Electron-Vibration-Vibration Two-Dimensional Infrared Spectroscopy: a tool to investigate interactions in plant proteins and mammalian peptides

Sophie Anne Sim

Doctor of Philosophy

Department of Chemistry

Imperial College London

2018

Abstract

The structural elucidation and characterisation of biomolecular complexes is crucial to both the agrichemical and pharmaceutical industries. Understanding the specific interactions that govern protein-inhibitor interactions enables more intelligent design of herbicide and drug candidates. While there are powerful tools currently available to obtain protein-inhibitor structures, they alone cannot answer the questions posed by protein-inhibitor binding. EVV 2DIR spectroscopy is a novel technique capable of detecting protein-inhibitor interactions that can provide structural information complementary to that of other methods (such as direct measurement of hydrogen bond formation) to build up a more complete picture of protein-inhibitor binding.

This thesis presents the first EVV 2DIR spectra of plant protein-inhibitor complexes, comprised of the herbicide target 4-Hydroxyphenylpyruvic acid dioxygenase (HPPD) bound to various herbicide candidates. Difference spectra of HPPD and project compound 329 consistently showed one cross-peak that could be tentatively assigned to coupled vibrational modes specific to protein-inhibitor binding. The data presented exhibits the capability of EVV 2DIR spectroscopy to detect protein-inhibitor binding for plant systems and demonstrates its potential as a tool for the screening of herbicide candidates.

The first EVV 2DIR spectra for the peptide amyloid beta (1-40) both with and without Cu(II) and Zn(II) are also presented. Comparison of the spectra reveals changes indicative of the literature observations that amyloid beta (1-40) forms rigid fibrils at an acidic pH, and that the addition of Cu(II) and Zn(II) disrupts and retards fibril formation. Tentative assignments were made for the resultant spectra, however a lack of corresponding calculated EVV 2DIR spectra (previously produced from vibrational modes elucidated via DFT calculations) made definitive assignment of such complex spectra not possible. There are currently no published structures of amyloid beta (1-40) with Cu(II), so accompanying calculated spectra enabling conclusive assignments may reveal significant structural insights into the coordination of copper by amyloid beta in the future.

Declaration of Originality

All the work presented in this thesis is my own, unless otherwise stated:

- Dr Alastair Magness collected the high-resolution microscope images and wrote the “stripe removal” MATLAB code
- Dr Hugh Sowley wrote the original versions of the remaining MATLAB codes I then adapted as required for data-processing.

Copyright Declaration

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

Acknowledgements

Thank you to my Imperial College London supervisors, David Klug and Laura Barter, for the opportunity to work on such a challenging and novel project. Thanks also to John Paul Evans, my original supervisor from Syngenta. Thank you especially to Nathan Kidley, who took over the role when JP left Syngenta mid-project and has been a huge help in terms of project input and connecting me to contacts at Syngenta.

Thank you to Tony Keeley, Shradha Singh, Mike Langford, Tim Hawkes, and everyone else from Syngenta involved in providing invaluable training, materials and knowledge for the project.

Thank you to the European Union Seventh Framework program for funding the NexGenAgriChem programme, and to all other agencies who provided other funding that went towards this project.

A big thanks to Hugh Sowley, for all his training, patience and friendship, and for sharing the equal parts of fun and despair that came with this project. Thank you also to Julia Davies, for her kind and thorough training during the early stages of my PhD project, without which I would have a much lesser understanding of 2DIR spectroscopy.

Many thanks to Alastair Magness, who helped with various aspects of this thesis and provided a great deal of laughs and encouragement during my entire PhD. A special mention also to Nate and Jim for their friendship and support over the past few years.

Lastly, a huge thank you to my parents Michelle and Dave, and siblings Lucy, Rosie, George and Tom for your ongoing support all the way from New Zealand. The past few years have been a challenging adventure in many ways (especially the difficulty of being away from you all during my PhD), and I would never have got to this point without all of you.

Table of Contents

Abstract ... 3

Declaration of Originality ... 4

Copyright Declaration ... 4

Acknowledgements ... 5

List of Abbreviations ... 10

List of Figures ... 12

List of Tables ... 16

1. Introduction ... 17

1.1 Detection and structural characterisation of protein-inhibitor interactions ... 17

1.2 HPPD-herbicide complexes ... 24

1.3 Amyloid beta ... 27

2. Background & theory ... 30

2.1 Introduction to infrared spectroscopy ... 30

2.1.1 Harmonic oscillator: Diatomic molecules ... 30

2.1.2 Selection rules in vibrational transitions ... 32

2.1.3 Anharmonicity & coupled vibrations ... 34

2.2 Introduction to nonlinear spectroscopy ... 37

2.2.1 Classical model ... 37

2.2.2 Quantum mechanical model ... 40

2.3 Electron-Vibration-Vibration Two-Dimensional Spectroscopy ... 42

2.3.1 Introduction to Electron-Vibration-Vibration Two-Dimensional Spectroscopy ... 42

2.3.2 Principles of Electron-Vibration-Vibration Two-Dimensional Spectroscopy ... 46

2.3.2.1 Diagrammatic representations: Coherence pathways, WMEL diagrams & Feynman diagrams ... 46

2.3.2.2 Phase matching ... 50

2.3.2.3 EVV 2DIR Spectra 54
2.3.3 Applications of EVV 2DIR 57
2.3.3.1 Complex formation 59
2.3.3.2 Investigation of biological molecules 61
2.3.3.3 Resonance enhancement 63
3. Experimental Methods 69
3.1 EVV 2DIR laser system 69
3.1.1 Experimental overview 69
3.1.2 Pulse Generation & Amplification 70
3.1.2.1 Pulse generation 70
3.1.2.2 Regenerative amplification 71
3.1.3 Optical Parametric Amplification 72
3.1.3.1 White Light Continuum Generation 74
3.1.3.2 Optical Parametric Generation (OPG) 74
a. White light amplification 75
b. Signal seed frequency selection 75
c. Signal amplification 76
3.1.3.3 Difference Frequency Mixing (DFM) 76
3.1.4 Nitrogen purge 78
3.1.5 Beam Alignment & Overlap 79
3.1.5.1 Temporal control: T0 Calibration 80
3.1.5.2 Spatial control 82
3.1.6 Signal Detection 83
3.1.7 Modification of IR Range 83
3.1.8 Polystyrene calibration 84
3.1.9 Data processing 84
3.1.9.1 Smoothing 84
3.1.9.2 Background removal 85

3.1.9.3 OPA energy profile correction 86
3.2 Sample preparation 88
3.2.1 HPPD assay 88
3.2.2 Preparation of gel-like sample spots 89
3.2.2.1 HPPD humidity cell preparation 92
3.2.2.2 Amyloid beta humidity cell preparation 93
4. Detection of intermolecular plant-herbicide interactions 94
4.1 HPPD activity: iron vs non-iron treatment 95
4.2 Standard frequency range experiments 96
4.2.1 Non-iron-treated HPPD experiments 96
4.2.1.1 Sample preparation: Methods & Materials 96
4.2.1.2 FTIR data collection 99
4.2.1.3 EVV 2DIR data collection 99
4.2.1.4 Data processing 100
4.2.1.5 EVV 2DIR spectra of HPPD & BSA 101
4.2.1.6 Difference spectra 104
4.2.2 Comparison of iron-treated vs non-iron-treated HPPD binding 115
4.3 Lower frequency range experiments 117
4.3.1 Sample preparation 119
4.3.2 EVV 2DIR data collection 119
4.3.3 Data processing 120
4.3.4 EVV 2DIR spectra of HPPD & BSA 121
4.3.5 HPPD+herbicide and difference spectra 122
4.4 Limitations 129
4.5 Chapter Summary 130
5. Detection of intermolecular interactions in amyloid beta (1-40) peptide 133
5.1 Methodology 133
5.1.1 Materials & Methods 133

5.1.2 Humidity optimisation of amyloid beta spots 134
5.1.3 FTIR data collection 134
5.1.4 EVV 2DIR data collection 135
5.1.5 Data processing 135
5.2 EVV 2DIR spectrum of Aβ(1-40) 136
5.3 EVV 2DIR spectra of Aβ(1-40) + Cu(II) 145
5.4 EVV 2DIR spectra of Aβ(1-40) + Zn(II) 148
5.5 Limitations 154
5.6 Chapter Summary 155
6. Conclusions & Future Work 156
6.1 Conclusions 156
6.2 Future work 157
6.2.1 Future work with HPPD 157
6.2.2 Future work with amyloid beta 158
6.3 Summary 158
Appendix A: MATLAB codes 167
Appendix B: *Arabidopsis thaliana* HPPD sequence 170
Appendix C: Amyloid beta 1-40 peptide sequence 170

List of Abbreviations

2D NMR	Two-dimensional nuclear magnetic resonance
2DIR	Two-dimensional infrared
MD	Molecular dynamics
AHA	Azidohomoalanine
EVV	Electron-Vibration-Vibration
HPPD	4-Hydroxyphenyl pyruvate dioxygenase
HPP	4-Hydroxyphenylpyruvate
HGA	Homogentisate
At	Arabidopsis thaliana
Aβ	Amyloid beta
PDB	Protein Data Bank
FWM	Four-wave mixing
DOVE	Doubly vibrationally enhanced
WMEL	Wave-mixing energy level
DFT	Density functional theory method
HF	Hartree-Fock method
MP2	Second order Møller–Plesset method
B3LYP	Becke, three-parameter, Lee-Yang-Parr hybrid functional
OPA	Optical parametric amplifier
OPG	Optical parametric generation
BBO	β-Barium Borate
DFM	Difference frequency mixing

MAA	Maleylacetoacetate
HGO	Homogentisate dioxygenase
BTP	Bis-tris propane
BSA	Bovine serum albumin
329	HPPD project inhibitor 329

List of Figures

Figure 1: (a) The typical pulse sequence, (b) non-collinear geometry and (c) pump-probe geometry for ultrafast 2DIR experiments (from Reference 25). Note that the experiments presented in this thesis employ geometry (b).....19
Figure 2: Typical beam geometry of a 2DIR photon-echo experiment (from Reference 27). 21
Figure 3: A schematic of parallel amyloid beta fibril formation.....28
Figure 4: A simple ball-and-spring diagram of a heteronuclear diatomic molecule 31
Figure 5: A schematic showing the potential energy (V) vs nuclear displacement (x) curve of a molecule displaying harmonic oscillation.....31
Figure 6: The asymmetrical curve for a molecule exhibiting anharmonic vibrations 36
Figure 7: The response of linear ($\chi 1$) and nonlinear ($\chi 2$ and $\chi 3$) polarisations P as a function of the applied electric field E in a system.....39
Figure 8: The EVV 2DIR FWM process.....43
Figure 9: (a) FTIR spectra, (b) EVV 2DIR spectra, and (c) chemical structure of L-arginine mono-hydrochloride. Spectra (a) & (b) are from Reference 96.....45
Figure 10: WMEL diagrams for the EVV-IR_1, EVV-IR_2 and EVV-Raman methods depict the succession of coherences for each process.....47
Figure 11: Feynman diagrams depicting the coherence pathways of the three EVV 2DIR FWM regimes.....48
Figure 12: The pulse ordering and time-relative timings in EVV 2DIR spectroscopy 49
Figure 13: Phase matching vector diagrams for EVV 2DIR spectroscopy.....52
Figure 14: A vector diagram showing the noncollinear phase matching geometry used in the EVV 2DIR experiments presented in this thesis.....53
Figure 15: The WMEL diagram for EVV 2DIR shown in further detail, including the two-quanta transition to the combination $a+b$ band. 55
Figure 16: A schematic of an EVV 2DIR spectrum over the typical frequency range.....56
Figure 17: (a-d) Four wave mixing scans of acetonitrile.....58
Figure 18: a) An experimental four wave mixing 2DIR spectrum of a mixture of acetonitrile, deuterated acetonitrile and benzene, and b) a model spectrum of the same system.....59
Figure 19: (a) The chemical structure of drug SU5402, and (b) the same drug complexed with its target protein, FGFR1 (a human growth factor receptor).....60

Figure 20: EVV 2DIR spectra of the SS3 DNA sequence (left) and corresponding Myc2345 quadruplex (right), from Reference 49 62
Figure 21: EVV 2DIR images of Hematoxylin (left) and protein CH_3 (middle) distribution in a cross-section of mouse kidney, and the corresponding microscope image (right) 63
Figure 22: WMEL diagrams for non-resonant vs resonant Raman scattering processes 64
Figure 23: A schematic of the near-resonance enhancement of the retinal chromophore in bacteriorhodopsin 65
Figure 24: EVV 2DIR difference spectra of (a) FGFR1/SU5402 and (b) BSA/SU5402 (as a control) 66
Figure 25: (a) The calculated EVV 2DIR spectrum of SU5402 and proximal amino acids & water molecule in the FGFR1 binding site, and (b) the corresponding experimental FGFR1/SU5402 difference spectrum for comparison. 67
Figure 26: The experimental setup of the EVV 2DIR laser system 69
Figure 27: The nonlinear frequency down-conversion process. 73
Figure 28: The process of OPG 75
Figure 29: The typical output wavelength ranges vs pulse energy for a) signal and idler beams in OPG, and b) signal produced via difference frequency mixing 76
Figure 30: The process of DFM showing conversion of the OPG signal and idler into a DFM signal 78
Figure 31: Typical pulse energy values as a function of OPA2 frequency when unpurged (left) and purged (right) by nitrogen gas. 79
Figure 32: A schematic showing the key elements of the EVV 2DIR system which enable beam overlap, delivery & signal detection. 80
Figure 33: Typical T0 calibration curves for a) OPA1 and b) OPA2, with fitted third-order polynomials. 82
Figure 34: A diagram demonstrating the “nearest neighbour” method of smoothing used in EVV 2DIR data processing. 85
Figure 35: Typical output pulse energy values for OPA1 (left) and OPA2 (right) as a function of frequency 87
Figure 36: The K_m/V_{max} curve for HPPD activity 88
Figure 37: A high-resolution 20x brightfield microscope image of a FGFR1 sample spot with coffee-ring edge 90
Figure 38: a) a schematic of the parts that make-up a humidity cell and b) an eagle-eye view of an actual humidity cell 91
Figure 39: A high-resolution 20x brightfield microscope image of an HPPD gel-like spot .. 93

Figure 40: The UV-Vis spectrum of HPPD inhibitors 329 (blue) and Nitisinone (red) in DMSO,.....97
Figure 41: The averaged EVV 2DIR spectra of a) non-iron-treated HPPD, and b) BSA over the standard frequency range. 102
Figure 42: The averaged EVV 2DIR spectra of a) non-iron-treated HPPD + 329 and b) BSA + 329 over the standard frequency range. 104
Figure 43: The chemical structure of HPPD herbicide Mesotrione 106
Figure 44: The difference spectra for a) non-iron-treated HPPD/Mesotrione and b) the corresponding BSA/MST spectrum.....107
Figure 45: The chemical structure of HPPD herbicide candidate 329.....109
Figure 46: The averaged EVV 2DIR difference spectra for a) non-iron-treated HPPD and 329 and b) control protein BSA and 329 over the standard frequency range. 110
Figure 47: Image of At HPPD binding site docking model showing binding of a the commercially available herbicide Tembotrione to HPPD.....112
Figure 48: The chemical structure of the HPPD inhibitor Nitisinone. 114
Figure 49: Averaged EVV 2DIR difference spectra for a) non-iron-treated HPPD and Nitisinone, and b) BSA with Nitisinone over the standard frequency range. 114
Figure 50: The averaged EVV 2DIR difference spectra for HPPD/Nitisinone showing a comparison between a) non-iron-treated HPPD and b) iron-treated HPPD. 116
Figure 51: The averaged EVV 2DIR difference spectra of a) non-iron-treated HPPD/329 and b) iron-treated HPPD/329. 117
Figure 52: Calculated EVV 2DIR spectra of all amino acid side chains.....118
Figure 53: Averaged EVV 2DIR spectra of iron-treated HPPD collected over the lower ω_β frequency range, shown a) before and b) after removal of the two horizontal “stripe” spectral artefacts.....121
Figure 54: The averaged EVV 2DIR spectra of a) iron-treated HPPD and b) BSA collected over the lower frequency range.....122
Figure 55: The averaged EVV 2DIR spectra for a) HPPD+329, and b) BSA+329 over the lower frequency range.....123
Figure 56: The averaged EVV 2DIR difference spectra for a) HPPD/329 and b) BSA/329 over the lower frequency range.....124
Figure 57: The averaged EVV 2DIR spectra for a) HPPD+Nitisinone and b) BSA+Nitisinone over the lower frequency range.....127
Figure 58: The EVV 2DIR difference spectra for a) HPPD/Nitisinone and b) BSA/Nitisinone over the lower frequency range.....128

Figure 59: The EVV 2DIR spectrum of Aβ(1-40) 137
Figure 60: A vertical cut-through of the Aβ(1-40) EVV 2DIR spectrum centred at ω_α=2885 cm^{-1}. 139
Figure 61: A vertical cut-through of the column of peaks centred at ω_α=1615 cm^{-1} present in the EVV 2DIR spectrum for Aβ(1-40) 140
Figure 62: A vertical cut-through of the column of peaks centred on ω_α=1700 cm^{-1} present in the EVV 2DIR spectrum for Aβ(1-40) 142
Figure 63: The cryo-electron microscopy structure of Aβ(1-42) with mapped electron density 143
Figure 64: EVV 2DIR spectra of a) Aβ(1-40) alone, and b) with Cu(II) added. 145
Figure 65: A comparison of the ω_α=1615 cm^{-1} vertical cut-through sections of the EVV 2DIR spectra for Aβ(1-40) alone (red) and Aβ(1-40) + Cu(II) (blue) 146
Figure 66: A comparison of the ω_α=1700 cm^{-1} vertical cut-through sections of the EVV 2DIR spectra for Aβ(1-40) alone (red) and Aβ(1-40) + Cu(II) (blue) 147
Figure 67: The EVV 2DIR spectra of a) Aβ(1-40) alone and b) with added zinc 149
Figure 68: A comparison of the ω_β=2820 cm^{-1} horizontal cut-through sections of the EVV 2DIR spectra for Aβ(1-40) alone (red) and Aβ(1-40) + Zn(II) (blue). 150
Figure 69: A comparison of the vertical cut-through sections of the cross-peaks at ω_α=1615 cm^{-1} for Aβ(1-40) (red) and Aβ(1-40)+Zn(II) (blue) present in their EVV 2DIR spectra. 151
Figure 70: The mechanism of zinc coordination in Aβ(1-40) 152
Figure 71: A comparison of the vertical cut-through sections of the cross-peaks at ω_α=1700 cm^{-1} for Aβ(1-40) (red) and Aβ(1-40)+Zn(II) (blue) present in their EVV 2DIR spectra. 153

List of Tables

Table 1: Coherence pathways for each of the three processes that it is possible to undergo in EVV 2DIR spectroscopy. 47
Table 2: The revised coherence pathway for the EVV-IR_1 process 54
Table 3: The list of saturated salt solutions and corresponding relative humidity values utilised in the sealed humidity cells used for sample preparation in this thesis 92
Table 4: General vibrational assignments of spectral features which appear in both the HPPD and BSA EVV 2DIR spectra in Figure 31. 103
Table 5: Tentative assignments for the non-specific binding peaks present in the a) HPPD/Mesotrione and b) BSA/Mesotrione EVV 2DIR difference spectra. 108
Table 6: Preliminary assignments for the coupled vibrational modes giving rise to the binding peak in the HPPD/329 difference spectrum. 111
Table 7: Tentative assignments for the non-specific interactions giving rise to the cross peak in the Nitisinone/protein difference spectra. 115
Table 8: Tentative assignments for vibrational modes giving rise to coupled cross-peaks indicative of HPPD-specific binding in the EVV 2DIR HPPD/329 difference spectrum. 125
Table 9: Group assignments for cross-peaks appearing in the EVV 2DIR spectrum of Aβ(1-40) 138
Table 10: Tentative assignments for the coupled vibrations giving rise to the column of peaks centred on ω_α=1615 cm^{-1}. 141
Table 11: Tentative assignments for the coupled vibrational modes giving rise to the cross peaks centred on ω_α=1700 cm^{-1}. 144

1. Introduction

1.1 Detection and structural characterisation of protein-inhibitor interactions

The importance of proteins to molecular biology is summarised well by Nobel Prize winning chemist Linus Pauling: "Proteins hold the key to the whole subject of the molecular basis of biological reactions"[1]. Central to these biological reactions are the formation of biomolecular complexes, and techniques that characterise the interactions that facilitate these have been the subject of a wide research effort.

Particularly, protein-inhibitor complexes have significant biological relevance with regards to understanding how drugs and herbicides bind to their target mammalian and plant proteins, respectively. This is hugely important to both the pharmaceutical and agrichemical industries, where several billion dollars is spent every year on the screening and development of new inhibitor candidates[2–5]. Specifically, the agrichemical sector is continually searching for new herbicide compounds to counteract the rising levels of herbicide resistance developed by many species of weeds[3,5,6].

To date, most protein-inhibitor structures have been elucidated using X-ray crystallography and two-dimensional nuclear magnetic resonance (2D NMR) techniques. While both are powerful methods facilitating significant insight into the specific interactions that govern protein-inhibitor complexes, both also have practical (and informational) limitations which must be considered. X-ray crystallography requires high-quality crystallisation of the system to be studied[7], which is not always possible, particularly for membrane proteins or proteins lacking a native structure[8]. The growth of such crystals can be hugely resource-consuming, both in terms of time and material required. Furthermore, protein crystals are often lacking the water molecules otherwise present in a hydrated state that are known to play a crucial role in their structure *in vivo* [9–12], and if water molecules are present they are often mistaken for the salts proteins are often co-crystallised with[13]. Additionally, X-ray diffraction measures a static form of a protein which may not represent the dynamic conformation seen in solution, as the protein would be expected in its native environment[14].

2D NMR spectroscopy overcomes the issue of protein flexibility, as it enables the determination of protein-inhibitor complexes in solution. NMR spectroscopy measures the

chemical shifts of a sample in the presence of an applied magnetic field, which are characteristic of the chemical environment of each functional group occurring in the sample, enabling the extraction of structural and dynamical information[15]. More specifically, 2D NMR produces correlated spectra revealing which parts of the system experience through-space spin-spin coupling, enabling determination of the structure and geometry of the protein-inhibitor complex[16]. Because experiments are performed in solution, 2D NMR spectra can be collected of the unstructured and membrane proteins often unable to be crystallised for X-ray diffraction.

While both 2D NMR and 2DIR methods apply sequential pulses to evaluate coherences inherent to the system of interest, the former occurs via micro or millisecond pulses eliciting nuclear spin-spin couplings that last seconds, whereas the latter utilises femto or picosecond pulses to produce picosecond coherences which measure the coupling of vibrational modes of said system[17]. Additionally, 2D NMR has sensitivity issues arising from its Boltzmann distribution, which describes the difference in population between nuclear spin states. In 2D NMR, this difference is very small at room temperature because the energy levels are so closely spaced, meaning ground and excited states are almost evenly occupied. In contrast, the Boltzmann distribution between ground and excited vibrational states at room temperature is much larger, meaning that the ground state is populated much more than the excited state. Therefore, the energy of transitions occurring in NMR is typically 10^6 lower than for IR spectroscopy[18]. Consequently, NMR spectroscopy is subject to poor sensitivity[19], therefore requires a large amount of material to achieve adequate signal to noise, hence is less suitable as a high-throughput screening tool in structural biology. Moreover, the timescale resolution in NMR is limited to several milliseconds[20], which is slower than the occurrence of many important biomolecular interactions (such as hydrogen-bond formation and amino acid side-chain reorientation, both of which occur on a picosecond timescale[21,22]). Thus, NMR spectroscopy gives only a time-averaged picture of protein dynamics. NMR spectroscopy is also limited by protein size[23,24], with the extent of protein-ligand structures able to be obtained via this method typically limited to ~50 kDa[25], and experimental time increasing as a function of system size rendering it unsuitable to study larger protein-inhibitor systems.

Nonetheless, X-ray crystallography and NMR spectroscopy are still the most effective tools to aid in the understanding of protein-inhibitor complexes. However, the need for techniques that give *complementary* structural and dynamical information about interactions occurring on rapid timescales (including the formation of hydrogen-bond

networks and rapid side-chain fluctuations[22]) has become clear, and this has initiated the advent of two-dimensional infrared (2DIR) spectroscopic techniques applied to protein-inhibitor research.

Ultrafast spectroscopic experiments can either be carried out in the time-domain or the frequency-domain, however more recent mixed time-frequency domain methods have been devised which are now widely used[26]. Traditionally, 2DIR experiments were performed using pump-probe methods, where a short ultrafast pulse is split into two parts, with the stronger (pump) used to excite the sample, and the weaker (probe) employed to monitor the induced changes in reflectivity or transmission of the sample[22] (outlined in Figure 1c[26]). The time delay between the two is controlled by either an optical delay line or calcium fluoride wedges.

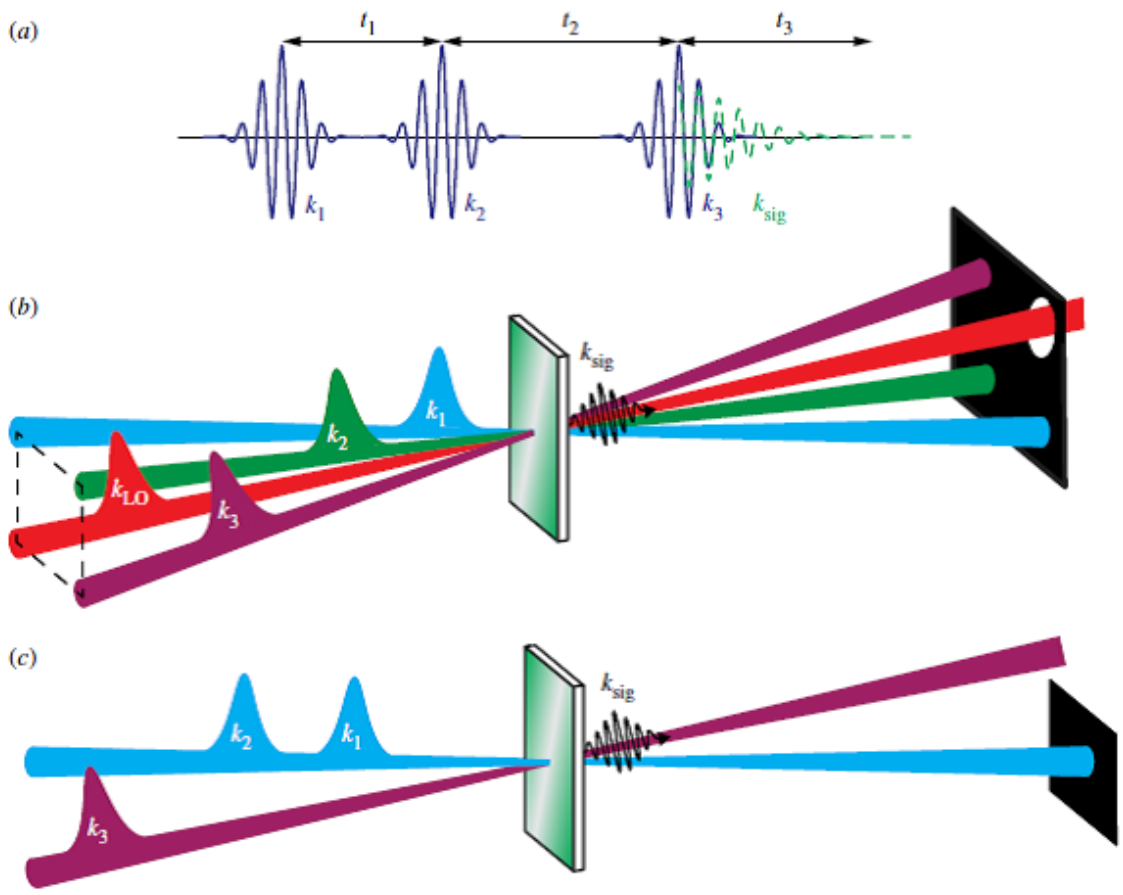


Figure 1: (a) The typical pulse sequence, (b) non-collinear geometry and (c) pump-probe geometry for ultrafast 2DIR experiments (from Reference 25). Note that the experiments presented in this thesis employ geometry (b).

Comparing the pump-probe geometry shown in Figure 1(c) with the four-wave mixing (FWM) non-collinear geometry (b) used in the experiments presented in this thesis, there are various advantages and disadvantages. Positives for the latter include; a background-free signal (estimated to improve signal-to-noise ~20 times more than pump-probe methods[27]), easy isolation of the generated outgoing signal from the incoming laser beams, and the potential to employ reversed polarisation schemes of the incoming pulses[26]. However, the key drawback to this geometry is that phase-matching must be maintained between the beams, which is difficult to implement experimentally, or distortions can appear in the resultant 2DIR spectra[22]. In contrast, 2DIR pump-probe methods overcome the phase-matching challenge when using two very different frequencies for the pump and probe beams (also known as "two-colour" experiments), but still need separate pump-probe measurements using the same set-up to measure the phase relationship between the beams[22,26].

Another traditional 2DIR method that has been utilised to study chemical and biomolecular structure and dynamics is the photon echo technique, detailed in Figure 2[28]. In this, the laser output is generally split into three beams, which are focused into sample such that they form three corners of a square after the sample, and the resultant FWM "echo" (signal) comprises the fourth corner, which is then overlapped with the (now weak) remaining laser output which measures the echo signal via heterodyne detection[22]. This is advantageous in that it amplifies the generated signal, and that the phase information of the signal is preserved.

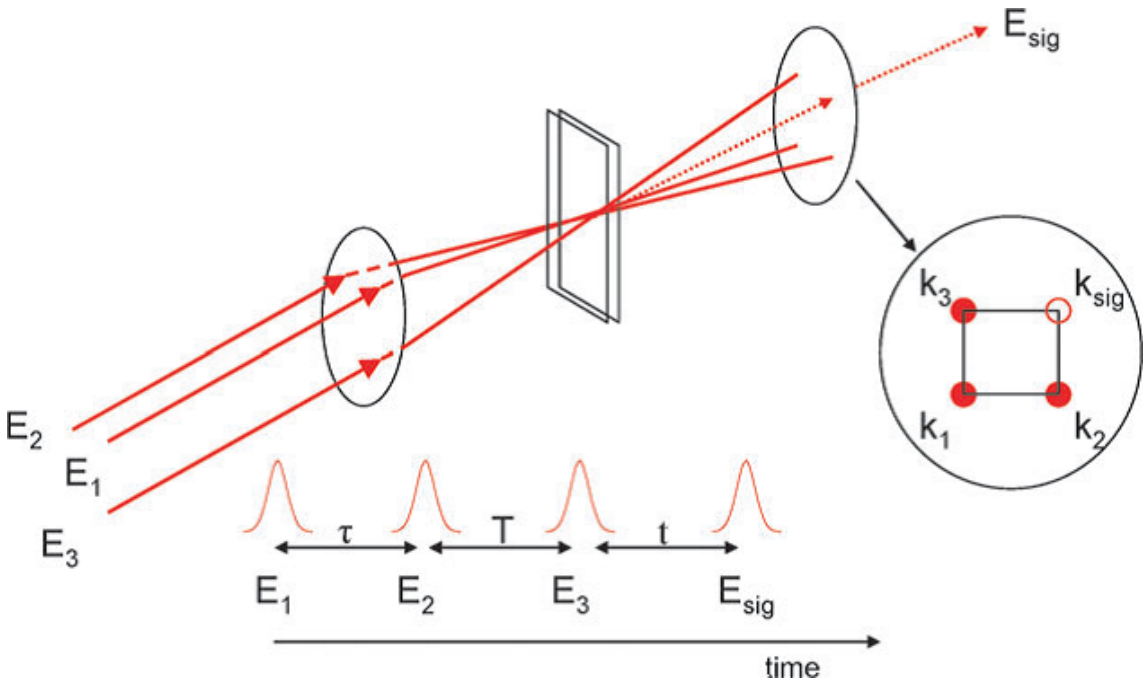


Figure 2: Typical beam geometry of a 2DIR photon-echo experiment (from Reference 27).

While these 2DIR methods are all coherent, FWM techniques, each has unique capabilities, pros and cons. These multidimensional ultrafast techniques (and many other variants, including 2D electronic and 2D electronic-vibrational spectroscopies) have been utilised to study the structure and dynamics of many chemical and biomolecular systems over the past decade, including quantum coherences in light-harvesting complexes[29–31], dynamics of vibrations in solid state systems[32–34], and protein folding[35,36] and structural dynamics[37–39], with applications ranging from solar energy to optoelectronics to pharmaceutical and agrichemical design. Some specific examples are discussed in more detail below, with particular emphasis on 2DIR experiments elucidating protein-ligand interactions.

With regards to protein-ligand interactions in particular, small groups with distinctive absorptions such as nitriles, azides and carbonyls have been introduced into proteins (or inhibitors) of interest to act as non-natural vibrational probes which give known 2DIR spectra, to elucidate structural and dynamical information[40,41]. The binding of anti-AIDS drug rilpivirine to its target protein HIV-1 reverse transcriptase has been examined using two-dimensional vibrational-echo spectroscopy. Rilpivirine has two arms comprised of a cyano moiety (one cyanovinyl, one benzonitrile), giving rise to a distinct nitrile mode able to be utilised as a vibrational probe. Initial experiments by Fang et al revealed that both arms are in chemically different environments, and that the cyanovinyl arm has a slower

relaxation time due to interactions with the HIV-1 transcriptase backbone and side-chains in the hydrophobic pocket of the active site[42]. Further experiments by the same group investigated the dynamics of HIV-1 reverse transcriptase and two clinically significant double mutants bound to rilpivirine using 2D vibrational-echo spectroscopy combined with molecular dynamics (MD) simulations. Interestingly, the nitrile group on one arm of the drug was discovered to be hydrogen bonded to a water molecule not shown in the previous X-ray crystal structure[43]. This case clearly shows that 2DIR spectroscopy can gain structural information which complements that of the more established technique of X-ray crystallography.

Azide-derived probes have also been incorporated into various protein systems as probes of structure and dynamics in 2DIR experiments. The Hamm group showed that the non-natural amino acid azidohomoalanine (AHA) could be used as a vibrational probe in 2DIR experiments to study binding interactions with the target peptide of the PDZ2 domain[44] (implicated in several signal transduction pathways). Various positions on the peptide were labelled with AHA, and the spectral changes observed (in terms of central frequency, line width and spectral diffusion) indicated that AHA induced only minimal perturbations to the binding pocket structure. Consequently, it was suggested that AHA could be incorporated into larger PDZ2 domain proteins (which have implications for various signal transduction pathways) as a vibrational probe for studies on protein structure and dynamics. The same group demonstrated this by recently performing further 2DIR spectroscopy experiments along with MD simulations on the PDZ2 domain of human tyrosine-phosphatase E labelled with AHA[45], covalently linked to a "photoswitch" which undergoes a light-driven isomerisation inducing a protein transition that mimics native ligand binding[46]. The binding of the probe to the wild-type protein in addition to six structurally important mutants was investigated. Only one of the six mutants showed a significant 2DIR difference signal, reflecting a change of intensity rather than frequency of signal. Additional quantum chemistry calculations suggested this was due to an electronic interaction involving charge transfer of the azido group with sulfonate groups from the photoswitch[45].

In addition to vibrational probes, isotopically labelled amino acid residues have been employed to probe the structure and dynamics of protein-inhibitor systems via 2DIR spectroscopy. This results in a frequency shift characteristic of local structure and environment which enables independent assignment (much like the chemical shift observed in deuterated NMR experiments)[47]. Ghosh et al examined the binding of drugs

rimantadine and 7,7-spiran amine to a double mutant of the M2 proton channel of the influenza A virus by isotopically labelling the carbonyl groups of Ala30 and Gly34 residues in the transmembrane domain[39]. 2DIR spectra revealed a significant increase in the mobility of the channel water attributed to a favourable increase in entropy, facilitated by drug-protein binding. The experiments also verified the binding mechanism of such drugs to the M2 channel, predicted by MD simulations. These experiments further validate the use of 2DIR as a technique to provide complementary structural and dynamical information on protein-inhibitor binding.

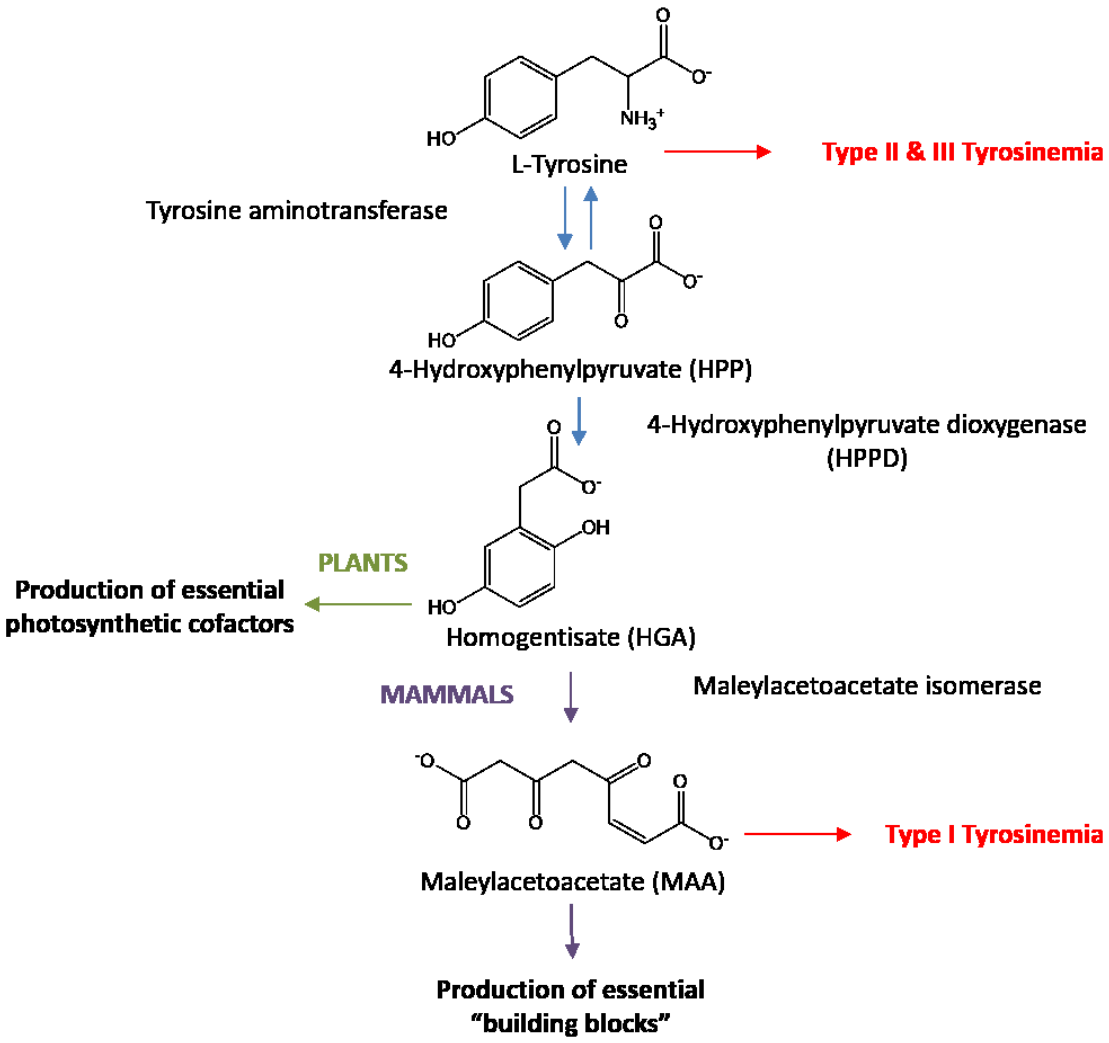
There are currently no fluorescent or isotopic label-free 2DIR spectroscopic methods being utilised elsewhere to study protein-inhibitor complexes. To address this, Electron-Vibration-Vibration 2-dimensional Infrared (EVV 2DIR) spectroscopy is applied in this thesis as a label-free tool to investigate plant protein-inhibitor interactions, with an inherent sensitivity to through-space interactions rendering it the capable of detecting herbicide-protein binding.

In brief, EVV 2DIR is a third order, non-linear, four-wave mixing, frequency-domain method that utilises a non-collinear geometry (as shown in Figure 1b) comprised of three picosecond pulsed laser beams (two mid-IR and one near-visible 790nm). When the frequencies of the mid-IR beams are resonant with coupled vibrations in the sample, a fourth beam is generated, the light intensity of which is measured and plotted against the mid-IR beam frequencies to produce a frequency-frequency 2DIR spectrum. In comparison with the other 2DIR techniques outlined earlier in this introduction, EVV 2DIR has relatively long pulses on the picosecond scale, as these are required to generate the narrow frequency profiles and long coherence dephasing times of the processes studied via this method[48]. For the application of protein-inhibitor (and peptide-peptide) binding examined in this thesis, a picosecond frequency-domain experiment such as EVV 2DIR enables detection of intermolecular interactions at a high sensitivity, such as hydrogen-bond formation and specific vibrational couplings occurring upon binding, elucidating structural information on the binding site. It does not give information on time-resolved system dynamics (as would be the case for the time-time or frequency-time domain ultrafast 2DIR methods discussed earlier), however this is not the goal for this project.

1.2 HPPD-herbicide complexes

While EVV 2DIR spectroscopy has previously been utilised to detect and characterise mammalian protein-drug binding interactions[49], this thesis presents the first example of such an investigation into plant protein-herbicide complexes. The system studied here is the enzyme 4-Hydroxyphenyl pyruvate dioxygenase (HPPD), a key herbicide target in the agrichemical industry[50], complexed with various corresponding herbicides. The primary aim of this research was to detect (and if possible assign) HPPD-herbicide interactions occurring upon complex formation, in order to validate EVV 2DIR spectroscopy as a potential method to study plant-herbicide binding.

HPPD is a non-heme, Fe(II)-dependent enzyme[51] that catalyses the first committed step in the tyrosine breakdown pathway: the conversion of 4-Hydroxyphenylpyruvate (HPP) to Homogentisate (HGA)[52], as shown in Scheme 1 below.



Scheme 1: Part of the tyrosine catabolic pathway, showing the first committed step catalysed by HPPD. Blue arrows indicate steps taken by all living organisms, while green and purple arrows show the utilisation of the pathway for plants and mammals, respectively. The red arrows indicate metabolic abnormalities leading to various forms of Tyrosinemia, a disease in mammals resulting from unregulated blood tyrosine levels.

Although HPPD is an agrichemical target, it is common to all aerobic life-forms, which vary in the way they catabolise tyrosine to generate anabolic processes[51]. Because HPPD is an enzyme common to both plants and animals, HPPD herbicides rely on a difference in metabolism between the two in order to prevent harm being inflicted on the latter. In plants, the resultant HGA is utilised in the biosynthesis of essential photosynthetic cofactors, such as plastoquinone and tocopherols[53]. Without plastoquinone, plants are unable to produce carotenoids which act as photo-protectants in the photosynthetic

process. Upon light exposure, plants lacking in carotenoids generate damaging singlet oxygen, leading to liberated chlorophyll which ultimately destroys leaf pigments. Consequently, inhibition of HPPD in plants by herbicides results in plant bleaching and eventually death.

In contrast, the breakdown of L-tyrosine catalysed by HPPD is required in animals for the regulation of blood tyrosine levels. An excess of L-tyrosine (due to a deficiency of the enzyme Tyrosine aminotransferase) causes Type I & II Tyrosinemia. Conversely, an excess of Maleylacetoacetate or its degraded product fumarylacetoacetate (not shown) caused by a deficiency in the enzyme fumarylacetoacetase leads to the deadly disease Type I Tyrosinemia. Many HPPD inhibitors can therefore function non-selectively between plant and mammalian HPPD species, as both herbicides and therapeutics[51]. Interestingly, the HPPD inhibitor Nitisinone (experiments involving which are discussed in Chapter 4 of this thesis) was originally developed as a herbicide candidate, but is now an orphan drug used to treat hereditary type I Tyrosinemia.

The first X-ray crystal structure of *Arabidopsis thaliana* (At) HPPD was published in 2004[54]. Like human HPPD, the structure is dimeric and made up of a C-terminus with conserved residues which contains the active site, and a flexible N-terminus. In the active form, the non-heme iron atom is coordinated to three highly conserved histidine residues in the C-terminus of the active site. The C-terminal loop (deemed H11) acts to sequester the iron from the cytoplasm and is proposed to function as a gating mechanism to the active site.

Subsequently in 2004, the X-ray crystal structures of rat and At HPPD both with and without the non-selective HPPD inhibitor benzoylpyrazole were also published[55]. There are a number of key similarities between the structures: both are again dimeric and show the non-heme iron in the base of the active site pocket coordinated by benzoylpyrazole, which displaces two water molecules that are otherwise bound to the iron in the absence of inhibitor. Upon inhibitor binding, 50 favourable π-π stacking interactions occur between benzoylpyrazole and two phenyl rings in the active site. In the same paper, structures of both enzymes co-crystallised with the highly plant-selective inhibitor candidate DAS645 were obtained. Comparing the two, one of the conserved phenylalanine residues is further away from the iron atom in the plant complex, indicating that steric factors may influence inhibitor binding affinity. Additionally, the C-terminal helix H7 is in a different position in both structures, further suggesting the importance of the C-terminus to inhibitor coordination.

Interestingly, although there is high sequence homology between the C-terminus of mammalian and plant HPPDs, the primary structure of the latter contains a 15-residue segment that differs from the former, close to the otherwise conserved C-terminus[54,56,57]. Multiple different conformations of this segment have been observed in both the presence and absence of HPPD inhibitors[58]. Additionally, removal of the segment does not prevent inhibitor binding but results in a decreased reaction rate[56]. In combination, these factors suggest that this sequence heterology may play a role in varying binding affinities between species[59], which provides motivation for this work. EVV 2DIR spectroscopy may be able to provide complimentary information in a comparison of inhibitor binding to mammalian and plant HPPD species leading to differing tyrosine catabolism.

In addition, because HPPD is a key herbicide target in both the agrichemical and pharmaceutical industries, there is a constant search for new HPPD inhibitors. While molecular docking models may aid in the prediction of HPPD-inhibitor binding, EVV 2DIR spectroscopy may provide a rapid, efficient method for testing HPPD-inhibitor docking models, and could deliver complementary information on HPPD-inhibitor systems that cannot be crystallised (or have poorly resolved regions) for X-ray crystallography or 2D NMR structures.

1.3 Amyloid beta

Another biomolecule that has been the focus of extensive research is the self-aggregating peptide amyloid beta (Aβ). Aβ(1-40) is the most abundant form of the amyloid beta family of peptides produced in the brain via proteolytic cleavage of the amyloid precursor protein via beta and gamma secretases[60]. The aggregation and accumulation of this 40-residue sequence leads to the formation of distinct Aβ plaques in the extracellular spaces of the brain[61,62] which along with the presence of tau-protein neurofibrillary tangles is one of the two key molecular hallmarks of Alzheimer's disease[63,64]. The enrichment of these plaques with metals such as copper, zinc and iron ions implies that they may influence the formation of these structures[65]. However, despite intensive research efforts, the precise mechanism underpinning the involvement of metal ions in the Aβ cascade remains elusive[66].

Of particular interest is the modification of amyloid beta fibrils upon the binding of copper and zinc, which are common neurotransmitters, released into the synaptic cleft during synaptic firing (from the post-synaptic and presynaptic cleft, respectively)[67,68]. Recent

studies show that the addition of copper to amyloid beta disrupts the formation of amyloid beta fibril structures[69], instead encouraging the formation of neurotoxic β-sheet rich oligomers[70,71]. In contrast, the coordination of zinc by amyloid beta retards fibril formation[72–74].

2DIR spectroscopy has been employed to study the structure of amyloid beta fibrils over the past decade. Kim et al isotopically labelled the carbonyl groups of Gly-38, Gly-33, Gly-29, Gly-25 and Ala-21 residues of amyloid beta (1-40) fibrils to investigate the amide-I interactions[75]. The vibrational frequencies obtained for the Gly-38, Gly-33, Gly-29, and Ala-21 labelled residues were indicative of a parallel, in-register β-sheet structures (as seen in Figure 3), while that of Gly-25 (in the loop region of the β-sheet) indicated a different structural distribution for this residue.

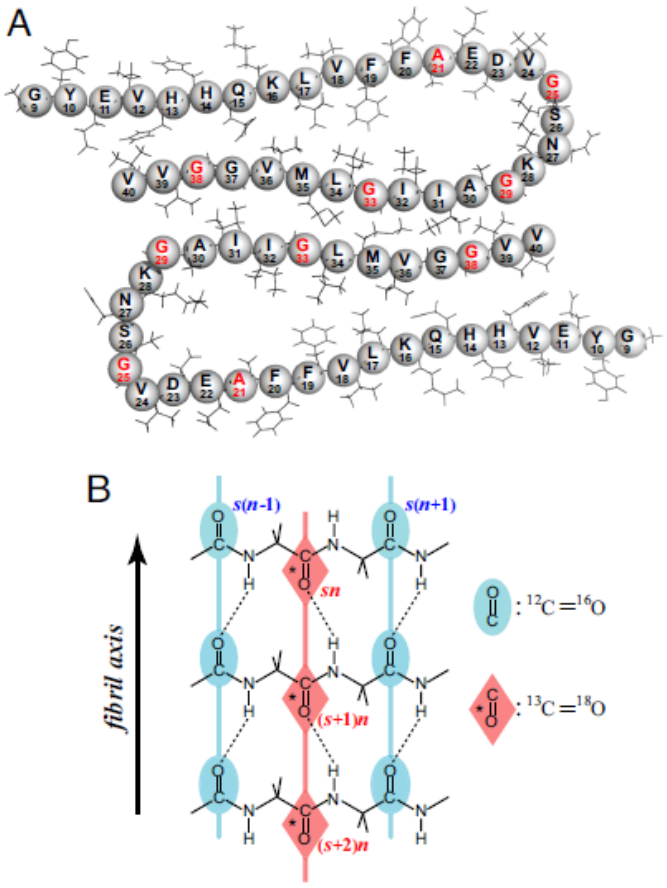


Figure 3: A schematic of parallel amyloid beta fibril formation, showing A) a cross-section of two laterally displaced layers of the amyloid beta fibril (only residues 9-40 are shown), and B) the idealised structure of a parallel β-sheet. Pink and blue coloured carbonyls indicate isotopic labelling. Image taken from reference 75.

