	5.40	3.16	5.19

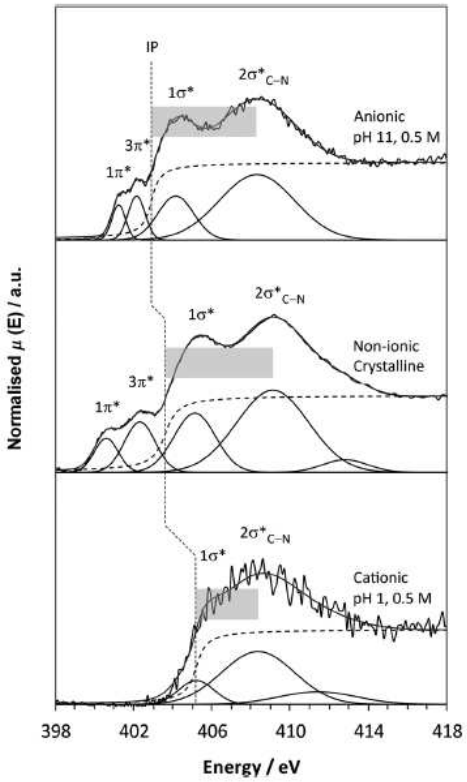


Figure 2. Nitrogen K-edge NEXAFS spectra for anionic (top), non-ionic solid-state β- (middle) and cationic PABA (bottom), together with the fits to the data; IPs are marked by vertical dashed lines, and the magnitude of the term value δ_{C-N} is indicated by shaded boxes.

amine groups through the aromatic ring leading to some orbital screening and redistribution of electron density occurring with change to the anionic form (Figure 1). Due to the absence of the lone electron pair on the nitrogen atom with NH_3^+, the cationic nitrogen will not be able to participate in the π MOs with the aromatic ring; therefore, no pre-edge π* peaks are visible in the pH 1 nitrogen K-edge (Figure 2), further signifying transformation to the cationic form. Interestingly, the π* resonances become closer to one another in the anionic form at pH 11 compared to the non-ionic form, which leads to a smaller Δπ* value for the anionic species (Figure 2, Table 1).

DFT calculations of the unoccupied MOs with nitrogen contributions allow us to pinpoint the electronic transitions leading to pre-edge π* nitrogen intensity in the anionic and non-ionic NEXAFS and to interpret the magnitude of the calculated energy shifts. Figure 3 shows the lowest unoccupied molecular orbitals (LUMOs) for non-ionic and anionic PABA, which reveals that the calculated gap between the two lowest-energy nitrogen π* orbitals (Δπ*) is over twice as large for non-ionic PABA as for the anionic form. In addition, the π* energies for anionic

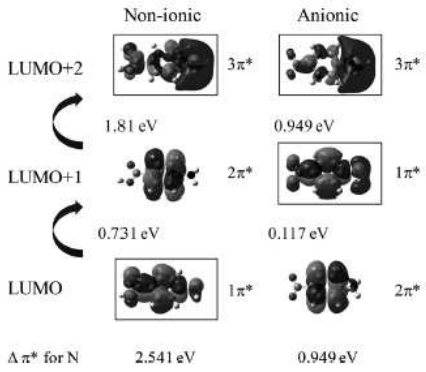


Figure 3. Lowest unoccupied molecular orbitals (LUMOs) for the non-ionic and anionic monomers of PABA. The π* orbitals are labelled 1π*, 2π* and 3π* from the lowest energy relative to non-ionic PABA and those with weight on the nitrogen absorber are marked with boxes; transitions from the N 1s core level to these marked orbitals are visible in the nitrogen K-edge NEXAFS.

PABA are raised above that for PABA, with the nitrogen 1π* becoming the LUMO+1 rather than the LUMO (Figure 3).

The experimental trend in the IP (Figure 2) for the different PABA species is also reproduced by the calculations (Figure 4),

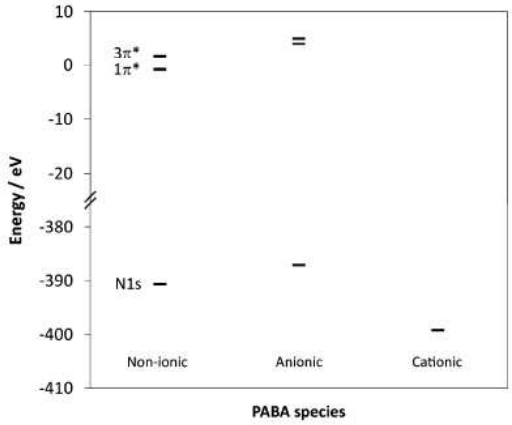


Figure 4. Calculated orbital energies for PABA showing the trends in nitrogen π* (top) and core 1s (bottom) orbital energies for non-ionic, anionic and cationic forms.

with a decrease in core-level N1s energy for cationic species (thus increased IP) and a slight increase for anionic relative to non-ionic PABA (decreased IP). The nitrogen π* LUMOs for the anionic form mirror the change in the core-level energy, with an increase in energy relative to non-ionic PABA (Figure 4). The energy gap between the 1π* and 3π* MOs also narrows for the anionic form (Figures 3 and 4); this decrease in the gap between the nitrogen π* resonances (Δπ*) for the anionic form

Table 2. Calculated orbital energies and shifts for different PABA species.

For N [eV]	Non-ionic	Anionic	Cationic	Zwitterionic
N 1s	−390.66	−387.10	−399.24	−396.02
N 1s→1π*	389.86	391.16	–	–
N 1s→3π*	392.40	392.11	–	–
Δπ* (1π*→3π*)	2.54	0.95	–	–

parallels that observed in the experimental NEXAFS (Figure 2), explaining it by a significant increase in the 1s→1π* energy and small decrease in the 1s→3π* energy relative to non-ionic PABA occurring with formation of the carboxylate form (Table 2). Comparison of calculations for the zwitterionic form (Table 2) illustrate that its presence would significantly raise the IP, albeit perhaps not quite as much as for the cationic form, alongside the loss of resolved pre-edge π* peaks through the positive charge on the nitrogen. Although the magnitude of the calculated energy shifts may not be accurate due to the absence of core-hole relaxation and screening from intermolecular interactions, the direction and trends observed correlate well with the experimental data, thus indicating that variations in ground-state (initial-state) properties allow the interpretation of electronic structure probed through core-level excitation, which tends to dominate over final-state effects.

C–N bond lengths

The experimental nitrogen K-edge NEXAFS spectra (Figure 2) also show variation in the position of the broader, post-edge σ* shape resonances relative to the IP for the non-ionic, anionic and cationic species. For a bond between two atoms X and Y, the term value δ_{X-Y} the difference between the energy of its σ^*_{X-Y} orbital and the IP of either X or Y, provides a highly sensitive measure of the bond length.[9,23] As can be seen in Figure 2, and from the term values (δ_{C-N}) in Table 1, the separation between the energy of the σ^*_{C-N} resonance and the IP is of a similar magnitude for anionic and non-ionic PABA. In contrast, the energy of σ^*_{C-N} and the IP are much closer for the cationic species, which reflects an increase in IP and decrease in the energy of the σ^*_{C-N} resonance following protonation of the amino group, and results in a substantially smaller δ_{C-N} value (Figure 2). As the C–N bond lengths are known from X-ray diffraction (XRD) crystal structures for the solid state (β non-ionic form[22] and cationic HCl salt,[24] as well as the additional non-ionic α form[20] previously investigated,[25] Table 3), these can be used to obtain a calibration plot for bond length determination in the solutions. Plotting the term value δ_{C-N} against the C–N bond lengths in the corresponding crystal structures indicates a linear dependence (Figure 5, black data points). The correlation between XRD-derived bond lengths and NEXAFS-derived term values is very high (Figure 5), which suggests sensitivity to bond length variations of less than 0.005 Å (Table 3). The term values for the solution species are similar to those found in their corresponding solid-state structures, thus suggesting that determination of the term values δ_{C-N} measures their bond length reliably (Table 3, Figure 5 grey data points) in both the solid state and in solution. We expect some variation in bond lengths as a result of the differences in local environment, for example, due to interaction with solvent molecules. The observed impact on the C–N bond length is on the order of less than tenths of an Ångstrom, but appears to be within the detection limit. Certainly the expected trend, overall longest C–N bonds for protonated nitrogen in the cationic species and shorter bonds for the neutral as well as the anionic species, is evident for both the solid and the liquid phase. This indicates strong potential for a generalisation of the σ* shape resonance analysis to other organic molecular solutes, and even "measuring" bond lengths in solution.

Table 3. Experimental C–N bond lengths from XRD and those derived from NEXAFS through δ_{C-N}.

Nitrogen	δ_{C-N} [eV]	XRD C–N [Å]	NEXAFS C–N [Å]
non-ionic α-PABA	6.31[a]	1.379(2)[b]	1.3822
non-ionic β-PABA	5.49	1.408(3)[c]	1.4029
anionic pH 11	5.40	[1.410(2)[d]]	1.4052
non-ionic methanol	5.19	–	1.4105
cationic pH 1	3.16	–	1.4617
cationic HCl salt	3.05	1.4626(5)[e]	1.4644

[a] Ref. [25]. [b] Ref. [20]. [c] Ref. [22]. [d] C–N bond length from the hydrated Na salt crystal structure[26] shown for comparison. [e] Ref. [24].

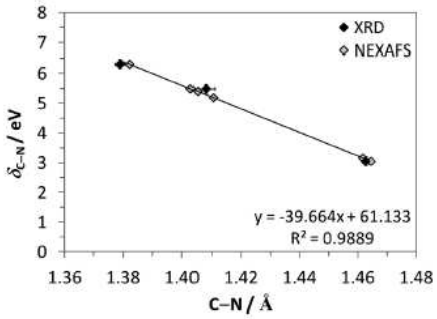


Figure 5. Correlation between the NEXAFS term value (δ_{C-N}) and C–N bond lengths from XRD (black),[20,22,24] and comparison with C–N bond lengths obtained from NEXAFS (grey) for the non-ionic and ionic PABA species (Table 3).

Methanol species

Comparison with the nitrogen K-edge NEXAFS for PABA obtained in methanol (Figure 6) reveals a shape in agreement with that of non-protonated nitrogen species (Figure 2). The IP and 1π* resonance (with their previously demonstrated sensitivity to the chemical species present) are consistent with those observed for the non-ionic form rather than anionic species (Table 1). If the molecules in the 0.5 M methanol solution were primarily zwitterionic (Figure 1), a spectrum resembling that of the cationic form (Figure 2) would be expected, with no pre-edge π* resonances and an increased IP resulting from

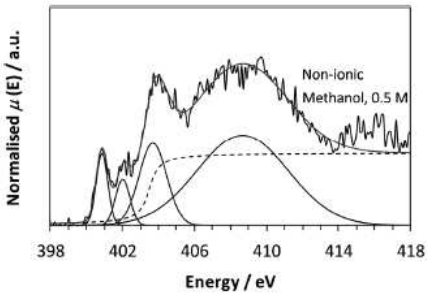


Figure 6. Nitrogen K-edge NEXAFS of PABA in methanol (0.5 M).

the positive charge on nitrogen. That this is not observed provides direct evidence that (at least the majority of) PABA molecules in methanol exist in the neutral, uncharged (non-ionic) form. The smaller δ_{C-N} for the methanol species relative to the solid state (Table 1) suggests that the C–N bond is slightly longer in the methanol solution, and applying the correlation above gives a predicted C–N bond length of 1.4105 Å (Table 3).

RIXS

While NEXAFS probes electron transitions from the core level to unoccupied orbitals, RIXS involves transitions from occupied valence orbitals to the core hole after excitation.[9] For nitrogen, this follows the 2p→1s transitions. The nitrogen RIXS spectra for the non-ionic and anionic species (methanol and pH 11 solutions, respectively) initially look similar, whereas that for pH 1 is clearly distinguished (Figure 7). Following protonation of nitrogen for the cationic species at low pH, the highest-energy peaks arising from the occupied 2p π valence orbitals→1s are absent due to the use of the nitrogen lone pair in forming an additional N–H bond and not part of the π MOs.

Closer inspection of the nitrogen RIXS for anionic and non-ionic species reveals a significant shift to lower energy for the anionic form (Figure 8, left). This is particularly noticeable towards the higher-energy region of the RIXS data. As the highest energy (inelastic) peak in the RIXS arises from the decay of electrons from the nitrogen HOMO to the core N1s level, this indicates a lower energy for this MO for the anionic species. Comparison of the relative energies for the HOMOs with nitrogen contributions for non-ionic and anionic species reveals that the HOMO→1s leads to the first RIXS peak for the non-ionic form (Figure 9). In contrast, this MO is lowered in energy for the anionic form, becoming the HOMO−3 (Figure 9), thereby elucidating the shift to lower energy seen experimentally (Figure 8).