More recently, 2DIR spectroscopy has been used to assign the frequency of β-sheets in amyloid beta fibrils[76]. In the amide I region of 1610-1640 cm^{-1} (characteristic of β-sheet vibrations), a distinctive peak at 1610 cm^{-1} was attributed to the formation of rigid amyloid beta fibrils, which could enable differentiation from other β-sheet rich structures such as amyloid beta oligomers[77]. Accompanying simulated 2DIR spectra produced from MATLAB calculations suggested that this downshifted frequency arose from a delocalised mode on the β-sheet backbone[77].

There has, to date, been no 2DIR spectra published for amyloid beta (1-40) complexed with Cu(II) or Zn(II). Additionally, there are, currently no structures of amyloid beta (1-40) complexed with copper available on the Protein Data Bank (PDB) database. This provides clear motivation for the research presented in this thesis, in which EVV 2DIR spectra of amyloid beta (1-40) both alone and with added Cu(II) and Zn(II) are exhibited.

2. Background & theory

2.1 Introduction to infrared spectroscopy

Infrared spectroscopy examines how infrared light interacts with a molecule. Molecules are not motionless, but constantly undergo vibrations which can be defined as the displacement and restoration of their nuclei relative to equilibrium positions (i.e. when the molecule is static)[78]. This nuclear displacement generates a dipole moment (an asymmetrical distribution of electron density[79]) which facilitates the absorption of infrared light, causing a transition from the ground state to a vibrationally excited state. In one-dimensional infrared spectroscopy, these vibrational modes will absorb infrared light at wavelengths specific to the structure and functional groups present in the molecule; therefore it is a useful tool to identify and characterise the structure of molecules[80].

While linear infrared spectroscopy is sufficient to analyse inorganic and small organic molecules, it is less effective at characterising larger biomolecules such as peptides and proteins, the many and complex vibrations of which lead to hugely congested IR spectra. Consequently, many two-dimensional infrared spectroscopic techniques have been developed which simplify the infrared spectra of nonlinear polyatomic molecules by showing only coupled vibrational modes. In EVV 2DIR spectroscopy, this vibrational coupling can occur via two anharmonic processes which act to perturb the ground state vibrations and induce molecular dipole-dipole interactions: namely, mechanical and electrical anharmonicity[81]. In order to understand the anharmonicities that accompany vibrational coupling, we must first briefly discuss the fundamental aspects of straightforward harmonic oscillation, where the motion of each vibrational mode relative to one another is considered as independent.

2.1.1 Harmonic oscillator: Diatomic molecules

To begin with, we consider only simple linear molecules. In motion, the nuclei of a linear molecule containing *n* atoms will undergo displacement from their equilibrium position in *3n-5* possible fundamental vibrations, termed *normal modes*[80]. Thus, for a heteronuclear diatomic molecule (shown in Figure 4), there is just one vibrational degree of freedom: the stretching and compressing of the chemical bond connecting the two atoms.



Figure 4: A simple ball-and-spring diagram of a heteronuclear diatomic molecule that displays a molecular vibration occurring via stretching or compression of the bond attaching the two atoms. This movement can be defined as nuclear displacement (x).

Electrons have a much smaller mass than nuclei, hence upon displacement of the nuclei they react to counteract that displacement[78]. The displacement of the nuclei, (x), from their equilibrium positions $(x = 0)$ results in an increase in absolute potential energy, V (as shown in Figure 5).

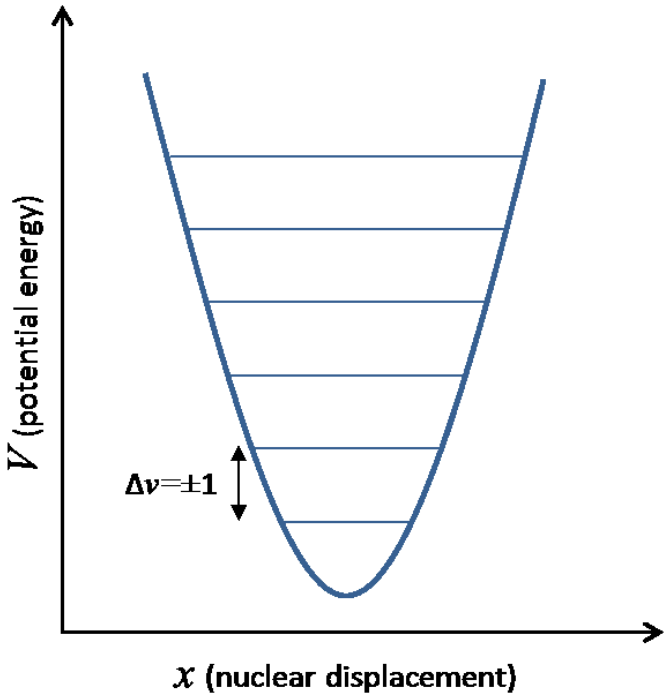


Figure 5: A schematic showing the potential energy (V) vs nuclear displacement (x) curve of a molecule displaying harmonic oscillation. Note that this leads to quantized (evenly spaced) vibrational transitions (v).

In a static diatomic molecule, the *Born-Oppenheimer* approximation allows us to overcome this issue by regarding the location of nuclei as fixed coordinates, then solving the Schrödinger equation for the electrons surrounding the nuclei[79]. When the nuclear displacement x is small, the potential energy of the displacement $V(x)$ can be described by the Taylor series expansion:

$$V(x) = V(0) + \left(\frac{dv}{dx}\right)_0 x + \frac{1}{2}\left(\frac{d^2v}{dx^2}\right)_0 x^2 + \frac{1}{6}\left(\frac{d^3v}{dx^3}\right)_0 x^3 + \cdots$$

Equation 1

where the nuclear displacement is considered negligible ($x = 0$)[78]. When the nuclei are at their equilibrium positions, the first term can be set to $V = 0$ and thought of as a constant (as only the displacement is relevant to the current discussion, and not the absolute potential energy). Given that $x = 0$ in Equation 1, the second term (describing the minimum of the potential energy curve in Figure 5) which depends directly on x will also be zero, and with little nuclear displacement the terms higher than second-order can be ignored. Thus, for small x the potential energy can be rewritten as:

$$V(x) = \frac{1}{2}\left(\frac{d^2v}{dx^2}\right)_0 x^2$$

Equation 2

and is a quadratic function of nuclear displacement (i.e. follows the parabolic curve with evenly spaced energy levels shown in Figure 5). This is known as the *harmonic oscillator* model and is an approximation that sufficiently describes simple systems.

2.1.2 Selection rules in vibrational transitions

For a linear heteronuclear diatomic molecule (as in Figure 4), the harmonic oscillator can also be used to describe the vibrational transitions within a molecule. The gross selection

rule for vibrational spectroscopy states that in order for a diatomic molecule to exhibit a vibrational spectrum, it must experience a dipole moment (μ) that fluctuates in the course of vibrational motion[80]. Essentially, for a harmonic oscillator transition, $\Delta v = \pm 1$ (shown by the evenly spaced energy levels in Figure 5). The rationale for this becomes clear when examining a vibrational transition $v \rightarrow v'$ within an electronic state ε in a heteronuclear diatomic molecule. Using the Born-Oppenheimer approximation, the dipole moment $\mu_{v'v}$ of the molecule in electronic state ε is given by the transition matrix element:

$$\mu_{v'v} = \langle v'|\mu|v\rangle$$

Equation 3

where μ is the dipole operator. Since the dipole moment of a molecule fluctuates with equilibrium bond extension, Equation 3 can be expanded (in much the same way as for the harmonic oscillator expression Equation 1) to describe this variation by the Taylor series:

$$\mu = \mu_0 + \left(\frac{d\mu}{dx}\right)_0 x + \frac{1}{2}\left(\frac{d^2\mu}{dx^2}\right)_0 x^2 + \cdots$$

Equation 4

where μ_0 is the dipole moment for a system at equilibrium (i.e. $x = 0$). Substituting Equation 3 into Equation 4, the transition matrix element for a system with $v \rightarrow v'$ transitions is subsequently given by:

$$\langle v'|\mu|v\rangle = \mu_0\langle v|v'\rangle + \left(\frac{d\mu}{dx}\right)_0 \langle v'|x|v\rangle + \frac{1}{2}\left(\frac{d^2\mu}{dx^2}\right)_0 \langle v'|x^2|v\rangle + \cdots$$

Equation 5

When $v \neq v'$, the dipole moment μ_0 is zero because the states are orthogonal and therefore the first term cancels. For small values of x, again the quadratic and higher terms can be ignored, and Equation 5 can be condensed to:

$$\langle v'|\mu|v\rangle = \left(\frac{d\mu}{dx}\right)_0 \langle v'|x|v\rangle$$

Equation 6

Consequently, homonuclear diatomic molecules such as H_2, N_2 and O_2 are not IR active, because their dipole moment is zero due to symmetrical stretching or compression about the one degree of vibrational freedom, and accordingly do not exhibit a vibrational spectrum.

2.1.3 Anharmonicity & coupled vibrations

In reality, molecules (especially large, complex ones such as proteins and peptides) often undergo significantly larger nuclear displacements which are not necessarily independent of one another, giving rise to the quadratic and higher order anharmonic terms seen in Equation 1. In EVV 2DIR spectroscopy, these anharmonicities can be considered perturbations affecting the harmonic ground states enabling them to mix[82] and can be categorised into two types: mechanical and electrical anharmonicity. The former is an intrinsic property of a molecule's potential energy surface that dominates through-bond processes. In contrast, the latter is the nonlinear variation of molecular electric dipole moment as a function of nuclear displacement and governs through-space interactions[81].

In nonlinear polyatomic molecules comprised of *n* atoms, there are *3n-6* degrees of vibrational freedom over which nuclear displacement can occur. The nuclear displacements in nonlinear polyatomic molecules are best described as the sum of collective displacements of atoms (vibrating at the same frequency and phase but with different amplitudes) along normal mode coordinates, Q[80]. In this way, Equation 4 can be rewritten as:

$$\mu = \mu_0 + \sum_i \left(\frac{\partial \mu}{\partial Q_i}\right)_0 Q_i + \frac{1}{2}\sum_{i,j}\left(\frac{\partial^2 \mu}{\partial Q_i \partial Q_j}\right)_0 Q_i Q_j + \frac{1}{3}\sum_{i,j,k}\left(\frac{\partial^3 \mu}{\partial Q_i \partial Q_j \partial Q_k}\right)_0 Q_i Q_j Q_k \cdots$$

Equation 7

where i, j & k are normal modes.

The presence or absence of terms expressed in Equation 7 determines the appearance of bands in an infrared spectrum due to different kinds of vibrational transitions. When nuclear displacements are small and $i \neq j$, the second-order and higher derivatives are zero and harmonic motion of vibrational modes is orthogonal. Consequently, *fundamental* transitions ($v'_i = v_i \pm 1$, for a single mode i) will occur. In comparison, when $i \neq j$ and the quadratic and higher terms are non-zero values, electrical anharmonicity dominates. It follows that the presence of cross-terms $Q_i Q_j$ renders the modes no longer independent of one another, resulting in a non-zero transition dipole moment μ for interacting vibrational modes. In this case, *combination* bands result in the infrared spectrum, arising from the simultaneous excitation of two or more fundamental modes[80]. Equally, if the higher-order derivatives are non-zero and $i = j$, bands from *overtone* transitions appear, due to the occurrence of multiple quanta transitions ($\Delta v = \pm i$ for x^i, and so on, where $i \geq 2$). Electrical anharmonicity affects the *intensity* of vibrational transitions, because it enables mixing of their wavefunctions such that they can take on properties of each other[78].

Similarly, mechanical anharmonicity allows the presence of combination and overtone transitions in infrared spectra (even in the absence of electrical anharmonicity). Recall that mechanical anharmonicity is the variation of molecular potential energy surface with changing nuclear displacement[81]. It follows that for nonlinear polyatomic molecules, Equation 1 can be rewritten in terms of normal coordinates to give:

$$V = V_0 + \sum_i \left(\frac{\partial V}{\partial Q_i}\right)_0 Q_i + \frac{1}{2}\sum_{i,j}\left(\frac{\partial^2 V}{\partial Q_i \partial Q_j}\right)_0 Q_i Q_j + \frac{1}{3}\sum_{i,j,k}\left(\frac{\partial^3 V}{\partial Q_i \partial Q_j \partial Q_k}\right)_0 Q_i Q_j Q_k \cdots$$

Equation 8

where i, j & k are normal modes. V_0 can be ignored (as for diatomic molecules). The presence of non-zero values for the third order and higher terms indicate the occurrence

of mechanical anharmonicity. In nonlinear polyatomic molecules there are multiple vibrations occurring simultaneously, thus the nuclear displacement of one mode may be affected by the restoration of another, and vice versa, allowing mixing of vibrational modes[78]. The consequence of this mechanical anharmonicity is that, for large values of vibrational quanta v (i.e. higher excited vibrational states), there is progressively less energy in each vibration, leading to increasingly dense energy levels until the molecule eventually dissociates[79]. The effect of anharmonicity on the vibrations of a molecule can be visualised by the asymmetrical curve in Figure 6 (known as the Morse potential).

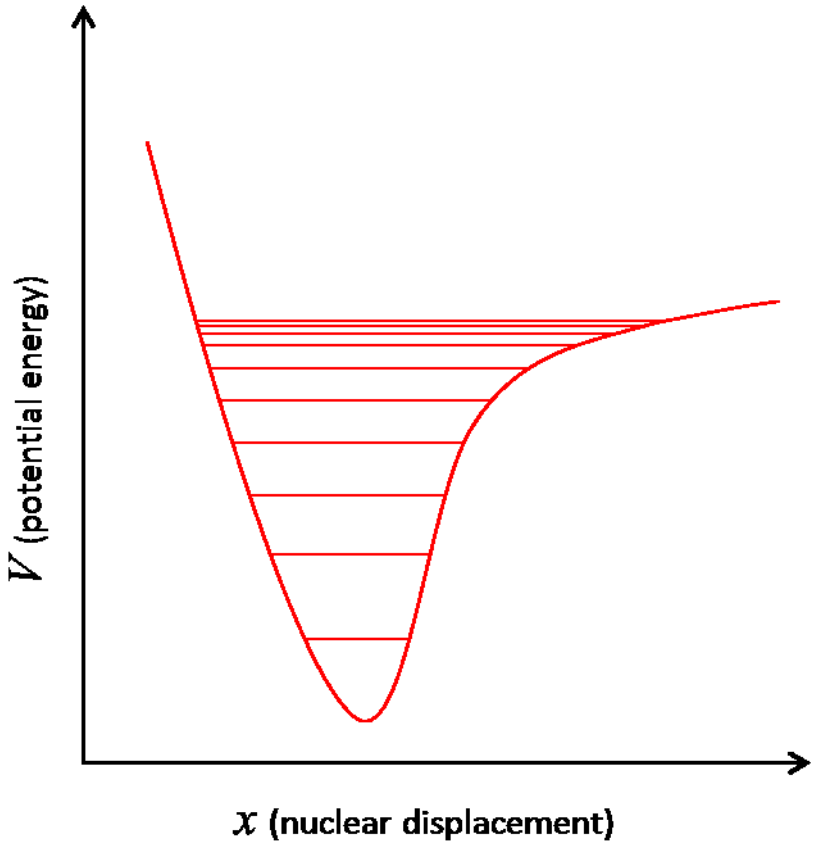


Figure 6: The asymmetrical curve for a molecule exhibiting anharmonic vibrations (known as the Morse potential) shows how the potential energy (v) varies as a function of nuclear displacement (x). Progressively excited vibrational levels become closer in energy until the molecule dissociates at finite x.

Electrical anharmonicity (arising from through-space interactions) can result in intra- or intermolecular vibrational coupling. This is useful when investigating both the coupling of vibrational modes *within* a molecule (for example, amino acid side chains in a protein) and

between molecules (such as between inhibitors and their target proteins). Previously published work indicates that EVV 2DIR spectroscopy is inherently sensitive to electrical anharmonicity[83]. Performing experiments with differing beam polarisation regimes enables separation of the electrical and mechanical anharmonicity contributions to coupling vibrations[84]. As such, these contributions can be used to calculate geometric parameters for molecules in close proximity, which is what forms the basis for this project.

2.2 Introduction to nonlinear spectroscopy

A brief description of nonlinear spectroscopy is provided in this section, with the aim of discussing only the concepts necessary to understand the nonlinear processes related to EVV 2DIR. Both the classical and quantum mechanical models can be used to describe nonlinear spectroscopy, and the aspects of each relevant to EVV 2DIR are conveyed briefly below. A more rigorous account of nonlinear spectroscopy is given by references [85–88].

2.2.1 Classical model

Spectroscopy can be defined as the interaction of electromagnetic radiation with matter[89]. An oscillating incident electromagnetic field generates a fluctuation in the electron density of a material, inducing a change in polarisation. This oscillating polarisation gives rise to a novel electromagnetic field which can radiate at a different frequency and intensity to that of the incident field[82]. This relationship between this induced polarisation and the applied electric field is given by:

$$P = \varepsilon_0 \chi^{(n)} E$$

Equation 9

Equation 9 describes this dependence, where P is the macroscopic polarisation (of the sum dipole moments of the molecules in the material), E is the electric field component of the incident electromagnetic field, $\chi^{(n)}$ is the electric susceptibility (i.e. how readily the

material is polarised in response to an electric field), and ε_0 is the vacuum permittivity constant.

When the intensity of the applied field is low, (as with most linear spectroscopic techniques), there is little distortion introduced in the material hence the relationship between the applied field and the induced polarisation can be considered linear. In this situation, $n = 1$ and consequently the $\chi^{(1)}$ term is regarded as the *linear* electric susceptibility, while the vacuum permittivity constant ε_0 is defined by:

$$\varepsilon_1 = \varepsilon_0(1 + \chi^{(1)})$$

Equation 10

where ε_1 is the dielectric constant[88]. In linear spectroscopy, the low strength of the incident field renders $\chi^{(1)}$ and ε_1 field-independent constants determined by the properties of the medium. However, linearity is an idealisation which often does not apply to most "real-life" experimental systems – thus, Equation 9 is only a good approximation when the strength of an incoming field is weak enough to make deviations from linearity negligible.

When the intensity of an applied electric field E is sufficiently high (such as that of pulsed laser beams), the distortion of charge distribution in the material is significant, and the resulting induced polarisation can be described by the Taylor series expansion (Equation 11):

$$P = \varepsilon_0(\chi^{(1)}E + \chi^{(2)}EE + \chi^{(3)}EEE \cdots)$$

Equation 11

where $\chi^{(n)}$ is the n^{th} order nonlinear susceptibility (and is a function of the incident field vectors E). Equation 11 shows that for the higher order terms, the relationship between the incident electric field E and polarisation P is nonlinear. Nonlinear susceptibilities of a higher order than cubic do exist; however, the scope of this thesis only involves 2nd and 3rd order nonlinear susceptibilities thus these will be discussed accordingly. When these

higher strength electric fields are applied to a medium, the nonlinear susceptibility and dielectric constant become functions of the applied electric field; $\chi^{(n)}(E^{(n)})$ and $\varepsilon(E)$, respectively[87]. Hence, the nonlinear susceptibility is frequency-dependent as it is contingent on the orientations and directions of the incident electric fields. Figure 7 demonstrates the linear and nonlinear responses ($\chi^{(2)}$ and $\chi^{(3)}$) of polarisation P in a material in response to an applied electric field E.

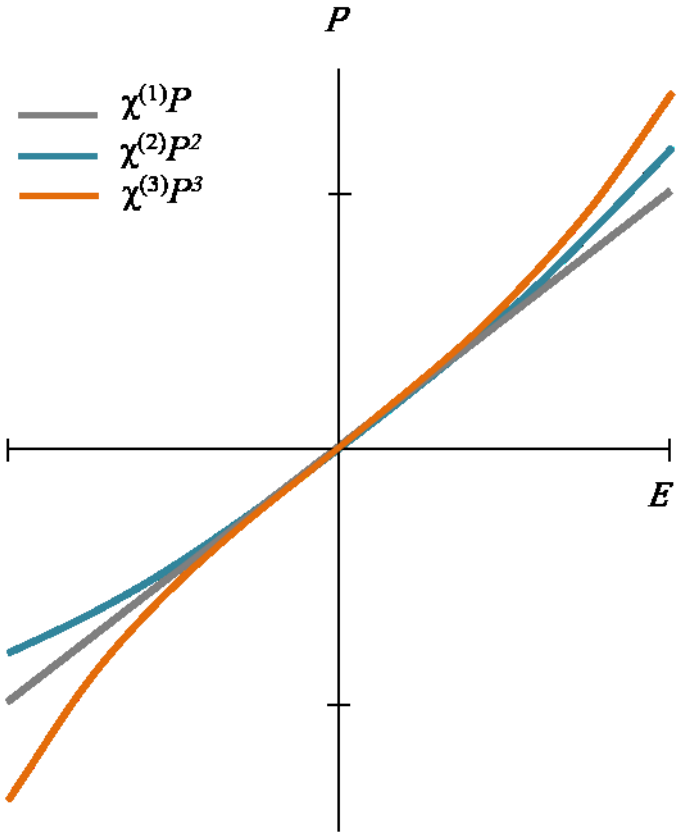


Figure 7: The response of linear ($\chi^{(1)}$) and nonlinear ($\chi^{(2)}$ and $\chi^{(3)}$) polarisations P as a function of the applied electric field E in a system (adapted from reference 88). Note that for quadratic nonlinear processes (blue line), the polarisation response is greater for one sign of the field than the other, (i.e. more polarisable with the electric field applied in one direction over another), hence the requirement for non-centrosymmetric crystals in $\chi^{(2)}$ nonlinear spectroscopy.

Each incident electric field contributing to the n^{th} order nonlinear susceptibility is a 3-dimensional vector generating 3^n possible combinations along the x, y & z axes, therefore $\chi^{(n)}$ is a tensor with 3^{n+1} components[88]. Many of these may be zero values (or identical by

symmetry) for certain materials and given tensor components can be selected for by adjusting the polarisation, direction and timing of the incident electric fields.

Quadratic nonlinear processes (such as second harmonic generation, sum and difference frequency mixing, and optical parametric amplification) can only occur in a non-centrosymmetric medium (i.e. one lacking inversion symmetry). The reason for this becomes clear when looking at the polarisation response for an applied $\chi^{(2)}$ field in Figure 7 – changing the sign of the electric field results in an uneven polarisation response, which suggests that the medium is more easily polarised when the electric field is applied in one direction rather than the other[90]. Consider once again Equation 9. When n is an odd number, the sign of P changes with the sign of E^n hence third-order nonlinear processes can occur in a centrosymmetric material. However for even numbers of n (i.e. $\chi^{(2)}$ and $\chi^{(4)}$ nonlinear processes), E^n remains unchanged whether E^n is negative or positive indicating $\chi^{(2)}$ and $\chi^{(4)}$ must be non-zero under normal inversion symmetry conditions[86]. The practical implications of this is that non-centrosymmetric birefringent crystals are required to facilitate the occurrence of even-ordered nonlinear processes. A relevant example of this is the necessity for a non-centrosymmetric β-Barium Borate crystal to carry out the quadratic nonlinear process of optical parametric amplification, discussed later in experimental terms in Section 3.1.3.

Practically, the nonlinear order can be thought of as the number of input fields applied to a sample, with the consequential new field being denoted as "$(n+1)$-wave mixing". Hence, as three laser beams are utilised in EVV 2DIR the term four-wave mixing (FWM) is employed to describe this nonlinear spectroscopic technique. FWM in EVV 2DIR experiments is characterised by a 3rd order nonlinear susceptibility $\chi^{(2)}$ while 2nd order nonlinear processes (such as optical parametric amplification, described later in Section 3.1.3) are characterised by an $\chi^{(2)}$nonlinear susceptibility.

2.2.2 Quantum mechanical model

Thus far, nonlinear spectroscopy has been described using a classical viewpoint. For the purposes of aiding in the understanding of EVV 2DIR, it is also useful to consider nonlinear spectroscopy from a quantum mechanical perspective. Again, the aim here is not to give a comprehensive discussion of quantum mechanics (as in references [78,79]), but to provide a

general description of the aspects relevant to EVV 2DIR processes (primarily, quantum mechanical coherences).

The laser-induced polarisation that arises in a material due to a transition dipole can be considered a coherence. A quantum mechanical coherence is defined as a linear superposition of states, generated when an applied electric field perturbs an initial state, mixing it with a second state and inducing a transition between the two states[48]. Over time, the mixing will increase if the electric field is resonant with the resultant transition. Consider a system with two quantum mechanical states, a and b. When a perturbation of the system is introduced by an electric field, the wavefunction that describes the subsequent mixing to produce a coherence of states a and b is given by Equation 12:

$$\Psi(x,t) = c_a(t)\psi_a(x)e^{(i\omega_a t)} + c_b(t)\psi_b(x)e^{(i\omega_b t)}$$

Equation 12

where $c_a(t)$ and $c_b(t)$ are the amplitudes of molecules in states a and b[91]. The proportion of these amplitudes relative to one another determines the coherences produced. The probability of finding the system in state a and b is given by $\langle\Psi|\Psi^*\rangle$. If $c_a(t) = 1$ and $c_b(t) = 0$ (or vice versa) the time-dependent terms in Equation 12 cancel and the states are stationary, with their populations defined as $c_a(t)c_{a^*}(t)$ and $c_b(t)c_{b^*}(t)$. However, if $c_a(t) = c_b(t)$, the maximum mixing of states a and b occurs, generating time-dependent coherences described by $c_a(t)c_{b^*}(t)e^{i(\omega_a-\omega_b)t}$ and $c_b(t)c_{a^*}(t)e^{i(\omega_b-\omega_a)t}$.

Additionally, it can be convenient to describe these coherences as observables of a density matrix, which for a two-level system involving states a and b is given by:

$$\rho \equiv \begin{bmatrix} aa & ab \\ ba & bb \end{bmatrix}$$

Equation 13

where the diagonal elements aa and bb denote the stationary populations, while the off-diagonal cross terms ab and ba represent the time-dependent coherences[92]. This is a

helpful notation to consider in the discussion of coherences relevant to EVV 2DIR and their coherence pathways in Section 2.3.2.1 below, where coherences are labelled as a combination of two populations, with each half of the wave-function in either g (ground state), a (first vibrational transition), b (second vibrational transition), e or e' (real or virtual excited electronic state).

Regardless of the notation used to describe them, the aforementioned time-dependent coherences resulting from electric field perturbation of a system will oscillate at the frequency difference of the two states, $\omega_b - \omega_a$. The initial coherence introduced by the first applied field can interact with a second electric field to create a second coherence, as a and/or b transition to other states. In third order nonlinear spectroscopy, a third field further mixes with this second coherence generating a third coherence. The nonlinear polarisation associated with the transition dipole of this final coherence radiates a fourth electromagnetic field, measured as the output signal.

2.3 Electron-Vibration-Vibration Two-Dimensional Spectroscopy

2.3.1 Introduction to Electron-Vibration-Vibration Two-Dimensional Spectroscopy

EVV 2DIR is a third order nonlinear spectroscopic technique utilising three high-intensity electric fields produced by picosecond pulsed laser beams, which undergo FWM to generate a third order nonlinear polarisation that radiates as a new field (as seen in Figure 8). Two of these are tuneable mid-IR beams (the properties of which are herein denoted by α and β, where β is the higher frequency of the two). The electric fields of the mid-IR beams E_α and E_β interact with the sample, and when their frequencies are resonant with coupled vibrational modes in the sample, a series of resonant vibrational coherences are generated. The third laser is an electron-polarising 790 nm beam (henceforth denoted by γ) comprised of the electric field component E_γ, which subsequently interacts with the mid-IR beams in the sample to induce a final non-resonant electronic coherence. In this way, the 790 nm beam can be thought of as a probe that "reads out" the coupled vibrations. Because the final coherence occurs via an electronic transition, it is emitted as a new field E_δ in the visible region, which is advantageous as the detection of photons is

more efficient in the visible region than the infrared[93] (as there is lower energy per photon in the IR region)[94].

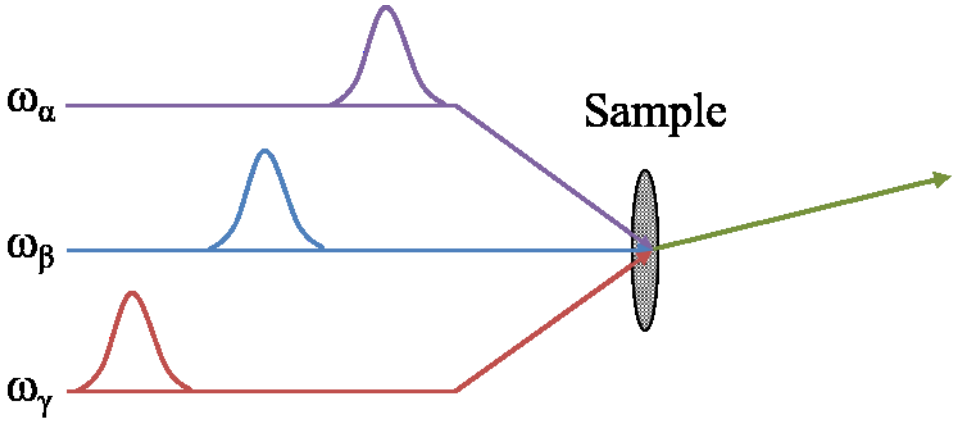


Figure 8: The EVV 2DIR FWM process that occurs when two mid-IR beams (ω_α and ω_β, in order of increasing frequency) and a 790 nm beam (ω_γ) are focussed into a sample. When the frequencies of the mid-IR beams are resonant with coupled vibrations in the sample, FWM occurs and a fourth signal is produced (ω_δ) which oscillates at the frequency defined by Equation 14.

Practically, an EVV 2DIR experiment is implemented by systematically scanning the frequency of the two mid-IR beams over the vibrational spectrum of a sample and measuring the intensity of the resultant FWM signal. This signal is only generated (and can be selectively enhanced) if the frequencies of the mid-IR beams are in resonance with coupled vibrational modes in the sample. The intensity of the detected FWM signal can then be plotted (z axis) against the frequencies of the two mid-IR beams (on the x and y axes) to produce an EVV 2DIR spectrum (such as that in Figure 16), where the cross-peaks correspond to the coupled vibrational modes probed at those frequencies. All three laser beams have a relatively narrow bandwidth of 25 cm^{-1} (full-width half-maximum) and a duration of 1.5 picoseconds. This timescale is relatively long compared with other ultrafast laser experiments, and is required to achieve the narrow frequency profiles and coherence dephasing times experienced in EVV 2DIR experiments.

A recurring issue with one-dimensional vibrational spectroscopies is substantial spectral congestion, leading to the development of a variety of two-dimensional techniques. EVV 2DIR spectroscopy overcomes this limitation, spreading the spectral information over two frequency domains, facilitating significant spectral decongestion. Further decongestion is

achieved by the fact that only coupled vibrational modes appear as cross-peaks on EVV 2DIR spectra, which enables structural information to be extracted about the system being probed. Additionally, input frequencies can be selectively tuned towards vibrational modes of interest such that resonance occurs, and the resulting signal is enhanced. Because this FWM signal is only generated when the mid-IR beams are resonant, the signal measured is background free, and is zero when the frequencies of the beams are not in resonance. EVV 2DIR is a particularly suitable technique for large biomolecules such as proteins and peptides, the spectra of which would be very difficult to gain meaningful structural information from using one-dimensional spectroscopic methods.

An example of spectral decongestion achieved using EVV 2DIR methods is shown in Figure 9, where the FTIR spectra of L-arginine mono-hydrochloride is shown (a) in contrast to the corresponding EVV 2DIR spectrum (b)[95]. The arrows on the FTIR spectra indicate the vibrational modes which couple to produce the cross-peaks seen in the EVV 2DIR spectrum. In particular, the feature at 1555 cm^{-1} (assigned to guanidyl group stretch motions[95]) is clearly discernible as a cross-peak coupled with a mode at 4460 cm^{-1} in the EVV 2DIR spectrum, while in the corresponding FTIR spectrum it is masked by a broad feature and is barely perceptible. Additionally, the FTIR spectrum does not give the information about the modes that the 1555 cm^{-1} couple to.

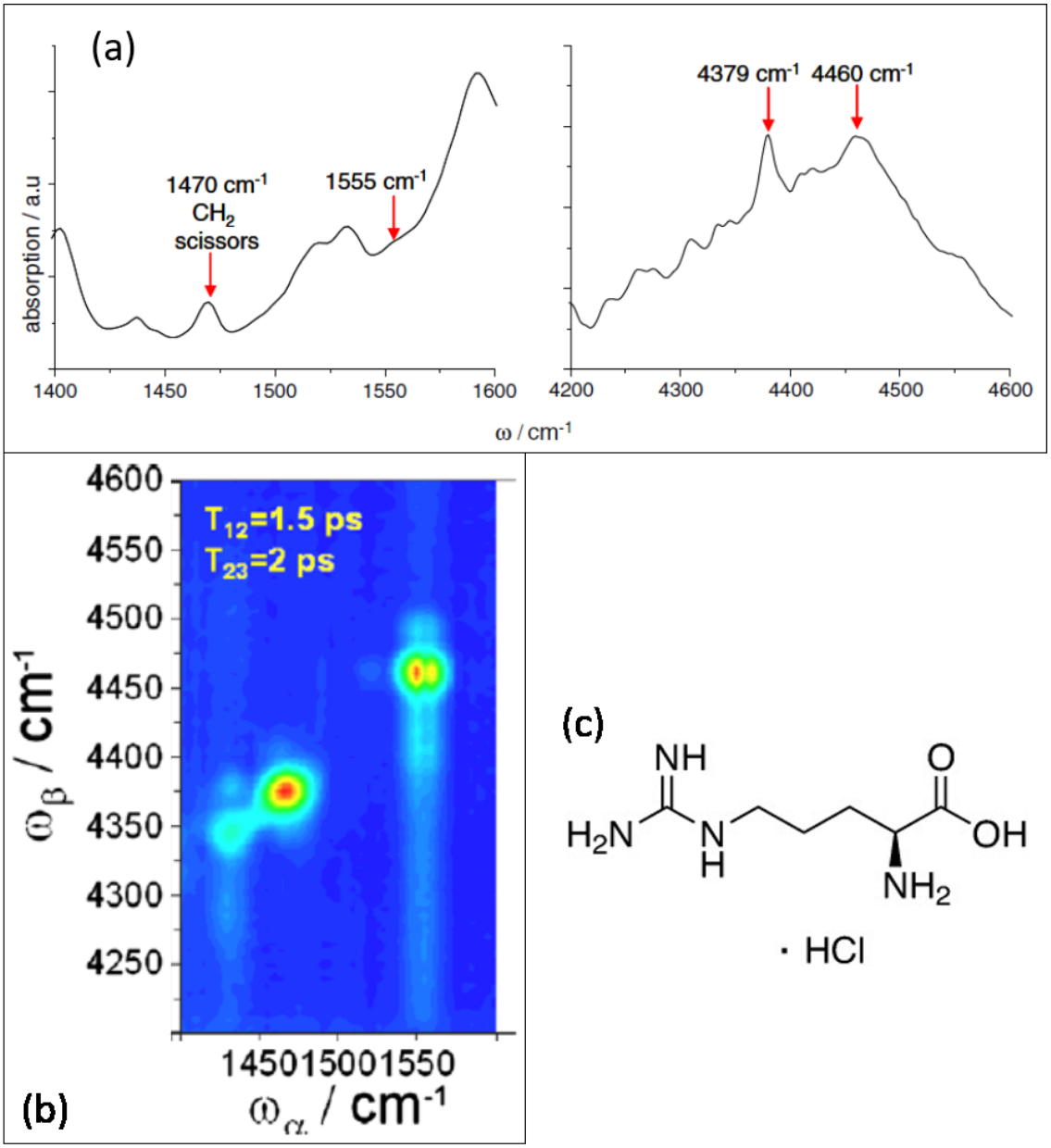


Figure 9: (a) FTIR spectra, (b) EVV 2DIR spectra, and (c) chemical structure of L-arginine mono-hydrochloride. Spectra (a) & (b) are from Reference 96.

This work clearly demonstrates the capabilities EVV 2DIR has for decongesting regions of IR spectra of biological molecules, the effect of which would be further pronounced for a larger, more complex biological system such as a protein-inhibitor complex.

2.3.2 Principles of Electron-Vibration-Vibration Two-Dimensional Spectroscopy

2.3.2.1 Diagrammatic representations: Coherence pathways, WMEL diagrams & Feynman diagrams

The FWM process in EVV 2DIR occurs via a coherence pathway, generated when the electric fields E_α, E_β and E_γ of three picosecond pulsed laser beams interact with the molecules in a sample. The resultant FWM signal is detected at the frequency ω_δ defined by Equation 14:

$$\omega_\delta = \omega_\beta - \omega_\alpha + \omega_\gamma$$

Equation 14

The 3rd order nonlinear polarisation which gives rise to the signal is now written as:

$$P = \chi^{(3)} E_\alpha E_\beta E_\gamma$$

Equation 15

A coherence pathway describes the evolution of coherences over time, with application of each pulse inducing a new coherence dependent on the previous one[96]. There are twelve possible coherence pathways that can be generated by the mixing of the three fields to produce the resonant coherences required for EVV 2DIR. However, only three of these satisfy the frequency relationship demonstrated in Equation 14. These three processes are known as EVV-IR_1, EVV-IR_2 and EVV-Raman, and the coherence pathways for each are shown in Table 1 below:

EVV-IR$_1$	$gg \rightarrow ga \rightarrow ba \rightarrow ea \rightarrow aa$
EVV-IR$_2$	$gg \rightarrow bg \rightarrow ba \rightarrow ea \rightarrow aa$
EVV-Raman	$gg \rightarrow bg \rightarrow ag \rightarrow e'g \rightarrow aa$

Table 1: Coherence pathways for each of the three processes that it is possible to undergo in EVV 2DIR spectroscopy. Each letter represents the state one half of the wave-function is in (bra or ket, using Dirac notation), and each pair of letters represents a coherence involving those two levels.

where g represents the ground state, a and b the first and second vibrationally excited states, and e denotes a non-resonant excited electronic state. In the case of the EVV-Raman pathway, e' is a *virtual* electronic state.

These can also be represented diagrammatically by wave-mixing energy level (WMEL) diagrams shown in Figure 10, where each horizontal line represents a state, and the arrows indicate a transition resulting in superposition of both quantum states involved:

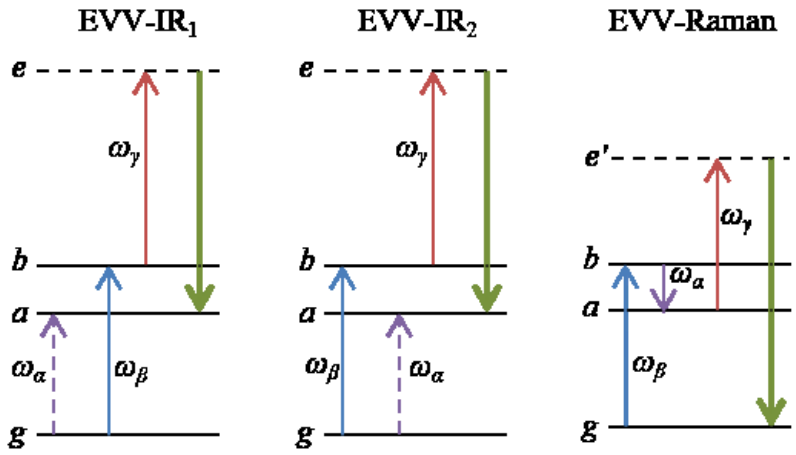


Figure 10: WMEL diagrams for the EVV-IR$_1$, EVV-IR$_2$ and EVV-Raman methods depict the succession of coherences for each process. As time increases (from left to right), the addition of each field E_α, E_β and E_γ induces a superposition of either states g (ground state), a, b (first and second vibrational transitions, respectively), e or e' (real or virtual electronic transition). Each arrow represents the generation or destruction of a coherence – the dotted line indicates a bra transition while the solid lines represent ket transitions (using Dirac notation). The final downwards arrow signifies the generation of the detected FWM field E_δ. The EVV-IR$_1$ pathway is the only EVV 2DIR FWM process employed in this thesis.

Double-sided Feynman diagrams provide another useful way to visualise these FWM processes. In contrast to the WMEL diagrams, they contain no information about the energy levels accessed in each of the coherences produced. *Bra* transitions are on the left of each diagram, while *ket* transitions are on the right-hand side. Inwards arrows represent a transition to a higher energy level (analogous with absorption), while outwards arrows denote the emission of a new field. Figure 11 contains the Feynman diagrams for the three EVV 2DIR pathways, where time increases from bottom to top.

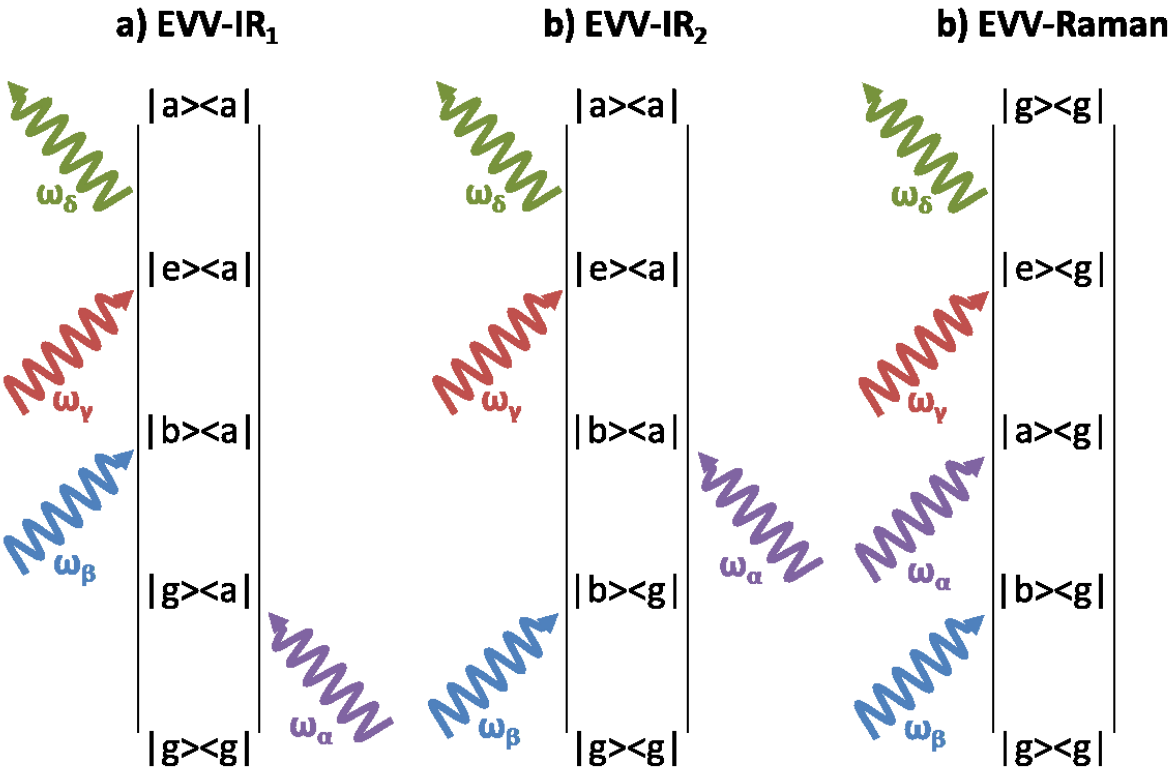


Figure 11: Feynman diagrams depicting the coherence pathways of the three EVV 2DIR FWM regimes. Here, each zig zag arrow represents creation (inwards arrows) or destruction of a coherence (outwards arrows), and time increases from the bottom (ground state) upwards. Bra and ket transitions are separated into the right and left-hand sides of the diagrams, respectively.

The EVV 2DIR experiments performed in this thesis are all via the EVV-IR$_1$ pathway. In this, the first mid-IR field E_α, induces a *bra* coherence in the sample from the ground state g to the first vibrationally excited state a, termed ga. Subsequently, the second mid-IR field E_β then generates a *ket* transition (i.e. interacts with the other side of the wave-

function) from the ground state g to the vibrationally excited state b, producing the coherence ba. Finally, the 790 nm E_γ field interacts with the previous coherence and causes a second *ket* transition from vibrational state b to a non-resonant electronic state e creating the coherence ae, which drives the FWM signal E_δ at the frequency ω_δ outlined by Equation 14. This process is demonstrated by the EVV-IR$_1$ coherence pathway, WMEL diagram and Feynman diagram in Table 1 and Figure 10 and Figure 11, respectively.

The EVV 2DIR pathway that will dictate the FWM signal generated at the sample is determined by the ordering of the three applied pulses E_α, E_β and E_γ. If all three pulses are overlapped in time, all three pathways will occur simultaneously and interfere[97]. As illustrated in Figure 10, both the EVV-IR$_2$ and EVV-Raman methods occur when the E_β mid-IR field is applied before the E_α field, which makes them experimentally difficult to differentiate. Neither of these pathways are utilised in the work presented in this thesis. The EVV-IR$_1$ pathway can be selected for by ensuring that the E_α pulse is applied first followed by the E_β and E_γ pulses, demonstrated in Figure 10. Practically, this is implemented by the use of a delay stage in the beam path of each pulse, which can be altered to manipulate the sequence and relative timings of the three pulses, as shown in Figure 12 below.

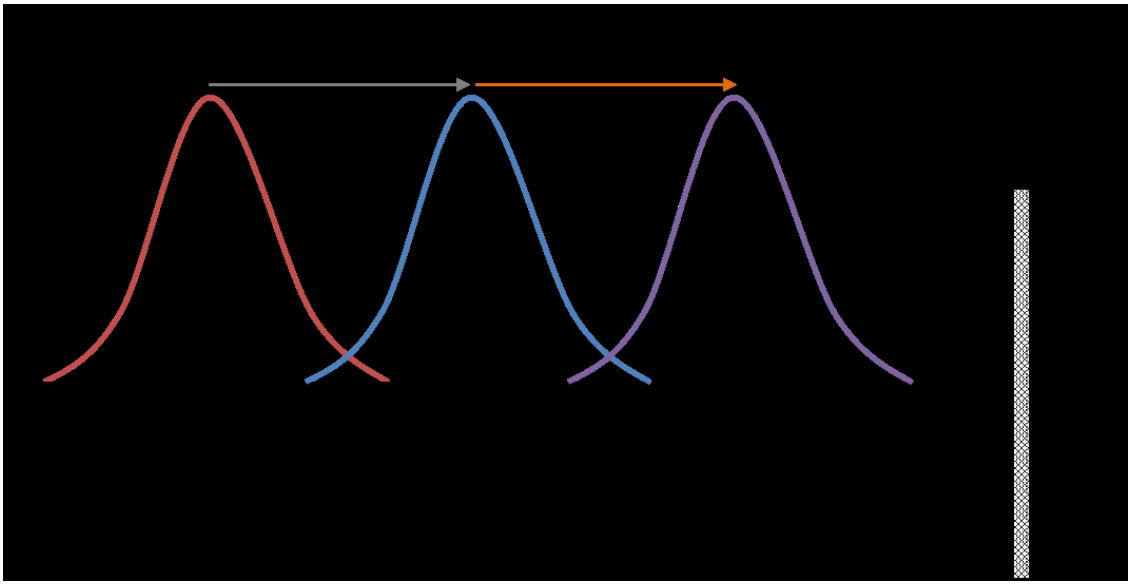


Figure 12: The pulse ordering and time-relative timings in EVV 2DIR spectroscopy, which are experimentally manipulated by individual delay stages to ensure the FWM signal is produced via the EVV-IR$_1$ pathway only (the pulse order and time delays of which are shown above). The delays $T_{\alpha\beta}$ and $T_{\beta\gamma}$ between the pulses are adapted such that unwanted non-resonant signal is minimised.

Prior to EVV 2DIR experiments, the time delays between pulses E_α & E_β (termed $T_{\alpha\beta}$) and E_β & E_γ (denoted $T_{\beta\gamma}$) are optimised to prevent the appearance of non-resonant signal in the resultant EVV 2DIR spectra. If the time delays are not optimal, the occurrence of unwanted nonlinear processes along with FWM can take place, leading to non-resonant spectral artefacts. Typically, time delays of $T_{\alpha\beta}$ = 1.75 ps and $T_{\beta\gamma}$ = 1 ps were used for the EVV 2DIR experiments presented in this thesis.