Further insight can be obtained through combination of the RIXS and NEXAFS data (Figure 8), which provides a model of the local density of valence states. The energy difference between the highest-energy RIXS peak and lowest-energy NEXAFS peak is a probe of the gap between the HOMO and LUMO per atomic chemical state (Figure 8); an alternative

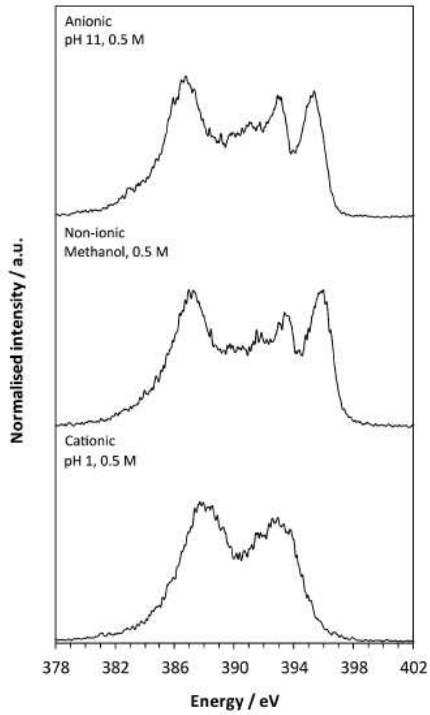


Figure 7. Nitrogen RIXS for anionic, non-ionic and cationic species in solution.

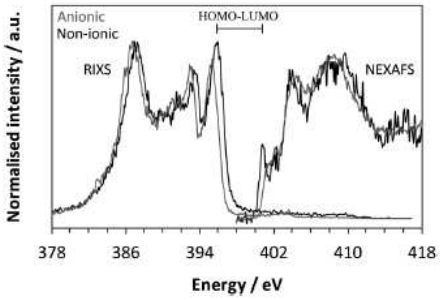


Figure 8. Nitrogen RIXS and NEXAFS showing the shift in energy between anionic and non-ionic solution species and the HOMO–LUMO gap.

method for the bandgap is to take the difference of the intersections between the first peak slopes and the backgrounds (tending to underestimate).[27] For nitrogen in the non-ionic form, this represents the HOMO↔LUMO gap, whereas for the anionic form, this is for HOMO−3↔LUMO+1 (Figure 9). Experimentally, this leads to energy gaps of 5.06 and 5.89 eV for the non-ionic and anionic PABA species, respectively, which compares favourably with the predicted values of 5.01 and 5.86 eV.

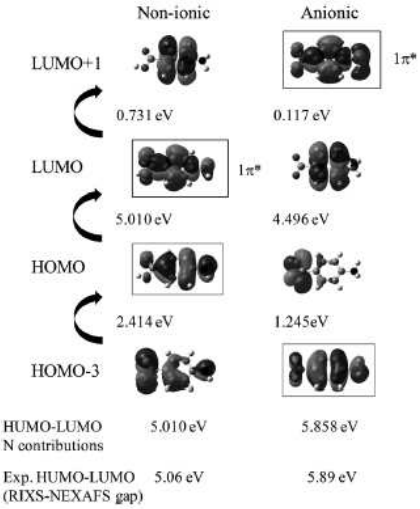


Figure 9. Highest occupied and lowest unoccupied molecular orbitals involved in the first RIXS (grey outline) and NEXAFS (black outline) resonances, respectively, with the experimental and calculated nitrogen HOMO–LUMO gap for non-ionic and anionic PABA.

Conclusion

Directly monitoring the core-level transitions of the amino group by nitrogen NEXAFS and RIXS as a function of pH successfully characterises the chemical and electronic state of PABA species in aqueous solution. Formation of the cationic species by protonation of the amino group at low pH leads to a significant shift in IP to higher energy, along with an absence of π^* resonances in NEXAFS and of π valence peaks in RIXS. Although the amino group is not protonated in the anionic and non-ionic forms, differences are observed in both the NEXAFS and RIXS. There is a slight shift to low energy for the IP and a widening of the HOMO–LUMO gap for the anionic species, in agreement with predictions of DFT calculations. In methanol, the NEXAFS resembles that of non-ionic PABA, with no indications of the presence of the zwitterionic form. Structural as well as chemical and electronic changes impact the NEXAFS spectra, with variation of the C–N bond length influencing the energy of the σ^*_{C-N} resonance relative to the IP, thereby providing access to bond length determination of solutes in solution by comparison with NEXAFS data of PABA species in the solid state.

Experimental Section

Solid-state NEXAFS

The solid-state PABA sample (β-PABA[22]) was formed through an aqueous slurry of the commercial form (>99%, Sigma–Aldrich, UK) at 5 °C. Solid-state NEXAFS measurements were performed at the U7A beamline of the National Synchrotron Light Source (NSLS) at Brookhaven National Laboratory, NY. Partial electron yield (PEY) spectra for the nitrogen K-edge were collected by a channeltron electron multiplier with the sample at the magic angle (54.7°) relative to the incident beam. An entrance grid bias of −150 V was used for PEY collection and a monochromator with a 600 L mm^{-1} grating, which provided energy resolution of about 0.15 eV. After collection, the spectra were normalised by the simultaneously recorded drain current from an in situ gold-coated, 90% transmission grid (I_0) placed in the incident X-ray beam to eliminate the effect of incident beam intensity fluctuations and beamline optics absorption features, and the monochromator energy was calibrated using the 400.6 eV first π^* feature of a titanium nitride grid.

Peak fitting and normalisation was performed using the Athena software,[28] with arctan steps for the IPs (edge steps) and Gaussian functions for the peaks.[9,25,29] For investigation of the relationship of C–N bond length with the NEXAFS term value δ_{C-N} (σ^*_{C-N}–IP), the standard deviation for the C–N bond lengths from single-crystal X-ray diffraction (XRD) was $\leq$0.003 Å[20,22,24] and good correspondence was observed for the fitted IP energy shifts with the XPS N 1s core binding energies ($\leq$0.15 eV standard deviation).[25] The correlation between the nitrogen NEXAFS and C–N bond length revealed that a 0.1 eV change in the term value δ_{C-N} corresponds to a 0.0025 Å alteration in C–N bond length.

Solution-state in situ NEXAFS

Solution-state nitrogen K-edge spectra were recorded with the LiXEdrom endstation[14] at the U41-PGM beamline of the BESSY II synchrotron at Helmholtz Zentrum Berlin (HZB), by using the liquid microjet technique. The 0.5 M acidic (HCl) and basic (NaOH) aqueous solutions and 0.5 M methanol solution of PABA were prepared and filtered at ambient pressure to remove impurities or undissolved crystals. Partial fluorescence yield data were recorded in scanning mode by a grating with line density 1200 lines mm^{-1} and radius 7.5 m dispersing the emitted photon energy from the sample,[14] and subsequently a detector consisting of a charge-coupled device, fluorescence screen and microchannel plate stack collecting the amplified signal. The sample, grating and detector were arranged in Rowland circle geometry for accurate focusing. (Although total fluorescence yield could be recorded with a GaAsP photodiode mounted in the vicinity of the liquid jet, there were problems with the signal recording and solute could crystallise on the photodiode, thus interfering with the signal.) Use of the microjet ensured fresh sample was probed by the X-ray beam, vastly minimising the potential for any X-ray-induced damage. An 18 μm diameter glass nozzle with 0.6 mL min^{-1} flow rate was used, and the measurement performed around 2 mm from the nozzle within the laminar part of the jet flow (droplets started to form after 3–5 mm, with the resulting frozen residues collected by a liquid nitrogen trap) with 10^{-5} mbar pressure in the main chamber. Beamline energy calibration was performed with N_2 gas (total electron yield X-ray absorption spectroscopy 1s→2p π transition) and the resolution was $\leq$0.1 eV.

Solution-state in situ RIXS

Microjet nitrogen RIXS spectra were recorded at BESSY II[14,30] by using the same setup and solutions as for the NEXAFS. Nitrogen RIXS data were recorded at multiple excitation energies corresponding to NEXAFS resonances. As the local valence region available for transitions by excited core electrons has predominantly the N 2p character, little valence electron excitation was expected due to the general weak correlation for the p orbital. In line with this, similar spectral features at constant emission energy with varying excitation energy were observed, and only changes in rela-

tive intensity occurred across the NEXAFS transition thresholds (excitation energies of 404.5, 410 and 408 eV for the anionic, non-ionic and cationic species, respectively, were used for further analysis).

DFT calculations

Non-ionic, anionic and cationic PABA monomers were optimised with the B3LYP functional and 6-31G* basis set in Gaussian 09[31] to obtain MOs[32] for the ground state. This permitted identification of electronic structure changes originating from chemical (initial-state) variations.[9] Strictly speaking, NEXAFS interpretation should also include the influence of relaxation effects due to the presence of the core hole, which were expected to affect the final state of the observed electronic transitions. We have recently included such effects in CASTEP calculations for crystalline PABA[25] and found that the inclusion of final-state effects did not alter the MO interpretation of the NEXAFS substantially. The reason for this insensitivity lies in a combination of the strong localised character of core-level excitations and the weakness of intermolecular interactions (with hydrogen bonding dominant) of organic molecules relative to internal covalent bonding and protonation effects. The ground-state calculations used for the interpretation of the NEXAFS are therefore expected to somewhat overestimate the absolute energies associated with the π^* transitions, but less so the relative energies and sequence of the unoccupied π^* states.

Acknowledgements

We thank Rachel Sullivan for providing the β-PABA sample, and gratefully acknowledge support for J.S.S., A.G. and S.L.M.S. through an ESPRC Critical Mass Grant (EP/1013563/1). Special thanks to Ervis Suljoti and Tim Brandenburg (BESSY II) for helping with the preparation for microjet experiments, and to Chemo Jaye (NSLS) for assistance with collecting the solid-state data. Use of the National Synchrotron Light Source, Brookhaven National Laboratory, was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. DE-AC02-98CH10886. Commercial names mentioned in this manuscript are for illustrative purposes and do not represent an endorsement by the National Institute of Standards and Technology. We acknowledge the Helmholtz-Zentrum Berlin for provision of synchrotron access at beamline U41-PGM of BESSY II. The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007–2013) under grant agreement no. 312284. E.F.A. was supported by the Helmholtz-Gemeinschaft through the VH-NG-635 grant and by the European Research Council through starting Grant No. 279344.

Keywords: ionization potentials · liquids · X-ray scattering · speciation · X-ray absorption spectroscopy

[1] M. N. Chatterjea, R. Shinde, *Textbook of Medical Biochemistry*, 8th ed., Jaypee Brothers Medical Publishers (P) Ltd, New Delhi, **2012**; J. M. Berg, J. L. Tymoczko, L. Stryer, *Biochemistry*, 5th ed., Freeman, San Francisco, **2002**.

[2] J. T. Carstensen, *Pharmaceutical Principles of Solid Dosage Forms*, Technomic Publishing Co., Lancaster, PA, **1993**.

[3] R. J. Davey, S. L. M. Schroeder, J. H. ter Horst, *Angew. Chem. Int. Ed.* **2013**, *52*, 2166–2179; *Angew. Chem.* **2013**, *125*, 2220–2234.

[4] R. J. Davey, K. Allen, N. Blagden, W. I. Cross, H. F. Lieberman, M. J. Quayle, S. Righini, L. Seton, G. J. T. Tiddy, *CrystEngComm* **2002**, *4*, 257–264; I. Weissbuch, M. Lahav, L. Leiserowitz, *Cryst. Growth Des.* **2003**, *3*, 125–150; D. Gebauer, H. Cölfen, *Nano Today* **2011**, *6*, 564–584; J. Anwar, D. Zahn, *Angew. Chem. Int. Ed.* **2011**, *50*, 1996–2013; *Angew. Chem.* **2011**, *123*, 2042–2061; D. Gebauer, M. Kellermeier, J. D. Gale, L. Bergstrom, H. Colfen, *Chem. Soc. Rev.* **2014**, *43*, 2348–2371.

[5] N. Pienack, W. Bensch, *Angew. Chem. Int. Ed.* **2011**, *50*, 2014–2034; *Angew. Chem.* **2011**, *123*, 2062–2083.

[6] D. Briggs, M. P. Seah, P. M. A. Sherwood, in *Practical Surface Analysis, Volume 1: Auger and X-ray Photoelectron Spectroscopy*, 2nd ed., (Eds.: D. Briggs, M. P. Seah), John Wiley & Sons, Chichester, **1990**.

[7] J. S. Stevens, S. J. Byard, C. C. Seaton, G. Sadiq, R. J. Davey, S. L. M. Schroeder, *Phys. Chem. Chem. Phys.* **2014**, *16*, 1150–1160; J. S. Stevens, S. J. Byard, C. C. Seaton, G. Sadiq, R. J. Davey, S. L. M. Schroeder, *Angew. Chem. Int. Ed.* **2011**, *50*, 9916–9918; *Angew. Chem.* **2011**, *123*, 10090–10092; D. Nolting, E. F. Aziz, N. Ottosson, M. Faubel, I. V. Hertel, B. Winter, *J. Am. Chem. Soc.* **2007**, *129*, 14068–14073.

[8] D. Nolting, N. Ottosson, M. Faubel, I. V. Hertel, B. Winter, *J. Am. Chem. Soc.* **2008**, *130*, 8150–8151.

[9] J. Stöhr, *NEXAFS Spectroscopy*, 2nd ed., Springer-Verlag, Berlin, **1992**.

[10] G. Hähner, *Chem. Soc. Rev.* **2006**, *35*, 1244–1255; J. Lehmann, D. Solomon, J. Brandes, H. Fleckenstein, C. Jacobson, J. Thieme, in *Biophysico-Chemical Processes Involving Natural Nonliving Organic Matter in Environmental Systems*, John Wiley & Sons, Hoboken, NJ, **2009**, pp. 729–781; P. Leinweber, J. Kruse, F. L. Walley, A. Gillespie, K.-U. Eckhardt, R. I. R. Blyth, T. Regier, *J. Synchrotron Radiat.* **2007**, *14*, 500–511; B. M. Messer, C. D. Cappa, J. D. Smith, K. R. Wilson, M. K. Gilles, R. C. Cohen, R. J. Saykally, *J. Phys. Chem. B* **2005**, *109*, 5375–5382; E. Otero, S. G. Urquhart, *J. Phys. Chem. A* **2006**, *110*, 12121–12128.

[11] D. Arvanitis, N. Haack, G. Ceballos, H. Wende, K. Baberschke, A. L. Ankudinov, J. J. Rehr, *J. Electron Spectrosc. Relat. Phenom.* **2000**, *113*, 57–65; N. Haack, G. Ceballos, H. Wende, K. Baberschke, D. Arvanitis, A. L. Ankudinov, J. J. Rehr, *Phys. Rev. Lett.* **2000**, *84*, 614–617.

[12] A. Nilsson, L. G. M. Pettersson, *Surf. Sci. Rep.* **2004**, *55*, 49–167.

[13] M. Faubel, B. Steiner, J. P. Toennies, *J. Chem. Phys.* **1997**, *106*, 9013–9031; K. M. Lange, R. Könnecke, M. Soldatov, R. Golnak, J.-E. Rubensson, A. Soldatov, E. F. Aziz, *Angew. Chem. Int. Ed.* **2011**, *50*, 10621–10625; *Angew. Chem.* **2011**, *123*, 10809–10813.

[14] K. M. Lange, R. Könnecke, S. Ghadimi, R. Golnak, M. A. Soldatov, K. F. Hodeck, A. Soldatov, E. F. Aziz, *Chem. Phys.* **2010**, *377*, 1–5.

[15] E. Ito, H. Oji, T. Araki, K. Oichi, H. Ishii, Y. Ouchi, T. Ohta, N. Kosugi, Y. Maruyama, T. Naito, T. Inabe, K. Seki, *J. Am. Chem. Soc.* **1997**, *119*, 6336–6344.

[16] A. F. Wagner, K. Folkers, *Vitamins and Coenzymes*, Interscience Publishers, New York, **1964**; E. Staunton, W. R. Todd, H. S. Mason, J. T. V. Bruggen, *Textbook of Biochemistry* The MacMillan Company, New York, **1967**.

[17] B. van de Graaf, A. J. Hoefnagel, B. M. Wepster, *J. Org. Chem.* **1981**, *46*, 653–657.

[18] R. A. Sullivan, R. J. Davey, G. Sadiq, G. Dent, K. R. Back, J. H. ter Horst, D. Toroz, R. B. Hammond, *Cryst. Growth Des.* **2014**, *14*, 2689–2696.

[19] M. L. Deviney, R. C. Anderson, W. A. Felsing, *J. Am. Chem. Soc.* **1957**, *79*, 2371–2373.

[20] T. F. Lai, R. E. Marsh, *Acta Crystallogr.* **1967**, *22*, 885–893.

[21] S. Athimoolam, S. Natarajan, *Acta Crystallogr.* **2007**, *C63*, o514-o517.

[22] S. Gracin, A. Fischer, *Acta Crystallogr.* **2005**, *E61*, o1242-o1244.

[23] F. Sette, J. Stöhr, A. P. Hitchcock, *J. Chem. Phys.* **1984**, *81*, 4906–4914.

[24] M. Colapietro, A. Domenicano, *Acta Crystallogr. B* **1980**, *36*, 354–358.

[25] J. S. Stevens, C. R. Seabourne, C. Jaye, D. A. Fischer, A. J. Scott, S. L. M. Schroeder, *J. Phys. Chem. B* **2014**, *118*, 12121–12128.

[26] G. Salas-Cimingo, *C. R. Seances Acad. Sci. Ser. C* **1968**, *267*, 1304–1305.

[27] R. G. Wilks, University of Saskatchewan (Saskatoon), **2009**.

[28] B. Ravel, M. Newville, *J. Synchrotron Radiat.* **2005**, *12*, 537–541; M. Newville, *J. Synchrotron Radiat.* **2001**, *8*, 322–324.

[29] B. Ravel, *ATHENA User Guide*, http://cars9.uchicago.edu/~ravel/software/doc/Athena/html/athena.pdf, **2009**.

[30] M. A. Soldatov, K. M. Lange, M. D. Gotz, N. Engel, R. Golnak, A. Kothe, E. F. Aziz, *Chem. Phys. Lett.* **2012**, *546*, 164–167; K. M. Lange, E. Suljoti, E. F. Aziz, *J. Electron Spectrosc. Relat. Phenom.* **2013**, *188*, 101–110.

[31] *Gaussian 09, Revision B.01*, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, Gaussian, Inc. Wallingford CT, **2009**.

[32] *GaussView, Version 5*, Roy Dennington, Todd Keith, and John Millam, Semichem Inc., Shawnee Mission, KS, **2009**.

Received: October 14, 2014
Published online on March 18, 2015

PCCP

PAPER

View Article Online
View Journal | View Issue

Cite this: *Phys. Chem. Chem. Phys.*, 2015, **17**, 29000

Intermolecular bonding of hemin in solution and in solid state probed by N K-edge X-ray spectroscopies

Ronny Golnak,[ab] Jie Xiao,*[a] Kaan Atak,[ac] Joanna S. Stevens,[d] Adrian Gainar,[d] Sven L. M. Schroeder[ef] and Emad F. Aziz*[ac]

Received 31st July 2015,
Accepted 6th October 2015

DOI: 10.1039/c5cp04529k

www.rsc.org/pccp

X-ray absorption/emission spectroscopy (XAS/XES) at the N K-edge of iron protoporphyrin IX chloride (FePPIX-Cl, or hemin) has been carried out for dissolved monomers in DMSO, dimers in water and for the solid state. This sequence of samples permits identification of characteristic spectral features associated with the hemin intermolecular bonding. These characteristic features are further analyzed and understood at the molecular orbital (MO) level based on the DFT calculations.

Introduction

Metalloporphyrins are functional materials with extensive applications in catalysis, solar energy conversion and biological systems.[1–5] In these systems, porphyrin molecules are often in a solution environment, in which self-association can lead to different types of non-monomeric species. The differences in solvation and solute–solute interactions will consequently affect the functionality of metalloporphyrins, requiring studies of speciation in solution to further our understanding of their functionality. For example, iron protoporphyrin IX chloride (FePPIX-Cl, or hemin, Fig. 1a) forms monomer species when dissolved in dimethyl sulfoxide (DMSO), while dimer structures are present in aqueous solution.[6–9] It is known that the electronic structure of hemin, especially the local electronic structure at the Fe center, which often serves as the major functional site, is altered due to dimerization.[7] Hemin oligomer species in various solvents have therefore been investigated previously by UV-Vis spectroscopy and X-ray absorption/emission spectroscopy (XAS/XES) at the Fe L-edge,[6–9] although the nature of the probed electronic transitions implied that the hemin intermolecular bonding interactions were probed rather indirectly. UV-Vis measures overlapping valence excitations of all components in the hemin solution, without differentiation between contributions from solute, solvent or solute–solvent interactions. The lack of exclusive probing of the π–π interactions expected for hemin dimerization left some ambiguity in the interpretation of the UV-Vis data.[9] XAS/XES measurements at the Fe L-edge, on the other hand, represent a local probe sensitive to the electronic structure changes at the Fe sites induced by hemin dimerization.[7] However, the information obtained for the dimerization is still fairly indirect, since hemin dimerization is primarily driven by π–π stacking of porphyrin rings, without intermolecular bonding by the Fe center, as illustrated in Fig. 1b.[7–9] The characteristic L-edge features at the Fe sites[7] do of course depend on their interaction with the N moieties of the porphyrin ring, but the Fe center in hemin is also coordinated by a chloride ligand. Moreover, there is a sixth vacant coordination site perpendicular to the porphyrin plane, available for possible coordination by molecules from the solvent. Even though contributions of Fe–solvent interaction have been determined to be insignificant,[7] they cannot be completely excluded. It is therefore difficult to unravel the different types of interactions at the Fe site for obtaining unambiguous information about the π–π interactions underlying dimerization.

Locally probing the N valence orbitals is expected to be a more explicit way of probing hemin dimerization, or the extent of hemin intermolecular interactions in general. The N moieties are part of the porphyrin ring and thus part of the valence system directly involved in the π–π stacking. Each N atom embedded in the porphyrin ring is fully coordinated with the metal center and C atoms of the porphyrin and consequently there is much less scope for strong interaction with solvent molecules. We have therefore targeted the local electronic structure at the N moieties

[a] *Joint Ultrafast Dynamics Lab in Solutions and at Interfaces (JULiq), Institute of Methods for Material Development, Helmholtz-Zentrum Berlin für Materialien und Energie, Albert-Einstein-Strasse 15, D-12489 Berlin, Germany. E-mail: jie.xiao@helmholtz-berlin.de, emad.aziz@helmholtz-berlin.de*
[b] *Fachbereich Chemie, Freie Universität Berlin, Takustrasse 3, D-14195 Berlin, Germany*
[c] *Fachbereich Physik, Freie Universität Berlin, Arnimallee 14, D-14195 Berlin, Germany*
[d] *School of Chemical Engineering and Analytical Science, The University of Manchester, Oxford Road, Manchester M13 9PL, UK*
[e] *School of Chemical and Process Engineering, University of Leeds, Leeds LS2 9JT, UK*
[f] *DIAMOND Light Source Ltd, Harwell Science and Innovation Campus, Chilton, Didcot OX11 0DE, UK*

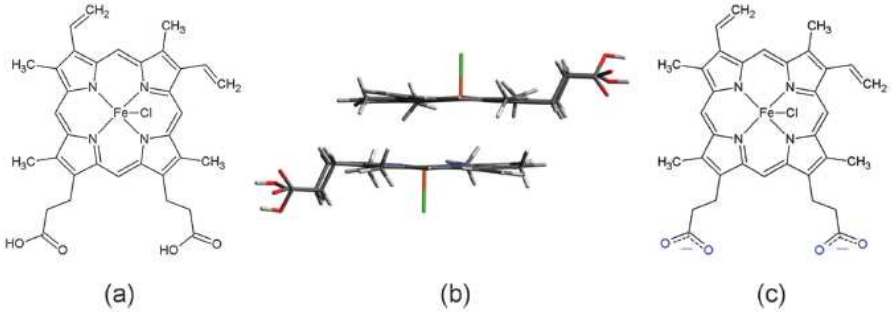


Fig. 1 (a) Schematic of hemin chemical structure with neutral carboxyl groups (COOH), (b) geometrical arrangement of hemin dimer, and (c) schematic of hemin chemical structure with anionic carboxylate groups (COO^-).

through X-ray absorption/emission spectroscopy (XAS/XES), to obtain electronic structure information about the unoccupied and occupied valence states of hemin in solution.[5,10,11] We will show that solvent-dependent XA/XE spectral differences at the N K-edge are observable, and certain XA features are identified as spectral characteristics associated with the extent of hemin intermolecular bonding.

DMSO and 0.5 M NaOH aqueous solution were chosen as solvents to dissolve hemin powder as in previous studies. These solutions contain hemin monomer and dimer species, respectively.[6,7] The high pH value (from NaOH) of the aqueous solution increases hemin solubility in water, but adds potential additional complexity due to formation of the carboxylate anionic form in basic solution, as shown in Fig. 1c. However, because the carboxylate groups are not conjugated with the π system of the porphyrin ring their influence on the N K-edge is expected to be minor, in contrast to the recently presented N K-edge spectra of *para*-aminobenzoic acid in solution, where the amine nitrogen felt the direct influence of the carboxylate group as both were part of the delocalised π system of the aromatic ring.[12]

Experimental methods

Hemin (iron protoporphyrin IX chloride, FePPIX-Cl) powder from bovine (purity >90%) was purchased from Sigma-Aldrich, and further dissolved in 0.5 M NaOH aqueous solution and dimethyl sulfoxide (DMSO) solvent to make 0.1 M and 0.05 M hemin solutions, respectively. The hemin solutions were introduced into a vacuum chamber by liquid micro-jet technique for the X-ray measurements.[13,14] Possible sample damage by intense X-ray beam was effectively avoided since the liquid sample was constantly refreshed. The solid powders were spread across a double-sided adhesive copper tape attached to a copper plate which was mounted at 45° with respect to the incident photon beam in the vacuum chamber. The possible radiation damage to the solid samples was minimized by reducing incident photon flux and frequent changing of the measurement spots.