2.3.2.2 Phase matching

Because it is a variant of *coherent* spectroscopy, the fields contributing to EVV 2DIR experiments must satisfy certain *phase-matching* requirements in order for FWM to occur efficiently in a sample. Phase matching can be thought of as a technique used in nonlinear spectroscopy to ensure that the FWM process satisfies the laws of energy and momentum conservation[98]. In EVV 2DIR, conservation of energy is achieved by the frequency relationship shown in Equation 14 above (and is also emphasised in the WMEL diagram in Figure 10). The conservation of momentum is more complex and has several components which require consideration. In EVV 2DIR experiments, the four fields involved in FWM can be considered as *wave vectors*, each with its own magnitude and direction of wave propagation[99]. Using the sum of these vectors, the conservation of momentum in EVV 2DIR can conveyed by the relationship:

$$\vec{k}_\delta = \vec{k}_\beta - \vec{k}_\alpha + \vec{k}_\gamma$$

Equation 16

where $\vec{k}_i$ is the wave vector for each field. This can be further defined by:

$$\vec{k}_i = \frac{\omega_i n_i}{c} \hat{k}_i$$

Equation 17

in which $\hat{k}_i$ is the unit vector for k_i, c is the speed of light, ω_i is the angular frequency and n_i is the refractive index of the material at that frequency[99]. The refractive index is frequency dependent, meaning it varies with varying input beam frequency, which has important experimental implications to be discussed later in this section. The unit vector $\hat{k}_i$ is determined by the relative angles of the beams, and if these are not optimised a phased mismatch (Δk) can occur, given by:

$$\Delta\vec{k} = \left(\vec{k}_\beta - \vec{k}_\alpha + \vec{k}_\gamma\right) - \vec{k}_\delta$$

Equation 18

Phase mismatch arises due to variation in refractive indices in a sample or material with varying frequencies of input beams[100]. The wavelength dependent refractive index (n_i) can be defined as the ratio of the velocity of light in a vacuum to its velocity in a specified medium. Variations in frequency lead to differences in the speed that a wave vector propagates through a sample. If the spatial phase of the induced polarisation (P) is out of sync with the field that it produces (E_δ) by an amount Δk the length over which FWM can efficiently occur in the sample is limited, and the field components generated along the length of sample interfere destructively, decreasing the resultant signal intensity[82].

The phenomenon of phase mismatch in dispersive processes like EVV 2DIR can be minimised by controlling the relative angles of the beams, as shown in Figure 13. Experimentally, the collinear geometry (Figure 13a) is not easily implemented, as the three input beams have different beam paths and it would be difficult to separate them from one another down the laser table. Alternatively, the non-collinear geometry exhibited in Figure 13c is used in these experiments, which enables all three beams to remain separate until just before the sample. Non-collinear phase matching enables specific nonlinear processes to be selected for[101]. It also produces the FWM signal at a different angle to the input fields, which is advantageous in detection as the output signal can be easily isolated from the input beams.

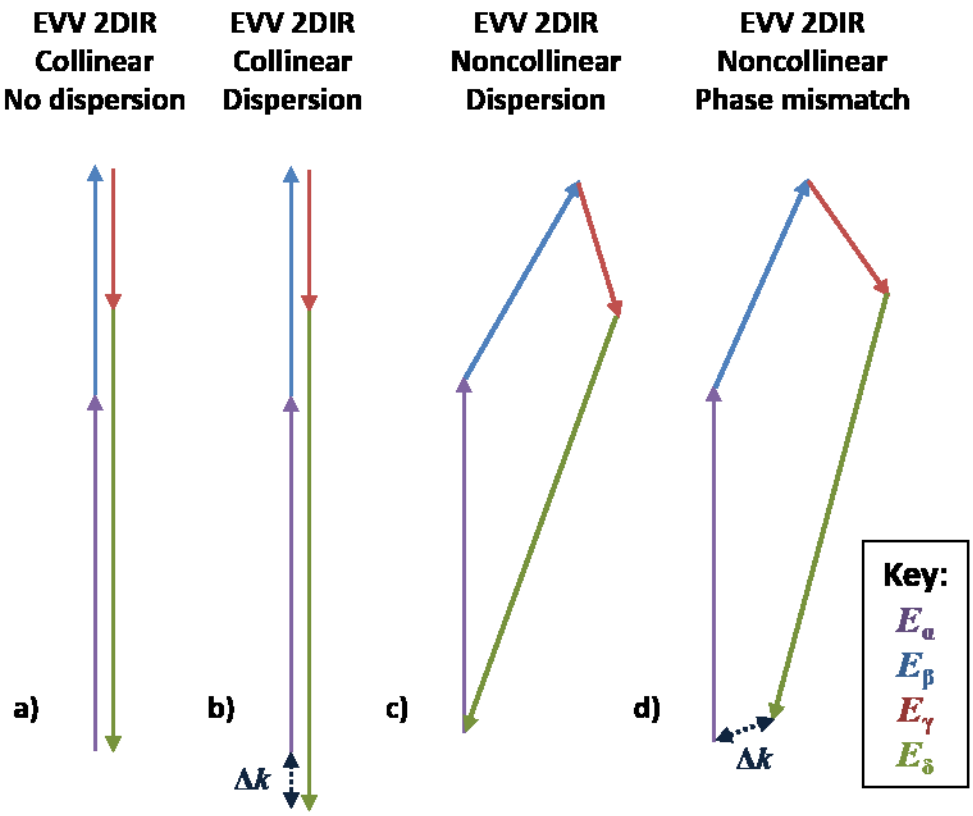


Figure 13: Phase matching vector diagrams for EVV 2DIR spectroscopy showing a) collinear no dispersion, b) collinear dispersion (with phase mismatch Δk), c) noncollinear dispersion and d) noncollinear dispersion (with phase mismatch Δk). Each arrow represents a wave vector for input fields E_α, E_β & E_γ (purple, blue, red, respectively) and the output FWM field E_δ. Vectors are not drawn to scale.

It is especially crucial to phase match the fields in EVV 2DIR, because each of the mid-IR beams cover a frequency range of around ~800 cm^{-1} during an experiment, and a set of angles which give $\Delta k = 0$ for a given group of incident frequencies will yield a slight phase mismatch for another. The most accurate way to account for this experimentally would be to set up the spectrometer such that the phase matching angles are altered stepwise with every frequency step (as with some coherent anti-Raman stokes setups) however this is impractical as it hugely increases experimental complexity. Instead, the phase matching angles are optimised to give $\Delta k = 0$ for the central frequency of each mid-IR beam (usually 1500 cm^{-1} E_α for and 3000 cm^{-1} for E_β), as seen in Figure 14. This ensures that phase mismatch is diminished particularly at these central regions where most of the absorptions of interest occur. To sustain phase matching at non-central frequencies, the input beams are tightly focused, which creates a range of input wave vectors. The

disadvantage to maintaining this tight focus is that the beam overlap of the input fields can be slightly reduced, leading to decreased intensity of the FWM signal produced[82].

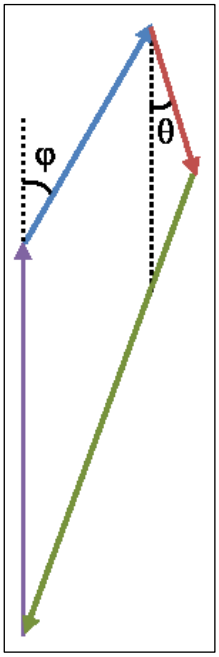


Figure 14: A vector diagram showing the noncollinear phase matching geometry used in the EVV 2DIR experiments presented in this thesis. Each arrow represents a wave vector for input fields E_α, E_β & E_γ and the FWM field E_δ (purple, blue, red & green, respectively. Vectors are not drawn to scale). The smallest possible angles for φ and θ able to satisfy this geometry were calculated to be 2° and 2.42°, respectively.

The EVV 2DIR experiments in this thesis were carried out using a non-collinear geometry, the vector diagram for which is shown in more detail in Figure 14. The geometry for angles φ and θ is optimised to give $\Delta k = 0$ at the set of central frequencies, which could be satisfied by multiple sets of angles. However, as discussed above, broad angles have the effect of decreasing beam overlap and hence intensity of the FWM signal, so the smallest practically implementable values of φ and θ (calculated using trigonometry to be 2° and 2.42° respectively) were used. Calculations are excluded here for simplicity.

2.3.2.3 EVV 2DIR Spectra

Section 2.3.2.1 above discusses the succession of coherences produced during EVV 2DIR, and the transitions a system undergoes to generate these can be categorised as *fundamental*, *combination* or *overtone* transitions. Similar to spin-spin transitions in 2D NMR spectra, fundamental vibrations are one quantum transitions, while combination or overtone transitions are double quanta transitions. A transition from ground state g to vibrational excited state a to create coherence ga is a *fundamental* (or one quantum) transition, while conversely one from ground state g to a second level vibrational excited state b is a two quanta transition. A coherence can also be formed via a transition from ground state gg to a *combination* band comprised of a mode where vibrational states a and b are coupled, labelled $g(a+b)$. In this way, the coherence pathway for EVV-IR_1 can be rewritten in Table 2:

EVV-IR_1 (revised)	$gg \rightarrow ga \rightarrow b(a+b) \rightarrow ea \rightarrow aa$

Table 2: The revised coherence pathway for the EVV-IR_1 process , showing the coherence involving the combination band $a+b$.

where the first and second transitions are that of fundamental and combination transitions, respectively. The WMEL diagram can also be redrawn to include a combination band as seen below in Figure 15:

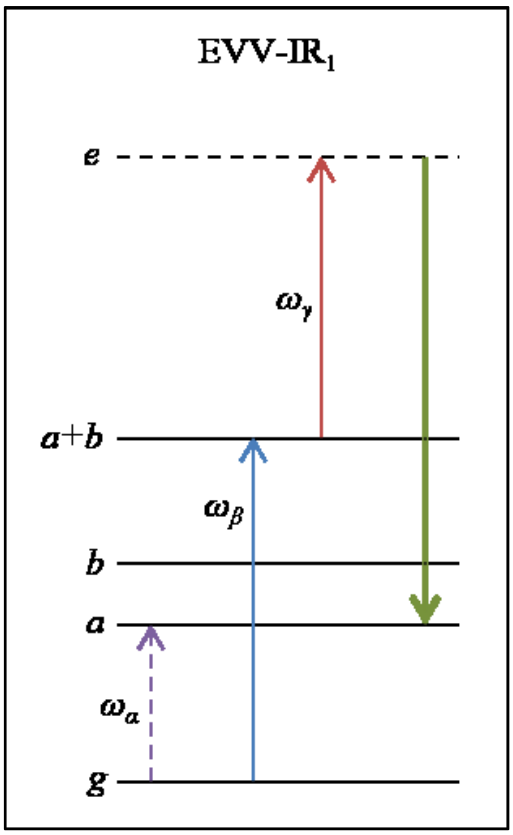


Figure 15: The WMEL diagram for EVV 2DIR shown in further detail, including the two-quanta transition to the combination *a+b* band. This results in the appearance of cross-peaks in the subsequent EVV 2DIR spectra which correspond to coupled vibrational modes between coherences involving the *a+b* energy level.

EVV 2DIR is often deemed the optical analogue of two-dimensional nuclear magnetic resonance (2D NMR) spectroscopy[102]. Like EVV 2DIR, 2D NMR can not only be considered in terms of quantum coherences (probing the coupling of nuclear spins as discussed in the introduction section, as opposed to vibrational mode coupling in EVV 2DIR), but also gives two-dimensional spectra where cross-peaks correspond to coupled modes[15]. As stated in Section 2.3.1, the intensity of the EVV 2DIR signal generated by FWM is plotted on the z-axis against the frequency ranges scanned by the mid-IR beams on the x- and y-axes, to produce a two-dimensional spectrum where cross-peaks correspond to resonant coupled vibrational modes in the sample. An example EVV 2DIR spectrum is exhibited in Figure 16. The cross-peaks shown in the left-hand schematic look Gaussian, however in reality, they are often distorted by anharmonic motions or overlapping with other peaks, leaving them imperfect as seen in the actual spectrum of polystyrene on the right.

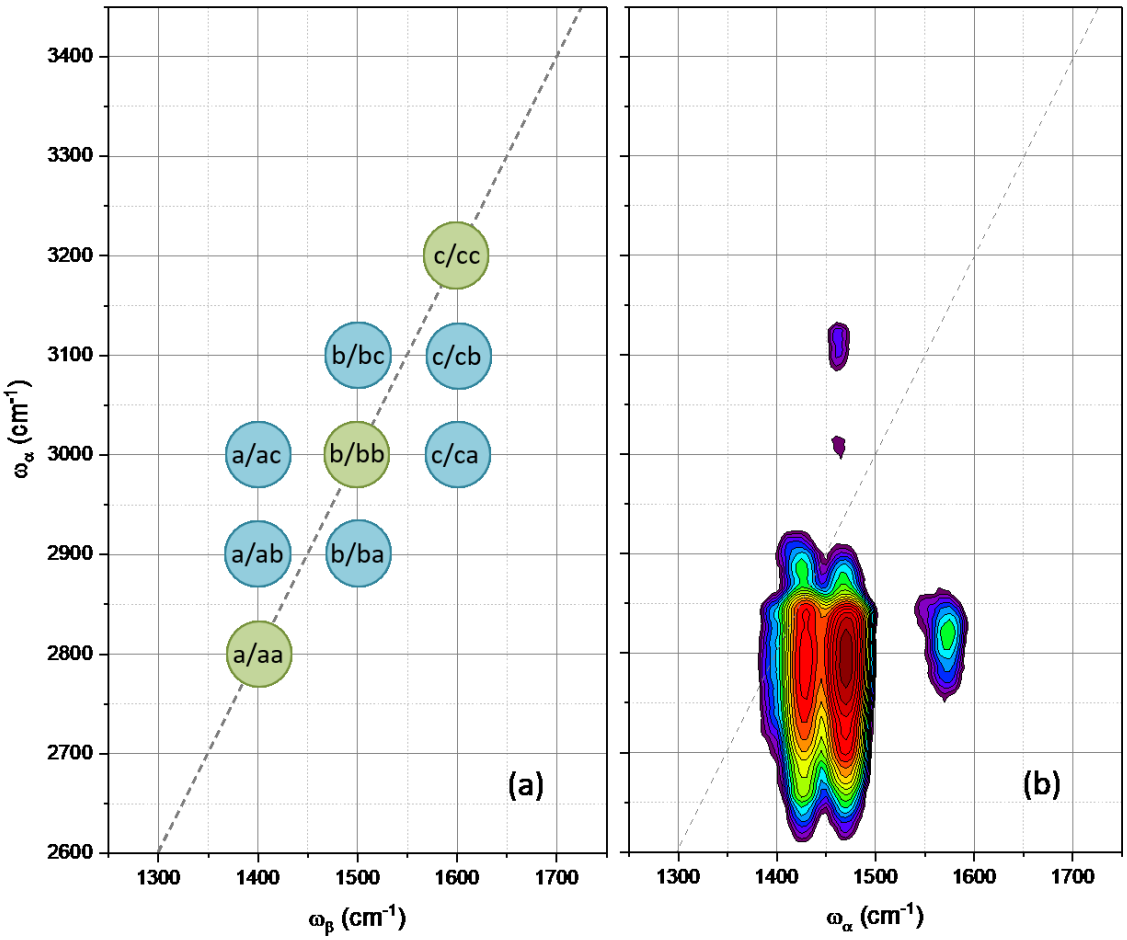


Figure 16: A schematic of an EVV 2DIR spectrum over the typical frequency range, showing a) vibrational coupling between fundamental modes a (1400 cm^{-1}), b (1500 cm^{-1}) & c (1600 cm^{-1}). b) an actual EVV 2DIR spectrum of polystyrene, showing coupled vibrations. The grey dashed line on a 2:1 gradient exhibits cross-peaks arising from overtone couplings. Note that the peaks in the schematic on the left are Gaussian, while the real peaks in the left-hand spectrum show imperfections and shifts arising from overlapping vibrational modes and/or anharmonicities.

As shown by the WMEL diagram in Figure 15 and coherence pathway in Table 2, EVV 2DIR involves the coupling of a pair of vibrational modes: one *fundamental* vibration coupled to a *combination* of both vibrations. The modes associated with combination bands are the strongest peaks seen in EVV 2DIR spectra[91]. Consider again a simple system with two vibrational modes, a and b, being probed by EVV 2DIR experiments. The resultant EVV 2DIR spectra shows cross-peaks that arise from $a/a+b$, and $b/b+a$ vibrational mode coupling (where / denotes the coupling of the two modes it separates). The fundamental

frequency is on the x-axis ($\omega_\beta = \omega_a$ and/or ω_b) and the y-axis gives the frequency for the fundamental coupled to the combination ($\omega_\alpha = \omega_{a+b}$). Hence, vibrations arising from mode *a* fall in columns, while vibrations due to mode *b* fall on a diagonal. Additionally, a vibrational mode can couple to a vibrational mode coupled to itself, called an *overtone* (in this case, *a*/*a*+*a* and *b*/*b*+*b*). The overtone line is a 2:1 gradient of self-coupled fundamental vibrations that presents itself in any EVV 2DIR spectra of this type. Figure 16 shows a schematic of an EVV 2DIR spectrum for a system comprised of vibrational modes *a*, *b* and *c*. It is important to note that the schematic in Figure 16a is idealised and frequency shifts due to anharmonicities or overlapping peaks have not been corrected for.

2.3.3 Applications of EVV 2DIR

EVV 2DIR spectroscopy is a doubly vibrational enhanced four-wave mixing (DOVE-FWM) method that was first established by Wright & colleagues in the late 1990s[103], however a more recent review is given by reference [96]. The Wright group used small, simple molecules in several validation experiments to show firstly that DOVE-FWM could be performed, and secondly to demonstrate its capabilities for spectral decongestion. The initial DOVE-FWM spectrum was of acetonitrile (deuterated by varying amounts) and showed only peaks resulting from vibrational mode coupling between the C=C and C≡N moieties, with the C-H peak (normally very strong in one-dimensional IR spectra) noticeably absent[103] as shown in Figure 17.

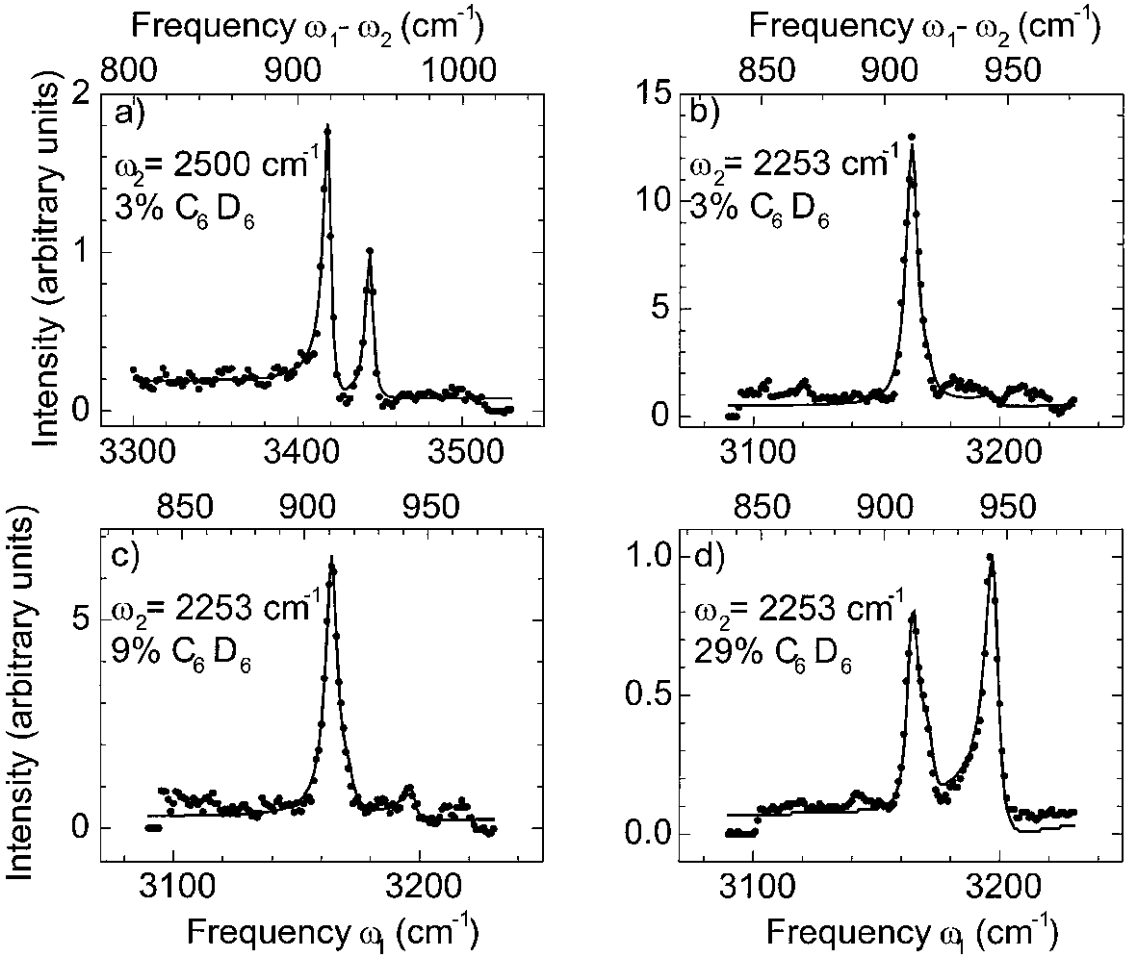


Figure 17: (a-d) Four wave mixing scans of acetonitrile, where ω_1 is denoted by the x-axes and ω_2 and the concentration of C_6D_6 is specified for each spectrum (from Reference 103). Peaks present are due to coupling between C=C and C≡N groups. Note the absence of the C-H stretch at >3200 cm^{-1}, which is normally present in 1D IR spectra and congests this region.

This proved that DOVE-FWM spectroscopy could be used to significantly decongest IR spectral information. Subsequently, additional DOVE-FWM spectra (shown in Figure 18) were recorded of an acetonitrile, deuterated acetonitrile and deuterated benzene (to act as an internal standard) mixture to further prove that cross-peaks only appeared between coupled vibrational modes[97,102]. To complement the spectral data, the corresponding model DOVE-FWM spectra were numerically calculated by *ab initio* methods[97,104]. The non-linear susceptibilities and molar absorption coefficients for each mode were calculated via density functional theory/Becke, three-parameter, Lee-Yang-Parr (DFT/B3LYP), Hartree-Fock (HF) and second order Møller–Plesset (MP2) methods all with the 6-311++G** basis set on GAUSSIAN 98, while experimental values were used for

vibrational dephasing constants. The resulting model spectra agreed well with the experimental spectra, and again showed only coupling between C=C and C≡N modes, with no coupling to C-H modes seen (as shown in Figure 18).

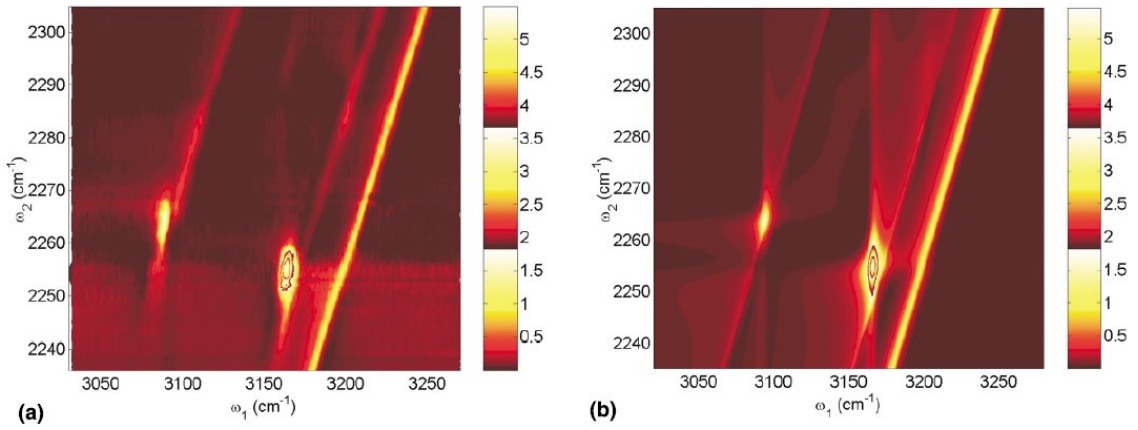


Figure 18: a) An experimental four wave mixing 2DIR spectrum of a mixture of acetonitrile, deuterated acetonitrile and benzene, and b) a model spectrum of the same system (from Reference 97). Cross-peaks again arise from coupling between C=C and C≡N groups, and features arising from C-H are absent.

Following the successes of Wright & colleagues, the Klug research group implemented the technique in 2007 and renamed it Electron-Vibration-Vibration Two-Dimensional Infrared (EVV 2DIR) spectroscopy to differentiate it from other 2DIR methods[82]. It is currently the only research group to practice the EVV 2DIR technique as outlined in this thesis.

2.3.3.1 Complex formation

Electrostatic through-space interactions (e.g. hydrogen-bonds) are crucial to the formation of intermolecular complexes, such as those formed in protein-inhibitor binding. In 2DIR spectroscopy, couplings between the vibrational modes belonging to each molecule are directly related to their proximity and relative orientation (enabling their dipole moments to interact)[105]. In 2009, the Klug research group demonstrated that EVV 2DIR spectroscopy has the capability to detect through-space intermolecular vibrational couplings derived from purely *electrical* anharmonicity, even in the absence of mechanical anharmonicity[84]. Guo et al compared the EVV 2DIR spectra generated from samples of

benzonitrile, phenylacetylene and a benzonitrile-phenylacetylene complex to demonstrate the presence of new cross-peaks in the latter due to purely electrical couplings. EVV 2DIR spectra of the benzonitrile-phenylacetylene mixture was performed using a standard polarisation for all three beams (PPP) and with the polarisation reversed for the third beam (PPS), as well as reversing the pulse ordering such that the EVV-IR2 pathway was selected for (see Figure 10). Moreover, because this coupling of transition dipoles is sensitive to their relative geometries, geometrical parameters such as intermolecular bond distances and angles could be calculated using the ratios of the resulting cross-peaks in each of the three regimes, which agreed relatively well with theoretical values.

This inherent sensitivity of EVV 2DIR to electrical anharmonicity (resulting in through-space vibrational couplings) is what forms the basis of this project. Inhibitors (such as herbicides) are often complexed with their target protein via electrostatic interactions such as hydrogen bonds, which EVV 2DIR has the capability to detect. This was first demonstrated for protein systems by Sowley in 2017, where the binding of the drug SU5402 to its target protein FGFR1 was detected[49] (the structure of SU5402 both alone and complexed with FGFR1 is shown in Figure 19).

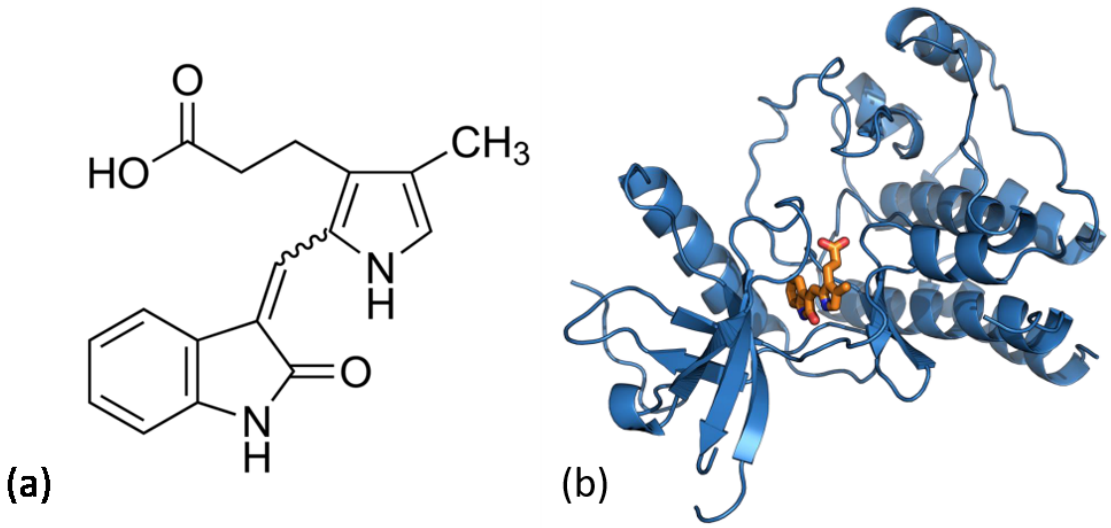


Figure 19: (a) The chemical structure of drug SU5402, and (b) the same drug complexed with its target protein, FGFR1 (a human growth factor receptor). Right image taken from Reference 49.

2.3.3.2 Investigation of biological molecules

EVV 2DIR has been previously utilised to research various biological molecules. Much like the "fingerprint region" that enables identification of certain compounds in routine one-dimensional infrared spectroscopy[80], EVV 2DIR has been applied in the optical fingerprinting of proteins. The relative ratios of the intensities of cross-peaks previously assigned[83] to phenylalanine, tyrosine and methylene were determined to be proportional to the actual ratios of these compounds within various amino acid sequences[106]. Additional experiments demonstrated the ability to distinguish 10 proteins by amino acid content, following the same process but using the extra residues of tryptophan, histidine and cysteine[107].

The quantification of biologically relevant post-translational modifications such as phosphorylation and nitration of tyrosine is another application made possible by EVV 2DIR spectroscopy. Prior work has shown that the post-translational phosphorylation of peptides (known to regulate various enzymatic processes) can be absolutely quantified[92,107]. Furthermore, Rezende Valim et al demonstrated that tyrosine nitration (implicated in various neurodegenerative and cardiovascular illnesses[108]) can be differentiated and characterised between the different states of 3-nitrotyrosine[109].

Sensitivity to DNA structure is yet another exciting biological application of EVV 2DIR spectroscopy that has been previously explored. Sowley showed that EVV 2DIR can be used to study the formation of G-quadruplexes in DNA[49]. G-quadruplexes are secondary structures formed from guanine-rich DNA and are biologically important due to implication in the control of genes associated with cancer[110]. EVV 2DIR spectra of duplexes and G-quadruplexes were compared to controls and included a multitude of cross-peaks that could be preliminarily assigned to Watson-Crick base pairing, glycosidic bond conformation and G-quadruplex formation. Figure 20 shows evidence of the latter, where the EVV 2DIR spectrum of the Myc2345 quadruplex (right) shows new spectral features (due to G-G base pairing on forming the structure), compared to the spectrum of single-stranded DNA sequence SS3 (left). The quadruplex is formed by annealing SS3 in the presence of a high concentration of potassium ions, so it acts as a standard.

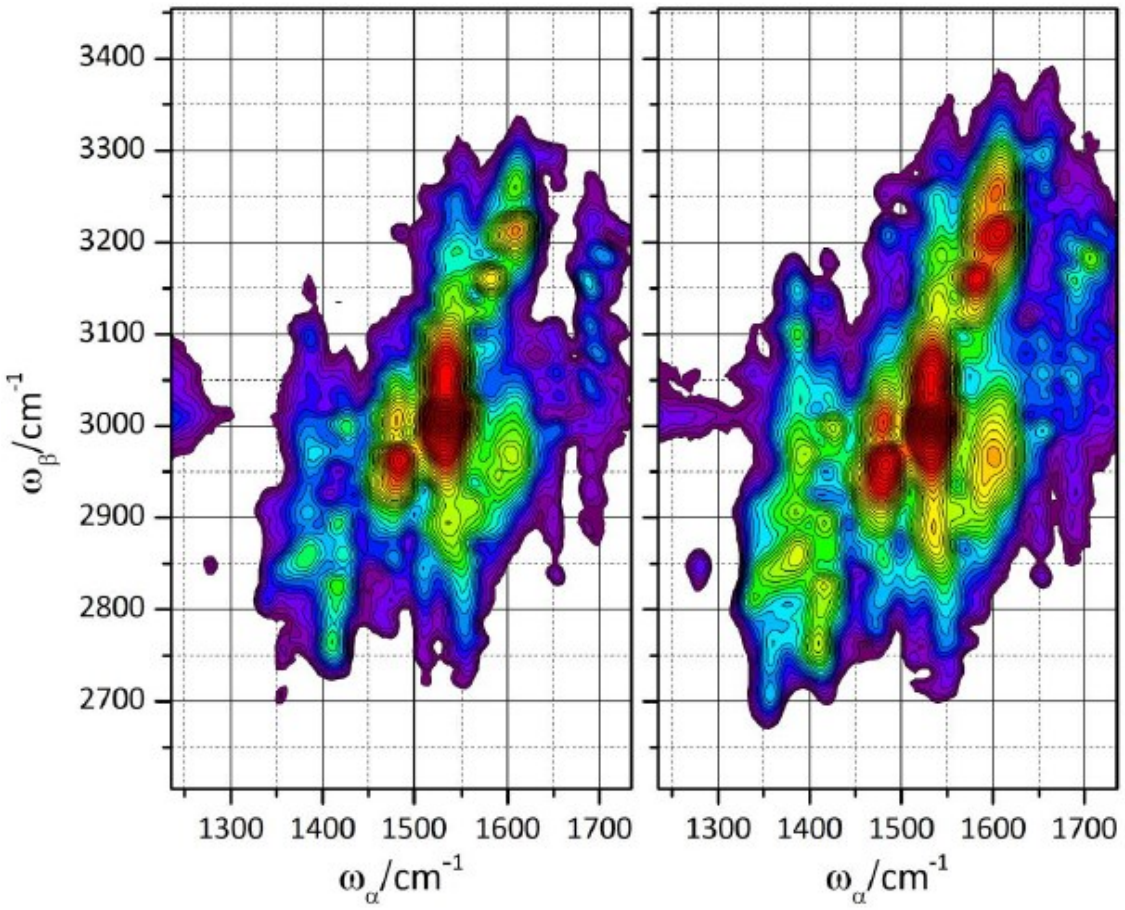


Figure 20: EVV 2DIR spectra of the SS3 DNA sequence (left) and corresponding Myc2345 quadruplex (right), from Reference 49.

EVV 2DIR spectroscopy has been further employed to perform label-free biological imaging with enhanced spectral contrast. Fournier et al utilised this method to investigate the spatial distribution of various chemical species (fixed by the stain haematoxylin) in mouse kidney tissue[111]. Images were obtained for cross-peaks corresponding to haematoxylin and protein CH_3 (as a control), and comparison of these showed distribution of haematoxylin indicative of different tissue types of the kidney cross-section (shown in Figure 21).

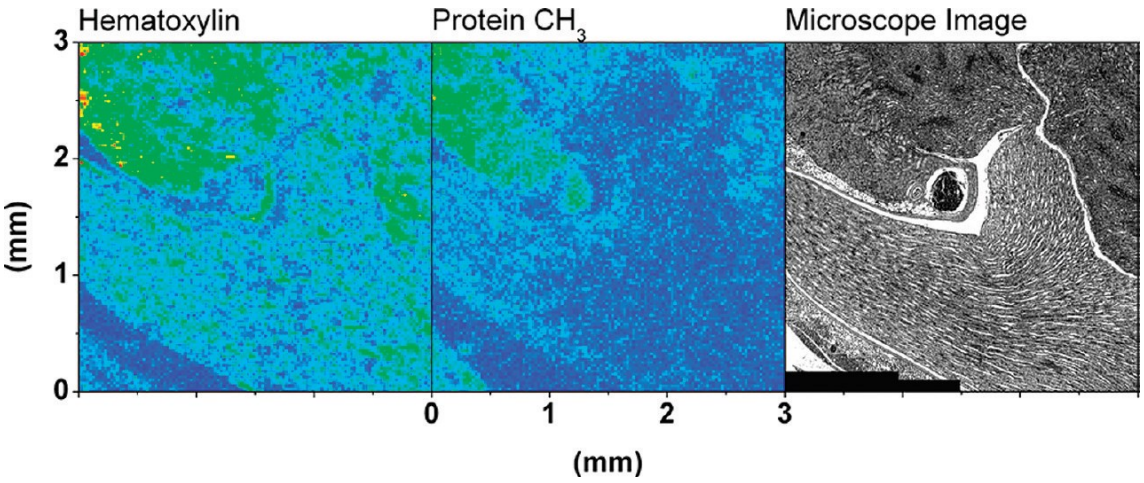


Figure 21: EVV 2DIR images of Hematoxylin (left) and protein CH_3 (middle) distribution in a cross-section of mouse kidney, and the corresponding microscope image (right). EVV 2DIR signal increases from blue to green to red, and indicates levels of hematoxylin and CH_3 corresponding to different tissues present in the cross-section. Image from Reference 111.

2.3.3.3 Resonance enhancement

In addition to the spectral decongestion achieved by spreading spectral information over two dimensions and examining only coupled vibrations, EVV 2DIR has the capability to increase sensitivity even further by *resonance enhancement*. EVV 2DIR comprises two resonant vibrational coherences, followed by a third coherence with a non-resonant electronic state. The final step of EVV 2DIR can be considered as half of a *coherent* non-resonant Raman scattering episode. The electronic characteristics of the sample therefore affect the resultant FWM signal. In comparison to the non-resonant Raman process, resonance Raman scattering occurs when the incident radiation (in this case, the input laser frequency) is close in energy to an electronic transition in the sample (shown by their WMEL diagrams in Figure 22). This provides enough energy to excite electrons in the sample to a higher electronic state, resulting in an increase of 10^2-10^6 scattering intensities[112].

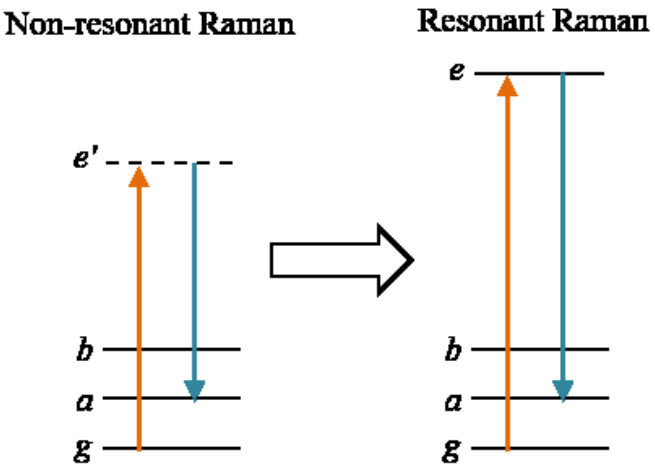


Figure 22: WMEL diagrams for non-resonant vs resonant Raman scattering processes. With non-resonant Raman (as with non-resonantly enhanced EVV 2DIR) a photon transitions to a virtual excited electronic state before scattering, whereas in resonant Raman the excitation source (the 790 nm beam in EVV 2DIR) is close to an electronic transition, so the photon transitions to an actual electronic state before scattering. This is not a binary process, and near-resonant enhancement can occur somewhere in between the two regimes.

Significantly increased signal size has been observed in spectra of highly conjugated molecules due to this phenomenon. Previously, this effect has been exploited to selectively detect the retinal chromophore CH_3 signal in the active site of the photosynthetic enzyme bacteriorhodopsin[113]. Near-resonance was achieved by tuning the 790 nm beam close to an electronic transition in the enzyme, resulting in a CH_3 signal enhancement 10^4 times larger than that for a CH_3 cross-peak in a non-conjugated system. Interestingly, the probe laser beam was tuned to 790nm whilst the UV-Visible (UV-Vis) spectrum of retinal has a maximum absorption at 562nm, which indicates that *on*-resonance does not have to occur – *near*-resonance is enough to see substantial signal increase. As shown in Figure 23, resonance enhancement does not begin at a sharply defined wavelength – if the exciting laser is within even a few hundred wavenumbers of the electronic transition in the sample, *near*-resonance enhancement of the signal can occur of several orders of magnitude.

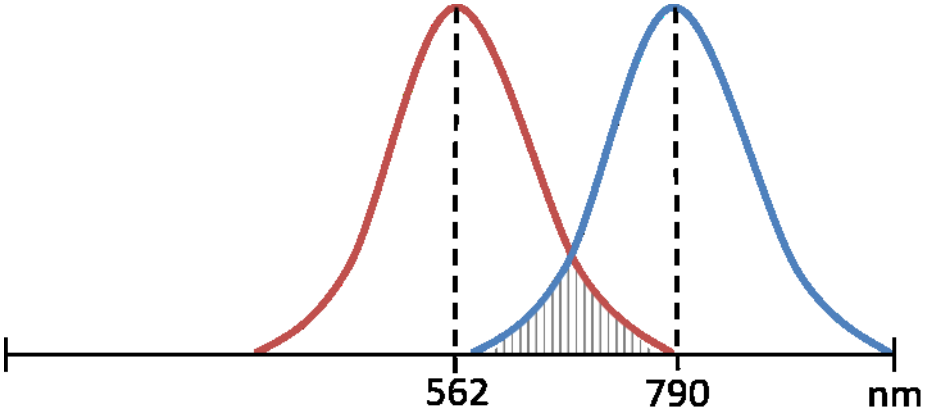


Figure 23: A schematic of the near-resonance enhancement of the retinal chromophore in bacteriorhodopsin that occurs when the applied 790 nm beam is close enough to a 562 nm electronic transition in the sample, increasing the likelihood of photons being excited to an electronic state before scattering, resulting in a significant increase in the CH_3 signal intensity. The closer the applied field is to the electronic transition, the greater the overlap between the two and hence higher probability of resonance enhancement.

Tuning the 790 nm beam closer to the electronic transition in a sample (by, for example, using a frequency doubling crystal) would theoretically result in greater resonance enhancement, and consequently larger signals. However, this has the practical implications of potentially introducing other undesirable processes (such as increased noise, or polarisation issues)[114] which may interfere with the EVV 2DIR FWM signal, hence it has been fixed at a frequency of 790 nm for the experiments presented in this thesis.

This resonance enhancement effect has great potential for biological applications such as inhibitor-protein binding studies, where the inhibitor is often vastly smaller than the target protein, and consequently signal generated from a non-conjugated inhibitor can be swamped by the much larger protein signal. Resonance enhancement was again utilised by Sowley in 2017 to validate the binding of bright orange drug SU-5402 to target protein FGFR1 (a mammalian growth factor receptor)[49]. Although SU-5402 is several hundred times smaller than FGFR1 (as seen in Figure 19), coupled vibrations arising due to drug binding were strongly observed. The resulting EVV 2DIR difference spectra (shown in Figure 24, produced from quadratically subtracting protein spectra from protein+inhibitor spectra, detailed in Section 4.2.1.6) formed the basis for this project (which is the plant-herbicide analogue of Sowley's pharmacological protein-inhibitor work).

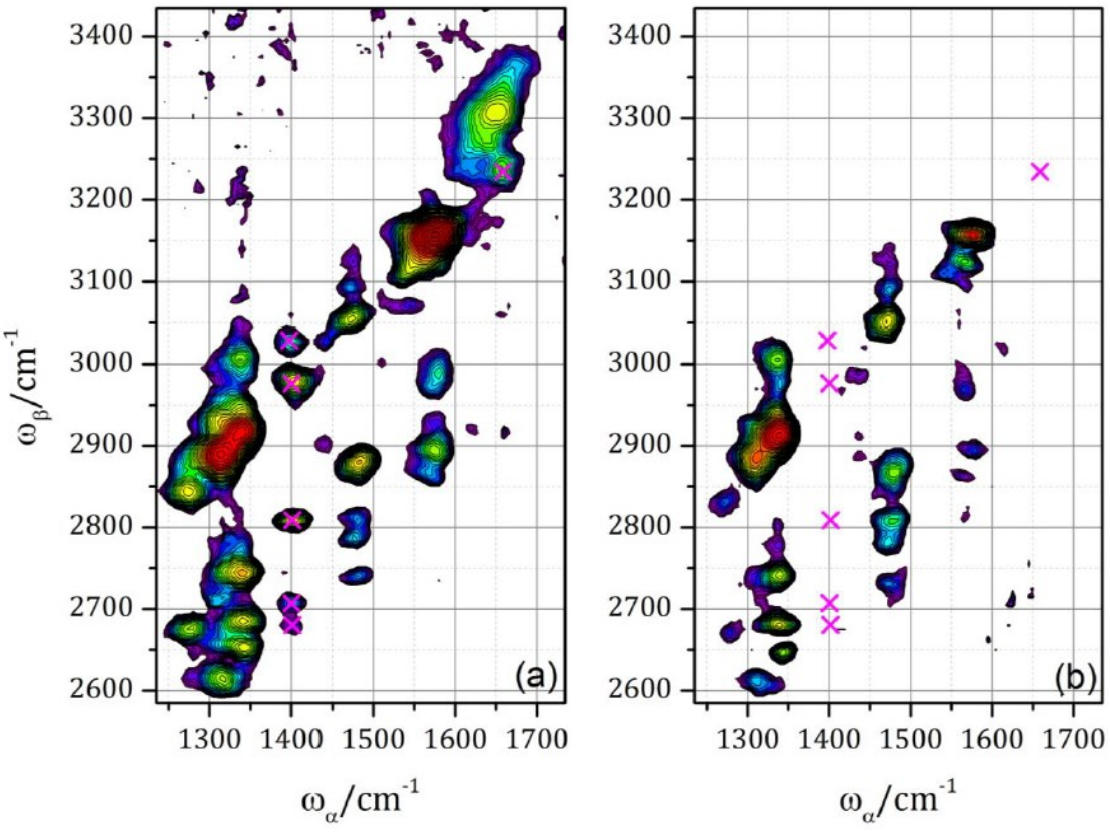


Figure 24: EVV 2DIR difference spectra of (a) FGFR1/SU5402 and (b) BSA/SU5402 (as a control), produced from quadratically subtracting protein spectra from protein+SU5402 spectra (from Reference 49). Pink crosses indicate cross-peaks arising from vibrational mode coupling due to specific protein-inhibitor binding.

In order to calculate the vibrational frequencies of SU5402 in the FGFR1 binding site, Sowley & colleagues also utilised *ab initio* DFT methods with a 6-31g (d,p) basis set in the Gaussian 09 program[115] (following the approach outlined by Kwak et al)[104]. Data from an existing co-crystal structure from the PDB was used[116], however as performing quantum calculations on an entire protein-inhibitor system is not feasible (due to time and computing resources required for calculating vibrational modes for such a complex system), only the protein residues (and one water molecule) within 5Å were included. The resultant EVV 2DIR spectrum is generated by representing each cross-peak as a 2D Gaussian function with a full width at half maximum (FWHM) of 25 cm^{-1} (to match the experimental FWHM values[49]) and a middle frequency corresponding with the harmonic frequencies of the implicated bands[109]. It is displayed in the left-hand side of Figure 25, with the corresponding experimental on the right, showing relatively good agreement.

Cross-peaks A-R are labelled to show equivalent spectral features between the calculated and experimental spectra, and binding-specific features are circled in pink. While the agreement is slightly imperfect, the few extra features in the calculated spectrum can be expected, as the calculations do not take into account the delays utilised between the three pulses and therefore vibrational couplings with differing coherence lifetimes may be present.

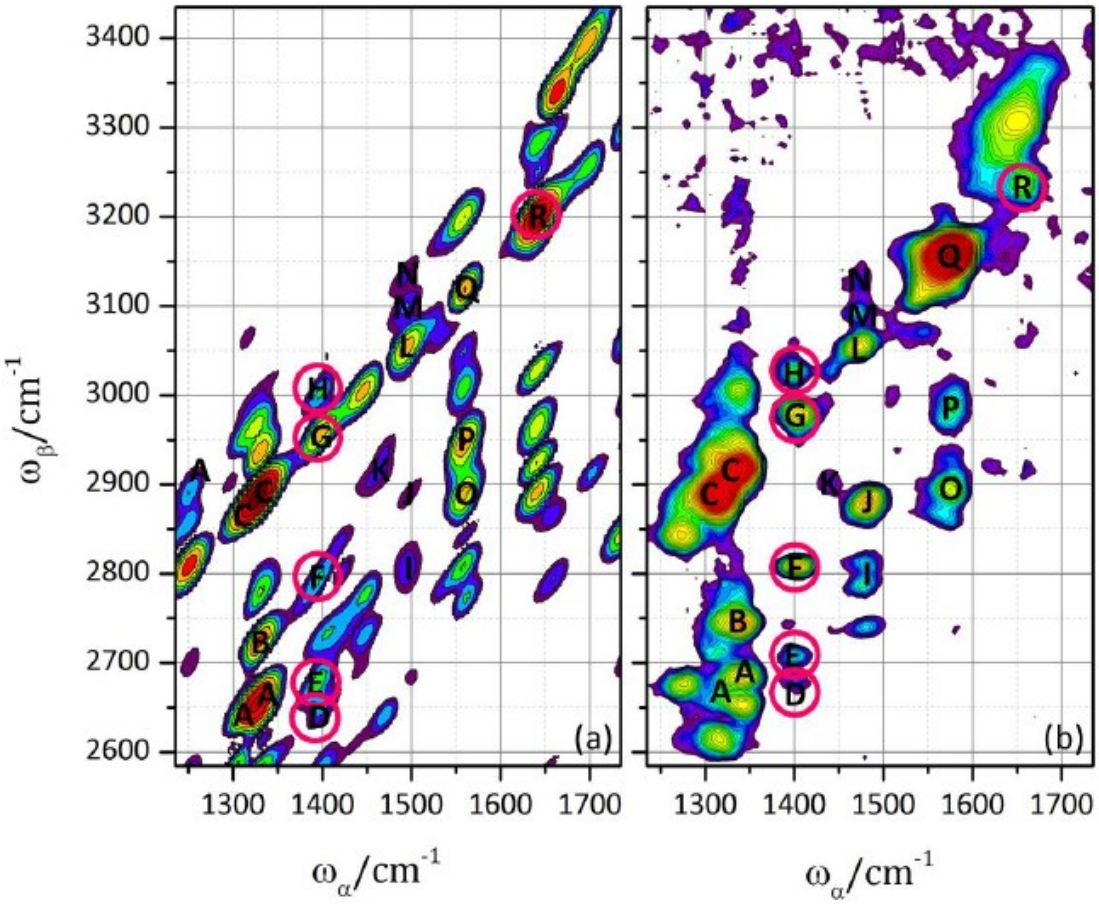


Figure 25: (a) The calculated EVV 2DIR spectrum of SU5402 and proximal amino acids & water molecule in the FGFR1 binding site, and (b) the corresponding experimental FGFR1/SU5402 difference spectrum for comparison.