The experiment was carried out at the U41-PGM undulator beamline and LiXEdrom endstation at the synchrotron facility BESSY II, Berlin. The U41-PGM provides horizontally linear polarized light with high photon flux ($\sim 10^{13}$ photon per s) and small focal spot ($\sim$40 μm of vertical size), which is crucial to our measurements. High flux is required for low-concentration samples. Small vertical focal size ($\sim$40 μm) assures high energy resolution for XES measurements. The photons emitted from the sample were collected along the polarization direction of the incident photon beam (to suppress the elastic peak) and subsequently dispersed by a spherical grating with 1200 lines per mm and 7.5 m radius. The dispersed photons were then detected by a microchannel plate (MCP)/fluorescence screen/CCD assembly. The sample (micro-jet), grating, and photon detector are arranged in a Rowland circle geometry for accurate focusing. The grating and detector chamber was kept at a pressure of 10^{-8} mbar or lower to protect the grating and MCP surfaces from contamination, while the sample chamber was at 10^{-5} mbar. The incident photon energy was tuned to the N K-edge, and the N 2p $\rightarrow$ 1s transition was selected for the detection of the emitted photon from both liquid and solid samples, giving rise to the XE spectra. Integration of a series of XE spectra acquired at various excitation photon energies over the emitted photon energy delivers one XA spectrum, which is referred as XAS in partial-fluorescence-yield (PFY) mode.

Theoretical calculations

The DFT calculations were conducted with the ORCA program package[15] to obtain the molecular orbital (MO) information of the hemin molecule in gas phase in the ground state.[6] Molecular geometry optimization was performed using the B3LYP DFT method together with the def2-TZVP(-f) basis set.[16–18] The N contributions to both occupied and unoccupied valence MO are acquired from Löwdin population analysis based on the DFT calculations.[19] A Gaussian-type broadening of 0.5 eV with the respective weight (according to the N contribution) is applied to each valence MO and then summed up, to simulate the XA/XE spectra obtained at the N K-edge.

Results and discussion

The obtained PFY-XA and XE spectra of hemin in DMSO (monomer) and in 0.5 M NaOH aqueous solution (dimer) are

shown in Fig. 2 as blue and red traces, respectively, along with the difference plot (black trace in the left panel) obtained from the subtraction of the monomer XA spectrum (blue) from that of the dimer (red). The XA spectra of the two hemin oligomer species are very similar but on closer inspection three regions with significant spectral differences can be identified, indicated by the three color-highlighted areas of the difference plot. The intense absorption feature at 398.5 eV, which represents the lowest unoccupied molecular orbital (LUMO) at the N sites, is associated with the porphyrin nitrogen in a N–Fe environment.[20,21] Such a low-lying energy feature is, therefore, completely absent in small metal-free N-containing molecules.[12,22,23] The metal-free porphyrin (2HTPP) usually has the first N absorption feature located more than 1 eV below the metal-N peak with much reduced intensity, arising from the inequivalence of the N species in the absence of the metal centre, leading to iminic =N– and pyrrolic –NH– nitrogen.[20,21] The slight energy shift between the two absorption edges at around 398 eV is made evident by the blue-highlighted area in the difference plot, which actually indicates a broadening, instead of an energy shift, of the first absorption feature for the hemin dimer species since its peak position at 398.5 eV does not exhibit observable shift when compared with that of hemin monomer (the broadening effect is discussed below). Nonetheless, the N–Fe features of the two hemin species are in general very similar, which is in line with the similarity of the LUMO observed in the Fe L-edge XA spectra,[7] indicating that the N–Fe bonding is not strongly influenced by the dimerization process, or affected by the different solvents significantly. The absorption features at 400–403 eV (green-highlighted) originate from N π^* orbitals as N is part of the aromatic ring system, while the arising absorption edges at the brown-highlighted area are in the region of the ionization potentials (IPs).[12,22] The N π^* orbitals and IP thus exhibit distinguishable characteristics for different hemin solutions. The exact origins of these differences will be discussed in detail for Fig. 3 below.

Besides the unoccupied valence orbitals revealed by XAS, slight differences between the occupied valence states at the N sites are also uncovered by the XE spectra of the two hemin solutions (Fig. 2). Emitted photon energies, instead of photon energy losses that are often observed in highly localized and correlated electron systems as in d or f orbitals, are shared by the N XE features measured at various excitation energies, marked by vertical dashed green lines in Fig. 2. The absence of the loss features indicates little electron correlation in the N 2p orbitals, as expected for generally delocalized orbitals like valence s or p. The XE spectra of hemin in NaOH aqueous solution (red traces) present generally a better resolution (narrower peaks) when compared with the spectra of hemin in DMSO (blue traces). This is in line with the previous observation at the

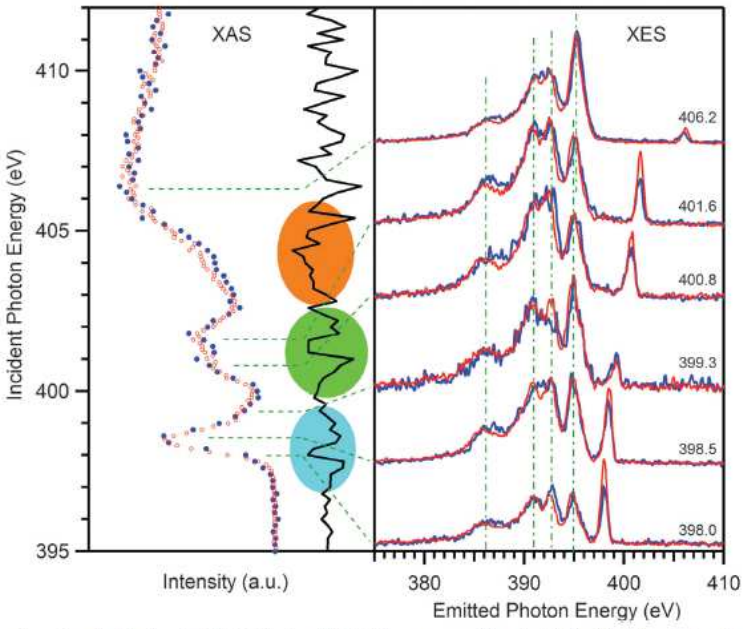


Fig. 2 PFY-XA and XE spectra of hemin dissolved in DMSO (blue) and in NaOH aqueous solution (red). The black trace is the difference plot of the two XA spectra, with color-highlighted areas indicating the differences observed in the PFY-XA spectra of the two hemin solutions. The excitation energies used for XES measurements are labeled on the right side of the figure, and also indicated by horizontal green dashed lines pointing to the corresponding PFY features in the left panel. The vertical green dashed lines in the right panel mark the major XE features at the N K-edge. All spectra were normalized to their respective background for intensity comparison.

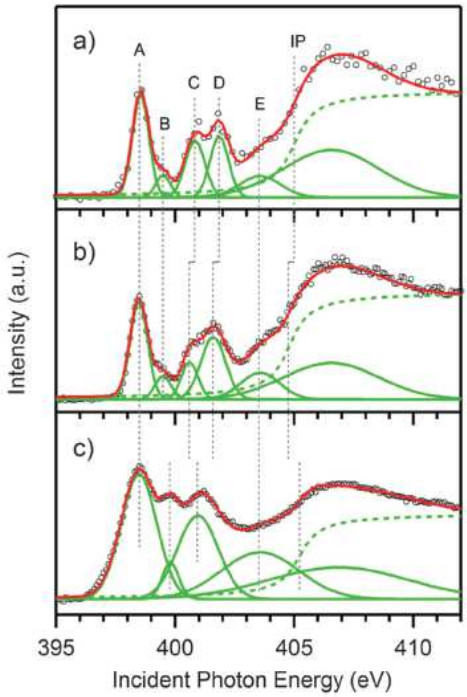


Fig. 3 N K-edge PFY-XA spectra, represented by circles ○, of (a) hemin in DMSO, (b) hemin in NaOH aqueous solution and (c) hemin powder in solid form, with their respective fitted components in green and the summation of these components in red. The ionization potential (IP) is fitted by step functions arctan, represented as green dashed traces, while the other components by Gaussian functions (solid green traces) with labels A, B, C, D and E. The energy positions of the IP and fitted Gaussian components, as well as their relative energy offsets, are marked by grey dashed lines.

Fe L-edge that the Fe XE peaks of hemin dimer species are narrower than those of hemin monomer, owing to the higher degree of orbital localization induced by hemin dimerization.[7] Different degrees of orbital delocalization (or localization) in the two hemin oligomer species also lead to the elastic peaks with different intensities. The elastic peak of each red XE spectrum (hemin dimer) in Fig. 2 shows higher intensity than the blue trace (hemin monomer) when the excitation energy tuned to the corresponding absorption peaks, indicating higher cross-sections of the N 2p → 1s transition for hemin dimer. It can be argued that the higher degree of orbital localization in hemin dimer preserves a density of states (DOS) with the p character of the N 2p orbital better than in the hemin monomer, resulting in more dipole-allowed 2p → 1s transitions. If the N 2p orbitals mixes more with Fe 3d orbitals, lower transition probability is expected since lower DOS with p character is present at the N sites due to the higher extent of N 2p orbital delocalization in the hemin monomer.

The more intense elastic peak and better resolved emission features in the N XE spectra of hemin in aqueous solution could be regarded as an indicator of hemin dimerization, similar to the observation of the local energy gap opening at the Fe L-edge used to indicate the hemin dimerization.[7] However, unlike the local gap opening at the Fe L-edge that involves the energy shift of the characteristic spectral feature, these observed spectral differences in the N XE spectra only contain intensity variations. Signal strength often largely depends on extrinsic parameters and thus cannot be used as an intrinsic and reliable indicator for hemin dimerization. The observed differences of the N π* orbitals and IP (green and brown areas) in the N XA spectra in Fig. 2 may, however, intrinsically differentiate between hemin monomer and dimer species. It is instructive to compare these N XA spectra of the small scale intermolecular bonding species (dimer) in Fig. 2 with that of a large scale case in solid form, to clearly identify the characteristic spectral evolution under the various extents of hemin aggregation. Such a comparison of the XA spectra is presented in Fig. 3, with their respective fitted components (green traces) as well as the summations of the components (red traces). Each individual fitted Gaussian component below IP is marked by vertical dashed lines and labeled as A, B, C, D and E. The IPs are fitted by arctan step functions with their energy positions marked by vertical lines as well.

The most striking difference between the solid state and the solution spectra is the considerable broadening of almost all the spectral features in the solid state spectrum. This is likely due to the extensive orbital overlapping in the solid state[24,25] and was also somewhat evident in the previous study of aminobenzoic acid.[12] Self-absorption distortions of the spectrum from the more dense solid state may also contribute to the observed broadening of features. This solid state broadening makes the component A, originating from the N–Fe interaction as discussed in Fig. 2, and component B resolvable in Fig. 3a and b after the Gaussian fitting appear to be irresolvable in Fig. 3c. The feature B may stem from a multi-electron excitation, *e.g.* a shake-up satellite to feature A, or reflect an electronic state that originates from solute–solvent interactions since the ground state calculation (vertical bars at the bottom of Fig. 4) does not give rise to any molecular orbitals in this energy region. Feature A keeps relatively similar peak widths in Fig. 3a and b because the N–Fe interaction is not directly involved in the dimerization process. However, its broadening in Fig. 3c suggests that the hemin oligomer species in the solid form may adopt different forms of intermolecular bonding that may involve the N–Fe interaction to some extent. Components C and D, constituting the green-highlighted difference in Fig. 2, exhibit monotonic energy shifts from the monomer (Fig. 3a) to the dimer (Fig. 3b), and appear to be even lower in the solid (Fig. 3c). It seems that these features are associated with the development of hemin–hemin interactions and can therefore be used to characterize the extent of intermolecular bonding. The shifts of the features C and D to lower energies indicate that the measured energy level distances between the N π* orbitals and the N 1s core level are systematically reduced when a larger scale of the hemin oligomer species is realized. The origin of this energy shift could be either an initial state effect in that either the valence N π* orbitals are lowered or the N 1s core level binding energy raised. Alternatively, a final

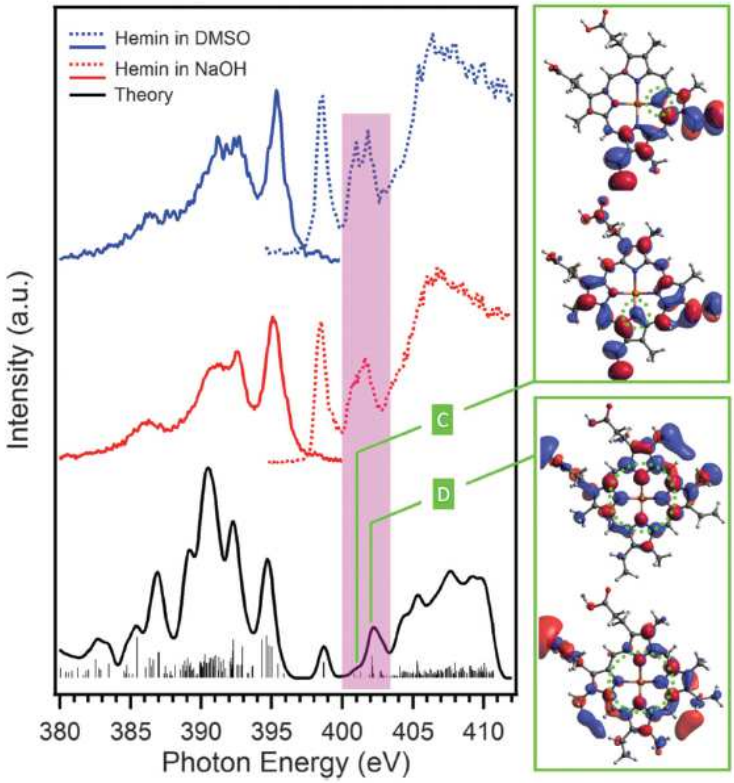


Fig. 4 Off-resonant XE spectra (solid blue and red traces) at the excitation energies of 420 eV for the hemin dissolved in DMSO and NaOH aqueous solution, respectively, combined with their respective PFY-XA spectra (dashed blue and red traces). The experimental spectra are identical to the corresponding spectra in Fig. 2. The theoretical DFT calculations and individual MOs weighted by N contributions are shown at the bottom as black trace and vertical bars, respectively. The characteristic features C and D identified in Fig. 3 are assigned to certain calculated MOs with visualized orbitals presented on the right side of the figure. The orbitals localized at the N sites are highlighted with dotted green circles.

state effect may contribute, perhaps reflecting differences in valence charge relaxation upon core-hole formation, analog to the relaxation shift of CO molecules going from gas phase to solid state observed in photoelectron spectroscopy.[26] However, recent studies of bipyridine acid salts and co-crystals indicated that the effect of final state effects on the relative positions of π^* features are negligible, even when comparing bipyridine nitrogen species in very different local chemical environments.[27] Further inspection into the individual components C and D uncovers that the feature C remains roughly constant in peak width from Fig. 3a to c, while the feature D becomes broader and relatively more intense upon hemin aggregation. The mechanism of this distinct peak evolution is likely due to the different extents of the orbital involvement in the hemin intermolecular bonding interactions which will be discussed in detail for Fig. 4 based on a molecular orbital (MO) picture derived from DFT calculations.

The evolution of feature E is very similar to feature A, both exhibiting solid state broadening. However, the feature E does not have a sole MO origin as for feature A where the N–Fe interaction can be assigned, because it consists of contributions of many MOs. A similar situation also occurs for the broadest fitted Gaussian components with their centers located above the IPs, representing contributions from a number of MOs, as illustrated by the vertical bars at about 404–411 eV at the bottom of Fig. 4. Strictly speaking, each spectral feature in Fig. 3 comprises multiple MOs, and therefore should not be represented by a single Gaussian function in principle. Nevertheless, the fittings of the components A, B, C and D with a single Gaussian function still hold the physical significance discussed above because the contributed MOs to each component are very few in number and also located closely to each other in energy, as demonstrated in Fig. 4. The characteristic spectral features that can distinguish various hemin oligomer species are therefore the features C and D in Fig. 3 exhibiting systematic energy shifts, and consequently associated with the extent of hemin intermolecular interactions.

The energy shifts of the IPs, however, do not follow the monotonic energy shift when going from dimer to solid, as

shown in Fig. 3b and c. The IP shift from Fig. 3a to b is in line with the energy shift of the features C and D, which seems to suggest they have the same origin – elevation of the N 1s level. However, the core level argument cannot explain the non-shifting features, such as feature A, and the opposite shifting of the IP in the spectrum of the solid sample. The other mechanism discussed above, *i.e.* the initial state effect of the valence orbital shifting, has to come to play.

With the help of the ground state DFT calculations, the origins of the characteristic features C and D that exhibit distinct spectral evolutions of the component width and intensity in Fig. 3 is uncovered at a MO level. The theoretically simulated spectrum at the occupied and unoccupied valence levels, obtained from the summation of individual N contributions with a universal Gaussian broadening, is plotted in Fig. 4, along with the combined XA and XE spectra of the hemin monomer (in DMSO) and dimer (in NaOH aqueous solution) for comparison. The N contribution to each calculated MO, shown as vertical bars at the bottom of Fig. 4, is acquired from the Löwdin population analysis based on the DFT calculations of the gas phase hemin in the ground state. Even though the experimental spectra are acquired from the core-hole excited final state (XAS) or initial state (XES), the simulated peak positions from the ground state calculation are in good agreement with experiment, allowing for accurate and reliable assignments of the calculated MOs to the experimental peaks. The characteristic features C and D in Fig. 3 can therefore be visualized as the MOs presented in Fig. 4. As illustrated in Fig. 4, the feature C mainly originates from the orbitals localized at a single N atom; while the feature D, on the other hand, is from the orbitals delocalized over all four N atoms, as highlighted by dotted green circles. With development of the hemin intermolecular bonding through π–π stacking of the porphyrin ring, the feature D that involves four N atoms will surely develop a higher extent of the peak width broadening as well as intensity enhancement, owing to the extensive orbital overlapping, when compared with the feature C that has only very localized orbitals involved in the hemin–hemin interaction. The localized nature of the feature C inhibits its orbital overlapping with neighbor orbitals and consequently suppresses the broadening and enhancement of the fitted peak, as demonstrated in Fig. 3.

Due to the computational challenges for large-scale hemin oligomer species – dimer and above – the DFT calculations are only performed on a geometry-optimized isolated hemin monomer. The proposed N 1s core level shift arising from the different hemin oligomer species is therefore unable to be probed theoretically. The DFT calculations are carried out on the ground state, so that the final state effect is not possible to investigate. The spectrum of the solid sample is not included in Fig. 4 for the comparison with the theoretical simulation because the strong solid state effect is expected to have significant influences on both initial state shifting and final state relaxation, which shifts certain experimental peaks significantly when compared with those of the isolated molecules, as has been demonstrated in Fig. 3 for the features C and D. Therefore, attempted assignments of the calculated MOs performed on the gas phase molecules to the solid state features may not possess high credibility.

Conclusions

X-ray absorption and emission spectroscopy (XAS/XES) have been performed on solvated hemin molecules dissolved in DMSO and NaOH aqueous solution, as well as on the solid state. Characteristic features in the N XA spectra are identified and associated with electronic structure changes arising from hemin intermolecular bonding interactions. The spectral evolutions of these features upon the development of hemin–hemin interactions are further analyzed at the molecular orbital level derived from the DFT calculations. The results show that N K-edge X-ray spectroscopies provide experimental access to porphyrin intermolecular interactions in solution and in the solid state by tracking the energy shifts of the characteristic N K-edge XA features, crucially permitting *in situ* detection for electronic structure investigations. The exact energy positions of these characteristic features also provide a reference to help identify hemin oligomer species in solution, *e.g.* monomer in DMSO or dimer in aqueous solution.

Acknowledgements

This work was supported by the European Research Council Grant No. 279344. Part of this work was financially supported by the Helmholtz-Gemeinschaft *via* the young investigator fund VH-NG-635. K.A. acknowledges the financial support of the Einstein Foundation Berlin for the postdoctoral scholarship. S.L.M.S., J.S.S. and A.G. acknowledge financial support from the EPSRC through the Critical Mass Grant EP/1013563/1. S.L.M.S. holds a Royal Academy of Engineering Chair supported by Infineum UK Ltd and DIAMOND Light Source Ltd.

References

1 M. Ethirajan, Y. Chen, P. Joshi and R. K. Pandey, *Chem. Soc. Rev.*, 2010, **40**, 340–362.
2 D. González-Rodríguez and G. Bottari, *J. Porphyrins Phthalocyanines*, 2009, **13**, 624–636.
3 L.-L. Li and E. W.-G. Diau, *Chem. Soc. Rev.*, 2012, **42**, 291–304.
4 A. Vecchi, P. Galloni, B. Floris and V. N. Nemykin, *J. Porphyrins Phthalocyanines*, 2013, **17**, 165–196.
5 E. F. Aziz, *J. Phys. Chem. Lett.*, 2011, **2**, 320–326.
6 K. Atak, R. Golnak, J. Xiao, E. Suljoti, M. Pflueger, T. Brandenburg, B. Winter and E. F. Aziz, *J. Phys. Chem. B*, 2014, **118**, 9938–9943.
7 R. Golnak, J. Xiao, K. Atak, M. Khan, E. Suljoti and E. F. Aziz, *J. Phys. Chem. B*, 2015, **119**, 3058–3062.
8 A. C. Maehly, Å. Åkeson, G. Bergson, A. Grönvall, B. Zaar and E. Diczfalusy, *Acta Chem. Scand.*, 1958, **12**, 1259–1273.
9 K. A. de Villiers, C. H. Kaschula, T. J. Egan and H. M. Marques, *JBIC, J. Biol. Inorg. Chem.*, 2007, **12**, 101–117.
10 K. M. Lange and E. F. Aziz, *Chem. Soc. Rev.*, 2013, **42**, 6840–6859.
11 E. Suljoti, R. Garcia-Diez, S. I. Bokarev, K. M. Lange, R. Schoch, B. Dierker, M. Dantz, K. Yamamoto, N. Engel and K. Atak, *et al.*, *Angew. Chem., Int. Ed.*, 2013, **52**, 9841–9844.

12 J. S. Stevens, A. Gainar, E. Suljoti, J. Xiao, R. Golnak, E. F. Aziz and S. L. M. Schroeder, *Chem. – Eur. J.*, 2015, **21**, 7256–7263.
13 K. M. Lange, R. Könnecke, S. Ghadimi, R. Golnak, M. A. Soldatov, K. F. Hodeck, A. Soldatov and E. F. Aziz, *Chem. Phys.*, 2010, **377**, 1–5.
14 R. Golnak, K. Atak, E. Suljoti, K. F. Hodeck, K. M. Lange, M. A. Soldatov, N. Engel and E. F. Aziz, *Phys. Chem. Chem. Phys.*, 2013, **15**, 8046–8049.
15 F. Neese, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2012, **2**, 73–78.
16 A. D. Becke, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1988, **38**, 3098–3100.
17 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648–5652.
18 F. Weigend and R. Ahlrichs, *Phys. Chem. Chem. Phys.*, 2005, **7**, 3297–3305.
19 K. Atak, R. Golnak, J. Xiao, M. Pflüger, T. Brandenburg, B. Winter and E. F. Aziz, *Phys. Chem. Chem. Phys.*, 2015, **17**, 3409–3414.
20 M. V. Nardi, R. Verucchi, L. Pasquali, A. Giglia, G. Fronzoni, M. Sambi, G. Mangione and M. Casarin, *Phys. Chem. Chem. Phys.*, 2015, **17**, 2001–2011.
21 N. Schmidt, R. Fink and W. Hieringer, *J. Chem. Phys.*, 2010, **133**, 054703.
22 J. S. Stevens, C. R. Seabourne, C. Jaye, D. A. Fischer, A. J. Scott and S. L. M. Schroeder, *J. Phys. Chem. B*, 2014, **118**, 12121–12129.
23 M. J. Thomason, C. R. Seabourne, B. M. Sattelle, G. A. Hembury, J. S. Stevens, A. J. Scott, E. F. Aziz and S. L. M. Schroeder, *Faraday Discuss.*, 2015, **179**, 269–289.
24 J. Xiao and P. A. Dowben, *J. Phys.: Condens. Matter*, 2009, **21**, 052001.
25 J. Xiao and P. A. Dowben, *J. Mater. Chem.*, 2009, **19**, 2172–2178.
26 H.-J. Freund and M. Neumann, *Appl. Phys. A: Mater. Sci. Process.*, 1988, **47**, 3–23.
27 J. S. Stevens, L. K. Newton, C. Jaye, C. A. Muryn, D. A. Fischer and S. L. M. Schroeder, *Cryst. Growth Des.*, 2015, **15**, 1776–1783.

The Structure of p-Aminobenzoic Acid in Water: Studies Combining UV-Vis, NEXAFS and RIXS Spectroscopies

A Gainar[1], J S Stevens[1], E Suljoti[2], J Xiao[2], R Golnak[2], E F Aziz[2,3], S L M Schroeder[1,4,5]

[1]School of Chemical Engineering and Analytical Science, The University of Manchester, Oxford Road, Manchester, M13 9PL, UK
[2]Joint Ultrafast Dynamics Lab in Solutions and at Interfaces (JULiq), Helmholtz-Zentrum Berlin für Materialien und Energie Institution Albert Einstein Strasse 15, 12489 Berlin, Germany
[3]Freie Universität Berlin, Arnimallee 14, 14195 Berlin, Germany
[4]School of Chemical and Process Engineering, University of Leeds, Leeds, LS2 9JT, UK
[5]DIAMOND Light Source Ltd, Chilton, Didcot, OX11 0QX, UK

E-mail: [1]adrian.gainar@postgrad.manchester.ac.uk; [1,4,5]s.l.m.schroeder@leeds.ac.uk

Abstract. NEXAFS-RIXS and home laboratory-based UV-Vis absorption spectroscopy are combined to examine the speciation and electronic structure of *para*-aminobenzoic acid (PABA) in aqueous solution as a function of pH. DFT and TD-DFT electronic structure calculations reproduce the experimental trends and provide a correlation between the experimental HOMO↔LUMO gap as well as the electronic transitions between molecular orbitals in the non-ionic, anionic and cationic forms of PABA.