In this chapter we have outlined the theory of EVV 2DIR, its relations to other spectroscopic techniques, and its advantages for the interrogation of biological systems. In doing so we have laid the necessary foundations for the physical interpretation of the

empirical spectra acquired during this project and which are the subject of the following chapters.

3. Experimental Methods

3.1 EVV 2DIR laser system

3.1.1 Experimental overview

EVV 2DIR spectroscopy utilises three picosecond pulsed laser beams, overlapped spatially and temporally in a sample to produce a four-wave mixing (FWM) signal. The experimental setup of the EVV 2DIR laser system (shown in Figure 26) is comprised of several components which function collectively to achieve FWM, and each can be categorised by one of four stages key to the overall process, namely: pulse generation & amplification, optical parametric amplification, beam alignment & overlap, and signal detection. This section discusses each of these stages in detail, as relevant.

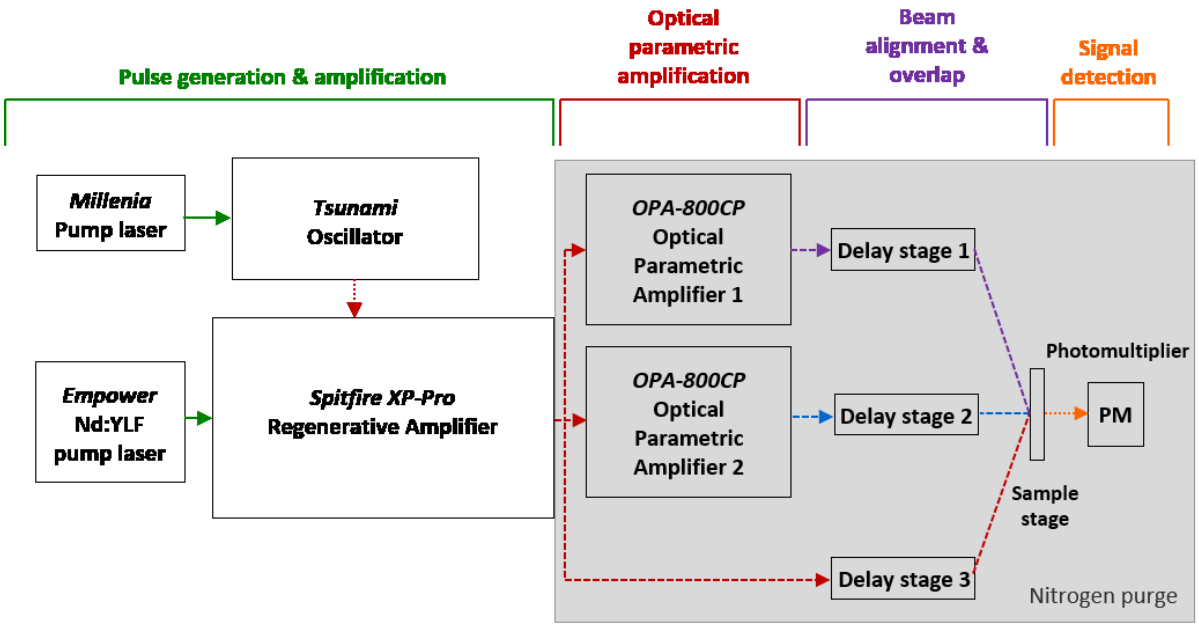


Figure 26: The experimental setup of the EVV 2DIR laser system , showing the various components classified under the role they play in the FWM process. The grey box indicates the system components which are enclosed and purged by nitrogen to prevent spectral artefacts due to any water vapour present in the air.

To summarise, a *Spitfire* laser seeded by a *Tsunami* oscillator (pumped by *Empower* and *Millenia* lasers, respectively, and all commercially available from *Spectra-Physics*) is used to generate an amplified output of 790 nm, 1.5 picosecond laser beam with 2 mJ pulses of 25 cm^{-1} bandwidth and 1KHz repetition rate. This output beam is split into three, with the first portion remaining at a fixed wavelength of 790 nm to form the near-visible probe beam (ω_γ). The remaining two sections are used to pump two *OPA-800CP* units to produce the two tuneable mid-infrared beams (ω_α and ω_β) required for EVV 2DIR experiments, by the process of optical parametric amplification. All three beams are delivered to the sample stage via a series of optics, including delay stages and a parabolic mirror to aid temporal control and telescopes for spatial control. The resultant FWM signal (ω_δ) is directed to the Hamamatsu photomultiplier via a second parabolic mirror and a filter stack which separates the three input beams and transmits only the FWM signal. The detected light intensity is then integrated by boxcar averaging and sent to LabView software and can subsequently be processed to produce an EVV 2DIR spectrum. From the *Spitfire* output to the detector, the laser system is enclosed and purged with nitrogen (signified by the grey box in Figure 26) to prevent beam distortion and water absorption from water vapour present in the air. This is especially important for the ω_α frequencies produced by OPA1 at ~1500 cm^{-1}, as this region contains a water vibrational mode which may obscure genuine coupled vibrational modes in an EVV 2DIR spectrum (further discussed in Section 3.1.4).

3.1.2 Pulse Generation & Amplification

3.1.2.1 Pulse generation

The EVV 2DIR spectroscopic system initially utilises two commercial lasers which work in conjunction to generate a 1.5 picosecond, 2 mJ pulsed 790nm laser beam with a repetition rate of 1KHz. The first of these is a mode-locked Titanium-doped Sapphire (Ti:Sapphire) *Tsunami* laser[117], which is pumped by a *Millenia Pro s-Series* diode-pumped continuous wave visible laser[118] (both from *Spectra-Physics*) to produce ~100 femtosecond pulses with a 100MHz repetition rate and bandwidth of 25 cm^{-1}. These ~100 femtosecond pulses are used to seed the *Spitfire Pro XP* regenerative amplifier[119], which is pumped by an *Empower* diode-pumped Nd:YLF laser[120] (again, both systems from *Spectra-Physics*).

3.1.2.2 Regenerative amplification

The process of regenerative amplification is required to produce the short, intense laser pulses required for EVV 2DIR spectroscopy[121]. An intense pulse passing through an amplifying crystal has the potential to "self-focus" (intensify further by altering the refractive index of the crystal) and cause permanent damage to the crystal[122]. Therefore, the peak power of each pulse must first be decreased by stretching the pulse temporally prior to allowing it to pass through the crystal for amplification.

Firstly, the *Tsunami* produces a mode-locked 100 fs 790nm seed pulse which enters the *Spitfire* stretcher. This is comprised of a diffraction grating which disperses the frequency and wavelength components of the pulse and is masked such that only a portion of the broad spectrum is reflected (narrowing the bandwidth to that of a picosecond pulse). This portion is retro-reflected vertically onto a second diffraction grating which is angled such that the blue component of the spectrum travels further than the red component, stretching the pulse temporally and decreasing the peak power.

A Ti:Sapphire crystal is the gain medium used to amplify the stretched low-energy seed pulses. To prevent potential heat damage arising from the high energy pulses passing through it, the crystal is cooled to ~-10°C by a thermoelectric cooling system. The cooling process can produce condensation which may also damage the crystal, therefore the crystal is housed in a sealed chamber and purged with nitrogen to maintain a constantly low humidity. Initially, the crystal is pumped by a high energy pulse from the *Empower* which excites it to a state of population inversion.

Immediately after, the stretched mode-locked seed pulse is passed though the crystal, and the majority of the energy from the resulting stimulated emission is retained in the crystal. The energy trapped in the medium is transferred to the pulse, amplifying it. The selected seed pulse travels around the *Spitfire* cavity and through the crystal several more times (determined by a user-controlled delay) until it has absorbed all the available energy from the crystal and saturated the population inversion, when it is ejected by the Pockels cell into the compressor.

The pulse is then recompressed by a similar process to stretching, however this time a horizontal retroflector is used. This switches the wavelengths causing the redder components to travel further than the bluer components, enabling the blue and red

wavelengths to come closer together in time and decreasing pulse duration. A telescope at the exit ensures that the resultant output is Gaussian and is straight leaving the *Spitfire*.

3.1.3 Optical Parametric Amplification

The regenerative amplifier ouputs a 1.5 ps pulsed 790 nm laser beam with a power of 2mJ. This is initially split into three parts by beam-splitters, to generate the three laser beams of varying wavelengths required for FWM in EVV 2DIR. The first (ω_γ) remains at a fixed wavelength of 790 nm and is delivered via a series of optics to the sample. In comparison, the latter two beams must undergo the nonlinear frequency conversion process of optical parametric amplification. This occurs via two separate *OPA800-CP* units (from *Spectra-Physics*, herein referred to as OPA1 and OPA2, or OPAs collectively) and is required to convert the beam wavelengths from 790nm to two different tuneable mid-IR frequency ranges necessary for EVV 2DIR (in the region of 1250-3450 cm^{-1}, usually centred around 1500 for ω_α and 3000 cm^{-1} for ω_β, respectively). Optical parametric amplification can be considered in two stages; namely the generation of a white light continuum, and optical parametric generation (OPG). Upon entering the OPA, the 790nm beam is again split into three parts; the former (just 4% of the original beam) being used for white light continuum generation, while the latter two are utilised as pre and post-amplification pump beams in the OPG process.

In contrast to a regenerative amplifier – which derives optical gain from a population inversion of energy states – an OPA attains amplification via a three-wave interactive down-conversion process through a nonlinear medium, outlined in Figure 27. This is nonetheless deemed amplification as a high-intensity pump beam is used to amplify a low energy seed, generating signal and idler beams[123].

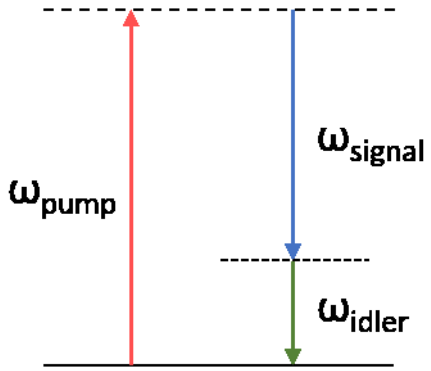


Figure 27: The nonlinear frequency down-conversion process. Note that the sum of the signal and idler frequencies is equal to the pump frequency, due to the principles of energy conservation.

A high energy input photon from the pump is split into two lower energy photons; namely the signal and the idler photons, respectively. Energy conservation (along with conservation of momentum) dictates the signal and idler output frequencies, and the frequency relationship between these and the pump is shown in Equation 19. There are many combinations of signal and idler frequencies which will satisfy Equation 19, which is what enables OPG to produce a wide bandwidth range.

$$\omega_{pump} = \omega_{signal} + \omega_{idler}$$

Equation 19

Highly efficient down-conversion also requires appropriate phase-matching between the interacting beams. As discussed earlier in Section 2.3.2.2, momentum conservation determines that a correct phase relationship must be maintained by the interacting beams as they propagate through the nonlinear medium in the same direction[98], which is shown by Equation 20:

$$\Delta k \equiv k_p - k_s - k_i = 0$$

Equation 20

where k_p, k_s and k_i are the phase matching wave vectors for the pump, signal and idler beams, respectively. Good phase-matching can be achieved by tuning the angle of the nonlinear crystal relative to the diffraction grating angle, (which again allows a wide range of wavelengths to be produced). Practically, this is implemented by having both the non-linear crystal and the diffraction grating mounted on a adjustable actuator.

Thus, the requirements of high pump intensity, a high second order nonlinear susceptibility in the nonlinear medium, and effective phase-matching must be considered when choosing a nonlinear medium for OPG. In the case of OPG in EVV 2DIR, β-Barium Borate (BBO) is a suitable nonlinear crystal as it is efficient at converting lower wavelength ps pulsed beams into higher wavelengths. The following description covers the practical aspects of optical parametric amplification using the EVV 2DIR system.

3.1.3.1 White Light Continuum Generation

By focusing several uJ of energy into a suitable material (in this case quartz), a pump laser light source with narrow bandwidth can be converted into a coherent white light continuum with a broad spectral range. This occurs through an assortment of high order non-linear processes, including self-phase modulation, Raman scattering, self-focusing, FWM and many other multi-photon processes[124]. While both solid-state and liquid materials can be utilised for white light continuum generation, quartz is the preferred material for EVV 2DIR, because unlike many liquid cells it is non-toxic and produces a more stable output. The extensive spectral bandwidth within the white light continuum provides an excellent seed for OPG, as it allows the user to conveniently select a wavelength range of choice.

3.1.3.2 Optical Parametric Generation (OPG)

Once a white light continuum has been generated, it can be used to seed the OPG process (an overview of which is given in Figure 28). This can be separated into three stages;

a. White light amplification

b. Signal seed frequency selection, and

c. Signal amplification.

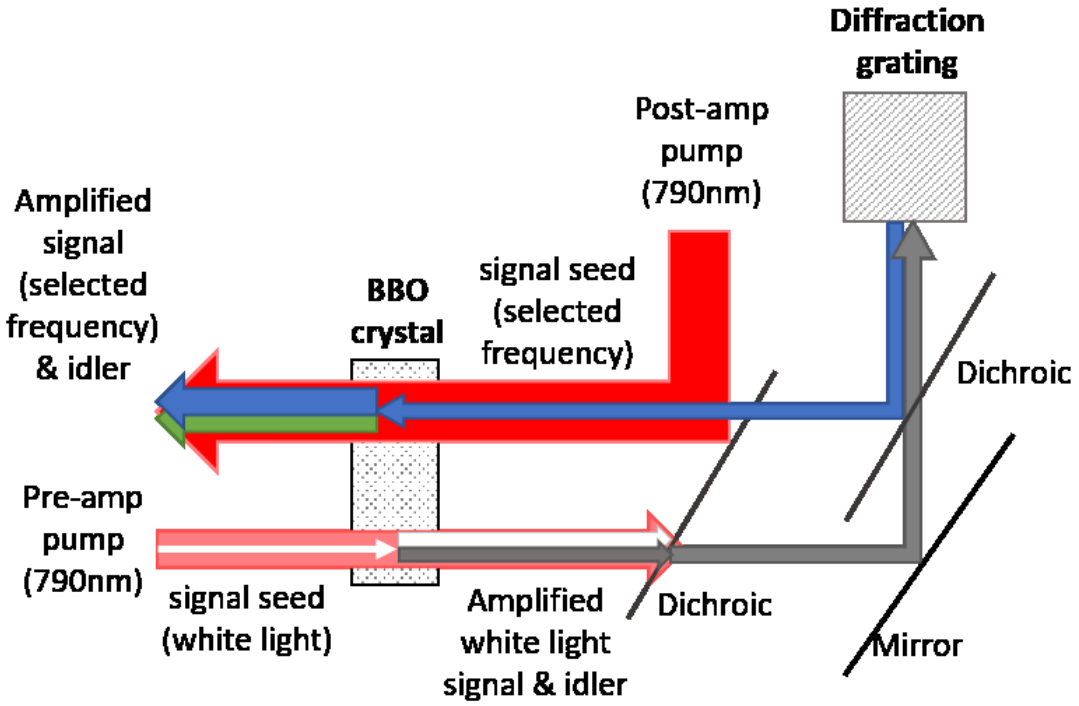


Figure 28: The process of OPG , showing the initial white light amplification followed by wavelength selection from the dispersed idler, and finally signal amplification of the selected frequency in the BBO crystal.

a. White light amplification

The white light continuum and pre-amplification pump beam are delivered via a series of optics to the BBO crystal and are overlapped spatially and temporally within. Temporal overlap is achieved by adjustable delay stages upstream in both the white light and pre-amplification pump beam paths, while the beam size of the pre-amp pump can be optimised via a focus telescope. An amplified white light signal and corresponding idler beam are produced which propagate collinearly through the BBO crystal.

b. Signal seed frequency selection

The resulting idler is then reflected to a tuneable diffraction grating which disperses the frequency components of the beam. The angle of the diffraction grating is adjusted such that the frequency of interest is reflected to the BBO crystal, where it is overlapped with the post-amplification pump beam for signal amplification.

c. Signal amplification

The selected frequency from the dispersed idler beam acts as the seed for signal amplification in the BBO crystal, where it is overlapped with the post-amplification pump beam (in the opposite direction to the white light amplification). Temporal overlap is again controlled via a delay stage earlier in the post-amplification pump beam path, and beam size by a focus telescope. For OPG to occur, the angle of the diffraction grating relative to that of the BBO crystal must be appropriately matched such that phase-matching can be achieved. The resultant output is an amplified signal of user-selected frequency and corresponding idler beam, both transmitting collinearly through the BBO crystal, with a bandwidth of ~25 cm^{-1}. The tuning range of the signal is 6250 - 9090 cm^{-1}, while that of the idler is 3333 - 6250 cm^{-1}. Figure 29a demonstrates the producible wavelength ranges for signal and idler in an *OPA-800CP* versus the maximum pulse energy that can be achieved at each of those wavelengths.

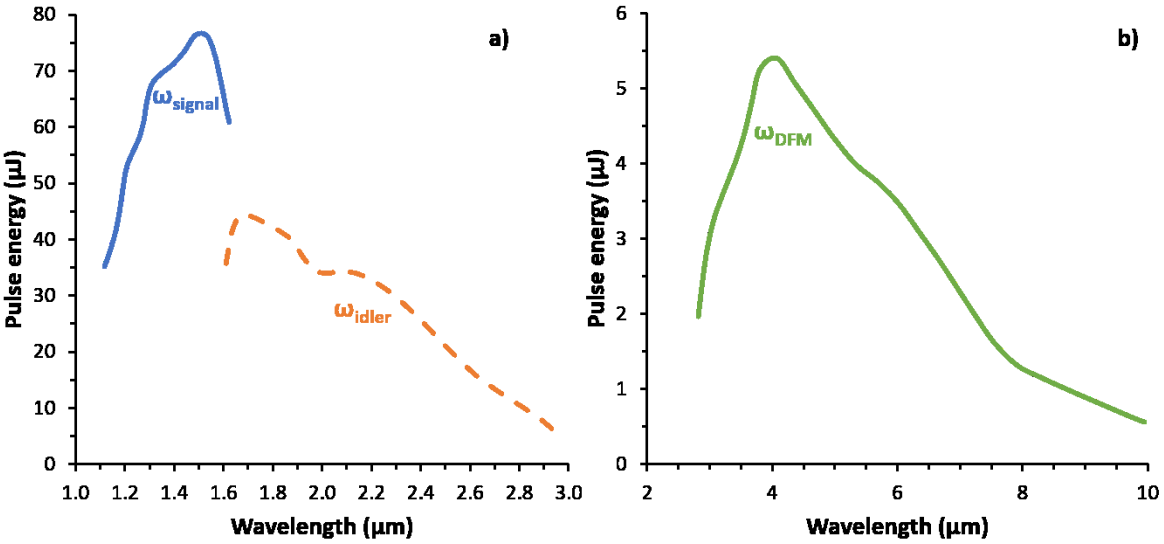


Figure 29: The typical output wavelength ranges vs pulse energy for a) signal and idler beams in OPG, and b) signal produced via difference frequency mixing (both adapted from reference 124)

3.1.3.3 Difference Frequency Mixing (DFM)

EVV 2DIR utilises FWM to probe the many (and important) coupled molecular vibrational transitions between two mid-IR beams. Thus, it is imperative that the signal and idler

beams produced in OPG finally undergo difference frequency mixing (DFM) to generate mid-IR light. DFM is a nonlinear conversion process comparable to OPG which mixes the signal and idler beams to produce a DFM signal, extending the wavelength from 3333 cm^{-1} down to 1000 cm^{-1}. Figure 29b shows the wavelength range and corresponding pulse energies achievable via DFM.

A DFM signal within the range of 1000 - 3333 cm^{-1} can be accomplished by collinearly propagating the OPG output amplified signal and corresponding idler beams into a nonlinear silver gallium sulphide ($AgGaS_2$) crystal. In contrast to OPG, the frequency of the output DFM signal is the *difference* between the incoming signal and idler frequencies, and the frequency relationship for this is shown in Equation 21:

$$\omega_{DFM} = \omega_{signal} - \omega_{idler}$$

Equation 21

The overall DFM process is illustrated in Figure 30. The signal and idler produced by OPG are delivered to the $AgGaS_2$ crystal. As with OPG, the $AgGaS_2$ crystal angle must be tuned for optimum phase-matching and is mounted on an adjustable actuator for this purpose. Temporal coordination is not required, as both beams are produced in parallel during OPG, and they do not need to be overlapped with a separate pump beam (as with white light amplification and OPG). The signal, idler and resultant DFM signal pass through the $AgGaS_2$ crystal collinearly to a unidirectional long-pass filter, which allows only the DFM signal to propagate through it. Hence a stable, tuneable, picosecond pulsed mid-IR beam with a bandwidth of ~25 cm^{-1} and pulse energy of 3 – 8 uJ is produced for EVV 2DIR via optical parametric amplification.

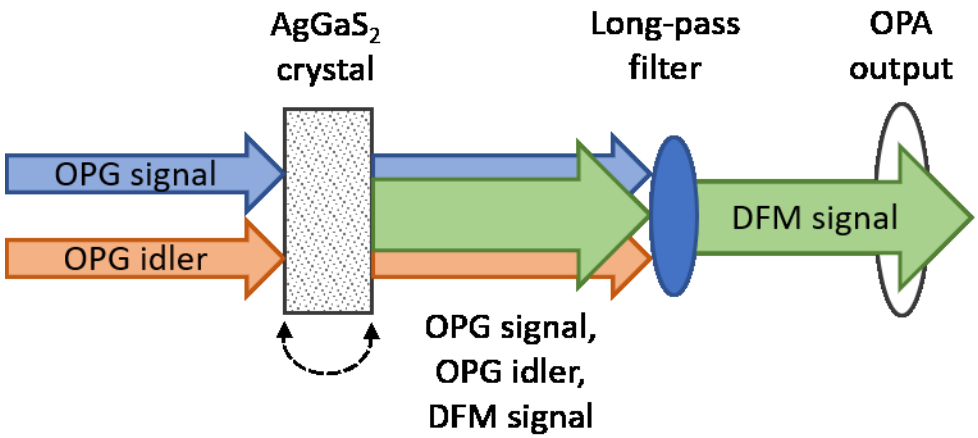


Figure 30: The process of DFM showing conversion of the OPG signal and idler into a DFM signal , the frequency of which is the difference of the original signal and idler. The result is a tuneable mid-IR beam which can be utilised for FWM in EVV 2DIR experiments.

3.1.4 Nitrogen purge

To keep out any water vapour present in the air, the EVV 2DIR apparatus is enclosed in an air-tight purge box from the Spitfire output to the detector, as illustrated in Figure 26, and is purged by nitrogen gas during experiments. The presence of water vapour in the apparatus may distort the laser beams leading to an attenuated signal and can significantly affect the quality of EVV 2DIR spectra. The OPA2 pulse energy can be particularly affected, as water vapour absorbs infrared light at ~1550 cm^{-1} resulting in a drop in pulse energy in this region[125]. This is evident when comparing the OPA2 energy profiles before and after purging with nitrogen, as shown in Figure 31:

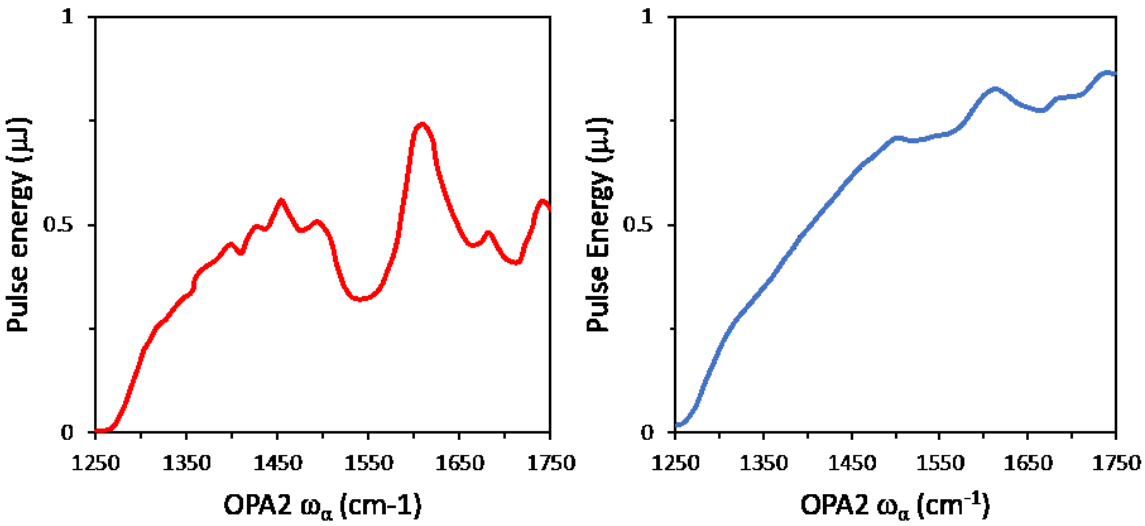


Figure 31: Typical pulse energy values as a function of OPA2 frequency when unpurged (left) and purged (right) by nitrogen gas. The attenuated pulse energies (particularly at ~1550 cm^{-1} due to an IR water vapour absorption) can significantly affect the EVV 2DIR data collected.

Consequently, the laser system must be kept tightly enclosed over the areas indicated in Figure 26 to prevent water vapour absorptions from interfering with the output OPA pulse energies.

3.1.5 Beam Alignment & Overlap

The two mid-IR beams produced in the OPA process, along with the original 790 nm beam from the regenerative amplifier, are then directed to the sample stage via a series of optics (the overall process of which is outlined in Figure 32).

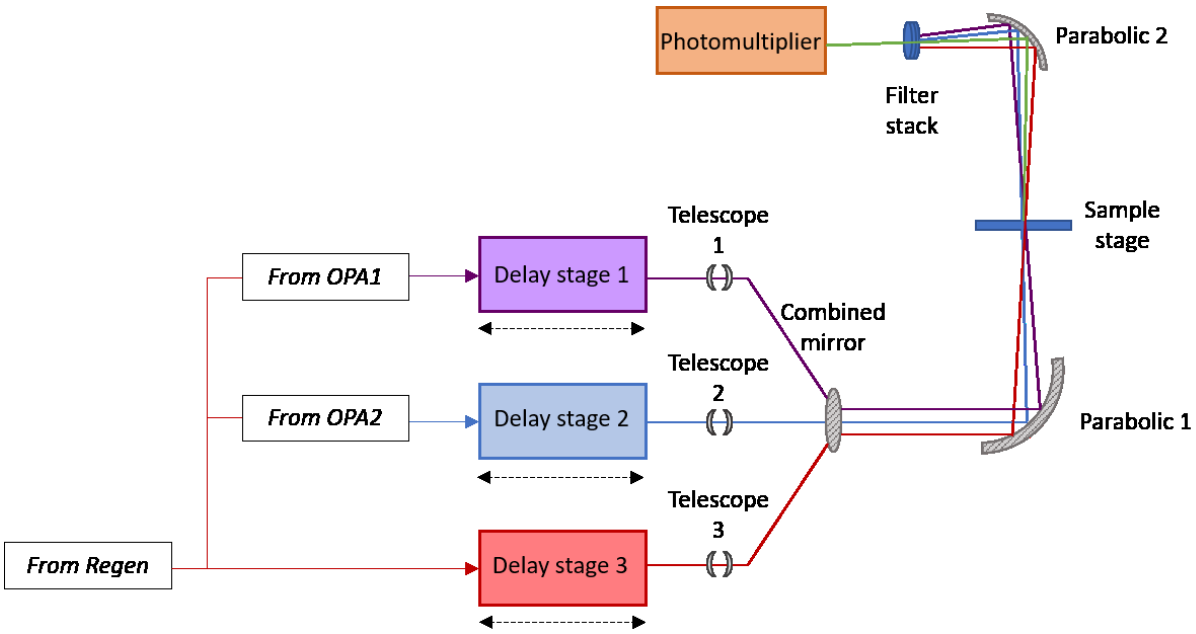


Figure 32: A schematic showing the key elements of the EVV 2DIR system which enable beam overlap, delivery & signal detection. Note that only the components relevant to discussion are shown - most of optics from the EVV 2DIR set up have been omitted for clarity.

The 790 nm beam and two mid-IR beams and 790 nm beam are delivered from the regenerative amplifier and two OPAs to the sample stage by a series of optics, where a FWM signal is generated and detected by a photomultiplier. A filter stack separates the FWM signal from the three beams utilised for FWM. Adjustable delay stages in each of the beam paths enable temporal control of pulse timings, while the use of telescopes and parabolic mirrors aid in correcting spatial alignment (both discussed in further detail in the following sections).

3.1.5.1 Temporal control: T0 Calibration

While the tuneable nature of the OPAs used in EVV 2DIR enables generation of the changeable mid-IR beams required for EVV 2DIR experiments, it can also produce artefacts which must be experimentally accounted for. In order for the phase matching conditions to be met, each output frequency exiting the OPA will have a different set of BBO crystal, diffraction grating and DFM crystal angles. This not only causes variation in beam path length, beam pointing and angles of the three EVV 2DIR beams (ω_α, ω_β & ω_γ), but also the relative pulse timings. Discrepancies of up to 0.8 picoseconds have been

previously observed between pulses for this system[49]. While these effects originally occur at the OPA output, they will be magnified further downstream at the sample, and can result in spectral artefacts such as non-resonant background.

To correct for this practically, a T0 calibration is carried out on both OPAs prior to each set of EVV 2DIR experiments. A delay stage is present in each of the three beam paths between the OPAs and the sample stage for this purpose (as seen in Figure 32). The T0 positions can be defined as the positions of each of the three delay stages relative to one another, which give a maximum non-resonant FWM signal from calcium fluoride at a central frequency (1500 cm^{-1} for OPA1 and 3000 cm^{-1} for OPA2). These can be experimentally determined by systematically scanning each stage in turn, sending them to the positions yielding optimum signal, then repeating until the signal no longer increases.

Calcium fluoride is an ideal material for infrared applications such as the T0 calibration as it has a low absorption, high damage threshold, and a low refractive index eliminating the need for any anti-reflective coatings[126]. The fact that the FWM signal is non-resonant means that it only occurs when the three beams are temporally synchronised. Experimentally it can be challenging to find an initial overlap of the pulses following a significant change in regenerative amplifier or OPA alignment, in which case a BBO up-conversion crystal can be used to find temporal coincidence between the 790 nm beam and one of the infrared beams[124]. The T0 position of the third beam can then be found by scanning the remaining IR beam delay stage.

In brief, the practical steps to performing the T0 calibration are as follows:

1. Find the T0 positions at the central frequency (1500 cm^{-1} for OPA1 and 3000 cm^{-1} for OPA2)
2. Systematically step the frequency across the entire OPA output frequency range (either side of the central frequency) and find the T0 positions at each frequency relative to the central T0 positions
3. Plot T0 vs frequency to create a calibration line fitted with a third-order polynomial so that the correction for any given frequency across the range is known (example T0 calibration curves are shown in Figure 33)
4. The T0 corrections are applied during experiments via LabView software as the spectra are collected.

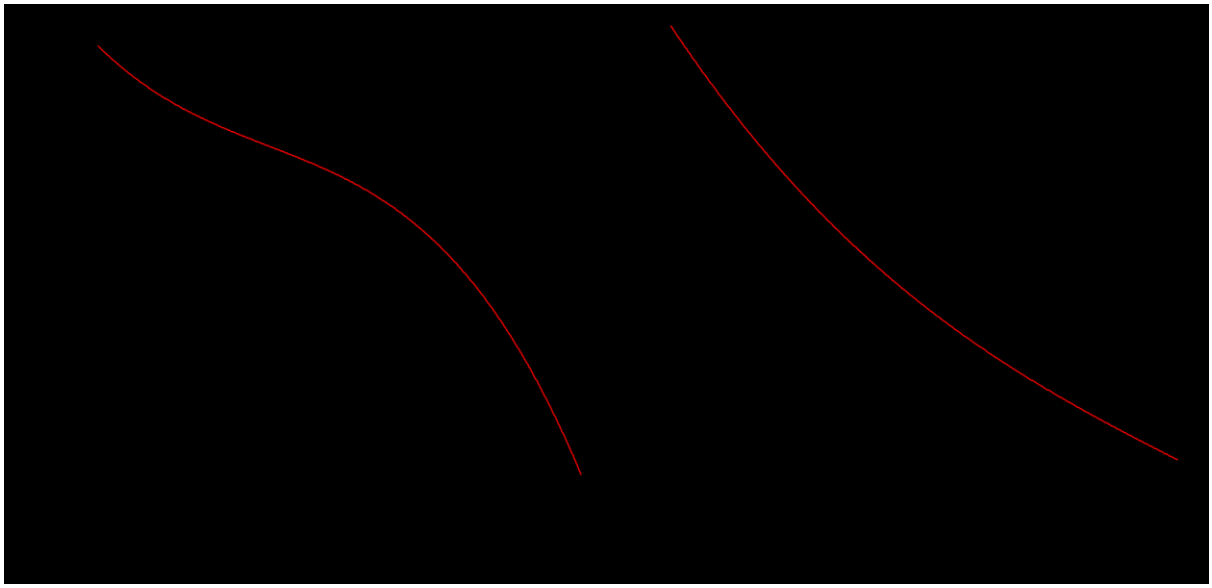


Figure 33: Typical T0 calibration curves for a) OPA1 and b) OPA2, with fitted third-order polynomials.

3.1.5.2 Spatial control

The variable beam pointing with varying frequency out of the OPAs can be corrected for using several optics in the beam paths. All three beams are directed to the sample stage via a series of flat, planar silver mirrors (most of which are omitted from Figure 32 for the sake of simplicity) which are systematically tweaked for optimum alignment down the beam path prior to experiments. Additionally, the beams are each passed through a telescope (as demonstrated in Figure 32) which ensures that the beams are collimated and traveling straight into the parabolic 1, and that the beams all have a common focus and are size-matched in the sample.

In order for a FWM signal to be generated, it is important that the beams reach the sample at angles appropriate for phase matching to occur, which is achieved by use of a large off-axis parabolic mirror. A second, smaller parabolic mirror (again off-axis) is required to deliver the FWM signal to the detector, as the beams are divergent when exiting the sample stage.

3.1.6 Signal Detection

Once the three beams are correctly overlapped both spatially and temporally in the sample, the resultant FWM signal is guided to the Hamamatsu photomultiplier via a second parabolic mirror and through a filter stack, where it is detected as a light intensity. The filter stack consists of several long-pass filters which transmit only the FWM signal and prevent the 790 nm and two infrared beams from passing through to the detector. The transmitted FWM signal is detected in the visible region. The FWM signal is detected as a voltage by the photomultiplier, then integrated via a boxcar average across a 2 ns gate centred on the signal. This integrated signal is then sent to the LabView software, which enables the user to optimise the optics and temporal control of the pulses for maximum intensity.

3.1.7 Modification of IR Range

Initially, EVV 2DIR experiments for this project were carried out over the spectral ranges of 1250-1750 cm^{-1} and 2600-3450 cm^{-1} for each dimension. However, later in the project the decision was made to collect spectra with the second-dimension spectral range altered to 2100-3100 cm^{-1}, as this region contains vibrational cross-peaks for several amino acids of interest (discussed further in Section 4.3). From a practical perspective, this requires realigning and recalibrating OPA2 to the lower frequency range, which has multiple implications. Due to the turning constraints of the DFM crystal mount, 2100 cm^{-1} was the lowest end of the wavelength range that we could possibly recalibrate the OPA unit. The key challenge when realigning the inside of OPA2 was ensuring that the DFM crystal could rotate enough to cover the required range without the signal and idler beams clipping on the side of the tuneable crystal mount. Another practical issue encountered was the appearance of two horizontal stripes of non-resonant background appearing in the 2DIR spectra for these experiments. It was initially thought that these were occurring due to an incorrect T0 calibration, however several attempts to address this proved ineffective. Instead, various combinations of filters were stacked before the photomultiplier to see if these were unknown nonlinear effects that could be separated from the FWM signal prior to detection. Unfortunately, none of these techniques were successful in preventing these non-resonant “stripes” from appearing, and the origin of the artefact remains unknown.

Instead, they were effectively removed in processing as discussed later in Section 4.3.3 (Image Processing).

3.1.8 Polystyrene calibration

Despite calibrating both OPA-800CP units, there is generally a consistent frequency offset of 5-20cm^{-1} for each OPA during a given set of EVV 2DIR experiments. To correct for this, the one-dimensional spectrum of a NIST polystyrene standard is measured for both OPA1 and OPA2 prior to each set of experiments. The resulting spectra are compared to the literature polystyrene spectrum (see references [127,128]) and an average of the differences between measured and known peaks is calculated for each OPA. This can then be subtracted from or added to the collected set of spectra in each dimension to apply the appropriate frequency offsets. It is highly important that these offsets are corrected for in the assignment of EVV 2DIR cross-peaks, as they cannot be compared to known literature values without the polystyrene calibration.

3.1.9 Data processing

Raw data collected during EVV 2DIR experiments requires processing in several stages, namely; transposing, smoothing, removing the electronic offset & non-resonant background, and correcting for the OPA energy profile variation. An additional stage was necessary for experiments where the range for OPA2 was altered to 2100-3100 cm^{-1}.

3.1.9.1 Smoothing

Nonlinear spectroscopic experiments are subject to sources of noise which requires specific treatment of the resulting data to minimise their impact. The two mid-IR beams used in these EVV 2DIR experiments are both produced by separate OPAs, each of which inherently introduces optical noise. The output mid-IR beams have a standard deviation of 0.5-1.5% over 100 laser shots/pulses, and additionally have a tendency for longer term fluctuations of the order 1% over several seconds. To mitigate this, data smoothing can be a valuable tool to reduce spectral noise during data processing.

In the experiments presented in this thesis, raw EVV 2DIR data is firstly transposed, then smoothed using the "nearest neighbour" method, coded in the MATLAB script given by Appendix A. This particular smoothing technique takes each pixel of the spectrum and smooths it as a function of the average of all the pixels surrounding it, as shown in Figure 34.

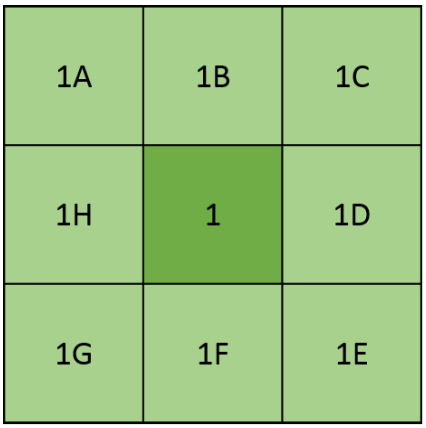


Figure 34: A diagram demonstrating the "nearest neighbour" method of smoothing used in EVV 2DIR data processing. Pixel 1 is smoothed by replacing it with an average of 1A-1G. This process is repeated for every pixel in an EVV 2DIR spectrum.

For this thesis, EVV 2DIR data is collected with a resolution of 5cm^{-1} per pixel, with the three laser beams comprised of a 25 cm^{-1} bandwidth at full-width half-maximum. Hence resolution is preserved, as data is smoothed with two nearest neighbours in each dimension (i.e. a 15 x 15 cm^{-1} average) which is less than the bandwidth of the beams (equivalent to 5 x 5 pixels, or 25 cm^{-1} in each dimension).

3.1.9.2 Background removal

Along with the laser noise produced by the OPAs, the EVV 2DIR apparatus has intrinsic background noise introduced by the electronics utilised in the system. In particular, the boxcar averager has a constant electronic background which contributes to a constant background level seen in all EVV 2DIR spectra collected for this thesis. Practically, this can be seen by blocking all three laser beams and measuring the electronic offset. The 790 nm beam also has a constant level of signal associated with some slight red leakage through

the filter stack before the photomultiplier. Experimentally, we can check for this by testing the signal with only the 790 nm beam unblocked.

Both backgrounds outlined above are *constant* and can be easily removed in data processing. To account for the resulting level of background, a section of 5 x 5 pixels (25 cm^{-1} in both dimensions) with a low and relatively constant level of intensity (standard deviation of ~1%) is averaged and removed from every pixel of each entire spectrum after smoothing. This method of background removal is executed using the MATLAB script in Appendix A.

Previously, conjugated molecules (such as the herbicides used in this thesis) have sometimes been found to show a low level of fluorescence during data collection[49]. To test for this, the two mid-IR beams are blocked and the residual signal measured. If the signal is higher than as would normally be expected for just the boxcar and 790 nm beam offsets, it is likely that fluorescence is occurring, and this needs to be removed as a separate offset in processing. However, a fluorescence background was not observed for any of the herbicide compounds investigated in this thesis.

3.1.9.3 OPA energy profile correction

The frequency ranges of the EVV 2DIR experiments performed using the system outlined in this thesis are limited by the specifications of the OPAs. Each OPA is optimised for a central frequency; usually 1500 cm^{-1} for OPA1 and 3000 cm^{-1} for OPA2, generating 3 uJ and 8 uJ of power, respectively. Because the DFM crystal must turn about such a large range of angles (spanning ~70°) during an experiment, the DFM process gets progressively less efficient as frequency deviates from the centre. Consequently, the energy of the OPA outputs as a function of frequency are uneven, which subsequently requires the data to be normalised by input energy in processing. Typical energy vs wavelength profiles for both OPAs are shown in Figure 35:

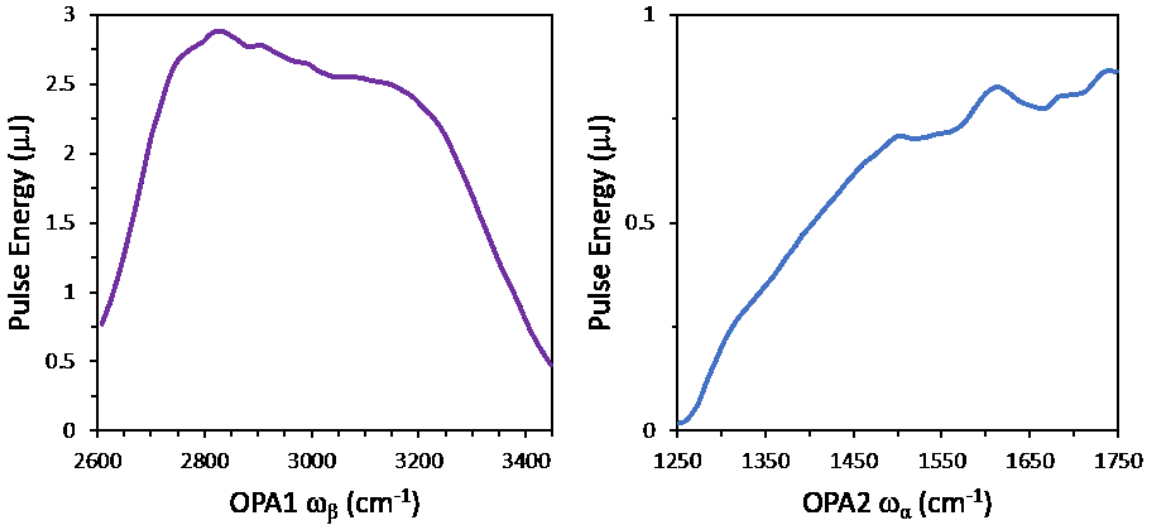


Figure 35: Typical output pulse energy values for OPA1 (left) and OPA2 (right) as a function of frequency , measured at the sample stage through a 150 μm pinhole.

Practically, the OPA energy profiles are measured by placing a 150 μm pinhole at the sample stage with a power meter behind it, and recording the energy produced as a function of varying OPA frequency over the full range. The pulse energy profiles can differ slightly with each new alignment of the laser system, so it is best to properly account for this in post data collection processing.

To correct for the variation in energy with changing wavelength, a matrix of the energy profiles is firstly produced in MATLAB by multiplying the data for the two OPA profiles together. Next, the data for each EVV 2DIR spectrum is normalised by the matrix, using the MATLAB normalisation code provided in Appendix A. This enables us to view the data with intensities proportional to the input energy at each frequency. However, it is important to be aware that the further a peak deviates from the central frequency, the higher the standard deviation, hence the resulting peak intensities should be considered accordingly.

3.2 Sample preparation

3.2.1 HPPD assay

Prior to carrying out protein-inhibitor experiments, it is vital to check that the enzyme of interest is functioning as expected. For this purpose, a functional UV assay was performed using a protocol provided by Syngenta based on the assay published by Yang et al[55]. In brief, HPPD produces Homogentisate (HGA), which is then converted to Maleylacetoacetate (MAA) by HGA dioxygenase (HGO), and the formation of MMA is measured by its UV absorption at 330 nm. HPPD protein was provided by Syngenta for all HPPD experiments, unless otherwise stated. 41 mg/ml *Arabidopsis thaliana* HPPD (~95% purity, in PBS pH 7.5) was assayed to determine the activity (expressed in terms of its K_m value) of the enzyme. The formation of MAA from the action of HGO was measured at 330 nm in 96-well UV transparent plates at room temperature. The 41 mg/ml *Arabidopsis thaliana* HPPD was diluted to 0.5 mg/ml in a buffer of 50 mM bis tris propane (BTP) and added to a mix containing 4 mg/ml catalase and 25 mM sodium ascorbate. The HPPD activity in this mixture is assessed at a variety of concentrations of HPP (0.78, 1.56, 6.25, 12.5, 25, 50 & 100 µM, prepared using a 2-fold series dilution) upon addition of 20 µg/ml HGO. A K_m value of 10 µM was determined for HPPD (Figure 36) which compares well with the literature value of 11 µM[55]. This indicates that the HPPD protein tested was functional and maintained the expected activity prior to EVV 2DIR experiments.

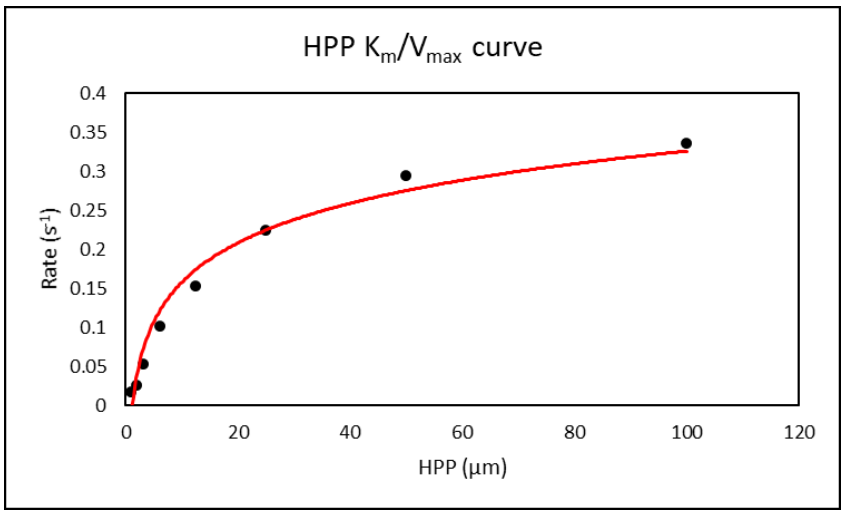


Figure 36: The K_m/V_{max} curve for HPPD activity , showing the concentration of substrate HPP levelling off at the maximum rate. K_m values are at half the V_{max} value, and when compared with literature values provide a good indication of protein activity.

While initially this assay is generally an adequate test of the HPPD activity and is a standard method used by biochemists, the limitations of this method became clear over the duration of the project. HPPD contains iron in the binding site, and in the purification stage of HPPD production process Syngenta usually passes the protein through a size-exclusion column which can remove some of the iron. It is not established in the literature the extent to which removing iron decreases the enzymatic activity of HPPD, but protein production experts at Syngenta estimate an overall activity of 50-70% following size-exclusion purification. Therefore, EVV 2DIR experiments have been performed on both HPPD that has had iron added following purification, and non-iron-treated HPPD protein.

Enzyme activity is especially important for this project, as these EVV 2DIR experiments rely on the fact that HPPD has a 1:1 stoichiometry of binding to its inhibitors. If the HPPD has significantly less than 100% active sites, it is difficult to know if there is excess inhibitor molecule unbound by active sites in the sample. This is discussed further in 4.2.2, where the binding spectra of iron-treated and non-iron-treated HPPD with two herbicides are compared.

3.2.2 Preparation of gel-like sample spots

In vivo, biomolecules are not dry, but are hydrated by surrounding water molecules that influence their structure and function[9], therefore to simulate their native environment it is desirable to perform experiments with proteins in a semi-native conformation. Due to the strong infrared absorptions of water (which often overlap with protein and peptide cross-peaks[14]) it would be practically challenging to implement EVV 2DIR experiments on aqueous samples of biomolecules. Similarly, diluting the samples (with the aim of decreasing the interfering water absorption) would significantly decrease the EVV 2DIR signal arising from the biomolecular sample. This has consequently required optimisation of the sample preparation of biological EVV 2DIR samples.

Previously, the Klug group allowed their samples to dry on cover glass slips open to the surrounding air, resulting in a structure deemed to be a “coffee-ring”, such as that shown in Figure 37. These donut-shaped forms occur when the protein dries in the highest concentration at the border of a droplet[111], due to surface-tension effects[129]. While collecting EVV 2DIR spectra on the coffee-rings may provide higher signal intensity due to pre-concentration of protein analyte[93], there are several limitations with this style of

sample preparation which may lead to variability in the resultant spectra. From a practical perspective, air-drying the samples is not hugely accurate, as the ambient temperature can change significantly with the weather, unless performed in a precisely temperature-controlled laboratory. In the same way, exposing the open samples to the nitrogen purge is likely to contribute to their deterioration over the course of an experiment. Additionally, the laser beams are subject to slight movement across the sample, due to changing frequency and pulse-timings over the duration of an experiment. With the coffee-ring structure, this slight movement may result in a significant change in thickness of the sample being probed, and because the EVV 2DIR signal intensity is directly proportional to the number of molecules in the beam[109], this could result in a sharp drop-off in signal intensity.

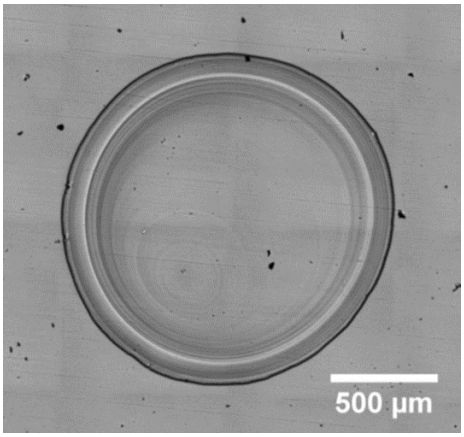


Figure 37: A high-resolution 20x brightfield microscope image of a FGFR1 sample spot with coffee-ring edge, taken from Reference 49.

The consequent development and optimisation of sealed humidity cells has enabled samples to form smooth, gel-like spots keeping the biomolecules in a semi-hydrated, *semi-native* conformation. The relative humidity levels inside each cell are regulated by saturated salt solutions that give known values of humidity[130]. It is well documented in the literature that various hygroscopic salt solutions can be used to achieve equilibrium in an enclosed space by inducing the desired chemical vapour pressure[130]. Gradual drying of 1 µL sample droplets that are deposited on a glass cover slip in an enclosed sample cell with controlled humidity conditions was observed to result in smooth, slightly dome-shaped

spots which give reasonably stable signal intensity when moving across the sample surface.

Each airtight sample cell is comprised of several components commercially available from Thorlabs. A circular glass cover slip (upon which spots of the sample and humidity solutions are deposited) is placed in a lens mount, followed by a rubber O-ring then a 2 mm thick calcium fluoride window. An airtight seal is achieved by use of two thin PTFE spacers placed in between the lens mount & cover glass, and the CaF_2 window & a metal ring which is carefully tightened to seal the cell. A graphic of a controlled humidity cell is shown in Figure 38, which gives the specifications of each component:

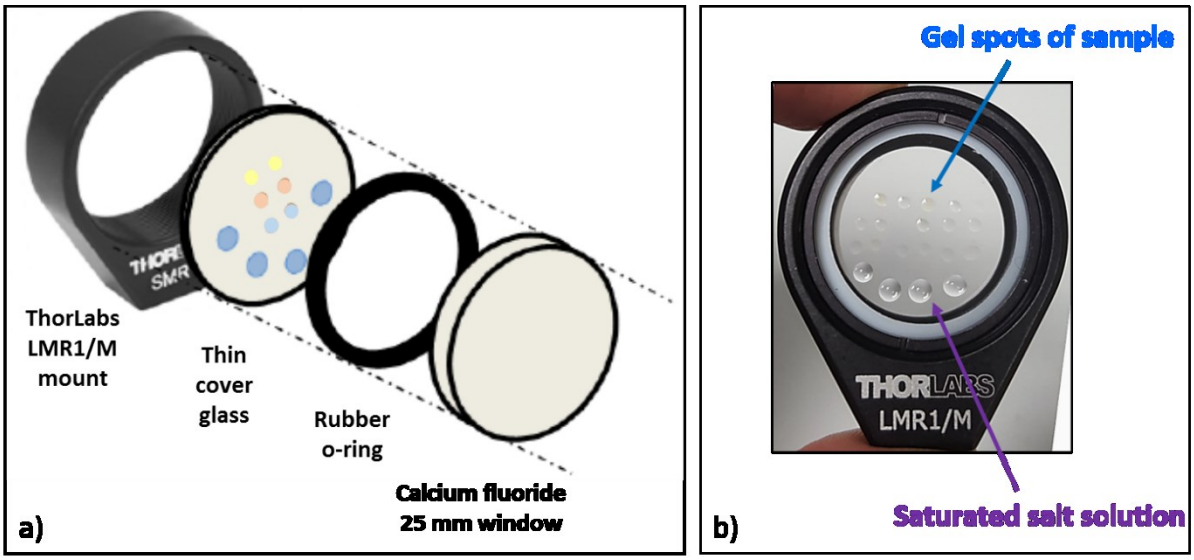


Figure 38: a) a schematic of the parts that make-up a humidity cell and b) an eagle-eye view of an actual humidity cell containing a series of samples to screen for herbicide binding. The white circular parts visible in b) but not shown in a) are thin PTFE spacers placed between the mount and glass, and mount and calcium fluoride window. The image in a) has been adapted from Reference 49.

The relative humidity levels inside each cell are regulated by saturated salt solutions that give known values of humidity[130]. The saturated salt solutions and corresponding humidity percentages (specific to a room temperature of 20°C, rounded to 1 decimal place for brevity) used in this thesis are shown in Table 3:

Saturated salt solution	Relative Humidity (%)
Potassium sulphate	97.6
Potassium nitrate	94.6
Potassium chloride	85.1
Sodium chloride	75.5
Magnesium nitrate	54.4
Potassium carbonate	43.2

Table 3: The list of saturated salt solutions and corresponding relative humidity values utilised in the sealed humidity cells used for sample preparation in this thesis , provided by reference 130.

Both the cover glass and calcium fluoride window are treated with spectroscopic grade acetone and propanol in turn before being rinsed with distilled water and dried with nitrogen. This is an important step to ensure that both windows are clear of dirt, dust or other irregularities which might otherwise impact on spot equilibration or cause unwanted light scattering during experiments[131].

3.2.2.1 HPPD humidity cell preparation

The preparation of gel-like protein and peptide spots presented in this thesis is based on the protocol previously devised Sowley[49]. In initial experiments, protein required buffer exchange prior to spot preparation. This entailed spinning down 50 uL aliquots 0.5 mM *Arabidopsis thaliana* (At) HPPD protein to half the original volume, to give a 25 uL aliquot of 1 mM HPPD. This was then split into three parts and mixed in a 1:1 stoichiometric ratio with highly concentrated solutions of various HPPD herbicides in DMSO, such that the end mixture contains only 2% DMSO. This amount has previously proven sufficient to dissolve the herbicide compounds without causing DMSO-related artefacts in the resulting EVV 2DIR spectra[49]. The high-resolution microscope image of a representative HPPD sample spot in Figure 39 shows an ideal gel-like spot attained from equilibration in a humidity cell containing NaCl to give a known relative humidity of 75.5%.

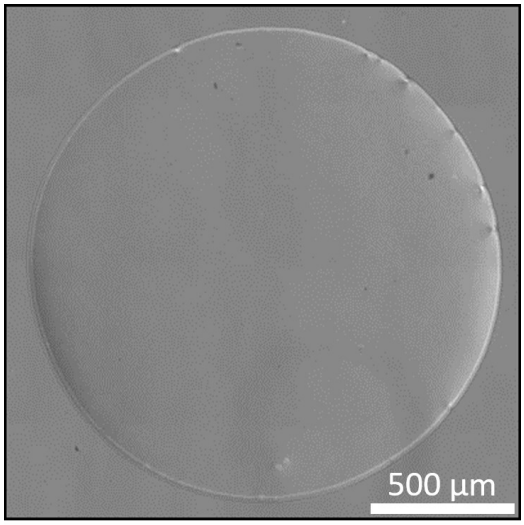


Figure 39: A high-resolution 20x brightfield microscope image of an HPPD gel-like spot after 24 hours equilibration in an air-tight sample cell with a controlled relative humidity of 75.5%.

3.2.2.2 Amyloid beta humidity cell preparation

Different biomolecules can display significant variability in their behaviour towards the local chemical microenvironment, therefore the controlled humidity within the sample cell requires optimisation for each biological system of interest prior to every set of EVV 2DIR experiments. The humidity cell optimisation of amyloid beta spots shows a tendency for equilibration at higher values of relative humidity. A series of airtight sample cells containing the range of saturated salts listed in Table 3 were prepared. The cell comprised of potassium nitrate giving a relative humidity of 94.6% produces the sample spots that look most like the HPPD gel spot shown in Figure 39. However, these appeared more powdery than the smooth, gel-like protein spots. Additionally, it was difficult to find a relative humidity which produced smooth sample spots for all three samples (amyloid beta alone, with copper II chloride and with zinc II chloride), which is discussed further in Section 5.1.2. It is important to perform a set of EVV 2DIR experiments with a common relative humidity used for each sample, as the literature has shown that varying the humidity of a biomolecule's environment can have a significant effect on its structure and function[132]. Thus, a relative humidity of 94.6% using saturated potassium nitrate solution was used.

4. Detection of intermolecular plant-herbicide interactions

In Chapter 2, we discussed the properties of EVV 2DIR that render it inherently sensitive to electrical, through-space couplings. To this end, the Klug group has demonstrated its ability to detect the intermolecular coupling of vibrational modes from two molecules in close proximity[84]. More specifically, EVV 2DIR has been recently used to detect the binding of the commercially available drug, SU54-02, with its target growth-factor receptor protein, FGFR1[49]. The detection of protein-inhibitor binding is a powerful tool in both drug and agrichemical discovery, where a great deal of time and resources is spent narrowing down potential candidates which may eventually prove to be poor at binding the target protein[3,4].

Due to their low cost and rapidity, computationally calculated molecular docking models are an increasingly attractive way to predict the binding of potential candidates to target proteins. Protein-inhibitor binding can be simulated using data from known X-ray or NMR crystal structures in the PDB, and molecular dynamics calculations (which allow the protein and inhibitor of interest to interact for a given time period) provide a dynamical view of the system. However, despite their usefulness in predicting candidate binding, models alone are not sufficient to determine actual protein-inhibitor binding in real-life systems. Thus, EVV 2DIR could prove an effective, rapid method to check that candidates suggested from computational docking models do, in fact, bind to their target proteins. EVV 2DIR is often dubbed the "optical analogue of 2D NMR spectroscopy"[96], the latter of which, along with X-ray crystallography, is frequently used in structural determination of protein-inhibitor complexes[23,133–135]. It is a technique that could provide complementary information to these methods, as it overcomes their requirements for crystallisation of the protein-inhibitor complex (which is sometimes not possible) and large quantities of compound[14]. Furthermore, EVV 2DIR spectroscopy can directly measure hydrogen bonds and other electrostatic interactions rather than infer them (as with 2D NMR and X-ray crystallographic techniques)[84]. Lastly, EVV 2DIR spectroscopy is label-free, unlike many other spectroscopic and crystallographic techniques utilised to elucidate protein-inhibitor binding information.

While EVV 2DIR has been previously utilised to study mammalian protein-inhibitor binding, this thesis presents the first investigation of plant protein-herbicide complexes. The target protein, 4-Hydroxypheylyruvic acid dioxygenase (HPPD), is a bleaching herbicide which remains a key target in the agrichemical industry. HPPD is a protein which catalyses the first step of tyrosine breakdown, which is required by plants as building blocks to produce carotenoids and photo-protectants necessary to perform photosynthesis[51]. Consequently, inhibition of HPPD causes the plant to eventually bleach and die.

Due to many crop species developing resistance to various herbicides[136], the agrichemical industry is constantly searching for new herbicide candidates and modes of action, hence exploring protein-herbicide interaction is very important. Moreover, the capability to detect plant-herbicide binding would expand the range of systems able to be studied by EVV 2DIR. The herbicides investigated in this work all have a significantly lower maximum UV absorption than the bright orange SU-5402 compound utilised in the recent work presented by Sowley[49]. Therefore, the detection of protein-inhibitor complexes with candidates with lower UV-Visible absorptions (such as those employed in this thesis) increases the scope of compounds capable of being examined via EVV 2DIR spectroscopy and shows that near-resonant enhancement can still occur between a probe and a sample with an electronic transition several hundred nm away.

The primary goal of this chapter was to detect the binding of herbicides to HPPD, and to preliminarily assign cross-peaks arising from binding to vibrational modes which may become coupled during this process. This was achieved by measuring EVV 2DIR spectra of HPPD with and without various herbicides, as well as the corresponding spectra for bovine serum albumin (BSA, as a control). The protein spectra were quadratically subtracted from the corresponding protein+herbicide spectra to give difference spectra showing only herbicide contributions. An additional aim of this chapter was to investigate the differences in binding of herbicides to iron-treated and non-iron-treated HPPD, the reasons for which are discussed below.

4.1 HPPD activity: iron vs non-iron treatment

HPPD is a heme-like protein which comprises an iron in the binding site to aid interactions between HPPD and corresponding herbicide compounds. Because HPPD has one iron

atom per binding site, one would expect a requirement of one iron atom per *active* binding site to bind each herbicide molecule in a 1:1 stoichiometry of binding[51]. This becomes important when considering the absolute activity of HPPD, with regards to sample preparation. In brief, the final step in the production of crystallographic grade HPPD by Syngenta involves filtration through a metal affinity column which could potentially remove some of the iron content. To mitigate this, the HPPD can be treated with iron to regain active site activity. It has been speculated that HPPD active sites missing an iron can still bind herbicides but the extent of this is not yet known. Therefore, it is of interest to determine the differences in the binding of herbicides by HPPD both treated with iron and not. Removal of the iron-treatment step could hugely save on resources if detection of HPPD-herbicide binding by EVV 2DIR were to eventually become a high-throughput method in the agrichemical industry.

Consequently, this chapter exhibits EVV 2DIR spectra for a variety of herbicides bound to both non-iron-treated and iron-treated HPPD, with the latter recorded over two different frequency ranges. The first EVV 2DIR spectrum of HPPD is presented and partially assigned. Difference spectra are presented (with preliminary assignments) for each set of experiments, generated by a quadratic subtraction of the HPPD spectrum from the HPPD + herbicide spectrum for each compound. Finally, the binding spectra of iron-treated HPPD complexed with two herbicides is compared to that of the non-iron-treated protein.

4.2 Standard frequency range experiments

4.2.1 Non-iron-treated HPPD experiments

4.2.1.1 Sample preparation: Methods & Materials

The herbicide candidates screened for HPPD binding were selected based on the compounds available which exhibited the UV-vis absorptions at the wavelengths closest to 790 nm, in order to facilitate near-resonance enhancement. Representative UV-vis spectra of two HPPD inhibitors 329 and Nitisinone in DMSO are shown in Figure 40, below:

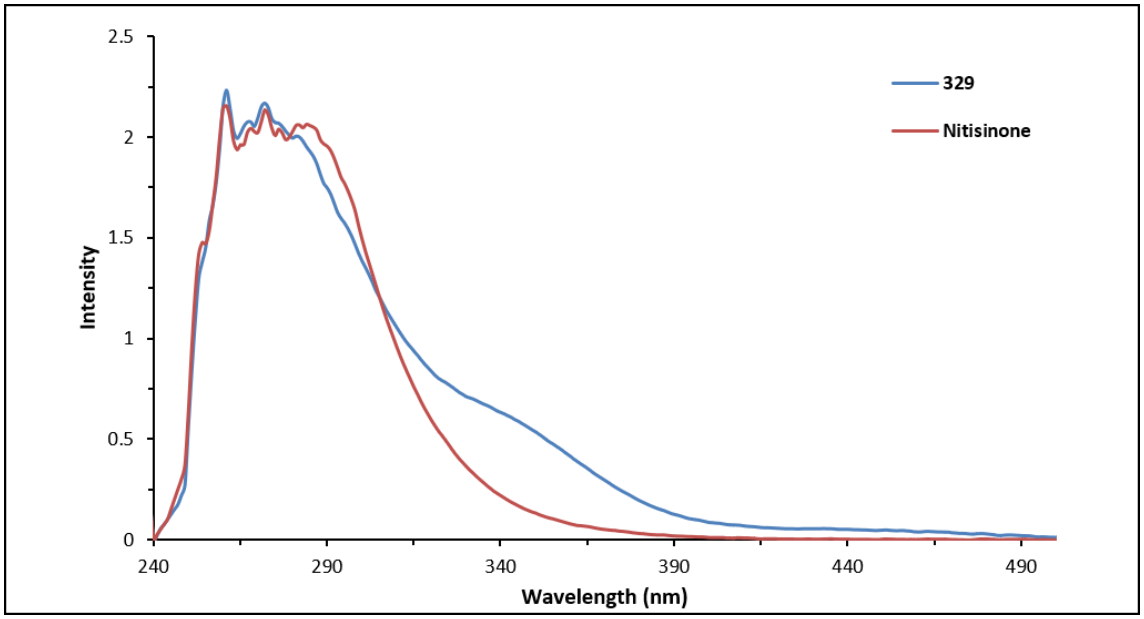


Figure 40: The UV-Vis spectrum of HPPD inhibitors 329 (blue) and Nitisinone (red) in DMSO, both showing maximum UV absorptions at ~285 nm.

The UV-Vis absorption maxima shown for both herbicides in Figure 40 are around 285 nm, which is significantly lower than the UV-Vis absorption maximum of SU-5402 at 430 nm (the drug which binds a growth factor receptor that was previously studied by EVV 2DIR)[49]. However, statistically significant cross-peaks are still observed (as discussed later in this chapter) that can be attributed to the presence of a herbicide which is >150 times smaller than the target protein, so some near-resonance enhancement is likely to still be occurring.

Preliminary data was collected at 10:1 molar ratios of herbicide:BSA to screen for candidates exhibiting potential herbicide or binding peaks, with the screened compounds including commercially available herbicides Tembotrione, Isoxaflutole, Pyrasulfatole, Mesotrione, Nitisinone and four HPPD project compounds provided by Syngenta. Following this, Mesotrione, Nitisinone and 329 (an HPPD project compound provided by Syngenta) were selected to be tested for binding with HPPD at a 1:1 molar ratio (which is the stoichiometry of binding between HPPD and its herbicides[137]).

Initial experiments were performed using non-iron-treated HPPD protein. This was provided by Syngenta as 0.5 mM *Arabidopsis thaliana* (At) HPPD in 20mM Tris pH7.5, 25mM NaCl, 5mM β-mercaptoethanol. It was produced via *E. coli* expression of the pET24_His_AtHPPD_delta28 construct with an intended sequence of 424 amino acids

(provided in Appendix B) and a mass of 48 kDa. The β-mercaptoethanol acts to prevent the protein from oxidation[138], while a high salt concentration is commonly used as a storage buffer to prevent protein aggregation[139]. However, these two components are not ideal for producing the HPPD sample spots required for EVV 2DIR, as the former can interfere with normal protein function while the latter may form crystals which cause laser beam scatter at the sample stage. Moreover, previous work has shown that final concentration of HPPD used in the humidity cells should be ~1 mM to produce an adequate EVV 2DIR signal[49]. Consequently, the protein provided by Syngenta was concentrated and buffer exchanged via spin column centrifugation, to produce 1 mM HPPD in 5 mM phosphate buffer + 5 mM NaCl (pH = 8.0).

Inhibitor compounds are often hydrophobic due to their inherent tendency to bind in pockets of proteins that are protected from water molecules[137] and consequently many display poor solubility in a variety of buffers. To counteract any potential solubility issues, HPPD herbicides were dissolved in a minimal amount of DMSO before addition to the protein solutions. This resulted in a final content of 2% DMSO by total volume, a concentration of which has been shown to interfere with neither the normal biochemistry of proteins nor the EVV 2DIR signal produced[49].

Following buffer exchange, solutions of HPPD and the control protein, bovine serum albumin (BSA), were added to the highly concentrated solutions of various herbicides in DMSO in a 1:1 ratio to give solutions containing either:

1. 1 mM At HPPD, 5 mM phosphate buffer, 5 mM NaCl, 2% DMSO (pH = 8.0)
2. 1 mM At HPPD, 1 mM herbicide, 5 mM phosphate buffer, 5 mM NaCl, 2% DMSO (pH = 8.0)
3. 1 mM BSA, 5 mM phosphate buffer, 5 mM NaCl, 2% DMSO (pH = 8.0)
4. 1 mM BSA, 1 mM herbicide, 5 mM phosphate buffer, 5 mM NaCl, 2% DMSO (pH = 8.0)

where the herbicides tested were the commercially available herbicide Mesotrione, 329 and Nitisinone (originally an HPPD herbicide candidate but now used as an orphan drug to treat Type I Tyrosinemia[140]. Thus, for the purposes of this thesis Nitisinone is referred to as a herbicide, although it is not strictly used to treat plants, currently.

Humidity cells were assembled using a variety of saturated salt solutions, with NaCl producing the most even gel-like sample spots. Thus, the EVV 2DIR experiments presented in this section were carried out in the gel phase at 75.5% relative humidity.

4.2.1.2 FTIR data collection

Attempts were made to collect FTIR spectra in both liquid and gel spot forms for all samples in this chapter, however these were either not successful or too congested for assignment, hence aren't included here. A representative biological spectrum is included in Figure 9 to illustrate this spectral congestion.

4.2.1.3 EVV 2DIR data collection

The non-iron-treated EVV 2DIR experiments were collected over the frequency range ω_α = 1250-1750 and ω_β = 2600-3450 cm^{-1}, with a resolution of 5 cm^{-1} per pixel and an average of 100 samples per pixel. This range was used in line with previous experiments designed to detect protein-inhibitor binding[49]. All three beams were set to the same polarisation (PPP). Following optimisation of relative pulse timings for maximum spectral resolution and minimum non-resonant background, time delays of $T_{\alpha\beta}$ = 1.75 ps and $T_{\beta\gamma}$ = 1 ps were selected to be used during the experiments. EVV 2DIR spectra of each sample spot in turn was collected until four repeats of each sample were obtained. Laser drift can sometimes occur throughout the duration of a set of experiments (each spectrum takes 100 minutes to collect, thus each set of experiments spans 1-2 days). Consequently, the sample position needs to be readjusted prior to beginning each spectrum to give optimum signal, thus each spectrum is collected on a different area of sample. Changing the sample position also ensures that data is not collected on an anomalous part of the sample spot.

The EVV 2DIR experiments in this section were performed in a series of separate sets, with each comprising six sample spots:

- iron-treated HPPD
- iron-treated HPPD + herbicide
- non-iron-treated HPPD
- non-iron-treated HPPD + herbicide
- BSA
- BSA + herbicide.

with protein:herbicide ratios of 1:1, for the protein+herbicide samples. The reason for dividing the experiments into sets is because laser stability starts to deteriorate after several days of experiments. Additionally, it would be difficult to fit all the sample spots needed into one humidity cell. It is also best to begin each set of experiments with a freshly prepared humidity cell to ensure that the protein is not dried out by the time it is measured.

4.2.1.4 Data processing

Data was processed following the protocols outlined in Section 3.1.9, unless otherwise specified. Spectra were normalised from 0 to 1 relative to the methyl peak (at ω_α=1470/ω_β=1495 cm^{-1} for HPPD spectra, and ω_α=1475/ω_β=1450 cm^{-1} for BSA spectra). Spectral normalisation is important for several reasons: it is necessary to be able to average the four repeats for each sample, and to enable comparison between what would otherwise be unitless intensity values of cross-peaks.

The methyl peak is an ideal point for the rest of the spectral features to be normalised against, because it is the most intense feature (meaning it has less fluctuations between repeats) and it is not near any of the potential herbicide binding peaks. The latter requirement is particularly important, because if the normalisation peak were to overlap with any peaks arising due to herbicide interactions, the intensities of the remaining normalised herbicide peaks would be *lower* than compared to the intensity of only the methyl peak.

Following the quadratic subtraction of the protein spectra from the corresponding protein+herbicide spectra, the resulting data was normalised again by *methyl content*, defined here as the ratio of amino acids containing methyl groups in the sequence for each

protein, which was determined to be 2.13:1 for BSA:HPPD. It is important to normalise by methyl content in order to compare the difference spectra of HPPD and BSA, as BSA has more methyl groups giving rise to the methyl signal. Because the methyl cross-peak was used to normalise the rest of the peaks in the spectrum, failing to normalise by methyl content would result in the HPPD cross-peaks having an artificially high (roughly two-fold) value of intensity. Thus, the final processing step in producing comparable BSA difference spectra was to multiply the data by 2.13.

4.2.1.5 EVV 2DIR spectra of HPPD & BSA

Figure 41 shows the first EVV 2DIR spectrum of HPPD (a) next to the spectrum of the control protein, BSA (b). Both spectra exhibit >50 peaks with an intensity of over 3σ (i.e. 99.7% confidence of being real peaks) where σ is defined as the standard deviation of the area of background removed from the average of each set of repeats prior to processing (cross-peaks below this level are omitted for clarity).

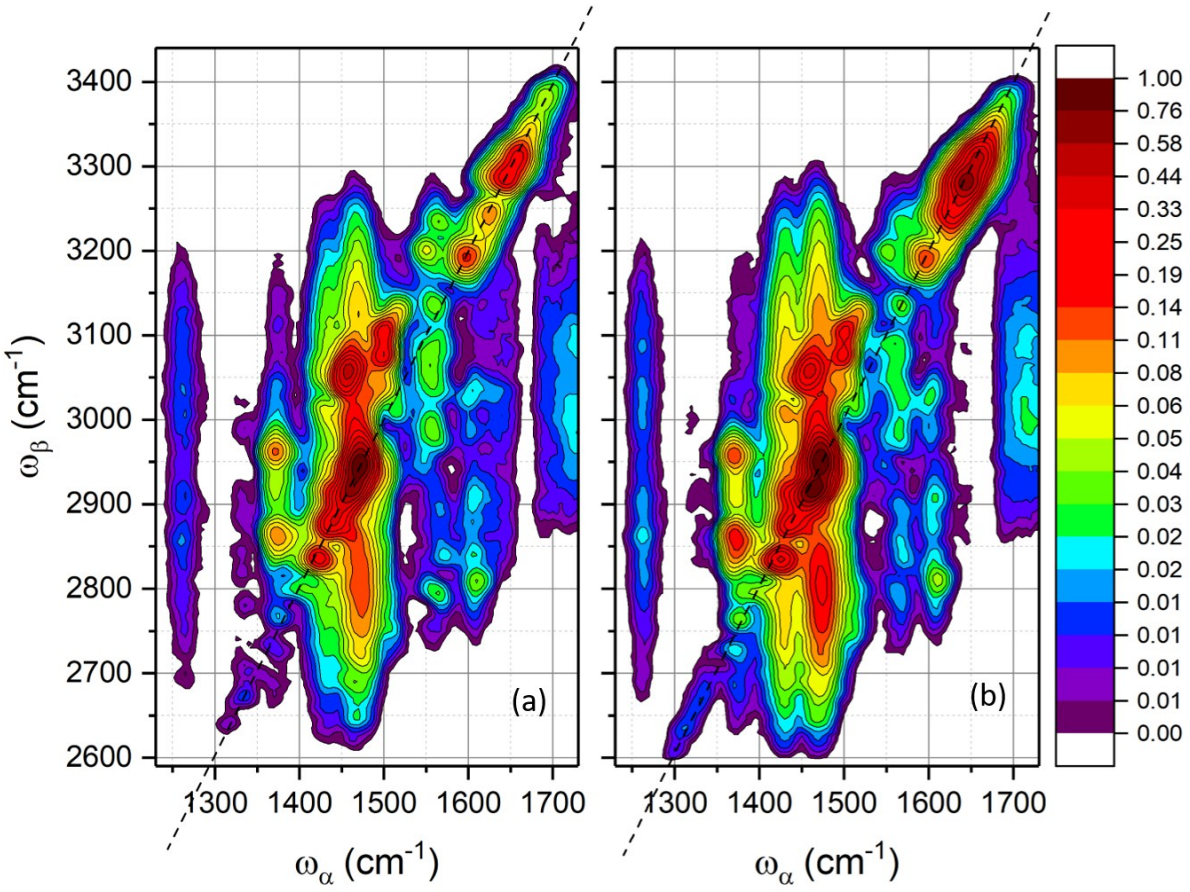


Figure 41: The averaged EVV 2DIR spectra of a) non-iron-treated HPPD, and b) BSA over the standard frequency range. Cross-peaks that fall on the dashed line arise from coupled overtone transitions. The dashed line represents cross-peaks arising from overtone transitions.

The Klug group has previously obtained and assigned spectra of amino acids and proteins, which enables assignment here of various cross-peaks present in both spectra to coupled vibrations arising from several amino acids, given in Table 4 below:

$\omega_\alpha/\omega_\beta$ cross-peak coordinates (cm^{-1})	Assignment
1470/2945	Methyl
1640/3300	Amide I
1470/3050	Phenylalanine
1480/3070	Phenylalanine
1530/3130	Tyrosine
1615/2810	Tyrosine

Table 4: General vibrational assignments of spectral features which appear in both the HPPD and BSA EVV 2DIR spectra in Figure 41.

Most of the spectral features are present in both spectra, the largest two of which are the methyl (centred at ω_α=1470/ ω_β=2945 cm^{-1}) and amide I regions (from ω_α=1620-1720/ω_β = 3200-3400 cm^{-1}). There are subtle differences in relative intensities and resolution in certain regions – mainly, the amide I region, which is known to reflect the tertiary structure of proteins as it arises from interactions involving the amide backbone[141]. Hence, variations between the protein spectra in this region indicate real structural differences between BSA and HPPD. The amide region of BSA (in Figure 41b) consists mainly of one intense broad feature centred at ω_α =1650/ ω_β = 3270 cm^{-1} which suggests the signal arises from a strong coupled vibration, and may obscure any underlying or closely neighbouring weaker cross-peaks. In comparison, the amide I region of HPPD (Figure 41a) shows a series of 5 resolvable peaks along the overtone line. The EVV 2DIR spectra shown in Figure 41 have not been fully assigned, as the main interest of this chapter lies with the protein + herbicide difference spectra, which is discussed in Section 4.2.1.6 below.

In the analogous work involving mammalian protein-drug interactions, there is an obvious difference between the protein only spectrum and the protein + drug spectrum[49]. In contrast, the protein + herbicide spectra for HPPD and its herbicide compounds presented in this thesis show subtler differences. The detection of potential drug and binding-specific peaks therefore relies on the *difference* spectra (calculated by a quadratic subtraction of the protein spectrum from the corresponding protein + herbicide spectrum), as shown

below in Section 4.2.1.6. A representative set of HPPD/BSA + herbicide EVV 2DIR spectra are shown in Figure 42, exhibiting a potential binding peak for 329 (circled).

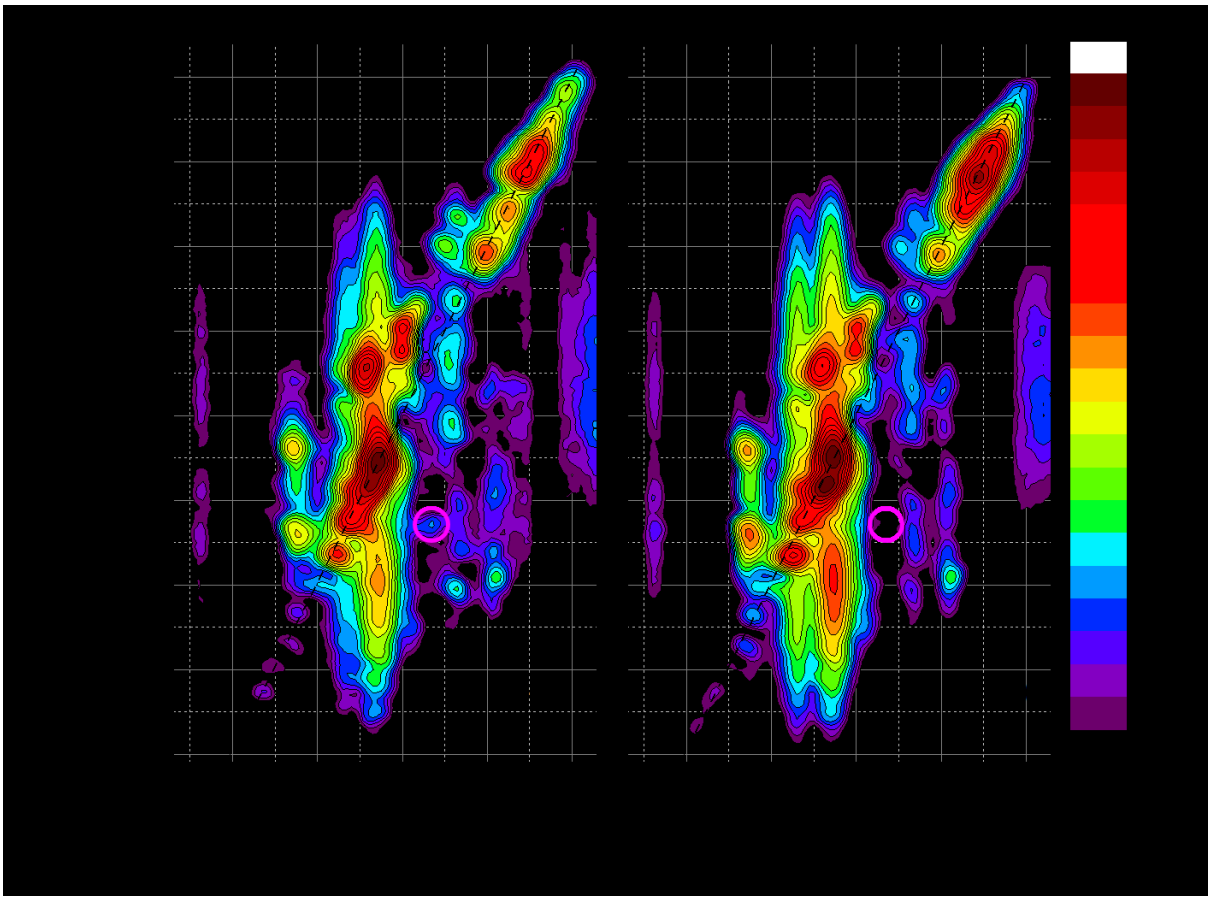


Figure 42: The averaged EVV 2DIR spectra of a) non-iron-treated HPPD + 329 and b) BSA + 329 over the standard frequency range. The presence of a new cross-peak (circled) in a) that is absent in b) is potentially indicative of herbicide binding.

4.2.1.6 Difference spectra

To properly calculate the vibrational contributions arising from a herbicide in a protein-herbicide complex, careful consideration is required when deciding how to subtract the protein signal from the protein+herbicide signal. An important property of EVV 2DIR spectroscopy is that it is a *homodyne* detection technique, which renders the signal produced from a sample as directly proportional to the number of molecules in the confocal ("in-focus") volume of the overlapping laser beams, squared:

$$Signal_A \propto [A]^2$$

Equation 22

It follows that the signal arising from a *complex* containing two components is given by:

$$Signal_{AB} \propto ([A]^2 + [B]^2)$$

Equation 23

which can be expanded to:

$$Signal_{AB} \propto [A]^2 + [B]^2 + 2[A][B]$$

Equation 24

If $Signal_A$ and $Signal_{AB}$ are known, the contribution arising from [B] can be calculated by subtracting the former from the latter. The presence of the cross-terms in Equation 24 indicates that a *linear* subtraction of $Signal_A$ from $Signal_{AB}$ would not correctly account for the contributions of [B] to the signal. Instead, a *quadratic* subtraction can be used by substituting Equation 23 into Equation 24, to give Equation 25:

$$([A] + [B])^2 - [A]^2 = [A]^2 + [B]^2 + 2[A][B] - [A]^2$$

Equation 25

Cancelling the $[A]^2$ terms on the right-hand side and rearranging gives Equation 26:

$$[B]^2 + 2[A][B] + [A]^2 - ([A] + [B])^2 = 0$$

Equation 26

In the EVV 2DIR experiments presented in this chapter, we measure the signal of a protein, $[A]^2$, and the signal of a protein+herbicide complex, $([A]+[B])^2$, thus it is possible to subtract the former from the latter as outlined above to give the spectrum of interactions arising from the addition of herbicide [B]. Practically, this is achieved by use of a MATLAB code (given in Appendix A) which subtracts the averaged protein spectrum from the averaged protein+herbicide spectrum (with each averaged spectrum generated four repeats), in this way.

Mathematically, the same difference spectra can be generated by subtracting the square root of the protein spectrum from the square root of the protein+herbicide spectrum. This can then be squared in order to be comparable with the original protein and protein+herbicide spectra.

Comparing the difference spectra of HPPD and its herbicides to the corresponding difference spectra for BSA enables determination of cross-peaks arising from *binding-specific* interactions. The appearance of a cross-peak in the HPPD spectra which is not present (or is significantly less intense) in the BSA counterpart suggests an interaction arising from the coupling of two vibrational modes associated with protein binding.

The first herbicide compound which showed herbicide induced cross-peaks was the commercially available diketonitrile Mesotrione, the structure of which is shown in Figure 43, below:

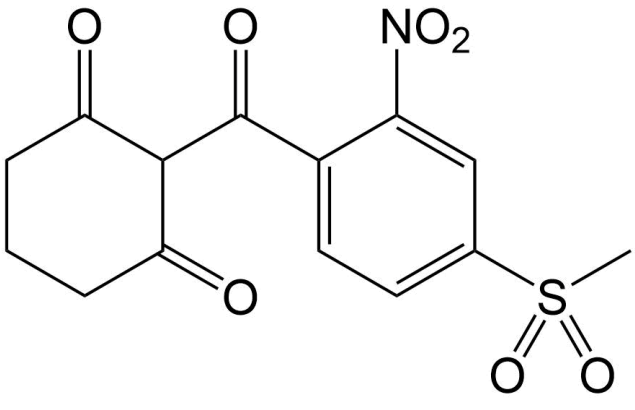


Figure 43: The chemical structure of HPPD herbicide Mesotrione , which is commercially available. Note the aromatic nitro- and sulfone moieties which usually give rise to strong absorptions at characteristic frequencies in 1D infrared spectroscopy.

Figure 44 shows the difference spectrum for non-iron-treated HPPD and Mesotrione on the left-hand side, compared with the corresponding BSA/Mesotrione control difference spectrum on the right. Note that all difference spectra in this thesis have been trimmed to a frequency range that includes all potential binding peaks, for clarity. Both difference spectra contain two cross-peaks above the 3σ level, where σ is defined as the standard deviation in the signal free non-resonant background. Thus, we can be 99.7% confident that these are real peaks.

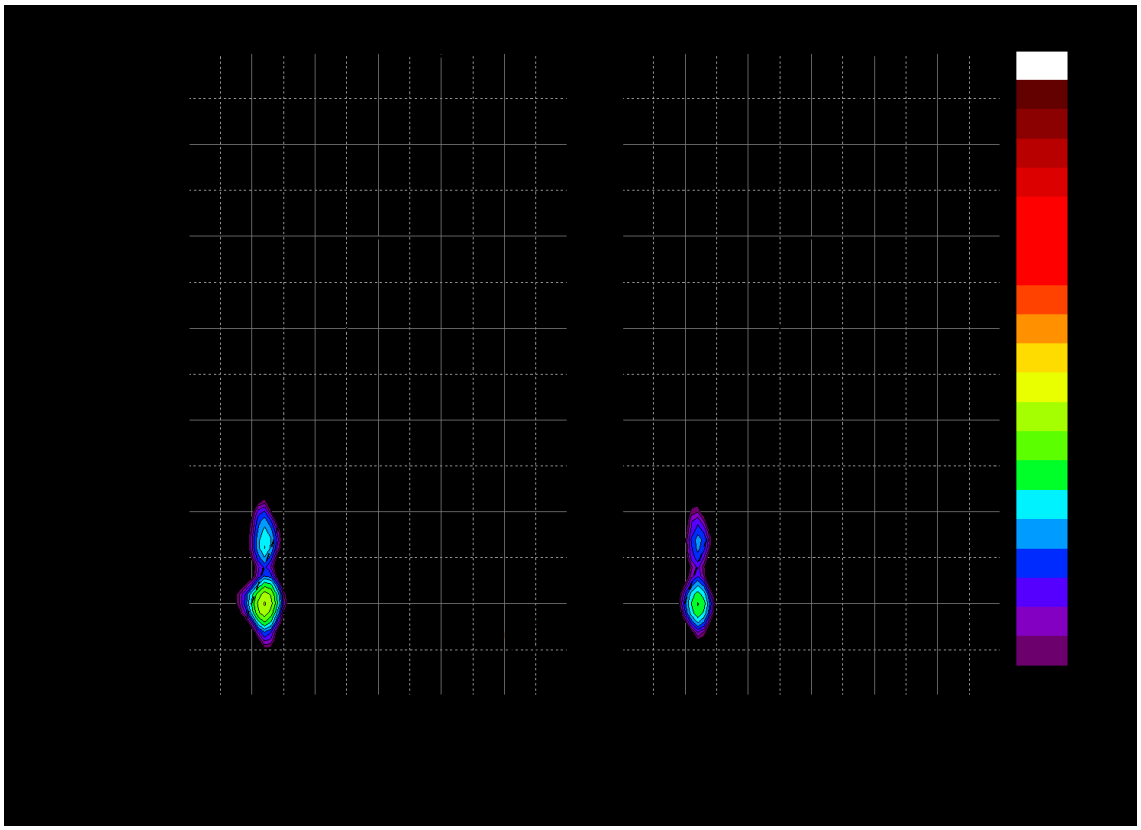


Figure 44: The difference spectra for a) non-iron-treated HPPD/Mesotrione and b) the corresponding BSA/MST spectrum, calculated by quadratically subtracting the protein spectrum from the protein+Mesotrione spectrum (each averaged from four repeats). Peaks that fall on the dashed line indicate overtone transitions. Presence of two cross-peaks in both difference spectra at similar intensities suggest interactions arising from non-specific binding.

The two cross-peaks are present in both difference spectra at similar intensities, which suggests that they arise from *non-specific* binding interactions. Essentially, they are interactions that occur in both HPPD and the control, hence are not likely to exclusively

occur upon binding to the target protein. Nonetheless, these peaks can still be assigned to vibrational modes occurring within either the protein or the herbicide molecule, which can be useful in assigning modes of *specifically bound* herbicides later in this section.

The presence of the aromatic nitro and sulfone functional groups aids in the assignments of coupled vibrational modes significantly. Both groups are known to produce strong absorptions at distinctive frequencies in 1D infrared spectroscopy[80]. The peaks are tentatively assigned in Table 5 (mode 2 is not assigned, as it is not specific to HPPD binding thus it is unclear whether it arises from a vibration common to both HPPD and BSA, or a second vibration occurring on the herbicide molecule).

ω_α (cm^{-1}) (mode 1)	**ω_β (cm^{-1})**	**ω_β-ω_α (cm^{-1}) (mode 2)**	**Intensity (3 sf)**	**Assignment (mode 1/mode 2)**
1360	2700	**1340**	a) 0.0611 b) 0.0401	NO symmetric stretch/unassigned
1360	2730	**1370**	a) 0.0273 b)0.02	NO symmetric stretch/unassigned

Table 5: Tentative assignments for the non-specific binding peaks present in the a) HPPD/Mesotrione and b) BSA/Mesotrione EVV 2DIR difference spectra.

The vibrational mode common to both peaks at 1360 cm^{-1} can be assigned to a symmetric NO stretch according to literature values (additionally considering that nitro frequencies are often downshifted when attached to an aromatic ring)[80]. Interestingly, neither NO asymmetric stretch nor either of the sulfone stretches (symmetric and asymmetric) are present in the spectra at a statistically significant level. This does not necessarily mean that coherences associated with these vibrational modes are not produced. Many coherences are formed and destroyed on the femtosecond (or shorter) timescale, thus the long dephasing times (of picoseconds) exhibited in EVV 2DIR may mask these. Additionally, vibrational modes for these moieties may still be coupled, but too weakly to see above the 3σ level. Alternatively, there may not be enough OPA1 energy in the region

of ≥1330 cm^{-1} typical of sulfone absorptions, because the energy decreases sharply at the edges of the spectral range (especially for OPA1).

The tentative assignment of a vibrational mode at 1360 cm^{-1} to a symmetric NO stretch in the protein/Mesotrione difference spectra in Figure 44 becomes useful later in Section 4.3 when assigning the potential binding peaks for herbicide candidate 329 over a larger spectral range, as they share the same functional groups. 329 is an HPPD project compound provided by Syngenta, and it also contains SO_2CH_3 and NO_2 functional groups, shown by the chemical structure in Figure 45.

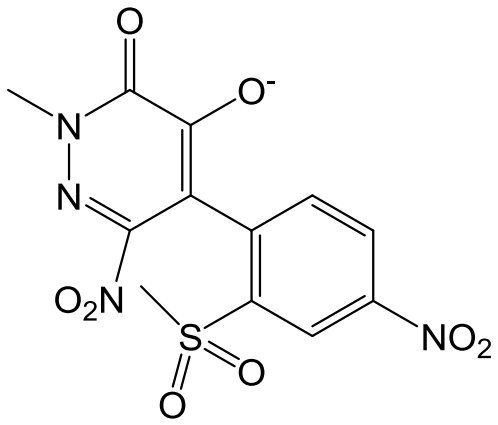


Figure 45: The chemical structure of HPPD herbicide candidate 329. Like Mesotrione, it contains NO_2 and SO_2CH_3 moieties.

The difference spectra for non-iron-treated HPPD and 329 (and the corresponding control difference spectrum for BSA) is shown in Figure 46.

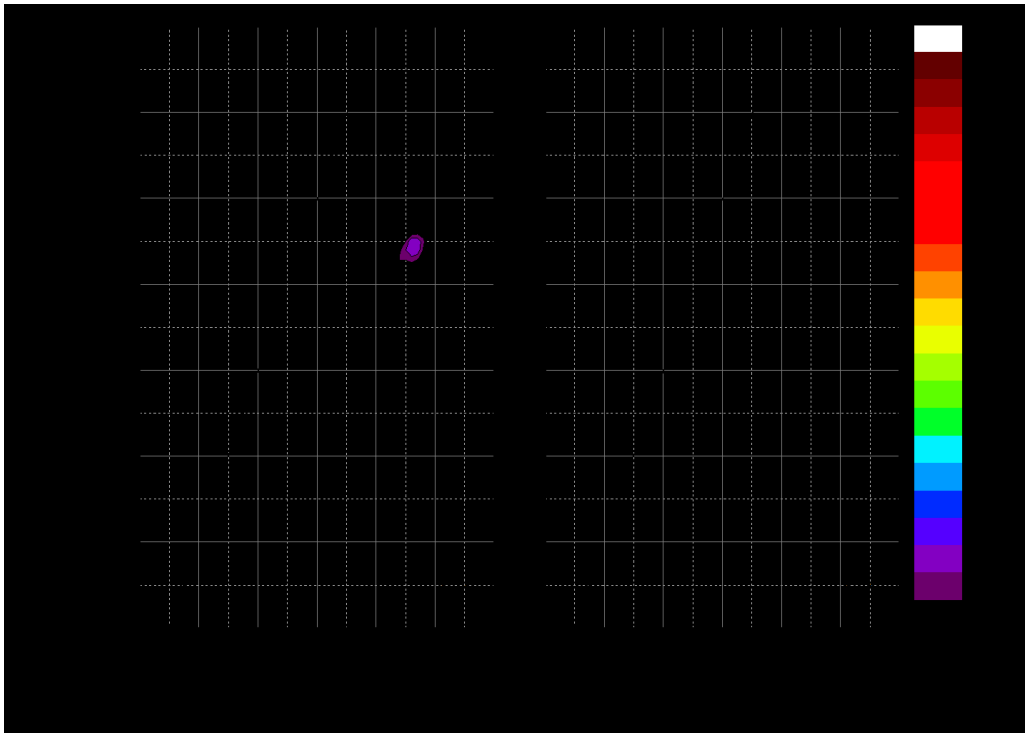


Figure 46: The averaged EVV 2DIR difference spectra for a) non-iron-treated HPPD and 329 and b) control protein BSA and 329 over the standard frequency range. The single peak present in spectrum a) indicates that the coupled vibrations giving rise to it are due to specific binding interactions.

There is a single statistically significant peak present at ω_α=1530/ω_β=2870 cm^{-1} over the 3σ level of noise in the HPPD/329 difference spectrum shown in Figure 46(a). The absence of the same peak in the BSA/329 difference spectrum exhibited in Figure 46(b) suggests that the peak arises from *binding-specific* interactions. The intensity and standard deviation (between the four difference spectra repeats) for the cross-peak and preliminary assignments for the vibrational modes giving rise to it are displayed in Table 6:

ω_α (cm^{-1}) (mode 1)	ω_β (cm^{-1})	ω_β-ω_α (cm^{-1}), (mode 2)	Intensity (3 sf)	Assignment (mode 1/mode 2)
1530	2870	**1340**	0.0103	NO asymmetric stretch/in-plane CH bend

Table 6: Preliminary assignments for the coupled vibrational modes giving rise to the binding peak in the HPPD/329 difference spectrum.