1. Introduction

Para-aminobenzoic acid (PABA) has therapeutic effects and sunscreen properties due to its capability to filtering out UV radiation through the electronic conjugation of the *para* orientated amino and carboxyl substituents through the aromatic ring. Core-level spectroscopies such as near-edge X-ray absorption fine structure (NEXAFS) and resonant inelastic X-ray scattering (RIXS) have given incisive insight on the electronic state of PABA species in solution and in the solid state including the anionic and cationic forms in water [1]. While NEXAFS involves excitation of electrons from 1s core orbitals to unoccupied molecular orbital (MO) levels (*e.g.* lowest unoccupied MO, LUMO), RIXS probes transitions from occupied MOs (*e.g.* highest occupied MO, HOMO) to 1s core orbitals. UV-Vis spectroscopy is quite similar to these core level techniques in that it probes electronic transitions, but from occupied MOs such as σ, n (non-participant) or π type orbitals to σ* or π* unoccupied MOs. As NEXAFS, RIXS and UV-Vis often probe the same final-state MOs, their combined measurement supported by electronic structure calculations (particularly density functional theory, DFT, and time-dependent density functional theory, TD-DFT) provides an opportunity to arrive at a deeper understanding of the local electronic structure landscape.

2. Experimental

2.1. Materials

PABA crystalline powder (99%, Sigma-Aldrich, UK) was dissolved in distilled, deionized water. pH of the solution was adjusted by dropwise addition of 1.2 M aqueous solution of NaOH or of HCl.

2.2. Transmission UV-Vis

A saturated stock solution of 0.03 M PABA in water was prepared and successively diluted to reach the detection range of the Beckman DU 520 spectrophotometer.

2.3. Gaussian calculations

Non-ionic, anionic and cationic PABA monomers were ground-state optimized with B3LYP/6-31G* in Gaussian09 to obtain the DFT molecular orbitals. Using these optimized monomers in a water continuum, TD-DFT polarizable continuum model (PCM) calculations were employed to generate the simulated spectra and the MOs involved in these transitions.

2.4. NEXAFS/RIXS

NEXAFS and RIXS were recorded at the N K-edge (resolution ≤ 0.1 eV) using the liquid microjet of the LiXEdrom endstation, U41-PGM beamline, BESSY II synchrotron. The common energy scale was calibrated using N_2 gas, measuring the $1s \rightarrow 2p\ \pi$ transitions in total electron yield mode.

3. Results and discussion

Depending on the pH of the aqueous solution, PABA is present in cationic, neutral or anionic form [1,2]. For low (acidic) pH solution (where PABA is almost entirely present as the cationic species), the main absorption band is centered at a wavelength of ~ 225 nm. For high pH, PABA is anionic and the main UV absorption peak occurs at ~ 265 nm, while the neutral species, at intermediate pH, absorbs at ~ 275 nm (Table 1 and Figure 1). TD-DFT calculations reproduce the observed trend, with the main absorption shifted to shorter wavelengths compared to non-ionic PABA (cationic < anionic < non-ionic, Table 1). Within this trend, it is noticeable that the calculated value for the anionic species is significantly lower than the experimentally observed value. It has been suggested previously that calculated charge transfer excitations can be less accurate for ionic π systems due to underestimated excitation energies of ionic components [3]. Perhaps it also plays a role here that the carboxylate group is less electron withdrawing and thus electrons are more consolidated around this group rather than being delocalized over the ring [4].

Table 1. Comparison of main absorption wavelengths in experimental UV-Vis and TD-DFT simulated spectra of PABA species in water, with primary transitions from HOMOs to LUMOs.

PABA species	Wavelength (experimental) / nm	Wavelength (simulated) / nm	Primary transition
Non-ionic	~ 275	266.96	HOMO→LUMO
Anionic	~ 265	228.26	HOMO–1→LUMO
Cationic	~ 225	218.95	HOMO–1→LUMO

Collating together the snapshots of the MOs generated from the TD-DFT calculations for the non-ionic, anionic and cationic monomers in water, Figure 1 reveals that the primary transition in the non-ionic PABA takes place between the orbital with annotation A as HOMO and the $1\pi^*$ orbital as LUMO (~ 267 nm, 4.64 eV), whereas the shift to shorter wavelength in the anionic PABA is due to widening energy separation between A and $1\pi^*$ (~ 228 nm, 5.43 eV), with A lowering in energy and becoming HOMO–1. Similarly, the noted shift to lower wavelength for cationic PABA (~ 219 nm, 5.66 eV) can be explained by the primary transition between A, which is now the HOMO–1, and $1\pi^*$.

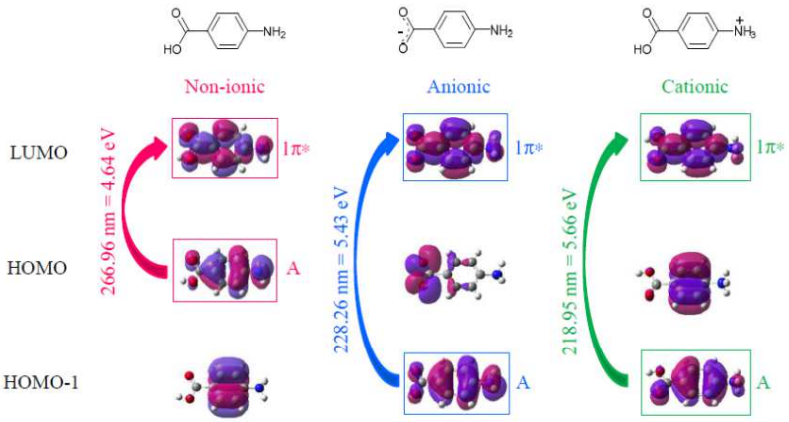


Figure 1. Transitions (arrows) between MOs of interest A and 1π* (marked with boxes) in TD-DFT calculations for non-ionic, anionic and cationic PABA monomers in a water continuum.

We can now correlate these observations with the MO information for these systems that we have previously obtained by RIXS and NEXAFS [1]. Figure 2 shows that the experimentally determined HOMO↔LUMO gap for nitrogen (which is not applicable to cationic PABA due to the protonation of nitrogen) is correlated with the calculated energy gap between orbitals A and 1π* from the DFT calculations for the non-ionic and anionic species. In the case of non-ionic PABA, orbital A corresponds to the HOMO and orbital 1π* to the LUMO, whilst for the anionic species, orbital A lowers down in energy becoming HOMO–3 and orbital 1π* raises up in energy as LUMO+1.

Although core hole relaxation effects are expected to affect the energetic position of features in the NEXAFS and RIXS data, the impact on taking the experimental RIXS-NEXAFS energy gap as a measure of the HOMO-LUMO is believed to be of the order of at most a few 100 meV; additionally, when such effects were included in CASTEP calculations for crystalline PABA [5] the inclusion of final-state effects did not change the MO interpretation significantly. For the DFT calculations, omitting relaxation effects may lead to a slight overestimation of transition energies, and thus of the RIXS-NEXAFS energy gap. The reason for the weak effects of relaxation are the strong localized core-level excitations and the weak nature of the intermolecular interactions (*i.e.* hydrogen bonding) compared to intramolecular interactions (*i.e.* covalent bonds and protonation effects).

Table 2. Values for the energy gap of interest A↔1π* reflected by the experimental UV-Vis, RIXS-NEXAFS, and by TD-DFT and DFT calculations.

Energy gap A↔1π*	Non-ionic PABA	Anionic PABA
UV-Vis (experimental)	~ 4.51 eV	~ 4.68 eV
TD-DFT (calculated)	4.64 eV	5.43 eV
RIXS-NEXAFS (experimental)	5.06 eV	5.89 eV
DFT (calculated)	5.01 eV	5.86 eV

The differences noted in the A↔1π* gap presented in Table 2 between UV-Vis and RIXS-NEXAFS spectroscopies may be explained through the presence of a hole in the valence band in the UV-Vis, that exerts a smaller effect over the final state than the core hole from RIXS-NEXAFS. The TD-DFT PCM calculations also involve a water continuum (thus PABA surface-polarized monomers

with slightly different MO energy levels), yet this rather simplistic model is sufficient to picture the increase of the A↔1π* energy gap when moving from non-ionic to anionic PABA species.

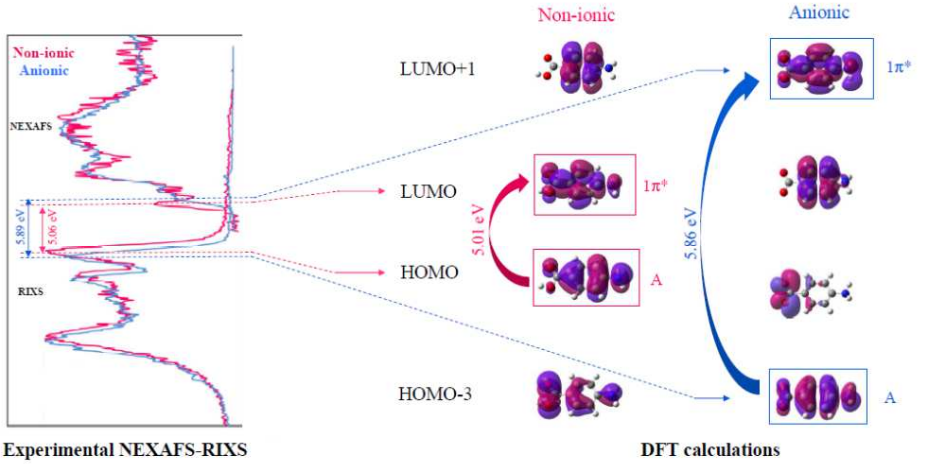


Figure 2. Nitrogen RIXS and NEXAFS showing the HOMO↔LUMO gap for non-ionic and anionic species, correlated with transitions (arrows) between DFT MOs of interest labelled as A and 1π*.

4. Conclusions

Monitoring the electronic transitions of PABA in the UV range in aqueous solution at intermediate, high and low pH, and interpreting the results with TD-DFT calculations provides a spectroscopic identification of PABA species that is complementary to the fine detail arising from the previously reported NEXAFS-RIXS combination. The effect of pH can be translated in the UV-Vis by the shifts to shorter wavelengths in the main absorption peaks encountered in the spectra of high and low pH solutions with respect to the intermediate pH solutions, which are reflected by the TD-DFT simulations. These shifts are explained by A↔1π* energy gap alterations with the changes in electronic structure accompanying the pH variation; the DFT calculations associated with the experimental RIXS-NEXAFS identify well with the orbital gap observed in the TD-DFT simulations that reflects the main absorption peaks in the UV-Vis, thus bridging the synchrotron X-ray core-level spectroscopies NEXAFS and RIXS with the laboratory UV-Vis technique.

Acknowledgements

This work was supported by the EPSRC Critical Mass Grant EP/I013563/1 and use of synchrotron at BESSY II (HZB) by European Community's Seventh Framework Programme, grant no. 312284.

References

[1] Stevens J S, Gainar A, Suljoti E, Xiao J, Golnak R, Aziz E F and Schroeder S L M 2015 *Chem. Eur. J.* **21** 7256–63.
[2] van de Graaf B, Hoefnagel A J and Wepster B M 1981 *J. Org. Chem.* **46** 653–57.
[3] Grimme S and Parac M 2003 *Chem. Phys. Chem.* **3** 292-95.
[4] Lampman G M, Pavia D L, Kriz G S and Vyvyan J R 2001 *Spectroscopy* (Belmont: Brooks/Cole Cengage Learning).
[5] Stevens J S, Seabourne C R, Jaye C, Fischer D A, Scott A J and Schroeder S L M 2014 *J. Phys. Chem. B* **118** 12121–28.