Consider again the structure of 329 (exhibited in Figure 45). There are two aromatic nitro-functional groups, known to have strong absorptions at ~1330 and ~1530 cm^{-1} (due to symmetric and asymmetric stretching vibrations, respectively)[80]. Additionally, theoretical EVV 2DIR spectra for 1,3-dinitrobenzene complexed with a tryptophan previously calculated has shown a cross-peak with a mode at 1530 cm^{-1} which was assigned to an asymmetric N=O stretch[142]. It therefore follows that the 1530 cm^{-1} vibrational mode contributing to the cross-peak on the ω_α axis can be tentatively assigned to an asymmetric N=O stretching motion. The fact that there are *two* nitrogen moieties per 329 molecule (which you might expect to give a larger EVV 2DIR intensity) further supports this assignment. This vibrational mode is coupled to another mode at 1340 cm^{-1}, which (if the peak is binding-specific, as proposed) must be attributable to an HPPD vibrational mode. The literature states that a vibration arising from a CH stretch from an amino acid side chain may occur at 1315-1350 cm^{-1}, therefore it seems reasonable to assign the mode at 1340 cm^{-1} to this vibration[143].

When assigning vibrational modes of protein amino acids more specifically than this in a protein-inhibitor complex, it is necessary to examine the interactions occurring between said protein and inhibitor in the binding site. While there is not currently a published crystal structure of the HPPD+329 complex, crystal structures are available for HPPD co-crystallised with very similar analogues, which Syngenta has computationally modified to create a model structure of HPPD complexed with 329. The binding site of this altered model can be used to assess potential HPPD-329 interactions by looking at which HPPD amino acid residues are near the 329 NO_2 groups. Unfortunately, this model is unable to be included here because it is commercially sensitive.

A possible assignment for the 1340 cm^{-1} mode is a histidine in-plane CH bend, which fits with the literature assignment and is structurally plausible as there are two conserved

histidine residues in the HPPD binding site that bind the Fe(II) ion and enable it to bind the iron, as shown below in Figure 47 (note that the image depicting herbicide binding has been taken from a docking model in reference [144]):

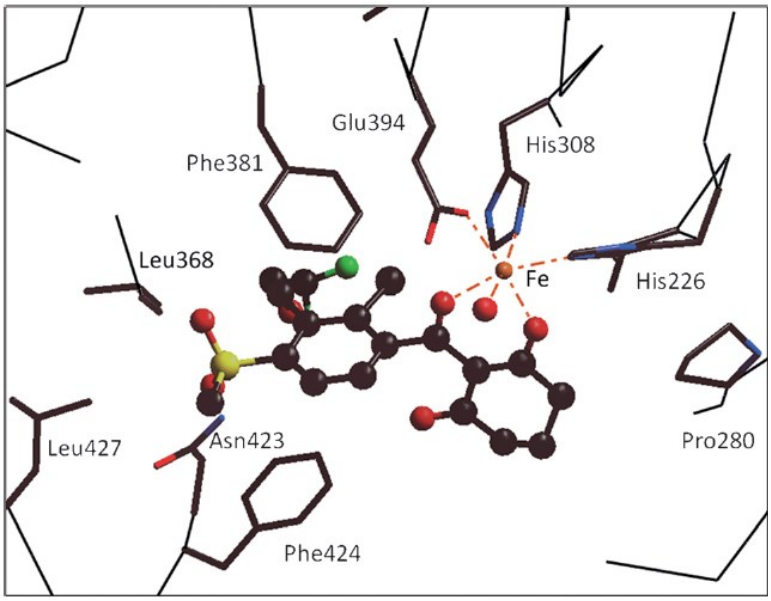


Figure 47: Image of At HPPD binding site docking model showing binding of a the commercially available herbicide Tembotrione to HPPD (taken from reference 144). Note the two conserved histidine residues which bind the Fe(II) ion.

Unfortunately, mutating the highly conserved histidine residues would not be possible without destabilising the HPPD protein, thus experimental verification for this assignment using mutants is not possible. Instead, quantum calculations of the vibrational frequencies of 329 and the surrounding protein residues in the HPPD binding site (using DFT methods in Gaussian 09[115], in accordance with Kwak et al[104] and Sowley[49], outlined in Section 2.3.3) are required to verify this potential assignment.

The structural model for the HPPD+329 complex shows various HPPD amino acids in the binding site that are in close proximity to 329 during binding; namely Q272, Q286, F360, F371, F403. The latter three phenylalanine residues have been proposed to be important for π-π stacking interactions upon herbicides binding to HPPD, so it would make sense if these were coupling with the nitro- moieties. However, literature assignment of amino acids at 1340 cm^{-1} suggests proline or alanine vibrations[145], neither of which are among

the residues suggested. As discussed in Section 2.1.3, the anharmonicity that governs intermolecular interactions in EVV 2DIR can result in coupled vibrational modes taking on characteristics of each other, which may cause frequency shifts. This, along with the lack of known EVV 2DIR assignments for protein-inhibitor couplings due to the novel nature of this work, makes it challenging to know how vibrational coupling in EVV 2DIR experiments might affect the 1D IR frequency shifts for amino acids compared to those specified in the literature.

It is difficult to definitively assign this in-plane CH bending vibrational mode to a specific amino acid residue side-chain without theoretically calculating the EVV 2DIR spectra for the HPPD-329 complex. One way to experimentally confirm or rule out possible amino acid residues is to express mutant HPPD, with each potential residue mutated in turn, and to then perform EVV 2DIR experiments on the mutant HPPD-329 complexes to see if the binding peak disappears or shifts in frequency. Based on the assignment of the binding peak to an asymmetric NO stretch and the observation of the 5 amino acid residues proximal to the 329 NO_2 groups in the active site of the HPPD-329 structural model, attempts were made to mutate the 5 residues listed above to alanine in 5 different HPPD variants. These attempts were unfortunately unsuccessful, and time constraints meant further attempts were not possible.

Again, it is interesting that this time interactions arising from the sulfone group (which were seen in the Mesotrione difference spectra) are absent in both 329 difference spectra in Figure 46, as typically SO stretching vibrations give strong IR absorptions at around 1350 cm^{-1}. This could possibly be due to the decreased OPA energy at the ω_α=1350 cm^{-1} region, as the energy is lowest at the lower end of the spectral frequency range (see Figure 35 in Section 3.1.9.3). Alternatively, it may be because the vibrations for the SO_2CH_3 group experience weaker coupling (or no coupling at all) with HPPD than the NO_2 group does upon binding.

The EVV 2DIR spectra of orphan drug Nitisinone (which originally started out as an HPPD candidate[51]) complexed with HPPD was also collected. Like Mesotrione and 329, the chemical structure of Nitisinone (shown in Figure 48) contains a nitrogen functional group, and additionally an electron withdrawing CF_3 moiety.

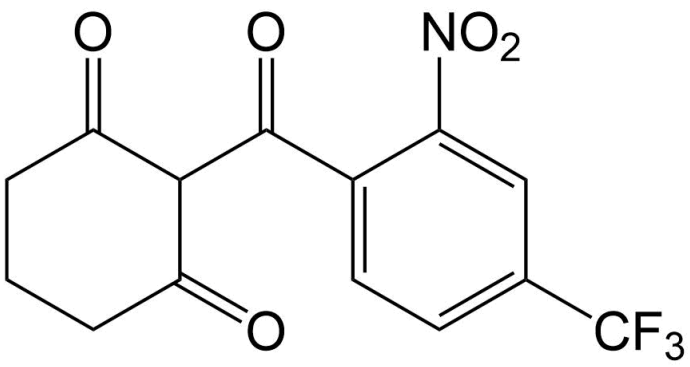


Figure 48: The chemical structure of the HPPD inhibitor Nitisinone. Of particular significance are the aromatic nitrogen and CF_3 groups (the former in common with Mesotrione and 329, and the latter in contrast).

The difference spectra for non-iron treated HPPD and Nitisinone (and that of the corresponding control BSA) is shown in Figure 49:

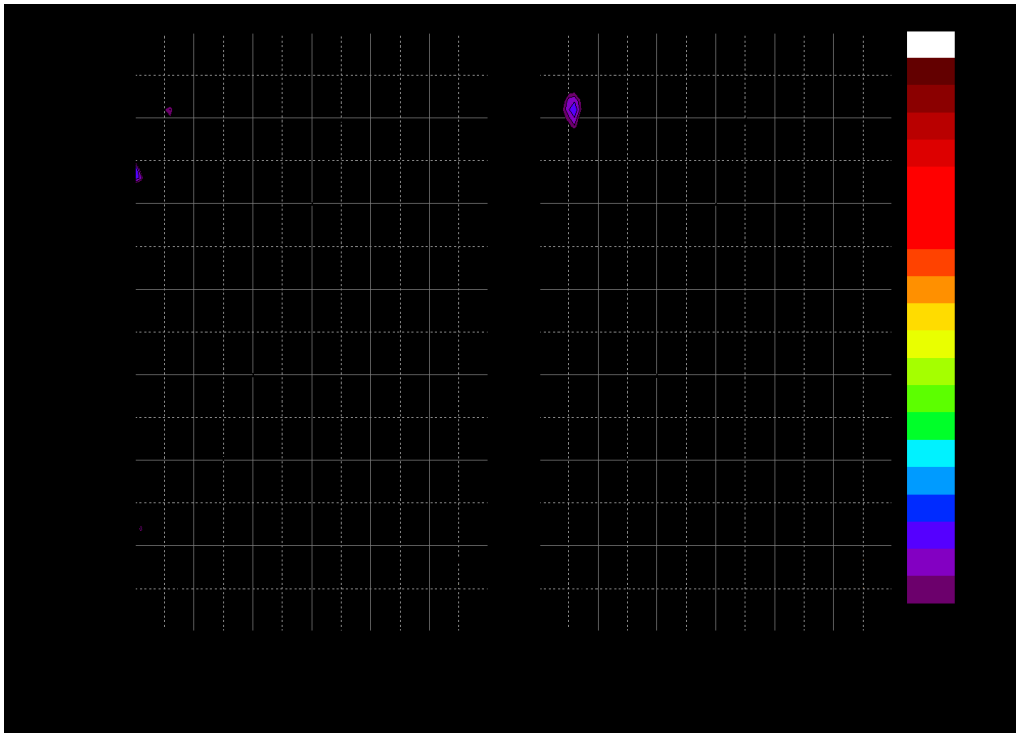


Figure 49: Averaged EVV 2DIR difference spectra for a) non-iron-treated HPPD and Nitisinone, and b) BSA with Nitisinone over the standard frequency range. The peak at ω_α=1330/ω_β=2955 cm^{-1} is present in both spectra, suggesting it is a peak arising from the coupling of vibrations associated with non-specific binding.

Because the same cross-peak appears in both the HPPD and BSA spectra in Figure 49, we can attribute this to a non-specific binding interaction that occurs either between Nitisinone and both proteins (or Nitisinone interacting with itself). The CF_3 moiety aids in assignment as a CF stretch is known to give a strong IR absorption at 1330 cm^{-1}. Tentative assignment is given below in Table 7 (mode 2 is not assigned, as it is not specific to HPPD binding and therefore we cannot know whether it arises from a non-specific protein vibration or a second vibration occurring in the herbicide molecule).

ω_α (cm^{-1}) (mode 1)	**ω_β (cm^{-1})**	**ω_β-ω_α (cm^{-1}), (mode 2)**	**Intensity (3 sf)**	**Assignment (mode 1/mode 2)**
1330	2955	**1625**	a) 0.00708 b) 0.0129	CF asymmetric stretch/Unassigned

Table 7: Tentative assignments for the non-specific interactions giving rise to the cross peak in the Nitisinone/protein difference spectra.

Interestingly, despite having the same aromatic NO_2 group as the 329 inhibitor, the difference spectra for Nitisinone does not show the same peak at 1530/2870 above 3σ arising from an asymmetric NO stretch coupled to the HPPD binding site. There are a number of reasons that could explain this; 329 has two NO_2 groups, thus could experience stronger coupling involving this mode. If there is a weaker interaction involving the NO_2 group in Nitisinone, there may not be enough energy in this spectral region to detect it above 3σ. Alternatively, the binding of 329 to HPPD may involve stronger coupling of its NO_2 groups than in the binding of Nitisinone, or the peak may be specific to 329 binding.

4.2.2 Comparison of iron-treated vs non-iron-treated HPPD binding

EVV 2DIR experiments were also performed on samples containing iron-treated HPPD added to Nitisinone and 329, to see if there was any difference in binding detection. Figure

50 shows a comparison of the difference spectra for both iron-treated and non-iron-treated HPPD and Nitisinone.

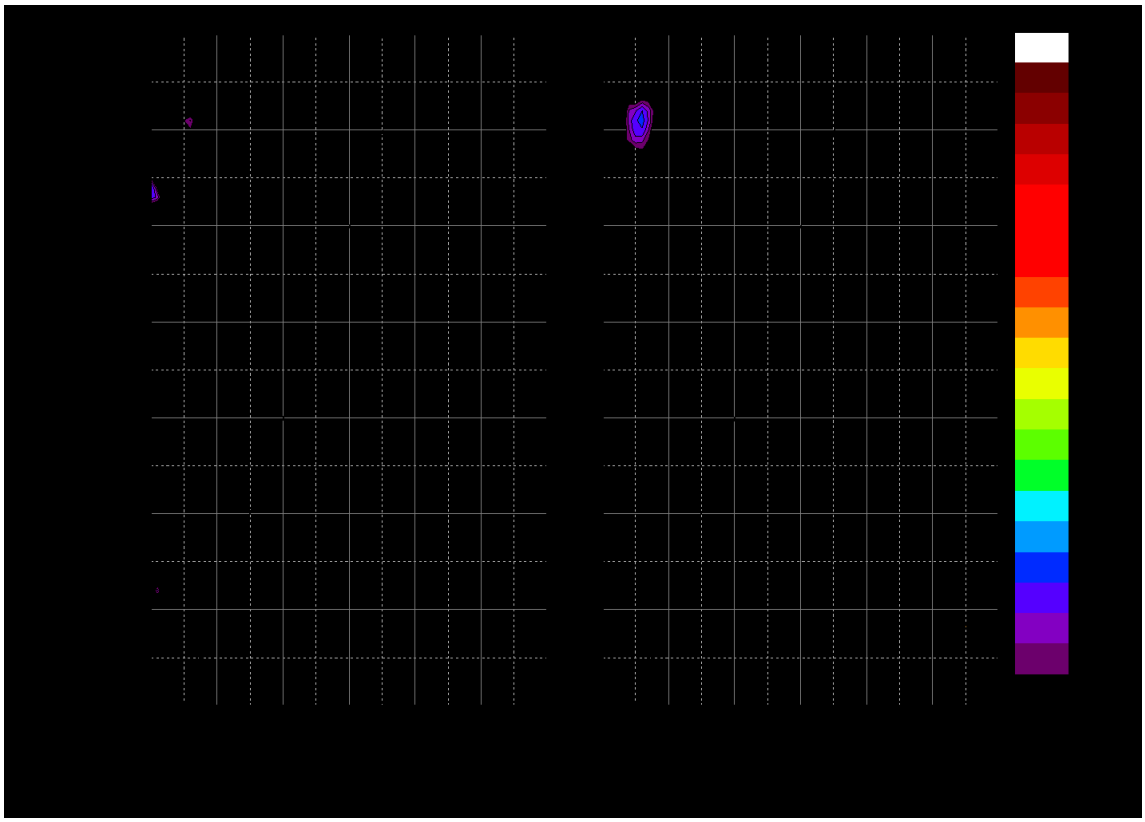


Figure 50: The averaged EVV 2DIR difference spectra for HPPD/Nitisinone showing a comparison between a) non-iron-treated HPPD and b) iron-treated HPPD.

As shown in Figure 50, the peak at ω_α=1330/ω_β=2955 cm^{-1} is broader and is just 1.6 times more intense for the iron-treated difference spectrum (b). However, as we have previously determined (from looking at Figure 49) that this peak is due to vibrational coupling arising from non-specific binding interactions, we would not necessarily expect the treatment of HPPD with iron would result in more intense binding. A more useful comparison is the difference spectra for 329 with non-iron-treated HPPD and with iron-treated HPPD, which is exhibited in Figure 51:

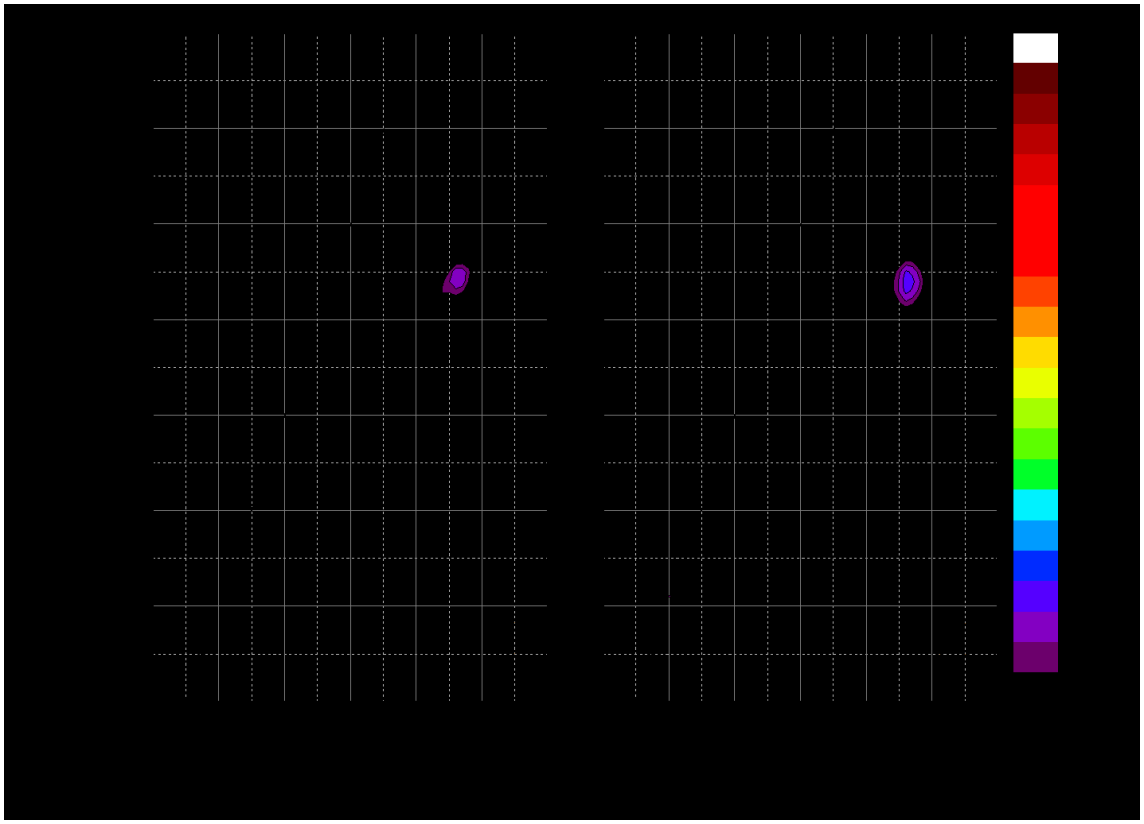


Figure 51: The averaged EVV 2DIR difference spectra of a) non-iron-treated HPPD/329 and b) iron-treated HPPD/329.

Again, the intensity of the cross-peak at ω_α=1350/ω_β=2870 cm^{-1} is very similar for both non-iron treated and iron-treated HPPD (0.0103 and 0.0108 to 3 sf, respectively). The results of the HPPD iron-treatment comparison are surprising, as HPPD is an iron-dependent dioxygenase and therefore iron-treatment would be expected to enhance inhibitor binding. However, this is not the case for the EVV 2DIR experiments presented in this thesis.

4.3 Lower frequency range experiments

So far, all EVV 2DIR data presented in this thesis has been collected over the spectral range ω_α = 1250-1750 and ω_β = 2600-3450 cm^{-1} (without considering any OPA calibration offsets). The primary reason for this was because this was the range that protein-inhibitor binding experiments had previously been performed over, so the OPAs were already set up for these values. This section presents the first EVV 2DIR data collected over a spectral

range of ω_α = 1250-1750 and ω_β = 2100-3100 cm^{-1} (again, before accounting for any OPA calibration offsets). A limiting factor in these experiments is the frequency range selected for by optimising each OPA around a central frequency. Additionally, the OPA energy profiles decrease at the edges of each frequency range. Therefore, additional EVV 2DIR experiments were carried out over a lower frequency range for the two most promising HPPD candidates; Nitisinone and 329. The previous experiments in the standard frequency range showed a potential binding peak at ω_α=1530/ω_β=2870 cm^{-1}. Previous work shows many calculated amino acid spectral features/cross-peaks in this new, lower frequency range, and that this region has many decongested areas (as demonstrated in Figure 52).

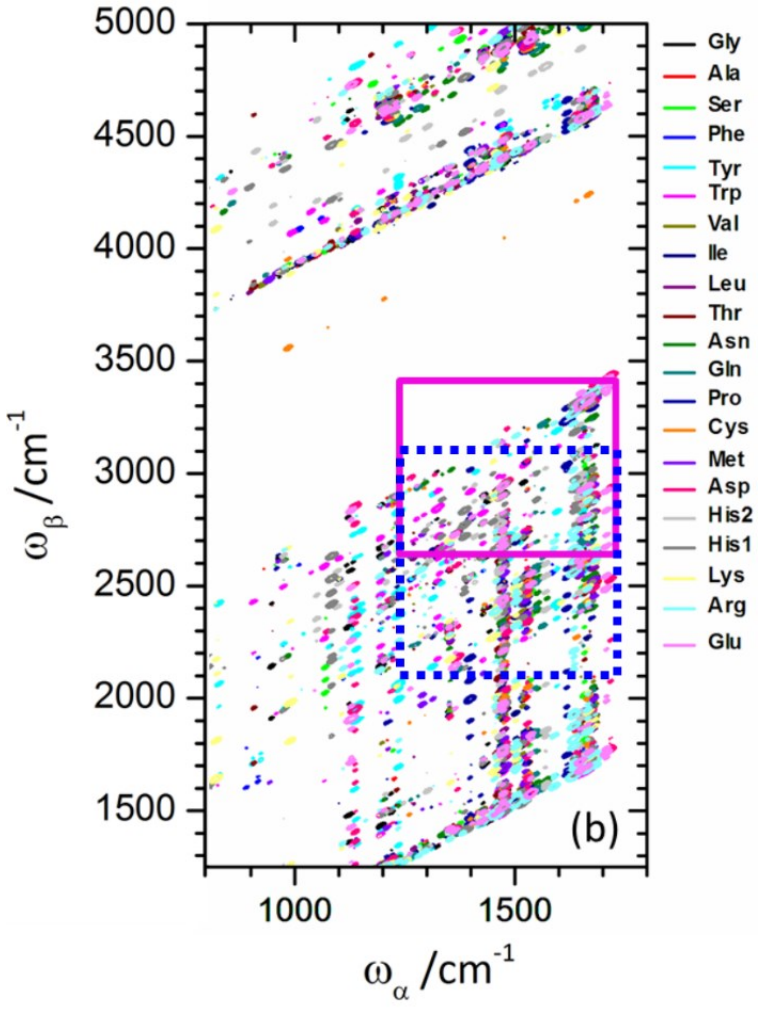


Figure 52: Calculated EVV 2DIR spectra of all amino acid side chains. The pink box shows the experiments performed over the previous frequency range, while the dotted blue box shows the new frequency range.

4.3.1 Sample preparation

Humidity cells were prepared for the lower frequency range experiments as for the previous frequency range, however only iron-treated HPPD was used in the HPPD+herbicide samples, giving six sample spots comprised of the following:

- 1 mM At HPPD, 5 mM phosphate buffer, 5 mM NaCl, 2% DMSO (pH = 8.0)
- 1 mM At HPPD + 1 mM 329, 5 mM phosphate buffer, 5 mM NaCl, 2% DMSO (pH = 8.0)
- 1 mM At HPPD + 1 mM Nitisinone, 5 mM phosphate buffer, 5 mM NaCl, 2% DMSO (pH = 8.0)
- 1 mM BSA, 5 mM phosphate buffer, 5 mM NaCl, 2% DMSO (pH = 8.0)
- 1 mM BSA + 1 mM 329, 5 mM phosphate buffer, 5 mM NaCl, 2% DMSO (pH = 8.0)
- 1 mM BSA + 1 mM Nitisinone, 5 mM phosphate buffer, 5 mM NaCl, 2% DMSO (pH = 8.0)

A relative humidity of 75.5% was again used to equilibrate the spots, by the presence of saturated NaCl spots in the air-tight sample cell.

4.3.2 EVV 2DIR data collection

As for the previous frequency range, the data was collected with a resolution of 5 cm^{-1} per pixel and an average of 100 samples per pixel. The same polarisation was applied to all three beams (PPP). Optimisation of relative pulse timings for maximum spectral resolution and minimum non-resonant background determined time delays of $T_{\alpha\beta}$ = 1.75 ps and $T_{\beta\gamma}$ = 1 ps should be used for the experiments. EVV 2DIR spectra of each sample spot in turn was collected to obtain four repeats of EVV 2DIR spectra for each sample.

4.3.3 Data processing

EVV 2DIR spectra of the HPPD system were collected over two separate frequency ranges. In addition to the standard processing steps laid out in Section 3.1.9, the HPPD data collected over the lower frequency range requires additional treatment due to the appearance of a pair of non-resonant horizontal "stripes" in every spectrum. Despite several attempts to remove this artefact in data collection by careful selection of filters in the PM filter stack, and optimisation of the OPAs, in the interests of time it was decided to remove it in post-collection processing.

To account for the possibility that the stripe was sample-dependent, the ω_α beam was shuttered and the first five columns of pixels preceding the start of the ω_α frequency range (i.e. ω_α =1225-1250/ ω_β=2100-3100 cm^{-1}) were measured. These were then averaged and removed from the corresponding spectrum in processing, using the MATLAB code in Appendix A. Several different methods were tested to remove the stripe in processing, and while none perfectly removed the stripe, the technique described above proved the best. A comparison of representative EVV 2DIR spectra before and after stripe removal in this manner can be seen in Figure 53:

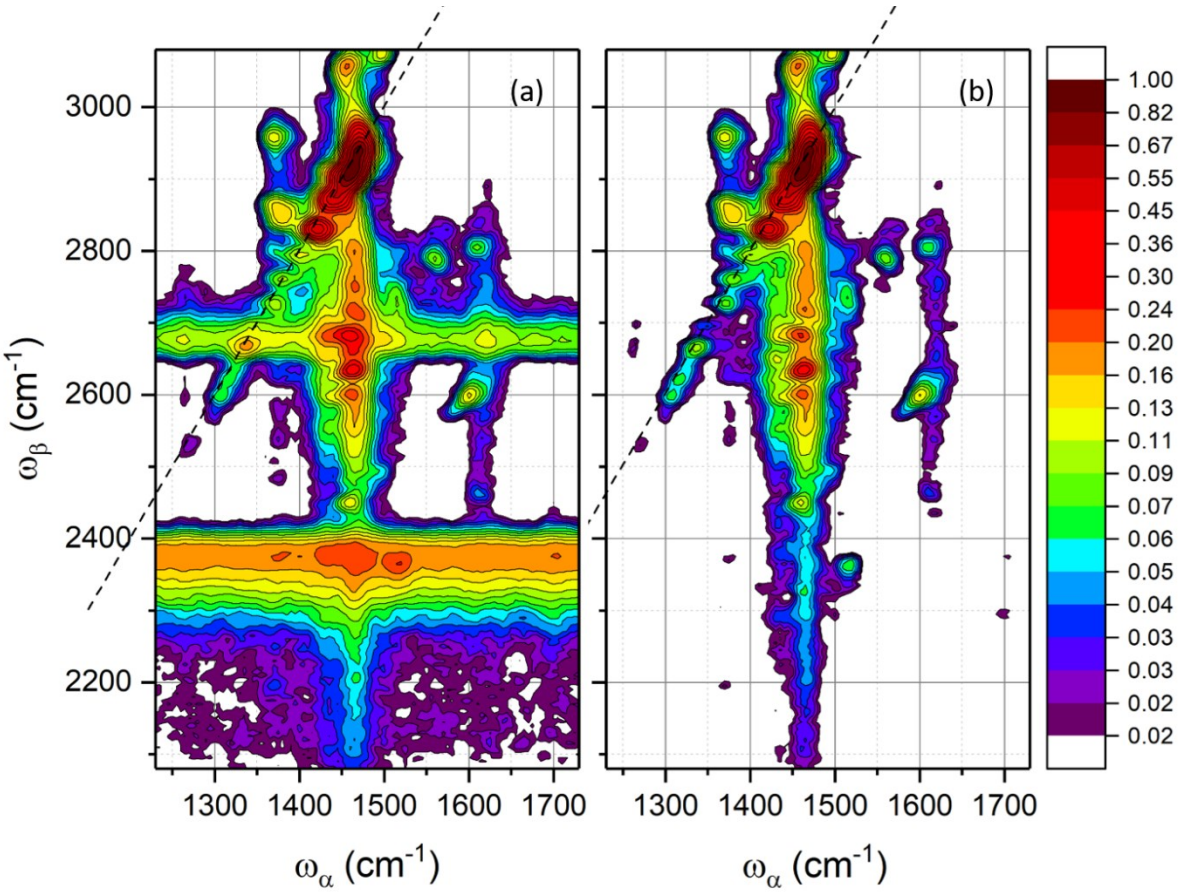


Figure 53: Averaged EVV 2DIR spectra of iron-treated HPPD collected over the lower ω_β frequency range, shown a) before and b) after removal of the two horizontal "stripe" spectral artefacts centred at ω_β=2400 & 2700 cm^{-1}.

This technique is appropriate for the purposes of the following section, as most of the key cross-peaks discussed in the following section are not in the "stripe" regions.

4.3.4 EVV 2DIR spectra of HPPD & BSA

The EVV 2DIR spectra of iron-treated HPPD and BSA are exhibited in Figure 54. Peaks shown are above the 3σ level, defined by the standard deviation in the signal-free region of the spectra. Both graphs show new spectral features compared with the previous spectral range. For these experiments, OPA2 has been optimised for ω_β=2550 cm^{-1} thus there is more energy in this region, resulting in an increased intensity of the cross-peaks due to overtone transitions seen at ω_α=1250-1400/ω_β=2470-2800 cm^{-1}.

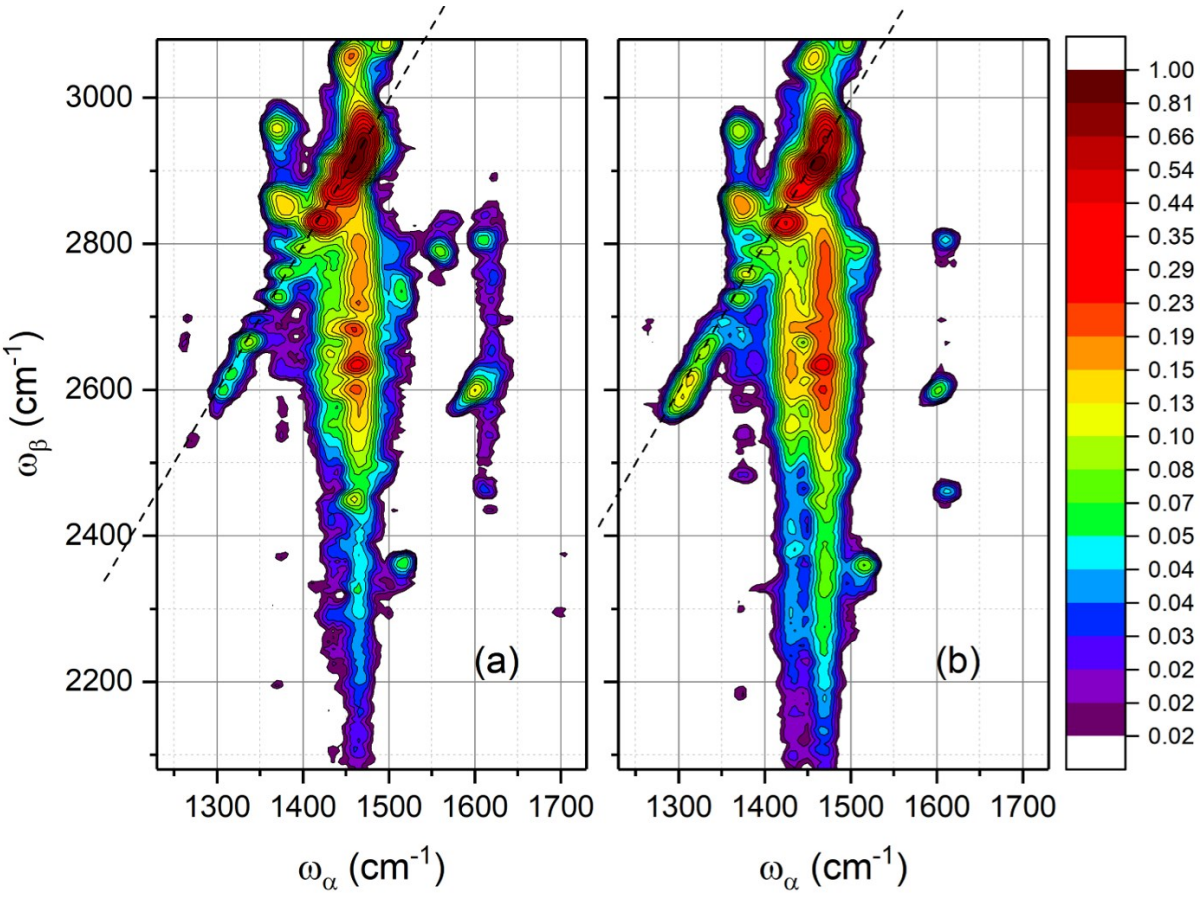


Figure 54: The averaged EVV 2DIR spectra of a) iron-treated HPPD and b) BSA collected over the lower frequency range.

4.3.5 HPPD+herbicide and difference spectra

As for the previous frequency range, EVV 2DIR spectra were collected with both HPPD and BSA complexed to each herbicide candidate. Difference spectra were then calculated by a quadratic subtraction of the protein only spectrum from the protein+herbicide spectrum (as described in Section 4.2.1.6), in order to assess whether any new peaks present arose from interactions specific to HPPD-herbicide binding.

The EVV 2DIR spectra of HPPD+329 and BSA+329 is exhibited in Figure 55. Circled are new peaks appearing upon the addition of 329 to HPPD and BSA, with pink circled peaks suggesting specific binding interactions and the black circled peaks (present in both) indicating non-specific binding.

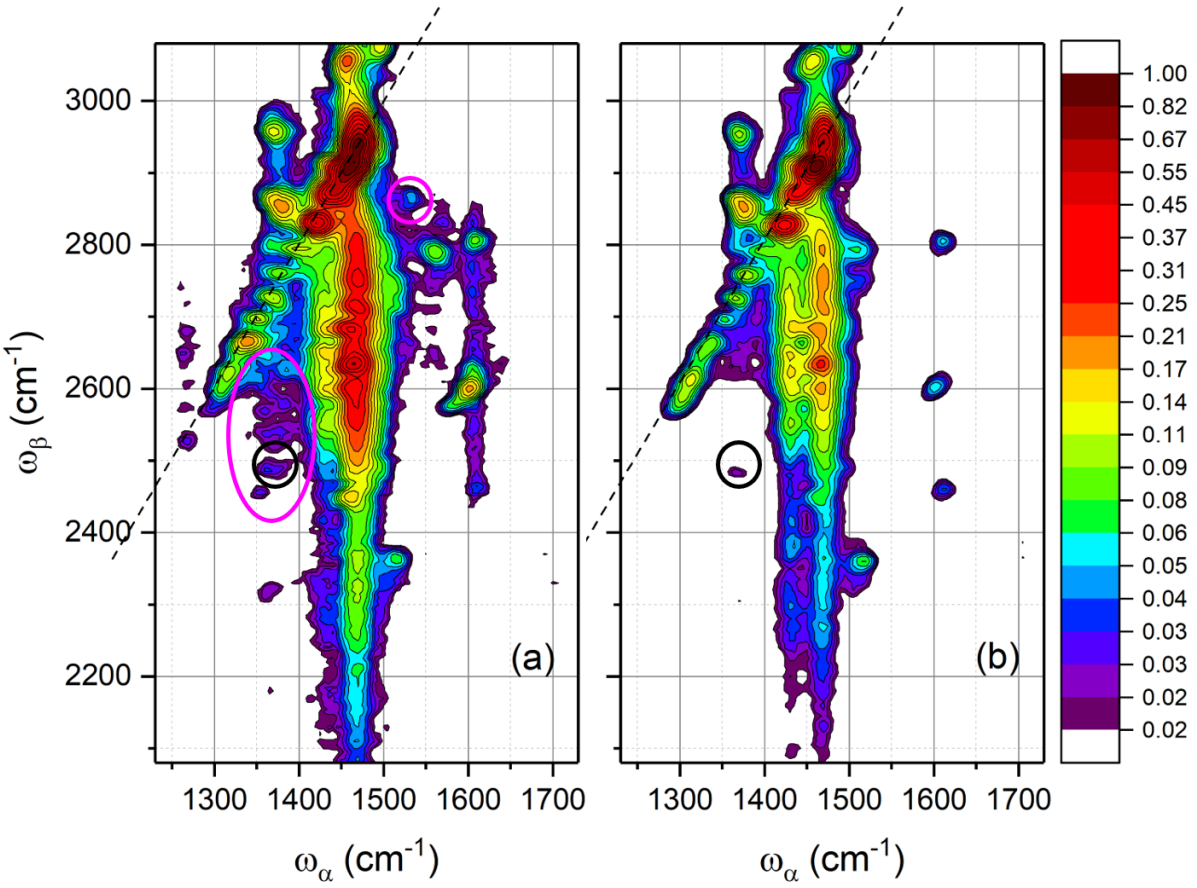


Figure 55: The averaged EVV 2DIR spectra for a) HPPD+329, and b) BSA+329 over the lower frequency range. Peaks are shown above the 2σ level. The peaks circled in pink are present in the HPPD+329 spectrum and absent in the BSA+329 and HPPD only spectra, suggesting they may arise from specific binding interactions, while the peak circled in black is present in both indicating coupling not specific to HPPD-binding.

To determine whether any of these new peaks were due to binding-specific vibrational couplings, the EVV 2DIR difference spectra of HPPD/329 and BSA/329 were calculated and are displayed in Figure 56 below. Note that all difference spectra in this section have been trimmed to show only the regions that displayed cross-peaks, for clarity.

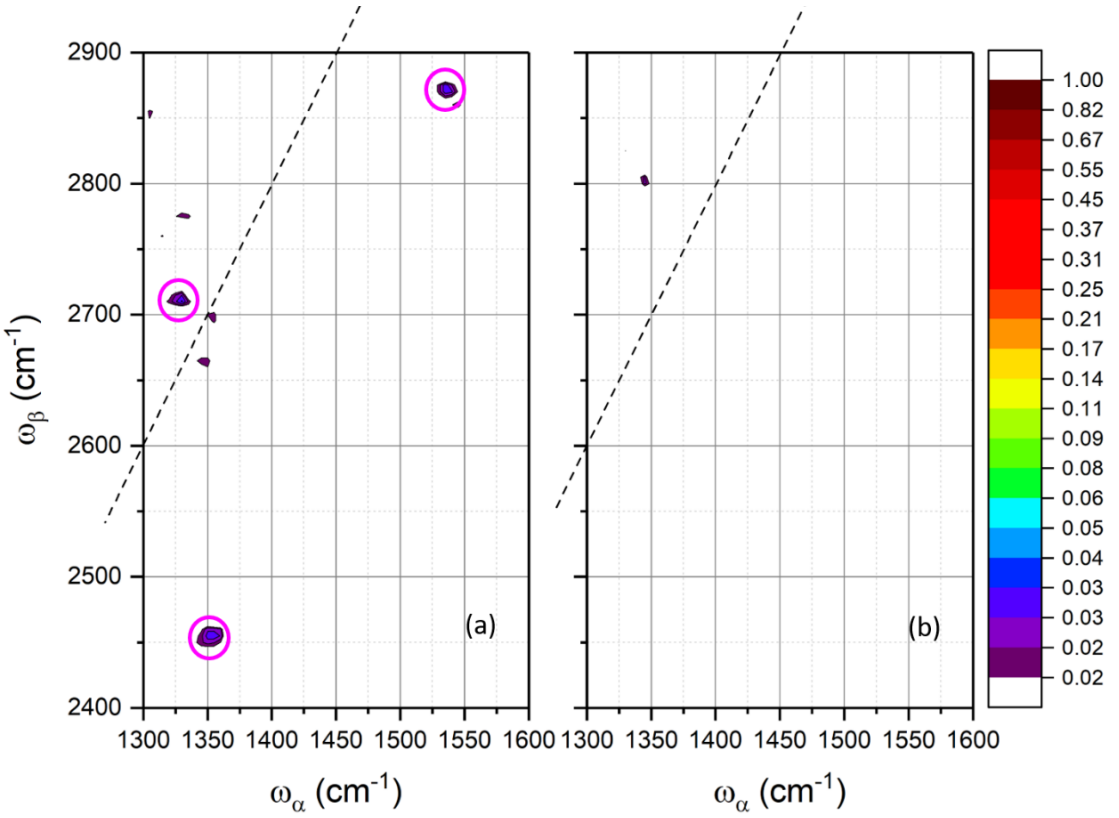


Figure 56: The averaged EVV 2DIR difference spectra for a) HPPD/329 and b) BSA/329 over the lower frequency range. Peaks are shown above the 2σ level.

It is important to note that the peaks are only visible over the 2σ level, as they are not seen at or above the 3σ level. The value of 3σ is ~5 times lower for the data obtained in the previous frequency range than the data presented in this section collected over the lower ω_β frequency range (0.00645 c.f. 0.0255 to 3 sf, respectively using HPPD+329 data as an example). This reflects a higher level of inherent noise in the laser system than the system, which may be associated with the laser system being less stable over the lower frequency range. Therefore, the experiments used to collect the data presented in this section should be optimised and repeated to improve laser system noise.

Nonetheless, the three peaks clearly appear in each of the four repeats used to make up the averaged difference spectrum above the 2σ level, which means that we can be 95% confident that they are real peaks.

The appearance of the three circled peaks present in the HPPD/329 difference spectrum appear and absence in the BSA/329 difference spectrum indicates that they arise from

binding-specific vibrational couplings. Tentative assignments for these three peaks are given in Table 8 below:

ω_α (cm^{-1}) (mode 1)	ω_β (cm^{-1})	ω_β-ω_α (cm^{-1}), (mode 2)	Intensity (3 sf)	Assignment (mode 1/mode 2)
1535	1340	**2875**	0.0292	NO asymmetric stretch/in-plane CH bend
1330	1380	**2710**	0.03221	SO asymmetric stretch/ unassigned
1355	1100	**2455**	0.0287	NO symmetric stretch/Histidine CN stretch, in-plane CH bend

Table 8: Tentative assignments for vibrational modes giving rise to coupled cross-peaks indicative of HPPD-specific binding in the EVV 2DIR HPPD/329 difference spectrum.

The cross-peak appearing at ω_α=1535/ω_β=2875 cm^{-1} is the same as that seen in the HPPD/329 difference spectra over the previous frequency range (within the spectral resolution of ±5 cm^{-1}). Therefore, it is easily assigned to the NO asymmetric stretch and in-plane CH bend vibrational modes occurring at 1535 and 1340 cm^{-1}, respectively. The fact that the peak is reproducible between the two sets of experiments collected at different times over different frequency regimes further supports the suggestion that it is a binding-specific peak.

The appearance of the cross-peak at ω_α=1355/ω_β=2455 cm^{-1} indicates coupling between two vibrational modes at 1355 and 1100 cm^{-1}. N-O stretched typically occur at ~1550 and ~1365 cm^{-1} (asymmetric and symmetric, respectively), however these are downshifted if the nitro group is attached to an aromatic ring[80], as in the structure of 329. Additionally, the difference spectra of Mesotrione showed two cross-peaks along a column arising from a common vibrational mode at 1360 cm^{-1}, which was tentatively assigned to a symmetric N-O stretch. Therefore, the HPPD/329 vibrational mode at 1355 cm^{-1} can be tentatively assigned to a symmetric N-O stretch.

In assigning the vibrational mode at 1100 cm^{-1} that couples with the 1355 cm^{-1} mode upon binding, it is useful to consider the information we gained from the HPPD/329 difference spectra over the previous frequency range. In Section 4.2.1.6, we postulated that the vibrational mode occurring at 1340 cm^{-1} giving rise to the cross-peak at ω_α=1530/ω_β=2870 cm^{-1} may be due to the nitro moiety on the 329 molecule interacting with a CH group on an amino acid side-chain. Several amino acids were suggested, particularly the histidine residue in proximity with the nitrogen group that is known to have a key role in binding. The literature for the infrared absorptions of amino acids indicates that histidine undergoes a motion involving a C-N stretch and in-plane C-H bend[143]. The fact that two cross-peaks for the HPPD/329 difference spectrum fit with observed frequency values for both aromatic NO_2 and histidine side-chains makes a compelling case for the binding interactions occurring between HPPD and 329. As discussed in Section 4.2.1.6, mutating the conserved histidine residues will destabilise the protein and may affect its activity, therefore unfortunately experimental verification of this assignment using mutated HPPD is not possible. Instead, quantum calculations of the coupled vibrational modes between 329 and the surrounding protein residues in the HPPD binding site (using DFT methods in Gaussian 09[115], in accordance with Kwak et al[104] and Sowley[49], outlined in Section 2.3.3) are necessary to simulate the EVV 2DIR spectra required to confirm this preliminary assignment.

Finally, the appearance of a new cross-peak at ω_α=1330/ω_β=2710 cm^{-1} arises from the coupling of two vibrations occurring at 1330 and 1380 cm^{-1}, respectively. The first can be assigned to an asymmetric SO stretch arising from the sulfone group, by consulting the literature values for SO stretch frequencies (expected to be downshifted by the attachment to an aromatic ring). If this cross-peak arises from binding-specific interactions as suggested, the vibrational mode at 1380 cm^{-1} which couples to the asymmetric SO stretch must arise from an amino acid vibration. Based on frequency alone, the literature suggests an in-plane CH_3 bend[143], however there are no amino acids near the SO_2CH_3 group that contain a CH_3 side-chain. Again, to assign this mode to a specific amino acid residue requires examination of the HPPD+329 structural model binding site, which shows a polar contact between an NH_2 group on the residue Gln293 and the one of the S=O moieties. Mutating the Gln293 residue to alanine and collecting EVV 2DIR data on the mutant HPPD complexed with 329 may enable assignment of this mode to a specific residue. Alternatively, calculated EVV 2DIR spectra could be used to confirm these preliminary assignments.

EVV 2DIR spectra for HPPD+Nitisinone and BSA+Nitisinone were also collected over this frequency range, and are shown in Figure 57:

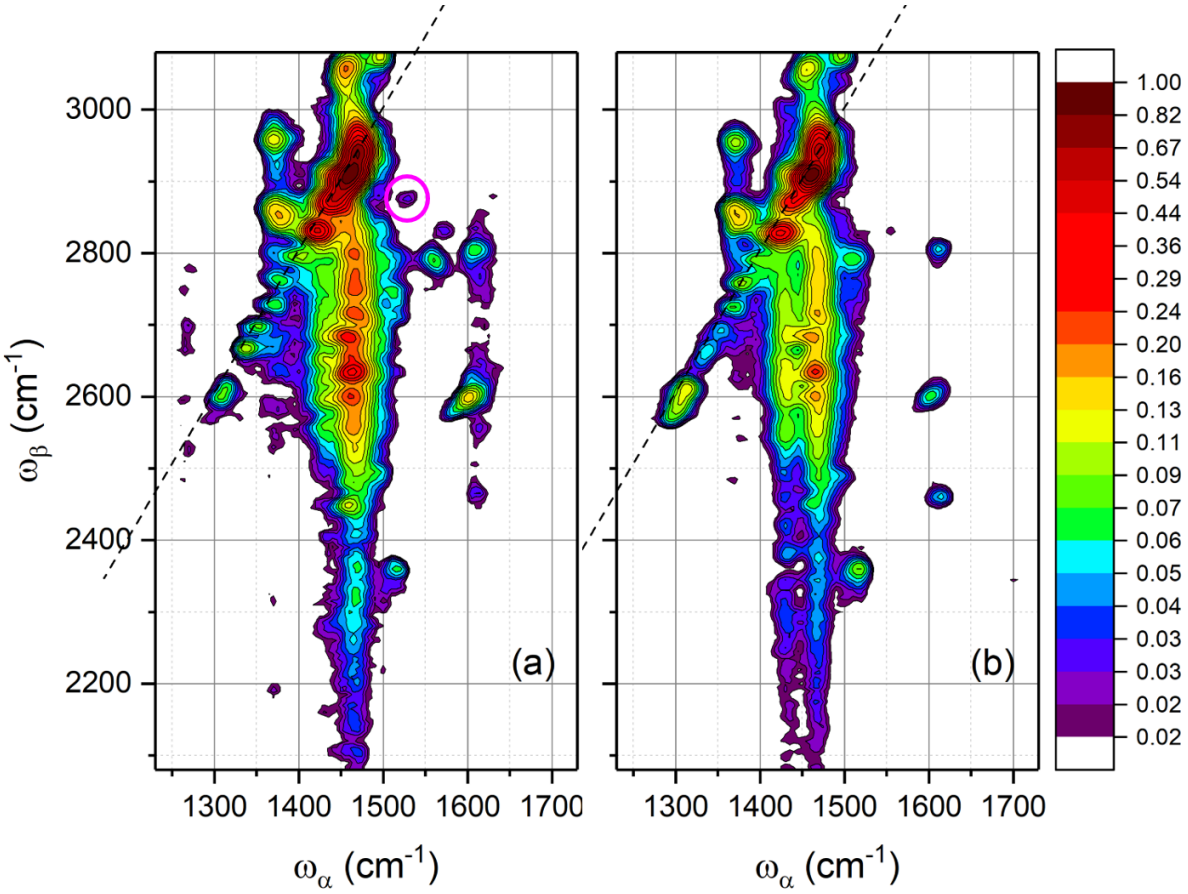


Figure 57: The averaged EVV 2DIR spectra for a) HPPD+Nitisinone and b) BSA+Nitisinone over the lower frequency range. Peaks are shown above the 2σ level. The circled peak present in (a) is absent in (b) and also absent in the HPPD only spectrum, thus is a potential binding peak.

The circled peak present in the HPPD+Nitisinone spectrum is absent in both the BSA+Nitisinone and HPPD only spectra, suggesting that it may arise from binding-specific interactions. Additionally, this peak occurs at ω_α=1535/ω_β=2875 cm^{-1} which is the same peak previously identified as a binding-specific peak for 329, which further suggests that it may be a potential binding peak.

The EVV 2DIR difference spectrum for HPPD/BSA and Nitisinone is shown in Figure 58:

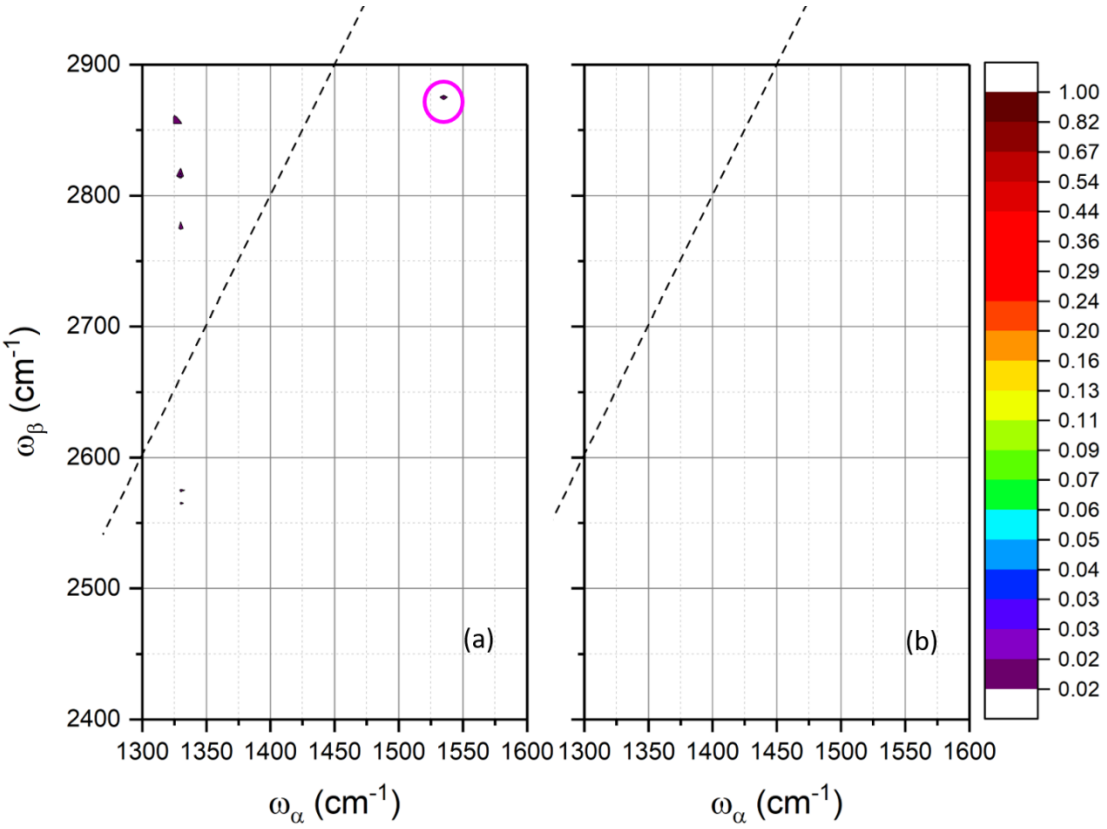


Figure 58: The EVV 2DIR difference spectra for a) HPPD/Nitisinone and b) BSA/Nitisinone over the lower frequency range. Only one peak (circled) is consistently present in each difference spectrum repeat for HPPD/Nitisinone and is absent in the BSA/Nitisinone difference spectrum, suggesting it is due to specific binding interactions.

There is just one peak at ω_α=1535/ω_β=2875 cm^{-1} that is present above the 2σ level in each of the four repeats of the HPPD/Nitisinone difference spectra, which is circled in Figure 58. It is important to note that the other peaks visible in the ω_α=~1330 cm^{-1} region of the HPPD/Nitisinone difference spectrum are not consistently seen above 2σ in every repeat, therefore we cannot be 95% confident that they are real, as with the circled peak, thus they are not included as potential binding peaks in this discussion.

The cross-peak at ω_α=1535/ω_β=2875 cm^{-1} has an intensity of 0.0194 (3 sf) and appears to be the same peak as the one previously assigned for the HPPD/329 difference spectra. Consequently, it is easily assignable to the same coupled vibrational modes, arising from an asymmetric NO stretch at 1535 cm^{-1} and an in-plane CH bend at 1340 cm^{-1}, respectively.

It is interesting that this cross-peak appears in the HPPD/Nitisinone difference spectra for the lower frequency range, because it did not appear in the difference spectrum for the same sample over the previous frequency range. Additionally, the non-specific binding peak seen in the previous frequency range at ω_α=1330/ω_β=2955 cm^{-1} is not reproduced in the lower frequency range difference spectra. One possible explanation for this could be that there is more OPA energy at ω_α=1535/ω_β=2875 cm^{-1} in the lower frequency range, however this is unlikely as this region is fairly central in both frequency ranges, and if this was the case then one would expect to still see the ω_α=1330/ω_β=2955 cm^{-1} peak.

The fact that this peak is identical to the HPPD/329 binding-specific cross-peak arising due to interactions between the herbicide NO_2 group and HPPD amino acid CH side-chains, and that both Nitisinone and 329 have NO_2 groups, suggests that the data from the lower frequency range experiments is more likely to be correct. However, this discrepancy between the peaks seen in the two Nitisinone data sets suggests a reproducibility issue for this sample over the two different frequency ranges, and both sets of experiments need to be repeated to determine whether the binding of Nitisinone to HPPD can, in fact, be detected.

4.4 Limitations

There are several limitations to the data presented in this chapter that should be considered. Firstly, the comparison of the binding of iron-treated and non-iron-treated HPPD to herbicides would have been aided by sending both samples of HPPD to be tested for iron-content.

Another key limitation of the HPPD experiments was the appearance of the non-resonant horizontal “stripes” in the lower frequency range data. While this appeared to be effectively removed in data processing, it would have been preferable to find the source of the artefact and remove it before collecting the data.

Another limitation of the lower frequency range data was the fact that the binding peaks were only visible above the 2σ level (instead of 3σ as for the data collected over the previous frequency range), despite recalculating the value for σ over various spectral regions free of peaks. This suggests that there is a higher source of noise occurring in the laser system, which needs to be addressed.

Finally, the assignments made here are only tentative due to a lack of accompanying calculated EVV 2DIR spectra which would enable assignment confirmation.

4.5 Chapter Summary

To investigate HPPD-herbicide binding interactions, EVV 2DIR spectra were collected for a range of herbicides complexed with both iron-treated and non-iron treated HPPD (as well as a control protein, BSA) over two different frequency ranges. By quadratically subtracting the protein spectra from the corresponding protein+herbicide spectra (both averaged from four repeats), difference spectra were obtained, which consistently showed statistically significant binding-specific cross-peaks (above the 3σ level) for the HPPD inhibitor 329.

In the upper frequency range (ω_α=1230-1730/ω_β=2590-3440 cm^{-1}) the single cross-peak present in the HPPD/329 difference spectra appears at ω_α=1530/ω_β=2870 cm^{-1}, which indicates coupling of vibrational modes occurring at 1530 and 1340 cm^{-1}. The chemical structure of 329 comprises two aromatic nitro- groups, which are known to show a strong IR absorption at 1530 cm^{-1}, thus enabling assignment of this mode to an asymmetric NO stretch[80]. Examining the binding site of the HPPD+329 complex structural model shows that the nitrogen groups are in close proximity with several amino acids. This enables assignment of the second vibrational mode (which one would assume is occurring on the protein, as the cross-peak is binding specific) to an HPPD in-plane CH bending vibration occurring upon herbicide binding. However, calculated EVV 2DIR spectra and/or experiments where HPPD has mutated amino acid residues in the binding site are needed to confirm specifically which residue the in-plane CH bend arises from. Preliminary attempts to produce mutant HPPDs for this purpose were unsuccessful.

In contrast, the difference spectra for both HPPD and BSA with Mesotrione and Nitisinone showed the same EVV 2DIR cross-peaks appearing at similar intensities, indicating that they arose from non-specific binding interactions (either protein-herbicide or herbicide-herbicide interactions not particular to HPPD-herbicide binding).

Interestingly, a comparison of the binding of non-iron-treated HPPD and iron-treated HPPD for both 329 and Nitisinone showed that iron-treatment in HPPD preparation does not significantly affect the intensity of EVV 2DIR cross-peaks in the spectra. This may have implications for the preparation of HPPD protein for the purposes of EVV 2DIR

experiments. Iron content testing should be performed on both HPPD samples to verify this.

The EVV 2DIR spectra of amino acids previously calculated has shown many additional cross-peaks in the region *below* ω_β=2600 cm^{-1}. Therefore, EVV 2DIR spectra for iron-treated HPPD both with and without the two most promising binding candidates (329 and Nitisinone) were also collected over a lower frequency range of ω_α=1230-1730/ω_β=2090-3090 cm^{-1}. The HPPD difference spectrum for 329 shows the same binding-specific cross-peak at ω_α=1535/ω_β=2875 cm^{-1}. Additionally, there are two new peaks present above the 2σ level at ω_α=1355/ω_β=2455 and ω_α=1330/ω_β=2710 cm^{-1}. Because two non-specific binding peaks appear at ω_α=1360 cm^{-1} in the difference spectra for Mesotrione (which also contains an aromatic NO_2 group) this facilitates preliminary assignment of the 1355 cm^{-1} vibrational mode to a symmetric NO stretch. Coupling of the vibrational mode at 1100 cm^{-1} is assigned to a histidine motion involving a C-N stretch and in-plane C-H bend. This is a probable assignment as it implies again that the NO_2 group (or groups) are interacting with one or both of the conserved histidine residues which play a key role in herbicide binding, and supports the preliminary assignment of the peak at ω_α=1535/ω_β=2875 cm^{-1} to an asymmetric NO stretch coupled to a histidine in-plane CH bend. The final peak at ω_α=1330/ω_β=2710 cm^{-1} is assigned to an asymmetric SO stretch at 1330 cm^{-1} and the second coupled vibration at 1340 cm^{-1} remains unassigned. The structural model of HPPD+329 shows that the 329 SO_2CH_3 moiety forms a polar contact with the NH_2 group of the residue Gln293. EVV 2DIR data of HPPD with this residue mutated complexed with 329 and/or calculated EVV 2DIR spectra are required to enable assignment of the 1340 cm^{-1} mode to a vibration specific to this glutamine residue.

The HPPD/Nitisinone difference spectra over the lower frequency range also shows the same binding peak at ω_α=1535/ω_β=2875 cm^{-1} as for HPPD/329. This is interesting, as it was not seen in the HPPD/Nitisinone difference spectra over the previous range. Considering that Nitisinone, like 329, has a NO_2 moiety, it would make sense that this peak is real. However, due to the discrepancies between the two HPPD/Nitisinone data sets, it is unclear whether HPPD-Nitisinone binding can be detected by EVV 2DIR spectroscopy, hence both sets of experiments should be repeated.

In conclusion, the research presented in this chapter fulfils the primary aim of this thesis: to detect plant-herbicide binding using EVV 2DIR spectroscopy. Additionally, it shows that treating HPPD with iron after protein purification does not significantly alter the outcome

of the EVV 2DIR spectra produced for a given sample. Further research is necessary to verify the preliminary assignments made here of HPPD-herbicide binding interactions.

5. Detection of intermolecular interactions in amyloid beta (1-40) peptide

While the primary aim of this thesis was to investigate plant protein-herbicide interactions, EVV 2DIR experiments were also performed to investigate the peptide amyloid beta (Aβ) and potential interactions with added metal ions. Peptides consist of less than 50 amino acids linked together and are the subunits that make up proteins[146]. Several biologically important peptides have been previously characterised using EVV 2DIR spectroscopy, as seen in references[92,109,111]. Additionally, they are more tractable to study using EVV 2DIR than proteins, as they are smaller in size hence often give greater signal to noise in the resultant EVV 2DIR spectra. Therefore, peptides are an excellent system to validate EVV 2DIR as a technique to study biological intermolecular interactions.

Amyloid beta (1-40), (herein referred to as Aβ(1-40)) is a 40-residue peptide that is implicated in the aetiology of Alzheimer's disease[69,147]. It is well established in the literature that copper (II) and zinc (II) bind to Aβ and cause peptide aggregation *in vitro*, though the mechanism that underpins this process is unknown[65,67,68,71]. This section presents the first EVV 2DIR data of Aβ(1-40) both alone and complexed with copper (II) and zinc (II).

5.1 Methodology

5.1.1 Materials & Methods

A series of 1 μL amyloid beta spots both with and without copper and zinc were prepared, as follows:

1. 5 mM Aβ(1-40), pH = 6.5
2. 5 mM Aβ(1-40) + 5mM copper chloride, pH = 6.5
3. 5 mM Aβ(1-40) + 5mM zinc chloride, pH = 6.5

The Aβ(1-40) peptide was purchased by Eurogentec (the sequence for which is provided in Appendix C). Prior to spotting down the sample spots, each of the three solutions was prepared using a protocol based on two published methods[148,149]. A high concentration of 5 mM of amyloid beta was used with the aim being to maximise EVV 2DIR signal to noise. Three vials of 0.3 mg of Aβ(1-40) were dissolved in 230 uL of 1,1,1,3,3,3-Hexafluoro-2-propanol (which is used to solubilise and monomerise Aβ peptides), vortexed and left for the excess solvent to evaporate overnight. 5 mM copper (II) chloride and zinc (II) chloride solutions with a pH of 6.5 were prepared in the same phosphate buffer used as for the protein experiments (pH adjusted using minimal HCl). Each vial of dried amyloid beta had 1:1 stoichiometric amounts of either the copper, zinc or phosphate buffer solution added to give the three end solutions outlined above and were vortexed prior to production of gel-like spots.

5.1.2 Humidity optimisation of amyloid beta spots

While the gel-like spots produced by HPPD favoured a relative humidity of 75.5%, the amyloid beta samples formed the best spots at higher values of relative humidity. 1 μL spots of the three solutions above were left to equilibrate for at least 36 hours in a humidity cell (as with the HPPD gel-like spots) containing potassium nitrate to give a controlled relative humidity of 94.6%. However, the amyloid beta spots produced were more crystalline than the protein counterparts, which might be expected for aggregated samples. This did not appear to affect the EVV 2DIR data collected, as the signal remained stable using these sample spots over the duration of the set of experiments.

5.1.3 FTIR data collection

As for the HPPD experiments, attempts were also made to collect FTIR spectra in both liquid and gel spot forms for all samples in this chapter. However, these again were either unsuccessful or too congested for assignment, hence aren't included here. A representative biological spectrum is included in Figure 9 to illustrate this spectral congestion.

5.1.4 EVV 2DIR data collection

EVV 2DIR data was collected over the frequency range ω_α = 1500-1750 cm^{-1} ω_β = 2700-3450 cm^{-1} (before applying any polystyrene calibration offsets) with a resolution of 5 cm^{-1} per pixel and an average of 100 samples per pixel. This reduced ω_α frequency range was chosen because preliminary amyloid beta data showed that most of the features of interest were in this region. Every 5 cm^{-1} in the ω_α dimension takes approximately 1 minute to accumulate, so reducing the frequency range significantly decreases the time taken to collect one spectrum.

All three beams were set to propagate with the same polarisation (PPP). Time delays were fixed at $T_{\alpha\beta}$ = 1.75 and $T_{\beta\gamma}$ = 1 ps as they gave the maximum spectral resolution with the minimum non-resonant background. Spectra were collected of each of the three different sample spots in turn, and this process was repeated four times. Because the full set of experiments is collected over a period of 10 hours, it is important to account for laser beam drift which may otherwise cause variation between repeats (as discussed in Section 4.2.1.2 above). Therefore, each repeat was carried out on a slightly different sample spot position to ensure none of the repeats were collected from an irregular part of the sample.

5.1.5 Data processing

Following the collection of four repeats of each sample, the data was processed via MATLAB using the protocols outlined in Section 4.3.3, unless otherwise stated. The EVV 2DIR data collected in this chapter was *not* normalised by OPA energy profiles (as for the HPPD data) because it was all collected as one set of experiments with the same OPA energy profiles. Therefore, comparison between spectra could be made without the need for normalising by OPA energy.

The polystyrene calibration performed prior to experiments indicated that OPA1 and OPA2 had an offset of 5 cm^{-1} and 20 cm^{-1} respectively, hence the first step in data processing was to subtract these from the raw data.

Prior to averaging, each individual repeat was normalised relative to the CH_2 peak which gave the highest intensity at ω_α=1480 cm^{-1} & ω_β=2945 cm^{-1}. Normalisation of intensity is necessary to enable averaging of repeats, and subsequent comparison between the

resulting averaged spectra for each sample type (i.e. amyloid beta with and without copper II or zinc II).

5.2 EVV 2DIR spectrum of Aβ(1-40)

The first EVV 2DIR spectrum[1] of Aβ(1-40) is presented in Figure 59 below:

[1] Aside from a BSc report written by a research group student Frank (Zehua) Hu containing a single preliminary Aβ(1-40) spectrum (which was also collected by me)

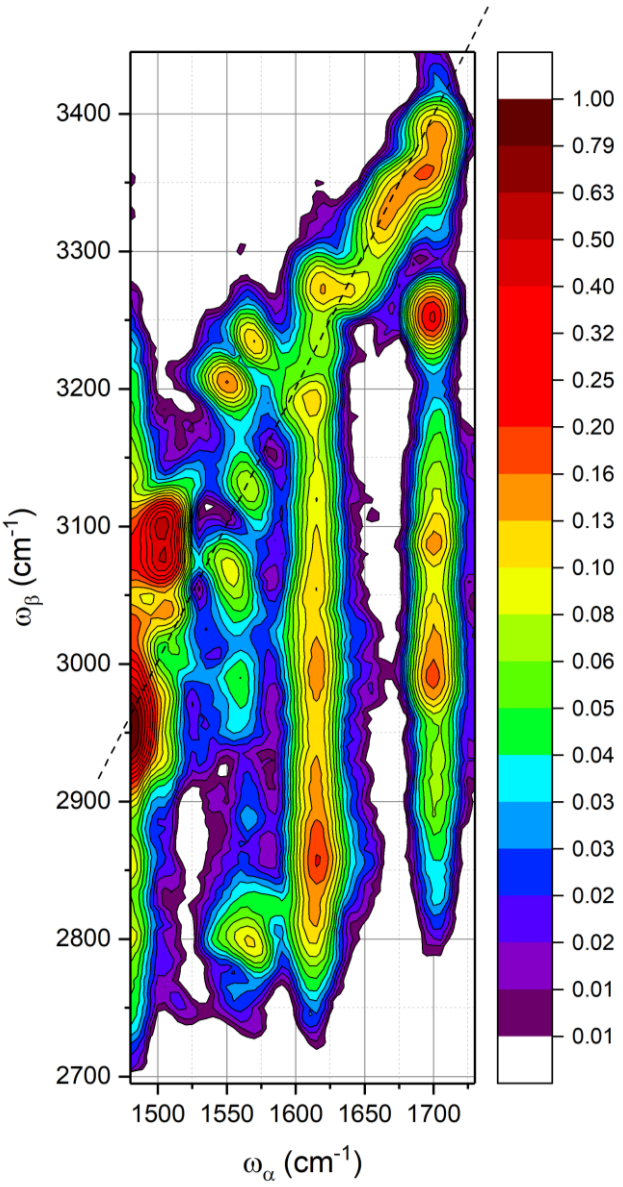


Figure 59: The EVV 2DIR spectrum of Aβ(1-40) , produced from an average of four repeats which were normalised relative to the most intense peak arising from CH_2 at 1480/2945 cm^{-1}. The colour bar on the right shows the scale of the contours following this normalisation. Peaks which fall on the black dashed line arise from overtone transitions.

Examining Figure 59, it is immediately clear that the spectrum is rich with information. There are over 35 cross-peaks with a statistically significant level of intensity. The criterion for this is being greater than 3σ, where σ is the standard deviation of the area of background removed from the average of each set of repeats prior to processing (cross-peaks below this level are omitted for clarity). Without the accompanying calculated EVV 2DIR spectra for this system (produced from vibrational frequencies of Aβ(1-40)

calculated by *ab initio* DFT methods, according to Kwak et al[104] and Sowley[49] and detailed in Section 2.3.3), it is difficult to interpret the spectra and make concrete spectral assignments. However, as discussed in Chapter 1, amyloid beta has been previously studied by other forms of 2DIR and FTIR, hence the literature can aid in tentative spectral assignment[76,77,150]. Additionally, Barth provides comprehensive mid-IR vibrational mode assignments of amino acids and secondary protein structures[141,143,145], which further assists preliminary peak assignment. Considering the novel nature of this work, it is beyond the scope of this thesis to provide full assignment of every peak in the spectrum. Nonetheless, assignments of several cross-peaks can be assigned based on previous EVV 2DIR spectra of biomolecules, which are outlined in Table 9 below.

$\omega_\alpha/\omega_\beta$ cross peak coordinates (cm^{-1})	**Group assignment**
1480/2945	Methylene
1620/3240	Tyrosine (overtone)
1520/3130	Tyrosine
1505/3100	Phenylalanine
1505/3075	Phenylalanine
1635/3270	Amide I: β-sheet (overtone)
1670/3335	Amide I: β-turn (overtone)

Table 9: Group assignments for cross-peaks appearing in the EVV 2DIR spectrum of Aβ(1-40) , based on previous literature assignments.

The key spectral features of interest in Figure 59 (which differ from features commonly seen in EVV 2DIR spectra of biomolecules) are not included in Table 9, but are instead discussed more in-depth in the following text. Notably, the spectrum in Figure 59 exhibits an intense column of cross-peaks at $\sim\omega_\alpha$=1615 cm^{-1}. A horizontal cut-through of the peaks centred on ω_β=2885 cm^{-1} is shown in Figure 60 below.

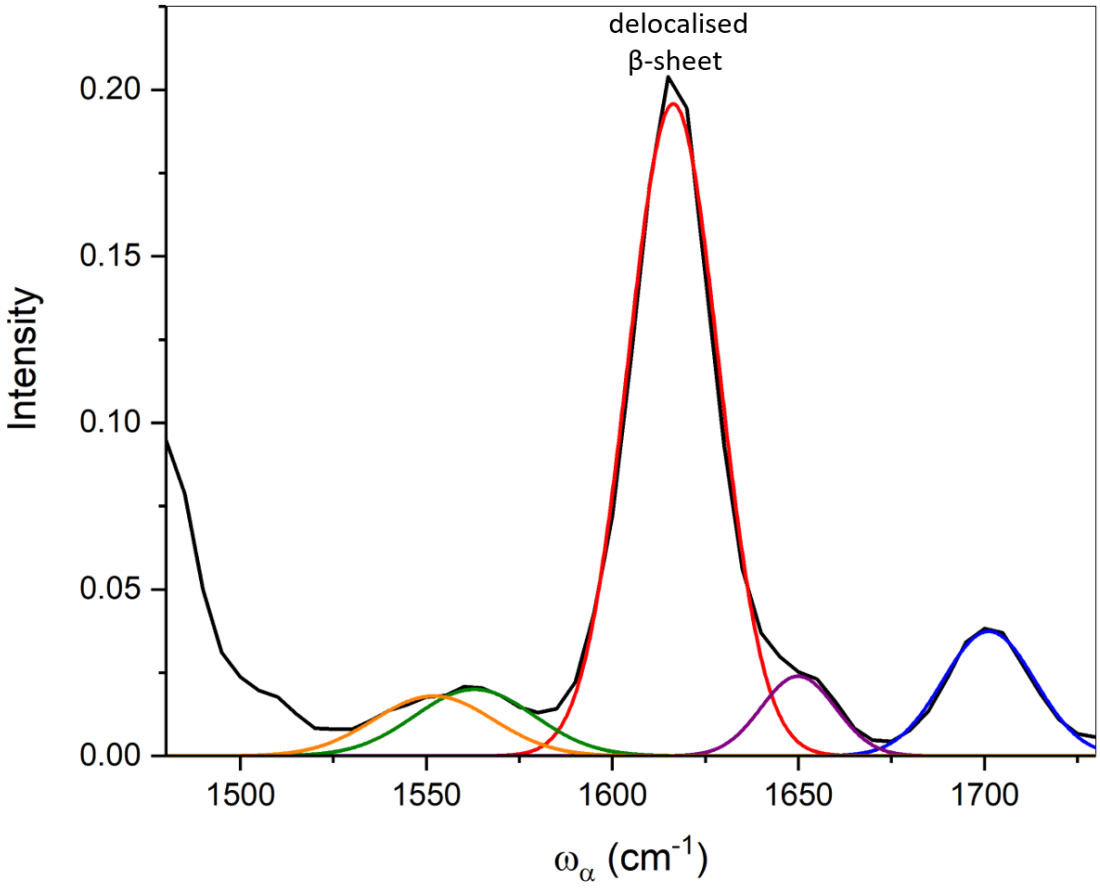


Figure 60: A vertical cut-through of the Aβ(1-40) EVV 2DIR spectrum centred at ω_α=2885 cm^{-1}. The unfitted data is shown by the black line, while with Gaussian fitting of the 5 main peaks occurring at this frequency is shown in various colours. Labels denote tentative vibrational mode assignments.

The peaks in Figure 60 have been Gaussian-fitted and show that the more accurate assignment is at ω_α=1616.3 cm^{-1}, however as the resolution of this technique is 5 cm^{-1} it will continue to be referred to as 1615 cm^{-1} (for simplicity when calculating the frequency of vibrations coupled to this mode). Recently published literature enables assignment of this to a delocalised β-sheet mode characteristic of amyloid beta fibril formation[76,77]. At a slightly acidic pH of 6.5 (such as the pH conditions used in these experiments), amyloid beta is more likely to form amyloid beta fibrils[151] with rigid β-sheets which further supports this assignment. 1D IR spectra of amyloid beta peptides do not exhibit this distinct peak, as it is overwhelmed by the broad peak at ω_α=1635 cm^{-1} traditionally assigned to β-sheets in amyloid beta[150]. Figure 60 indicates that the spectral feature at ω_α=1615 cm^{-1} appears to be overlapping with another frequency at ω_α=1650 cm^{-1}.

A vertical cut-through of the peaks centred at ω_α=1615 cm^{-1} is shown in Figure 61. The peaks were not able to be Gaussian-fitted here due to the overlapping spectral features. However, for the purposes of this discussion it is not important, as the main goal of the cut-through sections is to compare the how the intensities of each peak changes upon metal addition (discussed in the following sections). Additionally, Gaussian-fitting the peaks in Figure 60 only change the peak centres by ~1 cm^{-1}, therefore does not have much effect.

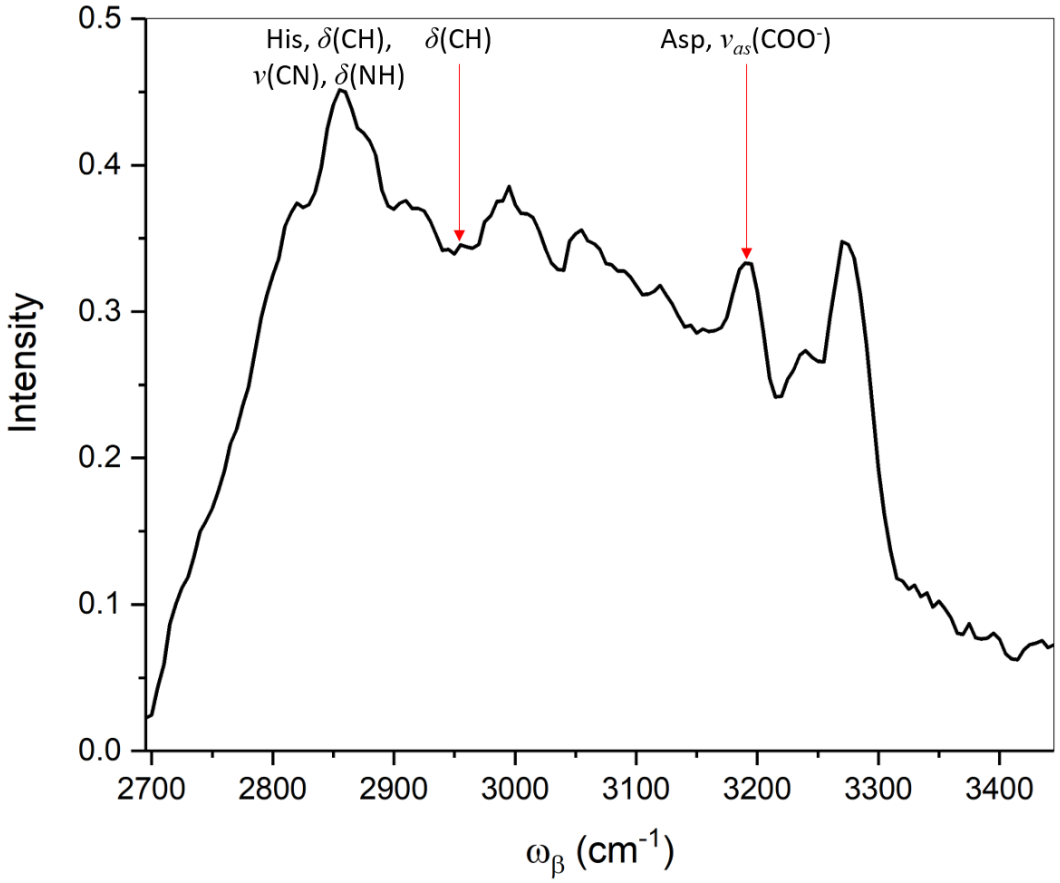


Figure 61: A vertical cut-through of the column of peaks centred at ω_α=1615 cm^{-1} present in the EVV 2DIR spectrum for Aβ(1-40). Labels denote tentative vibrational mode assignments.

Having assigned the 1615 cm^{-1} feature to a vibrational mode characteristic of Aβ(1-40) fibril β-sheets, we can now make tentative assignments for the vibrational modes coupled to it based on the literature of amino acid infrared absorptions. These are shown in Table 10 below:

Mode 1=ω_α (cm^{-1})	**ω_β (cm^{-1})**	**Mode 2= ω_β-ω_α (cm^{-1})**	**Mode 2 assignment**
1615	3190	1575	Asp asymmetric COO^- stretch
1615	2955	1340	in-plane CH bend
1615	2855	1240	Histidine in-plane CH bend, CN stretch, in-plane NH bend

Table 10: Tentative assignments for the coupled vibrations giving rise to the column of peaks centred on ω_α=1615 cm^{-1}.

An additional spectral feature of interest is the column of intense cross-peaks seen at ω_α=1700 cm^{-1} in the Aβ(1-40) spectrum shown in Figure 59. A vertical cut-through of the column of peaks centred on ω_α=1700 cm^{-1} is shown in Figure 62.

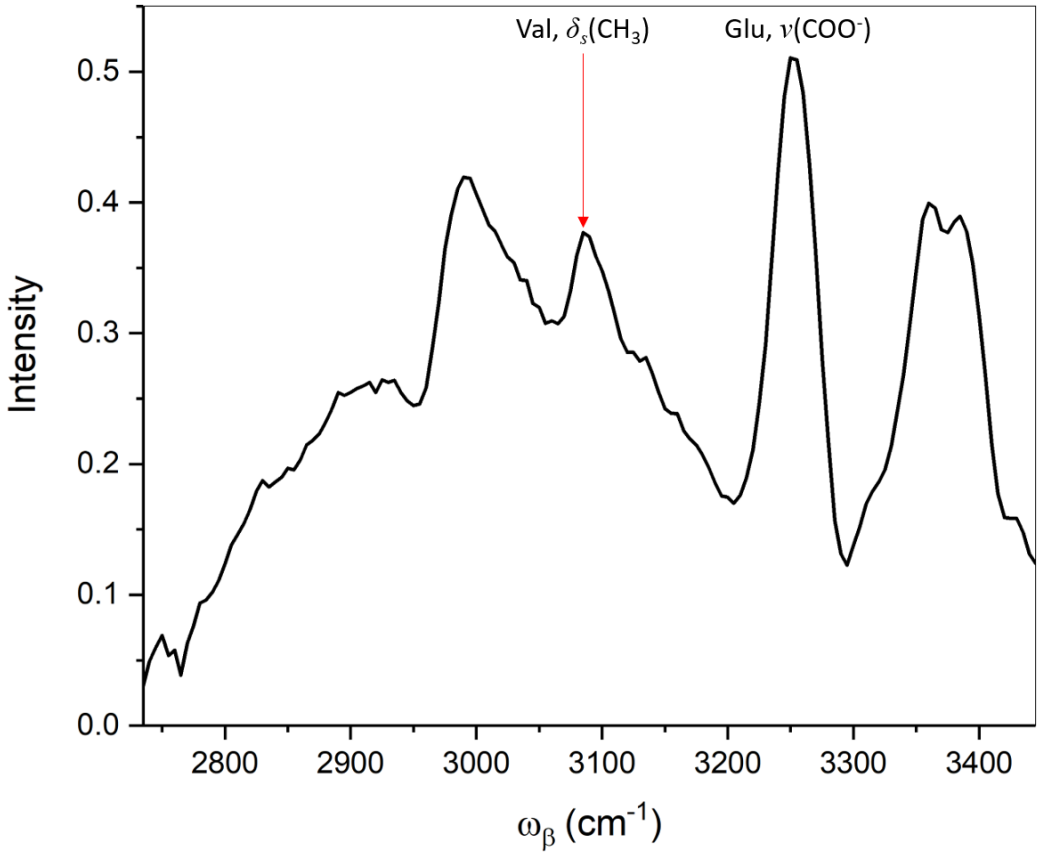


Figure 62: A vertical cut-through of the column of peaks centred on ω_α=1700 cm^{-1} present in the EVV 2DIR spectrum for Aβ(1-40). Labels denote tentative vibrational mode assignments.

This column could potentially be attributed to a vibrational mode arising from peptide backbone carbonyls interacting with the side-chains of various amino acid residues in the fibril structure. While it is difficult to assign the carbonyl mode at ω_α=1700 cm^{-1} to a *specific* amino acid residue without the aid of calculated spectra (generated by Aβ(1-40) vibrational frequency calculations via the *ab initio* DFT methods set out by Kwak et al[104] and Sowley[49], discussed in Section 2.3.3), the frequency of 1700 cm^{-1} suggests it is due to an Asp and/or Glu C=O stretch. Recent structures of Aβ peptides (1-42) and (1-40) (obtained via cryo-electron microscopy and molecular dynamics simulations, respectively) reveal that the Asp23 residue forms a salt-bridge with Lys28 in wild-type Aβ which stabilises β-sheets in the formation of parallel fibrils[152,153] as shown in Figure 63.

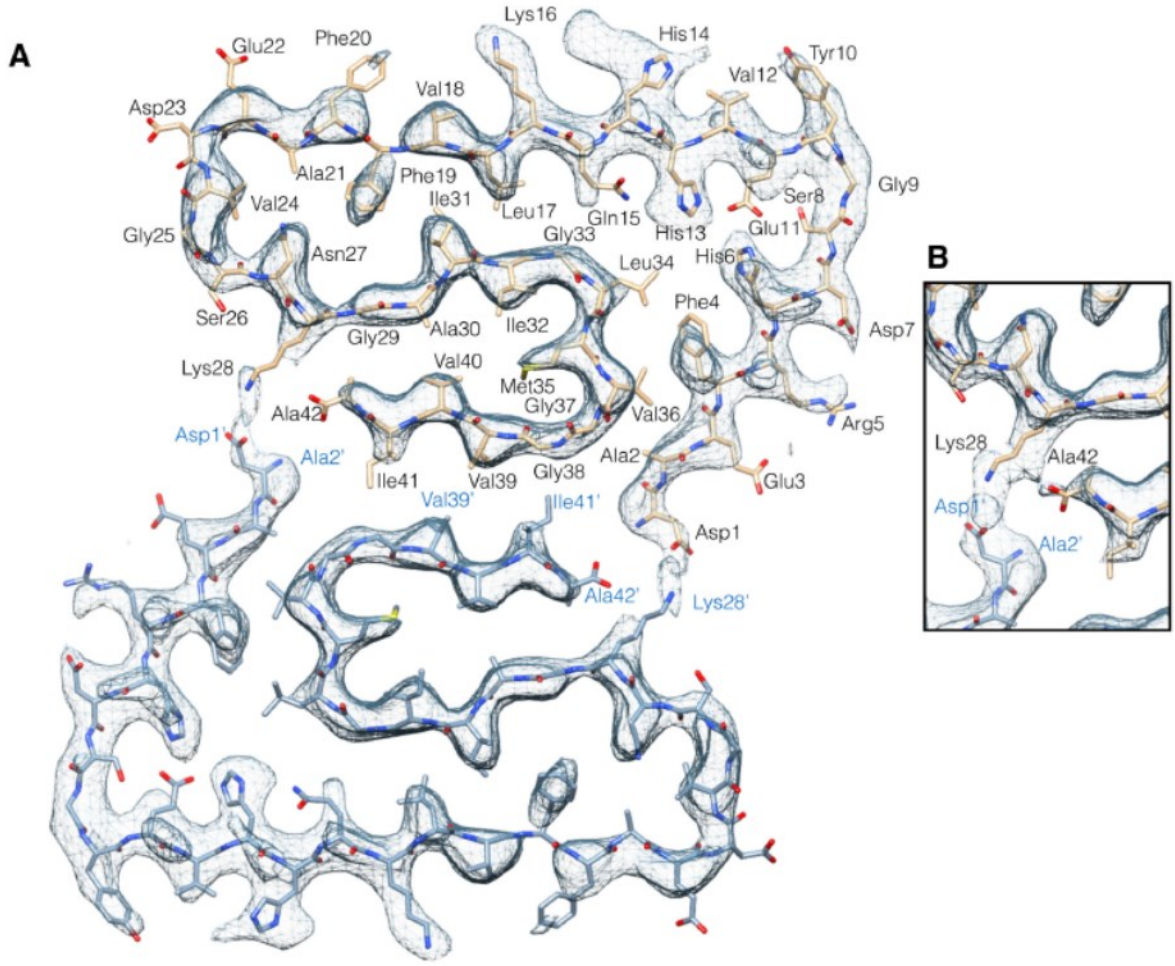


Figure 63: The cryo-electron microscopy structure of Aβ(1-42) with mapped electron density , showing A) the interactions between two β-sheets (one blue, one beige) in a parallel fibril, with B) the resulting Asp-Lys salt bridge formation shown in more detail (image taken from Reference 152).

Mutating the Asp23 residue to an asparagine is known to cause early onset Alzheimer's disease with increased neurotoxicity, and results in a conversion to less stable antiparallel fibrils, which involves a reorganisation of structure whereby monomers detach and reassemble[153]. This breakdown into smaller aggregates is proposed to be a source for toxic oligomer formation. Thus, Asp23 is a key residue in fibril formation, and it is therefore reasonable to tentatively assign the mode at ω_α=1700 cm^{-1} to the aspartic acid carbonyl stretch.

Having tentatively assigned the mode at ω_α=1700 cm^{-1} to the aspartic acid carbonyl stretch, Table 11 shows preliminary assignments for the vibrational modes coupled to it giving rise to the most intense cross-peaks:

Mode 1=ω_α (cm^{-1})	**ω_β (cm^{-1})**	**Mode 2= ω_β-ω_α (cm^{-1})**	**Mode 2 assignment**
1700	3250	1550	Glutamic acid COO^- stretch
1700	3085	1385	Valine symmetric in-plane CH_3 bend
1700	2990	1290	Unassigned

Table 11: Tentative assignments for the coupled vibrational modes giving rise to the cross peaks centred on ω_α=1700 cm^{-1}.

The intense cross-peak at ω_α=1700/ω_β=3250 cm^{-1} is tentatively assigned to an aspartic acid C=O stretch coupled with a glutamic acid asymmetric COO^- stretch. Examining the N-terminus of the parallel Aβ(1-40) fibril structure in Figure 63, Asp1 is curled around to interact with Lys28 on the adjoining β-sheet, which puts its C=O group closer to the side-chain of Glu 3, thus it is plausible that the two might interact. Additionally, the side-chains of Asp23 and Glu22 are near each other, potentially undergoing the same vibration and increasing the intensity of the resulting cross-peak.

The vibrational mode at ω_α=1700 cm^{-1} is also coupled to a mode at 1385 cm^{-1}, the latter of which can be tentatively assigned to a symmetric in-plane CH_3 bend occurring on a valine residue using literature values. This fits with the structure of Aβ(1-40), as residues Asp23 and Val24 are next to each other in the hairpin bend of the β-sheet structure, therefore their side-chains may well interact.

An assignment for the 1290 cm^{-1} mode listed in Table 11 is unable to be suggested based on the literature alone, as none of the amino acid residues that make up Aβ(1-40) have vibrations near this frequency. Again, calculated EVV 2DIR spectra (constructed from vibrational modes calculated using DFT methods, such as those utilised by Kwak et al[104] and Sowley[49]) is needed to make even a tentative assignment for this mode.

5.3 EVV 2DIR spectra of Aβ(1-40) + Cu(II)

In introducing the biological impact of Aβ, the aetiological effects of complexing with metals such as copper and zinc were discussed in Chapter 1. Below, Figure 64 presents the first EVV 2DIR spectra of Aβ1-40 complexed with (and without) Copper (II).

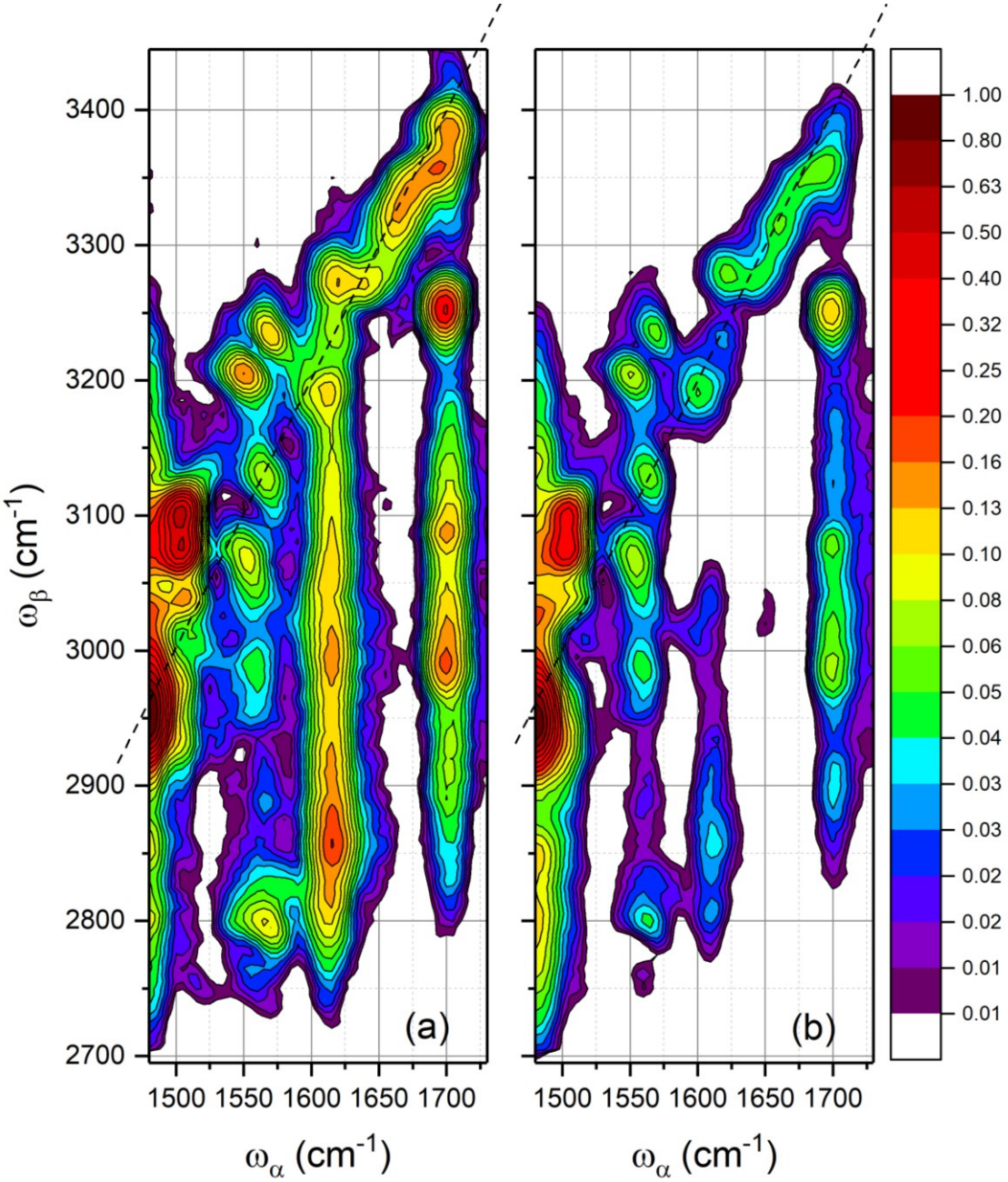


Figure 64: EVV 2DIR spectra of a) Aβ(1-40) alone, and b) with Cu(II) added. There are several key spectral differences, namely the decreased intensity of the columns of cross-peaks at ω_α=1615 cm^{-1} and ω_α=1700 cm^{-1}.

Examining Figure 64, there is a significant difference in the spectral features present upon the addition of copper. Figure 65 shows a vertical cut-through of the peaks centred at ω_α=1615 cm^{-1}, compared with that of Aβ only:

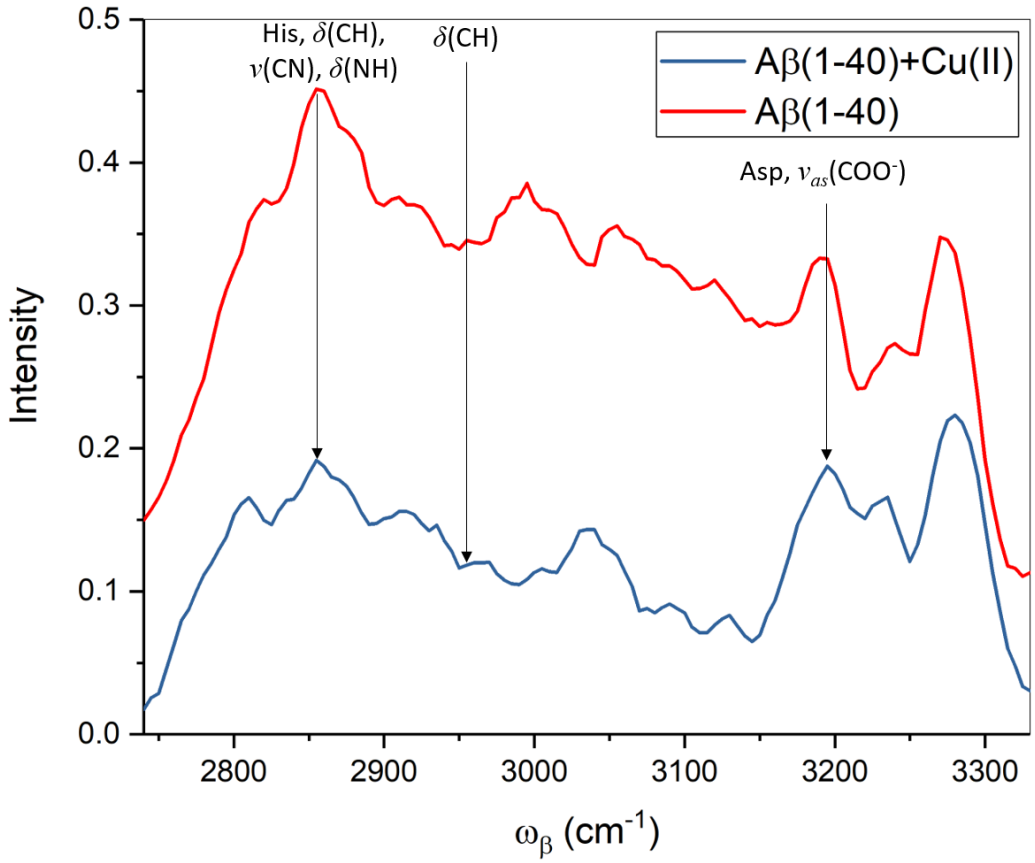


Figure 65: A comparison of the ω_α=1615 cm^{-1} vertical cut-through sections of the EVV 2DIR spectra for Aβ(1-40) alone (red) and Aβ(1-40) + Cu(II) (blue). Labels denote tentative vibrational mode assignments.

Figure 65 shows that the column of cross-peaks centred on ω_α=1615 cm^{-1} decreases in intensity significantly with copper binding. In particular, the signal intensity of the cross-peak centred at ω_α=1615/ω_β=2855 cm^{-1} decreases by a factor of 2.4. This had previously been assigned to the coupling of an amyloid fibril β-sheet mode with a mode histidine in-plane CH bend, CN stretch, in-plane NH bend. The presence of copper has been widely reported to disrupt the formation of amyloid fibrils[69]. Histidine is known to play a key role in the binding of Aβ(1-40). Thus, the decreased intensity of this peak may be due to reduced interactions between the amyloid fibril β-sheets (which are being partially prevented from forming) and histidine, which is instead binding the copper ion.