Journal of Physics: Conference Series **712** (2016) 012136 doi:10.1088/1742-6596/712/1/012136

NEXAFS Chemical State and Bond Lengths of p-Aminobenzoic Acid in Solution and Solid State

J S Stevens[1], A Gainar[1], E Suljoti[2], J Xiao[2], R Golnak[2], E F Aziz[2,3], S L M Schroeder[4,5]

[1]School of Chemical Engineering and Analytical Science, The University of Manchester, Oxford Road, Manchester, M13 9PL, UK
[2]Joint Ultrafast Dynamics Lab in Solutions and at Interfaces (JULiq), Helmholtz-Zentrum Berlin für Materialien und Energie, Albert Einstein Strasse 15, 12489 Berlin, Germany
[3]Freie Universität Berlin, Arnimallee 14, 14195 Berlin, Germany
[4]School of Chemical and Process Engineering, University of Leeds, Leeds, LS2 9JT, UK
[5]DIAMOND Light Source Ltd., Chilton, Didcot, OX11 0QX, UK

E-mail: joanna.stevens@manchester.ac.uk; s.l.m..schroeder@leeds.ac.uk

Abstract. Solid-state and solution pH-dependent NEXAFS studies allow direct observation of the electronic state of para-aminobenzoic acid (PABA) as a function of its chemical environment, revealing the chemical state and bonding of the chemical species. Variations in the ionization potential (IP) and 1s→π* resonances unequivocally identify the chemical species (neutral, cationic, or anionic) present and the varying local environment. Shifts in σ* shape resonances relative to the IP in the NEXAFS spectra vary with C–N bond length, and the important effect of minor alterations in bond length is confirmed with nitrogen FEFF calculations, leading to the possibility of bond length determination in solution.

1. Introduction

Through X-ray absorption and emission spectroscopies, the chemical, electronic and structural properties of organic species can be investigated. The nature of solution species and their interactions can play an important role in the activity of biological systems [1] and in the structural outcome and therefore properties of crystallization products [2]. Near-edge X-ray absorption fine structure (NEXAFS) measurements allow direct observation of the electronic state of *para*-aminobenzoic acid (PABA) as a function of its chemical environment [3]. While the IP and π* resonances are characteristic of the species and related to their chemical and electronic nature in the solution and solid state, the dominating influence of bond length on the σ* shape resonance proposed from experimental data is here investigated with FEFF calculations, exploring the ability and sensitivity of measuring bond lengths.

2. Methods

2.1. N K-edge Near Edge X-ray Absorption Fine Structure (NEXAFS) Spectroscopy

Solid-state partial electron yield (PEY) NEXAFS measurements were performed on the U7a beamline of the National Synchrotron Light Source (NSLS) at Brookhaven National Laboratory, NY with the sample at 54.7°, a 600 l/mm monochromator grating, and an entrance grid bias of -150 V [3]. Spectra were normalized by the simultaneously recorded total electron yield signal from a gold-coated, 90% transmission grid in the incident beam (I_0), and the monochromator energy scale was calibrated using the 400.6 eV 1st π* of a titanium nitride grid. Solution-state partial fluorescence yield (PFY) spectra were recorded using the liquid microjet of the LiXEdrom endstation at the U41-PGM beamline of the BESSY II synchrotron at Helmholtz Zentrum Berlin (HZB) using an 18 μm diameter glass nozzle with 0.6 mL min^{-1} flow rate [3]. Beamline energy calibration was performed with N_2 gas (total electron yield X-ray absorption spectroscopy 1s→2p π transition). Peak fitting and normalization were performed using the Athena software [4,5], with arctan steps for the ionization potentials (IPs, edge steps) and Gaussian functions for the peaks.

2.2. FEFF calculations

FEFF8.2 [6] was used to simulate the σ* shape resonance in the N K-edge NEXAFS for a geometry optimized monomer of non-ionic PABA. The SCF (7.0), XANES (4.0), and FMS (10.0) cards were used, and the calculated data were rigid shifted by +2.3 eV. To identify the influence of the C–N bond length, the C–N bond length of α-PABA was increased by 0.002-0.1 Å relative to the equilibrium value for the non-ionic monomer.

3. Results and Discussion

Nitrogen K-edge NEXAFS measurements reveal variations for the PABA species in both the solid state and in solution (Figure 1). Characteristic spectra are obtained for the different chemical species (non-ionic, cationic, anionic) present in solution with variation in pH, with a distinctive electronic signature for the cationic form following the change in nitrogen protonation state (NH_2 to NH_3^+). Shifts in the ionization potential (IP, edge step) and 1s→π* resonances unequivocally identify the chemical and electronic nature of the species present [3,7]. Further slight shifts and intensity variations are observed for the impact of differing local environments and interactions for the three non-ionic species investigated (α- and β-crystalline forms and the solutions species in methanol).

The energy of the σ* resonance associated with the C–N bond relative to the IP also varies across the series (term value $\delta_{C\text{-}N}$, Figure 1). The decrease in $\delta_{C\text{-}N}$ appears to correlate well with successively longer C–N bond lengths reported by X-ray diffraction (XRD) for the solid-state (Table 1, y = -39.664x + 61.133) [3]. This relationship can then be applied to the NEXAFS data, using it to provide an indicator of bond lengths for PABA species in solution.

Table 1. N K-edge NEXAFS features and bond lengths [3].

	$\delta_{C\text{-}N}$ / eV	$\Delta\delta_{C\text{-}N}$ / eV	XRD C–N / Å	XRD lengthening C–N / Å	NEXAFS C–N / Å[b]
Non-ionic α-PABA	6.31	0	1.379(2)	0	1.382
Non-ionic β-PABA	5.49	-0.82	1.408(3)	0.029	1.403
Anionic pH 11	5.40	-0.91	1.410(2)[a]	0.031	1.405
Non-ionic Methanol	5.19	-1.12	/	/	1.411
Cationic pH 1	3.16	-3.15	/	/	1.462
Cationic HCl salt	3.05	-3.26	1.4626(5)	0.0836	1.464

[a] C–N bond length from the hydrated Na salt shown for comparison
[b] y = -39.664x + 61.133 for linear relationship between $\delta_{C\text{-}N}$ and XRD C–N/Å [3].

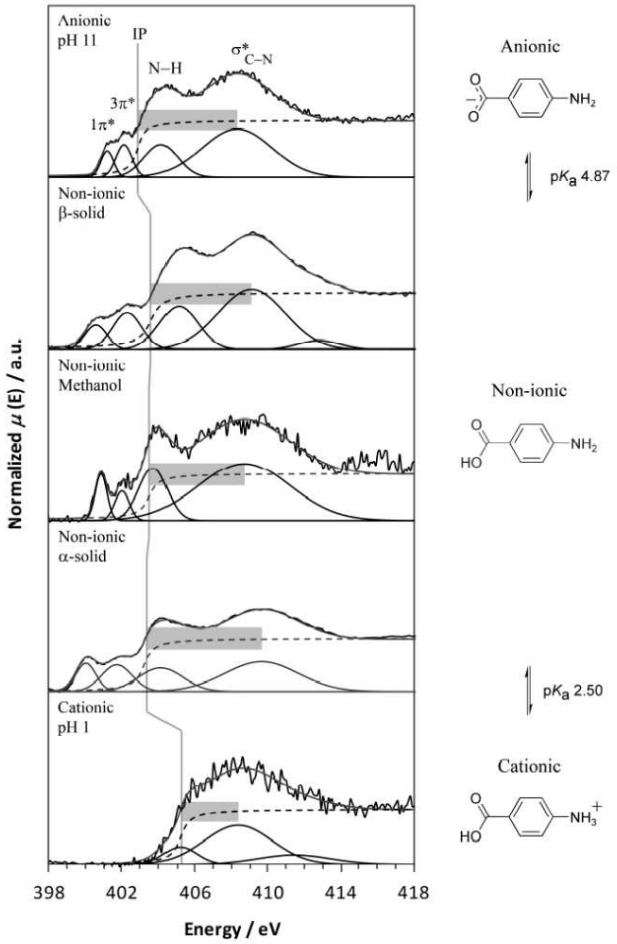


Figure 1. Nitrogen K-edge NEXAFS for anionic, non-ionic, and cationic PABA species in the solid and solution state, with the IPs marked by grey lines and the magnitude of the term value $\delta_{C\text{-}N}$ highlighted by grey boxes (adapted under CC BY license [3,7]).

To further investigate the impact of the C–N bond of PABA on the nitrogen NEXAFS, FEFF calculations were performed. Using an optimized non-ionic monomer as input, only the C–N length was varied, keeping all other parameters constant. This enables the effect of the C–N bond length to be isolated from other changes occurring between the species.

The C–N length was initially extended by 0.1 Ångstroms (Å), where a clear alteration and decrease in σ* energy of almost 2.5 eV is observed relative to the IP ($\Delta\delta_{C\text{-}N}$, Figure 2). The changes in C–N bond length in practice are smaller than 0.1 Å, ranging from 0.002 to 0.08 Å for the different PABA species (Table 1). Smaller increments were therefore tested, with lengthening by 0.03 Å (to mimic that occurring on moving from non-ionic α-PABA to β- or anionic PABA) resulting in a decrease in

energy by around 0.8 eV (Figure 2). This compares favourably with experimental NEXAFS, where a 0.82-0.91 eV change occurs (for cationic species, underestimation of FEFF $\Delta\delta_{C\text{-}N}$ values may be related to the alteration in electronic structure and geometry accompanying the positive charge on N, bent NH_2 → tetrahedral NH_3^+). The extreme sensitivity is further demonstrated for the case with just 0.005 Å longer C–N (Figure 2), and even the small shift occurring with 0.002 Å lengthening is observable, showing that even relatively subtle variations have a noticeable influence. The results suggest that bond length variations should be quantifiable to a precision of approximately 0.01 Å.

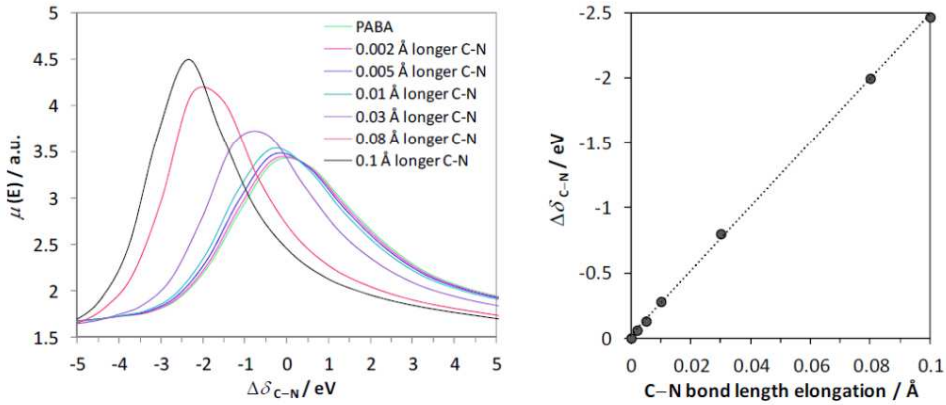


Figure 2. FEFF calculations for the nitrogen K-edge NEXAFS of PABA, demonstrating the effect of lengthening the C–N bond on the σ* energy relative to IP ($\delta^*_{C\text{-}N}$).

4. Conclusions

Nitrogen NEXAFS acts as a chemical and structural indicator for PABA, identifying the chemical species present in both the solid-state and in solution as well as providing a measure of C–N bond length. The ability to detect even minor variations in structure is verified by the FEFF calculations through incremental variation of bond length, with changes from as little as 0.002 Å alteration visible. NEXAFS can then potentially provide bond lengths for solution (or amorphous) species, such as demonstrated with PABA, and the combination of experiment and calculations is particularly strong.

Acknowledgements

Use of the NSLS at BNL was supported by U.S. Department of Energy Contract No. DE-AC02-98CH10886, and BESSY II at HZB by European Community's Seventh Framework Programme, grant no. 312284. JSS, AG and SLMS acknowledge support by EPSRC Critical Mass Grant EP/I013563/1.

References

[1] Chatterjea M N and Shinde R. 2012 *Textbook of Medical Biochemistry* (New Delhi: Jaypee Brothers Medical Publishers (P) Ltd.)
[2] Davey R. J., Schroeder S L M, and ter Horst J H 2013 *Angew. Chem., Int. Ed.* **52** 2166-2179
[3] Stevens J S, Gainar A, Suljoti E, Xiao J, Golnak R, Aziz E F and Schroeder S L M 2015 *Chem. - Eur. J.* **21** 7256-7263
[4] Ravel B and Newville M 2005 *J. Synchrotron Radiat.* **12** 537-541
[5] Newville M 2001 *J. Synchrotron Radiat.* **8** 322-324
[6] Ankudinov A L, Ravel B, Rehr J J and Conradson S D, 1998 *Phys. Rev. B: Condens. Matter Mater. Phys.* **58** 7565-7576
[7] Stevens J S, Seabourne C R, Jaye C, Fischer D A, Scott A J and Schroeder S L M 2014 *J. Phys. Chem. B* **118** 12121-12128

Journal of Physics: Conference Series **712** (2016) 012133 doi:10.1088/1742-6596/712/1/012133

NEXAFS and XPS of p-Aminobenzoic Acid Polymorphs: The Influence of Local Environment

J S Stevens[1], A Gainar[1], C Jaye[2], D A Fischer[2], S L M Schroeder[3,4]

[1]School of Chemical Engineering and Analytical Science, The University of Manchester, Oxford Road, Manchester, M13 9PL, UK
[2]Material Measurement Laboratory, National Institute of Standards and Technology, Gaithersburg, MD, 20899, USA
[3]School of Chemical and Process Engineering, University of Leeds, Leeds, LS2 9JT, UK
[4]DIAMOND Light Source Ltd., Chilton, Didcot, OX11 0QX, UK

E-mail: joanna.stevens@manchester.ac.uk; s.l.m.schroeder@leeds.ac.uk

Abstract. Nitrogen K-edge XPS and NEXAFS of the two polymorphic forms of para-aminobenzoic acid (PABA) are significantly different reflecting variation in hydrogen bonding. Alteration in hydrogen bonding at the amino group leads to a shift to high energy for both the XPS N 1s core level and the 3π* NEXAFS resonance with β-PABA. Participation of the amine group in the aromatic system causes the 1π* resonance to be sensitive to the nature of the intermolecular bonding at the *para*-carboxylic acid group, and a shift to low energy for α-PABA is observed due to hydrogen-bonded carboxylic acid dimer formation. FEFF calculations also successfully reproduce both the energy and intensity variations observed for the σ* shape resonance associated with the C–N bond, with the majority of the decrease in energy observed for β-PABA arising from the longer C–N bond.

1. Introduction

The local 3-dimensional environment and interactions of a molecule in the solid-state can have a substantial influence on physical and chemical properties, such as stability, solubility, melting temperature, and reactivity [1]. The ability to characterize and identify such forms (polymorphs) with identical chemical composition but variations in crystal structure (structural packing, interactions, and bonding) is thus extremely important. Understanding or insight into the electronic nature of these materials could provide the basis for investigating how they form in solution, and therefore ultimately how to control the outcome of crystallization [2]. Near-edge X-ray absorption fine structure (NEXAFS) spectroscopy and X-ray photoelectron spectroscopy (XPS) are ideal for this purpose, as they intrinsically probe the local environment of X-ray absorbing atoms. The effects of the local environment and interactions were initially investigated for *para*-aminobenzoic acid (PABA), an intermediate for folic acid synthesis with pharmaceutical applications, by nitrogen NEXAFS and *ab initio* density functional theory (DFT) [3]. Comparison of the crystalline α-form and an isolated molecule revealed substantial differences for the extreme case of removing all effects of the surroundings, and this was especially notable for the calculated N K-edge NEXAFS [3]. XPS and NEXAFS of the nitrogen moiety in PABA have therefore been collected for the two different crystalline polymorphs (α- and β-, Figure 1) [4,5] to investigate their sensitivity to more subtle

variations in local environment and intermolecular interactions. This is complemented by investigation into the influence of relatively minor alteration in bond length with FEFF calculations.

2. Methods

2.1. Materials

α-PABA was used as supplied (>99% purity, Sigma-Aldrich UK) and β-PABA was formed through an aqueous slurry of α-PABA at 5°C. Powder samples were used for analysis.

2.2. X-ray Photoelectron Spectroscopy (XPS)

N 1s XP spectra were recorded with a Kratos Axis Ultra instrument, employing a monochromatic Al K_α source (1486.69 eV) operating at 180 W (15 kV and 12 mA) [3]. Samples were fixed using double-sided tape and measurements were obtained in constant analysis energy mode below 10^{-8} mbar, with a pass energy of 20 eV, 0.1 eV steps, and 500 ms dwell time per data point. Analysis of the data was carried out with CasaXPS software using a Shirley background and GL(30) lineshape [6]. Samples were referenced to the C 1s photoemission C=C at 284.8 eV.

2.3. Near Edge X-ray Absorption Fine Structure (NEXAFS) Spectroscopy

N K-edge partial electron yield (PEY) NEXAFS measurements were performed on the U7a beamline of the National Synchrotron Light Source (NSLS) at Brookhaven National Laboratory, NY with the sample at 54.7°, a 600 l/mm monochromator grating, and an entrance grid bias of -150 V [3]. Spectra were normalized by the simultaneously recorded I_0, and the monochromator energy scale was calibrated using the 400.6 eV 1st π* of a titanium nitride grid using the Athena software [7,8]. Repeatability of resonance positions is <0.1 eV.

2.4. FEFF calculations

FEFF8.2 [9] was used to simulate the σ* shape resonance in the N K-edge NEXAFS for a molecule from the crystal structure of α- and β-PABA following geometry optimization of the crystal structure with CASTEP [3]. The SCF (7.0), XANES (4.0), and FMS (10.0) cards were used, and the calculated data were rigid shifted by +2.3 eV. To identify the influence of the C–N bond length, calculations were also run in which the average C–N bond length of α-PABA was increased by 0.035Å relative to the equilibrium value in the crystal structure so that it matched that of β-PABA (1.399 Å).

3. Results and Discussion

The nitrogen XPS and NEXAFS for the α- and β-PABA polymorphs yield characteristic spectra (Figures 1 and 2), with NEXAFS clearly differentiating between the two forms. For N 1s XPS, a small yet reproducible increase of around 0.2 eV in binding energy is observed for β-PABA (Figure 1). This follows the change in intermolecular hydrogen bonding with the additional hydrogen-acceptor effect at the amine nitrogen in β-PABA (Figure 1), reducing the electron density on nitrogen.

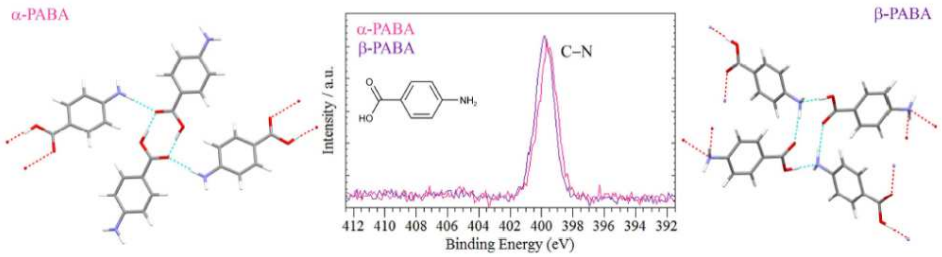


Figure 1. Nitrogen 1s XPS and intermolecular interactions of the α- and β-PABA polymorphs.

In the N K-edge NEXAFS spectra, both energy and intensity shifts occur for each of the peaks (Figure 2), reflecting the influence of variations in the local environment (intermolecular H-bonding, electron density distribution) and geometry (such as bond lengths) between the polymorphs. Visualization of the associated unoccupied π* molecular orbitals (MOs) shows the first energy resonance (1π*) involves promotion from the core N 1s level into the lowest unoccupied molecular orbital (LUMO), which is delocalized across the amine, aromatic ring and *para*-carboxylic acid group (Figure 2) [3]. The chemical state and interactions at the opposite end of the ring to nitrogen appear to influence the position of this 1π* resonance, with successively increasing absorption energy reflecting the carboxylic acid dimer in α-PABA, hydrogen bonding to the amino group of another molecule in β-PABA (Figure 1), and the carboxylate COO^- in the anionic form [10]. In contrast, the second resonance (3π*) involves the LUMO+2 (Figure 2), which is much more concentrated around the nitrogen atom [3]. This leads to it being more sensitive to the nature of the interactions at nitrogen, with the additional hydrogen-acceptor effect at the amine nitrogen in β-PABA (Figure 1) increasing the energy of this 3π* MO in addition to that of the N 1s core level (Table 1). The O K-edge π* resonances are correspondingly decreased in energy with the accompanying additional hydrogen-donor at oxygen.

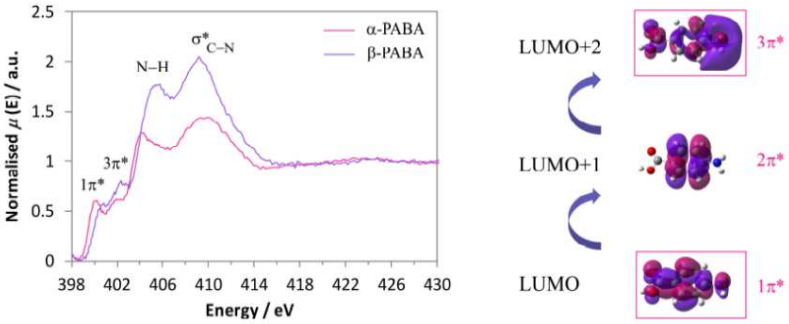


Figure 2. Nitrogen K-edge NEXAFS for the α- and β-PABA polymorphs, and unoccupied MOs involved in the first two resonances (outlined with boxes). Significant promotion to LUMO+1 (2π*) is not expected as it has no density at N.

Table 1. Nitrogen K-edge NEXAFS features

NEXAFS features	α-PABA / eV	β-PABA / eV
IP (aromatic NH_2)	403.30	403.60
N 1s → 1π*	400.00	400.58
N 1s → 3π*	401.70	402.30
N 1s → N–H	404.09	405.11
N 1s → $\sigma^*_{C\text{-}N}$	409.61	409.09

Following the edge step (ionization potential, IP) and third resonance associated with N–H, the broader post-edge fourth resonance is a σ* shape resonance for the C–N bond [3,10]. This $\sigma^*_{C\text{-}N}$ peak shows a shift to lower energy for β-PABA, in contrast to the increased π* and IP positions (Table 1). This is investigated for the two polymorphs using the FEFF code (as it predicts the post-continuum resonances) [9]. The results for α- and β-PABA show significant differences in both energy and intensity for the two polymorphs (Figure 3). Comparison of the FEFF output with the experimental data (Figure 3) shows excellent correlation, successfully predicting both the relative energy and intensity variations.

The dependence of this σ* resonance on the C–N bond was further investigated through variation of the C–N bond length for α-PABA. Lengthening of the C–N bond to that for β-PABA (+0.035 Å), leads to a decrease in σ* energy relative to the IP, accounting for the majority of the energy and intensity shifts between α- and β-PABA (Figure 3).

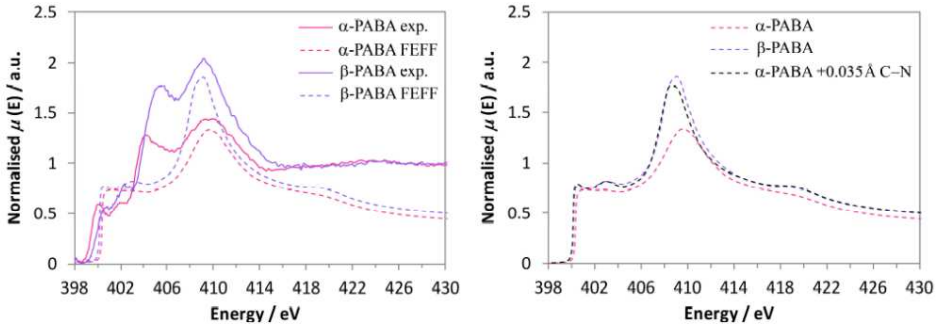


Figure 3. Comparison of experimental (solid line) and FEFF-calculated (dashed line) nitrogen K-edge NEXAFS for the two polymorphs of PABA, and the impact of variation in C–N bond length.

4. Conclusions

Insights into the electronic and structural nature of the two polymorphs of PABA are obtained with nitrogen K-edge NEXAFS and XPS. A small shift in core level N 1s energy is attributed to the variation in intermolecular hydrogen bonding. Significantly different, highly characteristic NEXAFS spectra arise from the changes in local surroundings and bonding, with the π* shifts related to the nature of the MOs and local environment, and the σ* shape resonance primarily influenced by the C–N bond length. This demonstrates the sensitivity of XPS and NEXAFS to the relatively minor changes in local environment despite their identical chemical composition.

Acknowledgements

Use of the NSLS at BNL was supported by U.S. Department of Energy under Contract No. DE-AC02-98CH10886. JSS, AG and SLMS acknowledge support by EPSRC Critical Mass Grant EP/I013563/1.

References

[1] Cartensen J T *Pharmaceutical Principles of Solid Dosage Forms* 1993 (Lancaster, PA: Technomic Publishing Co.)
[2] Davey R. J., Schroeder S L M, and ter Horst J H 2103 *Angew. Chem., Int. Ed.* **52** 2166-2179
[3] Stevens J S, Seabourne C R, Jaye C, Fischer D A, Scott A J and Schroeder S L M 2014 *J. Phys. Chem. B* **118** 12121-12128
[4] Lai T F and Marsh R E 1967 *Acta Crystallogr.* **22** 885-893
[5] Gracin S and Fischer A 2005 *Acta Crystallogr., Sect. E: Struct. Rep. Online* **61** o1242.
[6] Fairley N and Carrick A 2005 *The Casa Cookbook - Part 1: Recipes for XPS Data Processing. Acolyte Science* (Knutsford, Cheshire: Acolyte Science)
[7] Ravel B and Newville M 2005 *J. Synchrotron Radiat.* **12** 537-541
[8] Newville M 2001 *J. Synchrotron Radiat.* **8** 322-324
[9] Ankudinov A L, Ravel B, Rehr J J and Conradson S D, 1998 *Phys. Rev. B: Condens. Matter Mater. Phys.* **58** 7565-7576
[10] Stevens J S, Gainar A, Suljoti E, Xiao J, Golnak R, Aziz E F and Schroeder S L M 2015 *Chem. - Eur. J.* **21** 7256-7263