Figure 64 also shows differences in the carbonyl region with the addition of Cu(II) to Aβ(1-40). A comparison between the vertical cut-through sections of the column of cross-peaks centred at ω_α=1700 cm^{-1} for Aβ(1-40) both alone and with Cu(II) is given in Figure 66 below:

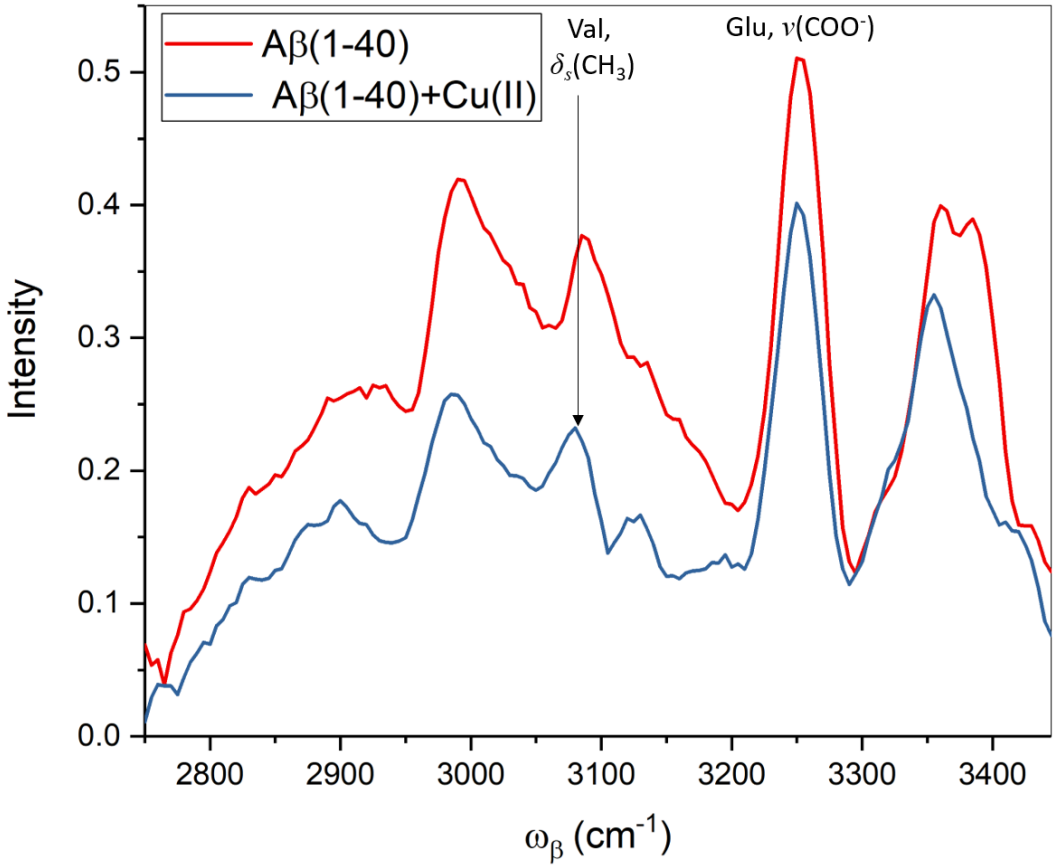


Figure 66: A comparison of the ω_α=1700 cm^{-1} vertical cut-through sections of the EVV 2DIR spectra for Aβ(1-40) alone (red) and Aβ(1-40) + Cu(II) (blue). Labels denote tentative vibrational mode assignments.

The signal intensity for the cross-peaks at ω_α=1700/$\omega\beta$=2990 cm^{-1} and ω_α=1700/$\omega\beta$=3085 cm^{-1} decreases by 0.2-fold for both upon the binding of copper to Aβ(1-40). While the second peak remains unassigned, the first peak was previously assigned to coupling between an aspartic acid C=O stretching mode and valine symmetric in-plane CH_3 bend mode. Examining the structure of the parallel Aβ(1-40) fibril subunit in Figure 63, the side-chains of C-terminal Val36 residue comes into proximity with the carbonyl backbone of the N-terminal Asp1 residue when the C-terminal end of the peptide

curls around to form the rigid β-sheet structure characteristic of fibrils. Thus, if copper is preventing fibril sheet formation as proposed in the literature, this could explain the decreased intensity for this peak upon copper binding.

5.4 EVV 2DIR spectra of Aβ(1-40) + Zn(II)

Further spectral differences are evident when comparing the EVV 2DIR spectra for Aβ(1-40) alone and with zinc (II), as seen in Figure 67 below.

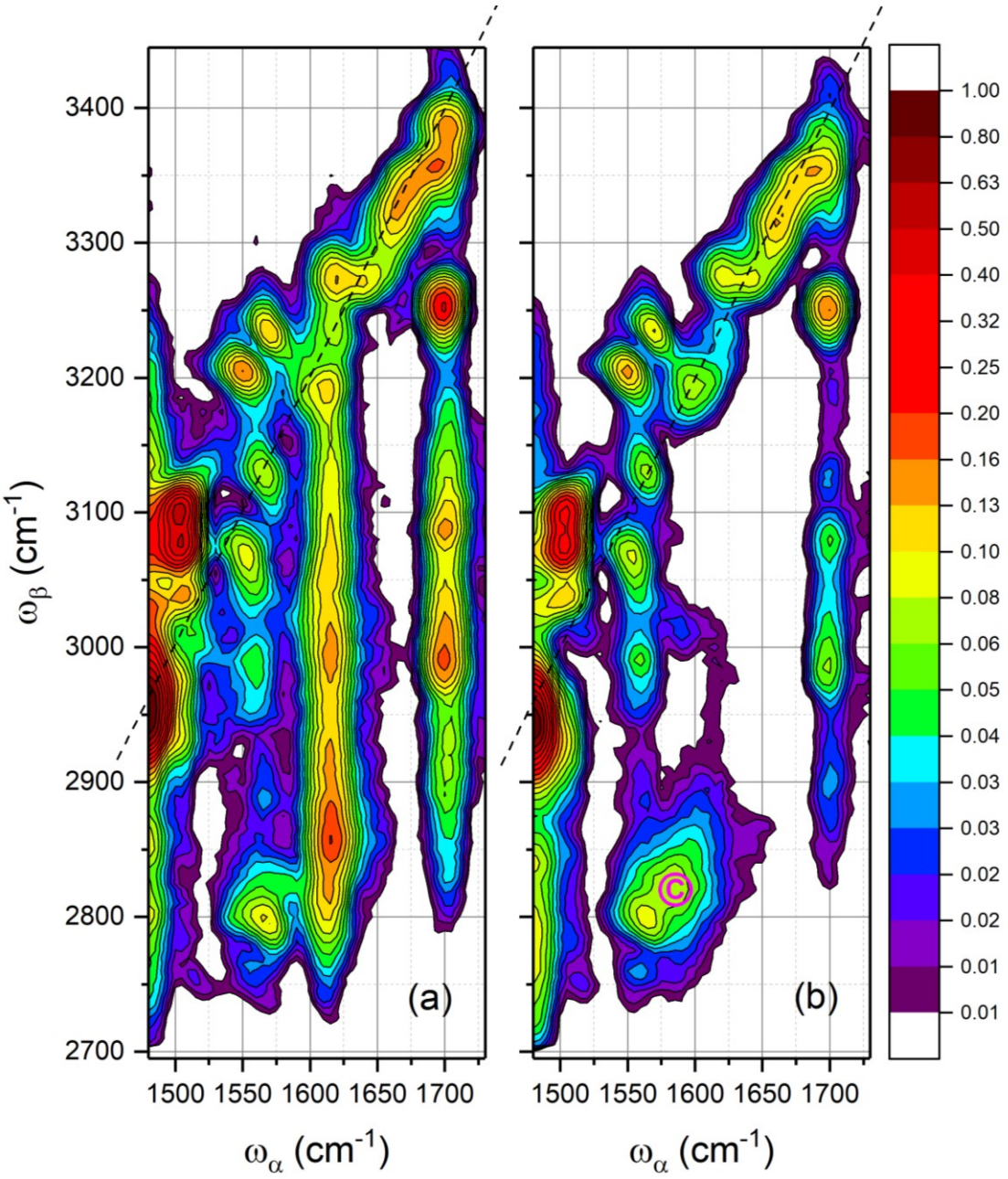


Figure 67: The EVV 2DIR spectra of a) Aβ(1-40) alone and b) with added zinc , averaged from four repeats of each. There is a new overlapping peak at ω_α=1585 cm^{-1}, labelled C, which appears due to binding of Zn(II).

There is a broadening and overlapping of spectral features from ω_α=1550-1650/ω_β=2750-2900 cm^{-1}, due to the appearance of a new peak at ω_α=1585/ωβ=2820 cm^{-1} (labelled "C" in Figure 67), which is exhibited more clearly in the horizontal cut-through centred on ω_β=2820 cm^{-1} shown for both Aβ(1-40) alone and with Zn(II) in

Figure 68.

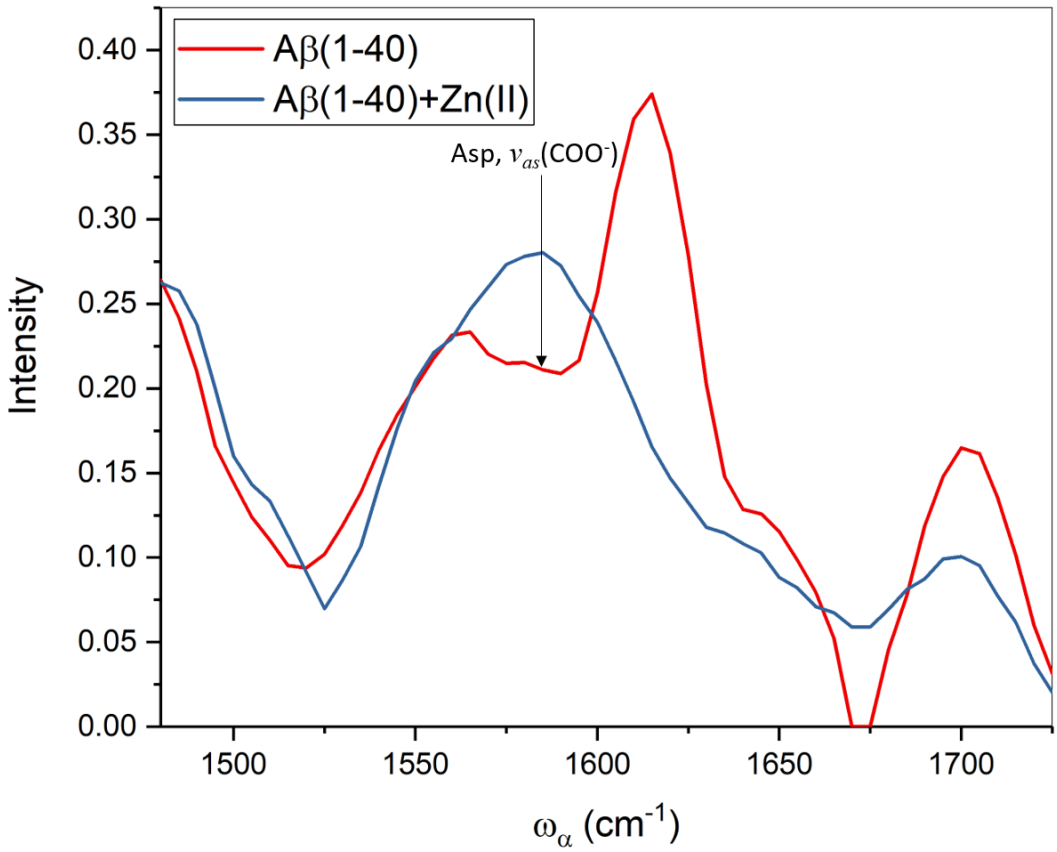


Figure 68: A comparison of the ω_β=2820 cm^{-1} horizontal cut-through sections of the EVV 2DIR spectra for Aβ(1-40) alone (red) and Aβ(1-40) + Zn(II) (blue). Labels denote tentative vibrational mode assignments.

A comparison of the vertical cut-through sections between Aβ(1-40) and Aβ(1-40)+Zn(II) for the column of peaks centred on ω_α=1615 cm^{-1} is given in Figure 69 below:

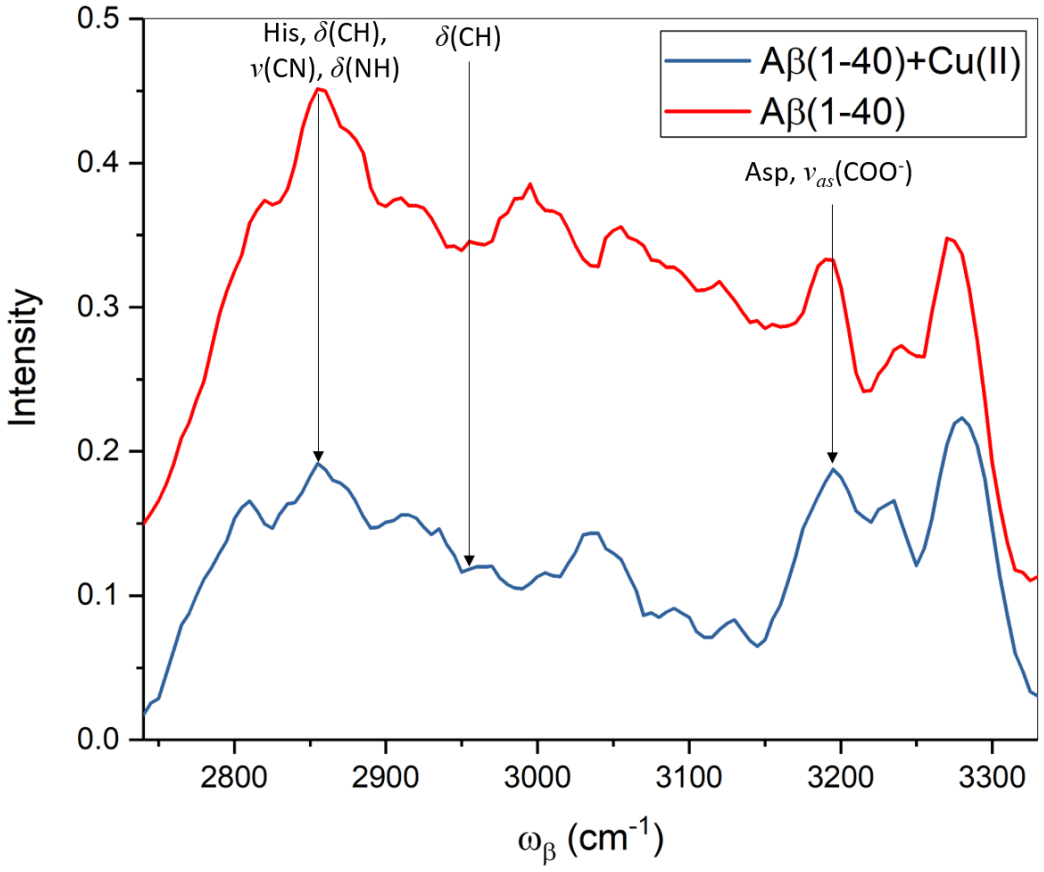


Figure 69: A comparison of the vertical cut-through sections of the cross-peaks at ω_α=1615 cm^{-1} for Aβ(1-40) (red) and Aβ(1-40)+Zn(II) (blue) present in their EVV 2DIR spectra. Labels denote tentative vibrational mode assignments.

As shown in Figure 69, the intensity of the cross-peaks at ω_α=1615/ω_β=2855 and ω_α=1615/ω_β=2955 cm^{-1} decrease significantly upon the binding of Zn(II) to Aβ(1-40) (2.8 and 3.4 times, respectively). The changes in this region of the spectra can again be related to amyloid fibril structure disruption upon addition of metal. In contrast with the situation with copper addition, adding zinc to Aβ does not *prevent* fibril formation but instead *retards* it[72,74].

As mentioned earlier, the most accepted view of zinc binding occurs at the N-terminus of Aβ(1-40) (defined by residues 1-16) at a 1:1 ratio mainly interacting with Asp1, Glu11 and His6, His13 & His14 as shown in Figure 70[67,68,74,154]. While much emphasis is placed on the involvement of histidine residues due to the favourable π-π stacking interactions their rings form in Aβ fibrils (thought to be disrupted by metal binding)[155], less is known about the role of the N-terminal Asp1.

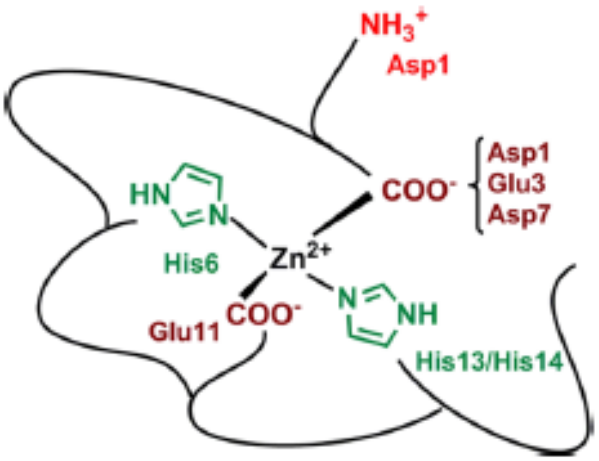


Figure 70: The mechanism of zinc coordination in Aβ(1-40) (image taken from Reference 154).

X-ray absorption spectroscopy[73] and modelling studies (based on experimental Aβ data)[74] have shown that zinc is able to simultaneously coordinate both intra- and intermolecularly (i.e. along or between β-sheets in a fibril structure, respectively), enabling the bridging of two peptides. Zinc is proposed to form these intermolecular interactions by coordinating Asp7 in the N-terminal region, and structural distances indicate that the side-chains of Asp7 replace backbone carbonyl group upon zinc binding[74]. Consequently, the vibrational mode at ω_α=1585 cm^{-1} may be tentatively assigned to an aspartic acid asymmetric COO^- stretch, which fits with the literature frequency assignment[143].

Figure 71 shows a comparison between the vertical cut-through sections of the column of peaks centred on ω_α=1700 cm^{-1} in the Aβ(1-40) and Aβ(1-40)+Zn(II) EVV 2DIR spectra. The signal intensity for the cross-peaks at ω_α=1700/$\omega\beta$=2990 cm^{-1} and ω_α=1700/$\omega\beta$=3085 cm^{-1} decreases by 1.7-fold for both upon zinc binding. The decreased intensity in this column correlates with the proposed replacement of some of the carbonyl backbone interactions with Asp7 side-chain COO^- interactions[74,143]. This further supports the assignment of the cross-peaks centred on ω_α=1700 cm^{-1}, previously assigned to backbone carbonyl interactions involving Asp or Glu, coupling with an aspartic acid asymmetric COO^- stretch mode.

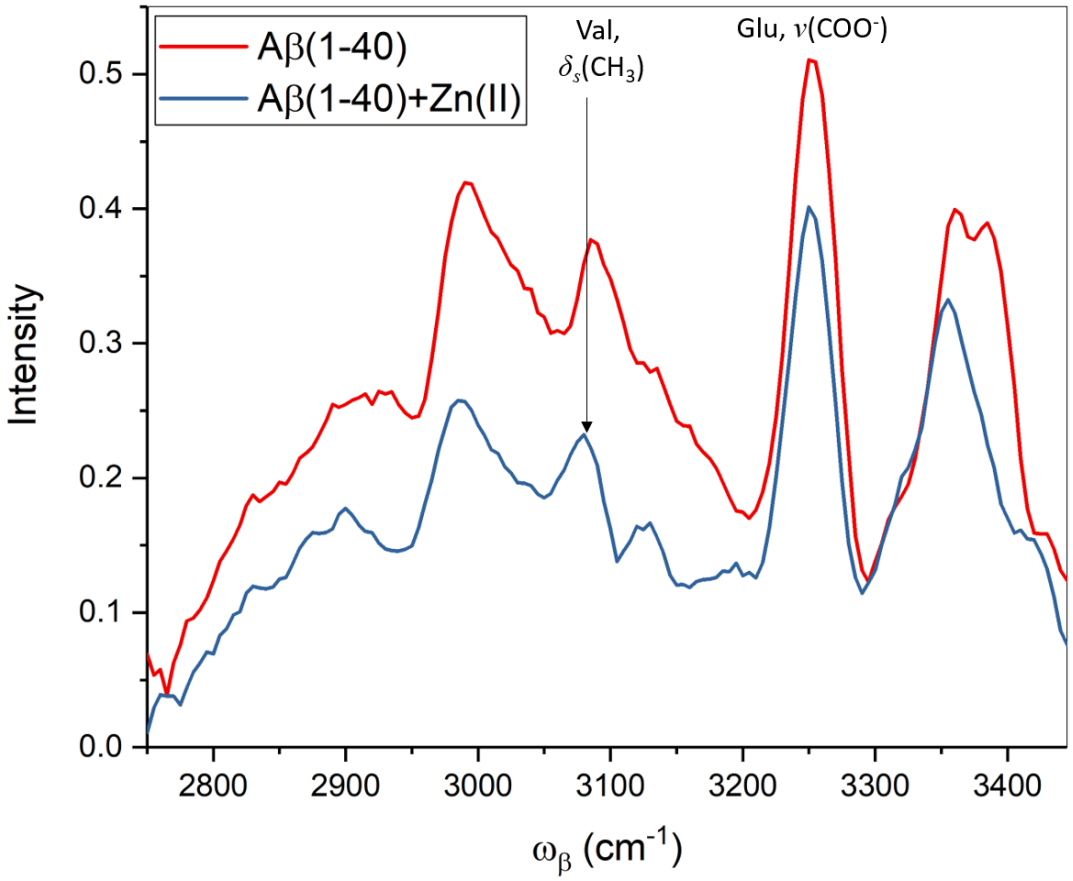


Figure 71: A comparison of the vertical cut-through sections of the cross-peaks at ω_α=1700 cm^{-1} for Aβ(1-40) (red) and Aβ(1-40)+Zn(II) (blue) present in their EVV 2DIR spectra. Labels denote tentative vibrational mode assignments.

Because peak C appears at ω_α=1585/$\omega\beta$=2820 cm^{-1}, this means that the vibrational mode at ω_α=1585 cm^{-1} is coupled to another mode at 1235 cm^{-1}. According to amino acid assignment literature, this can be attributed to a histidine motion comprising in-plane C-H bending, CN stretching and in-plane N-H bending vibrations. Examining the coordination of zinc by Aβ(1-40) exhibited in Figure 70, zinc is bound to Asp1/Glu3/Asp7, Glu11 and His6, His13/14. The fact that zinc coordination brings the Asp and His groups closer together (coupled with the ability zinc has to form intra- and intermolecular bridging of peptides) suggests a likely interaction between the two, resulting in the cross-peak at ω_α=1585/$\omega\beta$=2820 cm^{-1}.

Alternatively, the feature at ω_α=1585/$\omega\beta$=2820 cm^{-1} could arise due to the vibrational coupling of a histidine C=C stretching mode at 1585 cm^{-1} with another with the histidine

in-plane C-H bending, CN stretching and in-plane N-H bending vibration, which fits with literature frequency assignments of amino acid side chain vibrations. However, this is less likely, as histidine is also known to bind copper therefore we would expect to see the same cross-peak in the Cu(II) spectra. The absence of this peak in the Aβ(1-40)+Cu(II) spectra suggests an interaction occurring characteristic of Zn(II) binding.

5.5 Limitations

A key limitation of this work is that the samples were not tested for fibril formation. Unfortunately, due to time constraints high definition microscope images were not collected of each sample prior to performing EVV 2DIR experiments. These may have given us more idea of the morphology of the Aβ(1-40) in each sample (rather than relying solely on 2DIR spectroscopic assignments). However, the fact that the samples were prepared with an acidic pH of 6.5 and the presence of the strong distinct column of cross-peaks at ω_α=1615 cm^{-1} in the Aβ(1-40) spectrum in Figure 59 (and the decreased intensity of some of these peaks in the Aβ+metals spectra) means we can be fairly confident of fibril formation.

As with the HPPD experiments in Chapter 4, the sample preparation for this work requires further optimisation. It was difficult to find a common relative humidity which provided smooth sample spots which gave a decent EVV 2DIR signal for all three samples. Additionally, initial attempts to form samples using a pH of 7 produced powdery sample spots which gave poor EVV 2DIR signal. Thus, sample preparation for this system needs future consideration.

A major constraint in assigning and interpreting the EVV 2DIR spectra presented in this chapter is the lack of calculated EVV 2DIR spectra to confirm the preliminary spectral assignments made above. Accompanying quantum calculations need to be performed for each system in Gaussian 09[115], using DFT methods (according the technique detailed by Kwak et al[104] and Sowley[49]) to produce simulated EVV 2DIR spectra, as discussed in Section 2.3.3. This will enable more definitive assignments of the cross-peaks to vibrational modes which reflect the physiological processes occurring in each sample.

5.6 Chapter Summary

A series of EVV 2DIR spectra have been collected of Aβ(1-40) samples both with and without added copper (II) and zinc (II) metals at pH=6.5, each made up of an average of 4 repeats. There are clear differences between the features in each spectrum, which can be related to the effects of metal binding on the morphology of Aβ. The assignments made are corroborated by the literature, demonstrating the effects of the addition of metal ions to amyloid beta structure.

Specifically, the spectrum of Aβ(1-40) exhibits a column of intense cross-peaks at ω_α=1615 cm^{-1} characteristic of rigid β-sheets formed in fibrils present in the Aβ(1-40) spectrum (Figure 59) that are show decreased intensity in the EVV 2DIR spectra of Aβ(1-40) with copper (Figure 64) and zinc (Figure 67). This is consistent with observations made in the literature that copper binding disrupts the formation of amyloid beta fibrils, while zinc binding retards fibril formation. There are currently no published structures of Aβ(1-40) with Cu(II) on the PDB, so combining the Aβ(1-40)+Cu(II) EVV 2DIR data with corresponding calculations could give significant structural insight into the metal-binding process that occurs in Aβ.

The appearance of a new peak at ω_α=1585/ω_β=2820 cm^{-1} in the Aβ(1-40)+Zn(II) EVV 2DIR spectrum concurs with the latter observation. This can be attributed to an aspartic acid COO^- stretch coupled to a histidine motion comprising in-plane C-H bending, CN stretching and in-plane N-H bending vibrations, which fits with both literature frequency values and the likely interactions between aspartic acid and histidine when they bind Cu(II).

The preliminary data presented in this chapter once again displays the potential of EVV 2DIR spectroscopy for investigating intermolecular interactions in biological systems, and the versatility it has in exploring both plant protein and mammalian peptide systems.

6. Conclusions & Future Work

6.1 Conclusions

The research reported in this thesis demonstrates the capability of EVV 2DIR spectroscopy to detect herbicides binding to their target protein at a high level of sensitivity. The first EVV 2DIR spectra of key industrial target HPPD both alone and with various herbicides is presented. Difference spectra for the herbicides 329 over the standard frequency range exhibits a statistically significant cross-peak at ω_α=1530/ω_β=2870 cm^{-1} indicative of binding interactions, at 1% of the maximum EVV 2DIR signal. This cross-peak is tentatively assigned to specific vibrational modes that couple upon binding; namely an asymmetric NO stretch at 1530 cm^{-1} on the herbicide, and an in-plane CH bend at 1340 cm^{-1} occurring on the amino acid side-chains of the protein.

The same cross-peak was reproduced over a shifted frequency range which further corroborated the suggested vibrational mode assignments, and new binding peaks were detected at ω_α=1355/ω_β=2455 cm^{-1} and ω_α=1330/ω_β=2710 cm^{-1}. The vibrational modes contributing to the former cross-peak are preliminarily assigned to a symmetric NO stretch on the herbicide coupled with a histidine CN stretch, in-plane bend on the protein, while the latter is comprised of an asymmetric SO stretch at 1330 cm^{-1} on the herbicide coupled with a mode at 1340 cm^{-1} speculatively due to an unassigned vibration occurring on a proximal glutamine residue on the protein.

Additionally, a comparison of difference spectra for HPPD both treated and not treated with iron after purification during protein expression shows the same binding peak at similar intensities for both herbicides, which suggests that treating the HPPD with iron doesn't significantly affect the activity in terms of active binding sites required for HPPD inhibition.

A series of EVV 2DIR spectra is also presented of mammalian peptide amyloid beta (1-40), both alone and with copper and zinc ions added. There are several clear spectral differences when comparing the EVV 2DIR spectrum of Aβ(1-40) alone with that of it added to Cu(II) or to Zn(II). Additionally, the signal intensity of some of the column of cross-peaks centred at ω_α=1615 cm^{-1} attributed to amyloid fibril rigid β-sheet formation decreases upon addition of copper. This suggests that the copper disrupts amyloid fibril formation upon binding Aβ(1-40), in accordance with literature observations. Moreover,

the appearance of a new, cross peak at ω_α=1585/ω_β=2820 cm^{-1} overlapping with the ω_α=1615 cm^{-1} feature can be attributed to the presence of zinc retarding amyloid fibril formation, again agreeing with the literature.

6.2 Future work

6.2.1 Future work with HPPD

To confirm the tentative assignments made here for the vibrational modes giving rise to HPPD-herbicide binding specific cross-peaks, the binding of herbicides to HPPD protein with mutated residues could be investigated. Mutant HPPD protein (with residues selected for successive mutation by their proximity to the nitro- functional groups) can be produced. If the EVV 2DIR difference spectra cross-peaks previously assigned to binding-specific interactions are not present in any such mutant HPPD-herbicide complexes, this will enable confirmed assignment of which vibrational mode (arising from a specific amino acid residue) interacts with the nitro- groups. Calculated spectra will also aid in confirming the assignments.

Once the assignments have been confirmed, EVV 2DIR experiments with a reversed polarisation of the third beam should be performed along with corresponding calculated spectra. By comparing the peak intensity ratios in the two different polarisation regimes, experimental spectra can be compared with calculated spectra to validate the geometry of the HPPD-herbicide complexes.

EVV 2DIR spectra of mammalian HPPD (such as the rat variant) should also be collected, to enable a comparison between mammalian and plant HPPD. If spectral differences can be attributed to differences in the C-terminus, this may give information relating to differences in binding affinities between the two species of enzyme. Finally, further HPPD candidates can be screened for binding interactions through EVV 2DIR experiments and accompanying calculations.

6.2.2 Future work with amyloid beta

In the work presented in this thesis, we have proposed that spectral features appearing or weakening from the addition of copper and zinc to amyloid beta are due to the prevention or retardation of Aβ fibrils, respectively. To verify this, EVV 2DIR experiments should be performed on samples of amyloid beta that has been mutated such that it is unable to form fibrils. This will enable comparison between the spectra of non-fibril forming Aβ(1-40) and Aβ(1-40) bound to metals. Short 5-residue peptides with the sequence KLVFF (which comprise residues 1-16 in full-length amyloid beta) can bind to Aβ(1-40) and prevent it from assembling into fibrils[156]. Additionally, a truncated version of the native amyloid beta peptide containing only the N-terminal 1-16 residues has been shown to maintain the capability to bind copper[157]. EVV 2DIR spectra of both these AB variants should be collected in order to confirm the tentative assignments made here.

Additionally, calculated EVV 2DIR spectra should be obtained to further confirm experimental-based assignments. At the time of writing, no X-ray crystallography or 2D NMR structures for any AB-Cu complexes have been published in the PDB. Combining calculated EVV 2DIR spectra with the amyloid beta data presented may therefore reveal significant structural insights into the coordination of copper by amyloid beta.

Once the assignments proposed for the spectra in this thesis have been confirmed, EVV 2DIR spectra of amyloid beta (1-40) with and without metals can be collected with varying conditions, such as pH, concentration and ratios of Aβ:metals. These three conditions are known to affect amyloid beta fibril formation.

6.3 Summary

This thesis fulfils its primary aim to detect – and make preliminary assignments of – binding interactions between the herbicide target HPPD and its inhibitors, along with intermolecular interactions between the potential therapeutic target amyloid beta and metal ions known to affect its morphology. While the work presented has limitations and requires further work to optimise the experimental methods and verify the preliminary assignments made, it demonstrates the potential for EVV 2DIR as a probe to study intermolecular interactions in a range of biomolecular systems.

Bibliography

1 L. Pauling, *Electron. Med. Dig.*, 1949, 35–36.

2 The Association of the British Pharmaceutical Industry, *abpi.org.uk*, 2018.

3 W. Rueegg, M. Quadranti and A. Zoschke, *Weed Res.*, 2007, **47**, 271–275.

4 S. Begemann, *Agweb.com*, 2017.

5 T. C. Sparks and B. A. Lorsbach, *Pest Manag. Sci.*, 2017, **73**, 672–677.

6 S. O. Duke, *Pest Manag. Sci.*, 2012, **68**, 505–12.

7 M. S. Smyth and J. H. J. Martin, *J. Clin. Pathol. - Mol. Pathol.*, 2000, **53**, 8–14.

8 M. Pellecchia, I. Bertini, D. Cowburn, C. Dalvit, E. Giralt, W. Jahnke, T. L. James, S. W. Homans, H. Kessler, C. Luchinat, B. Meyer, H. Oschkinat, J. Peng, H. Schwalbe and G. Siegal, *Nat. Rev. Drug Discov.*, 2008, **7**, 738–745.

9 A. C. Fogarty and D. Laage, *J. Phys. Chem.*, 2013, **118**, 7715–7729.

10 S. Martini, C. Bonechi, A. Foletti and C. Rossi, *Sci. World J.*, 2013, **2013**.

11 R. G. Bryant, *Annu. Rev. Biophys. Biomol. Struct*, 1996, **25**, 29–53.

12 Y. Qin, L. Wang and D. Zhong, *Proc. Natl. Acad. Sci.*, 2016, **113**, 8424–8429.

13 K. R. Acharya and M. D. Lloyd, *Trends Pharmacol. Sci.*, 2005, **26**, 10–14.

14 P. I. Haris and F. Severcan, *J. Mol. Catal. - B Enzym.*, 1999, **7**, 207–221.

15 H. Friebolin, *Basic one- and two-dimensional NMR spectroscopy*, VCH Verlagsgesellschaft, 1991.

16 V. Kanelis, J. D. Forman-kay and L. E. Kay, *IUBMB Life*, 2001, **52**, 291–302.

17 M. T. Zanni, *Proc. Natl. Acad. Sci.*, 2016, **113**, 4890–4891.

18 P. M. Skrabal, *Spectroscopy: An interdisciplinary integral description of spectroscopy from UV to NMR.*, vdf Hochschulverlag AG an der ETH Zürich, 2012.

19 J. H. Lee, Y. Okuno and S. Cavagnero, *J Magn Reson*, 2015, **241**, 18–31.

20 P. Hamm, J. Helbing and J. Bredenbeck, *Annu. Rev. Phys. Chem.*, 2008, **59**, 291–317.

21 S. Woutersen, Y. Mu, G. Stock and P. Hamm, *Chem. Phys.*, 2001, **266**, 137–147.

22 N. T. Hunt, *Chem. Soc. Rev.*, 2009, **38**, 1837–1848.

23 R. Ishima, *Appl NMR Spectrosc*, 2015, **1**, 143–181.

24 Bruker BioSpin, *theresonance.com*, 2018.

25 W. Becker, K. C. Bhattiprolu, N. Gubensäk and K. Zangger, *ChemPhysChem*, 2018, **19**, 895–906.

26 T. A. A. Oliver, *R. Soc. open sci.*, 2018, **5**, 171425.

27 F. D. Fuller, D. E. Wilcox and J. P. Ogilvie, *Opt. Express*, 2014, **22**, 1018.

28 M. Khalil, N. Demirdöven and A. Tokmakoff, *J. Phys. Chem. A*, 2003, **107**, 5258–5279.

29 G. S. Engel, T. R. Calhoun, E. L. Read, T.-K. Ahn, T. Mančal, Y.-C. Cheng, R. E. Blankenship and G. R. Fleming, *Nature*, 2007, **446**, 782–786.

30 T. R. Calhoun, N. S. Ginsberg, G. S. Schlau-Cohen, Y.-C. Cheng, M. Ballottari, R. Bassi and G. R. Fleming, *J. Phys. Chem. B*, 2009, **113**, 16291–16295.

31 G. D. Scholes, G. Fleming, L. X. chen, alán aspuru-Guzik, andreas Buchleitner, D. F. coker, G. S. Engel, rienk van Grondelle, akihito ishizaki, D. M. Jonas, J. S. Lundeen, J. K. Mccusker, S. Mukamel, J. P. Ogilvie, alexandra Olaya-castro, M. ratner, F. Spano, K. Birgitta whaley and X. Zhu, 2017.

32 S. T. Cundiff, A. D. Bristow, M. Siemens, H. Li, G. Moody, D. Karaiskaj, X. Dai and T. Zhang, *IEEE J. Sel. Top. Quantum Electron.*, 2012, **18**, 318–328.

33 V. M. Huxter, T. A. A. Oliver, D. Budker and G. R. Fleming, *Nat. Phys.*, 2013, **9**, 744–749.

34 A. A. Bakulin, O. Selig, H. J. Bakker, Y. L. A. Rezus, C. Müller, T. Glaser, R. Lovrincic, Z. Sun, Z. Chen, A. Walsh, J. M. Frost and T. L. C. Jansen, *J. Phys. Chem. Lett.*, 2015, **6**, 3663–3669.

35 H. S. Chung, Z. Ganim, K. C. Jones and A. Tokmakoff, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 14237–42.

36 S.-H. Shim, R. Gupta, Y. L. Ling, D. B. Strasfeld, D. P. Raleigh and M. T. Zanni, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 6614–9.

37 J. K. Chung, M. C. Thielges and M. D. Fayer, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 3578–83.

38 P. Pagano, Q. Guo, A. Kohen and C. M. Cheatum, *J. Phys. Chem. Lett.*, 2016, **7**, 2507–2511.

39 A. Ghosh, J. Wang, Y. S. Moroz, I. V. Korendovych, M. Zanni, W. F. Degrado, F. Gai and R. M. Hochstrasser, *J. Chem. Phys.*, 2014, **140**, 1–9.

40 S. McCoy and W. S. Caughey, *Biochemistry*, 1970, **9**, 2387–2393.

41 A. Ghosh, J. S. Ostrander and M. T. Zanni, *Chem Rev.*, 2017, **117**, 10726–10759.

42 C. Fang, J. D. Bauman, K. Das, A. Remorino, E. Arnold and R. M. Hochstrasser, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 1472–1477.

43 D. G. Kuroda, J. D. Bauman, J. R. R. Challa, D. Patel, T. Troxler, K. Das, E. Arnold and R. M. Hochstrasser, *Nat Chem.*, 2013, **5**, 174–181.

44 R. Bloem, K. Koziol, S. A. Waldauer, B. Buchli, R. Walser, B. Samatanga, I. Jelesarov and P. Hamm, *J. Phys. Chem. B*, 2012, **116**, 13705–13712.

45 B. Stucki-Buchli, P. J. M. Johnson, O. Bozovic, C. Zanobini, K. L. Koziol, P. Hamm, A. Gulzar, S. Wolf, S. Buchenberg and G. Stock, *J. Phys. Chem. A*, 2017, **121**, 9435–9445.

46 B. Buchli, S. A. Waldauer, R. Walser, M. L. Donten, R. Pfister, N. Blochliger, S. Steiner, A. Caflisch, O. Zerbe and P. Hamm, *Proc. Natl. Acad. Sci.*, 2013, **110**, 11725–11730.

47 J. Ma, I. M. Pazos, W. Zhang, R. M. Culik and F. Gai, *Annu Rev Phys Chem.*, 2015, **66**, 357–377.

48 J. C. Wright, *Int. Rev. Phys. Chem.*, 2002, **21**, 185–255.

49 H. R. Sowley, PhD thesis, Imperial College London, 2017.

50 A. van Almsick, *Outlooks Pest Manag.*, 2009, **20**, 27–30.

51 G. R. Moran, *Arch. Biochem. Biophys.*, 2005, **433**, 117–28.

52 W. E. Knox, *Methods Enzymol.*, 1955, **2**, 287–300.

53 T. W. Goodwin and E. I. Mercer, *Introduction to plant biochemistry*, Pergamon Press, Sydney, 2nd edn., 1983.

54 I. M. Fritze, L. Linden, J. Freigang, G. Auerbach, R. Huber and S. Steinbacher, *Plant Physiol.*, 2004, **134**, 1388–1400.

55 C. Yang, J. W. Pflugrath, D. L. Camper, M. L. Foster, D. J. Pernich and T. a. Walsh, *Biochemistry*, 2004, **43**, 10414–10423.

56 G. R. Moran, *Arch. Biochem. Biophys.*, 2014, **544**, 58–68.

57 T. Borowski, A. Bassan and P. E. M. Siegbahn, *Biochemistry*, 2004, **43**, 12331–12342.

58 J. Brownlee, P. He, G. R. Moran and D. H. T. Harrison, *Biochemistry*, 2008, **47**, 2002–2013.

59 F. Ndikuryayo, B. Moosavi, W. C. Yang and G. F. Yang, *J. Agric. Food Chem.*, 2017, **65**, 8523–8537.

60 C. Cheignon, M. Tomas, D. Bonnefont-Rousselot, P. Faller, C. Hureau and F. Collin, *Redox Biol.*, 2018, **14**, 450–464.

61 C. C. Curtain, F. E. Ali, D. G. Smith, A. I. Bush, C. L. Masters and K. J. Barnham, *J. Biol. Chem.*, 2003, **278**, 2977–2982.

62 J. Nasica-Labouze, P. H. Nguyen, F. Sterpone, O. Berthoumieu, N. V. Buchete, S. Coté, A. De Simone, A. J. Doig, P. Faller, A. Garcia, A. Laio, M. S. Li, S. Melchionna, N. Mousseau, Y. Mu, A. Paravastu, S. Pasquali, D. J. Rosenman, B. Strodel, B. Tarus, J. H. Viles, T. Zhang, C. Wang and P. Derreumaux, *Chem. Rev.*, 2015, **115**, 3518–3563.

63 J.-X. Lu, W. Qiang, W.-M. Yau, C. D. Schwieters, S. C. Meredtih, R. Tycko, S. C. Meredith and R. Tycko, *Cell*, 2013, **154**, 1257–1268.

64 S. A. Gravina, L. Ho, C. B. Eckman, K. E. Long, L. Otvos, L. H. Younkin, N. Suzuki and S. G. Younkin, *J. Biol. Chem.*, 1995, 270, 7013–7016.

65 M. . Lovell, J. . Robertson, W. . Teesdale, J. . Campbell and W. . Markesbery, *J. Neurol. Sci.*, 1998, **158**, 47–52.

66 R. Aleksis, F. Oleskovs, K. Jaudzems, J. Pahnke and H. Biverstål, *Biochimie*, 2017, **140**, 176–192.

67 K. Wild, A. August, C. U. Pietrzik and S. Kins, *Front. Mol. Neurosci.*, 2017, **10**, 1–12.

68 P. Faller, C. Hureau and O. Berthoumieu, *Inorg. Chem.*, 2013, **52**, 12193–12206.

69 M. Mold, L. Ouro-Gnao, B. M. Wieckowski and C. Exley, *Sci. Rep.*, 2013, **3**, 1–6.

70 R. Kayed, E. Head, J. L. Thompson, T. Mcintire, S. C. Milton, C. W. Cotman and C. G. Glabe, *Science (80-.).*, 2003, **300**, 486–489.

71 M. Innocenti, E. Salvietti, M. Guidotti, A. Casini, S. Bellandi, M. L. Foresti, C. Gabbiani, A. Pozzi, P. Zatta and L. Messori, *J. Alzheimer's Dis.*, 2010, **19**, 1323–1329.

72 A. Abelein, A. Gräslund and J. Danielsson, *Proc. Natl. Acad. Sci.*, 2015, **112**, 5407–5412.

73 J. Dong, J. E. Shokes, R. A. Scott and D. G. Lynn, *J Am Chem Soc*, 2006, **128**, 3540–3542.

74 Y. Miller, B. Ma and R. Nussinov, *Proc. Natl. Acad. Sci.*, 2010, **107**, 9490–9495.

75 Y. S. Kim, L. Liu, P. H. Axelsen and R. M. Hochstrasser, *Proc. Natl. Acad. Sci.*, 2008, **105**, 7720–7725.

76 J. P. Lomont, J. S. Ostrander, J. J. Ho, M. K. Petti and M. T. Zanni, *J. Phys. Chem. B*, 2017, **121**, 8935–8945.

77 J. P. Lomont, K. L. Rich, M. Maj, J. J. Ho, J. S. Ostrander and M. T. Zanni, *J. Phys. Chem. B*, 2018, **122**, 144–153.

78 P. Atkins and R. Friedman, *Molecular Quantum Mechanics*, Oxford University Press, Oxford, 5th edn., 2011.

79 S. M. Blinder, *Introduction to quantum mechanics : in chemistry, materials science, and biology*, Elsevier, 2004.

80 B. H. Stuart, *Infrared spectroscopy : fundamentals and applications*, Weinheim : VCH, 1991.

81 S. Califano, *Vibrational states*, Wiley, 1976.

82 P. M. Donaldson, PhD thesis, Imperial College London, 2007.

83 P. M. Donaldson, R. Guo, F. Fournier, E. M. Gardner, L. M. C. Barter, C. J. Barnett, I. R. Gould, D. R. Klug, D. J. Palmer and K. R. Willison, *J. Chem. Phys.*, 2007, **127**, 114513.1-114512.10.

84 R. Guo, F. Fournier, P. M. Donaldson, E. M. Gardner, I. R. Gould and D. R. Klug, *Phys. Chem. Chem. Phys.*, 2009, **11**, 8417–21.

85 S. (Shaul) Mukamel, *Principles of nonlinear optical spectroscopy*, Oxford University Press, New York, 1st edn., 1995.

86 Y. R. Shen, *The principles of nonlinear optics*, Wiley, New York, 1984.

87 C. L. Tang, in *Handbook of Optics: Devices, Measurements, & properties*, ed. M. Bass, McGraw Hill, Inc., 2nd edn., 1995, vol. 2, p. 38.1-38.26.

88 G. New, ed. Michael Bass, Cambridge University Press, Cambridge, 2011, pp. 1–18.

89 J. D. Graybeal, G. S. Hurst, J. O. Stoner and S. Chu, *Encycl. Br.*, 2018.

90 P. A. Franken and J. F. Ward, *Rev. Mod. Phys.*, 1963, **35**, 23–29.

91 J. C. Wright, W. Zhao, K. Murdoch, D. Besemann, N. Condon and K. Meyer, in *Handbook of Vibrational Spectroscopy*, eds. P. Griffiths and J. M. Chalmers, John Wiley & Sons, Ltd, 2006, pp. 853–865.

92 M. S. Y. Cheung, PhD thesis, Imperial College London, 2014.

93 E. M. Gardner, PhD thesis, Imperial College London, 2009.

94 R. C. Jone, *Adv. Electron. Electron Phys.*, 1959, **11**, 87–183.

95 P. M. Donaldson, R. Guo, F. Fournier, E. M. Gardner, I. R. Gould and D. R. Klug, *Chem. Phys.*, 2008, **350**, 201–211.

96 J. C. Wright, *Annu. Rev. Phys. Chem.*, 2011, **62**, 209–230.

97 K. M. Murdoch, N. J. Condon, W. Zhao, D. M. Besemann, K. A. Meyer and J. C. Wright, *Chem. Phys. Lett.*, 2001, **335**, 349–356.

98 R. Paschotta, Phase Matching, *Encycl. laser Phys. Technol.* 1. ed October 2008, Wiley-VCH, ISBN 978-3-527-40828-3, 2008.

99 G. Huo, H. Chen and M. Zhang, *J. Lasers, Opt. Photonics*, 2016, **03**, 1–5.

100 H. Suchowski, K. O'Brien, Z. J. Wong, A. Salandrino, X. Yin and X. Zhang, *Science (80-.).*, 2013, **342**, 1223–1226.

101 N. Boeuf, D. Branning, I. Chaperot, E. Dauler, S. Geurin, G. Jaeger, A. Muller and A. Migdal, *Opt. Eng.*, 2000, **39**, 1016–1024.

102 W. Zhao and J. C. Wright, *Phys. Rev. Lett.*, 2000, **84**, 1411–1414.

103 W. Zhao and J. C. Wright, *Phys. Rev. Lett.*, 1999, **83**, 1950–1953.

104 K. Kwak, S. Cha, M. Cho and J. C. Wright, *J. Chem. Phys.*, 2003, **118**, 2968.

105 P. Hamm and M. T. Zanni, *Concepts and Methods of 2D Infrared Spectroscopy*, Cambridge University Press, 2011.

106 F. Fournier, E. M. Gardner, R. Guo, P. M. Donaldson, L. M. C. Barter, D. J. Palmer, C. J. Barnett, K. R. Willison, I. R. Gould and D. R. Klug, *Anal. Biochem.*, 2008, **374**, 358–365.

107 F. Fournier, E. M. Gardner, D. a Kedra, P. M. Donaldson, R. Guo, S. a Butcher, I. R. Gould, K. R. Willison and D. R. Klug, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 15352–15357.

108 J. R. Lee, J. K. Kim, S. J. Lee and K. P. Kim, *Arch. Pharm. Res.*, 2009, **32**, 1109–1118.

109 L. Rezende Valim, J. A. Davies, K. Tveen Jensen, R. Guo, K. R. Willison, C. M. Spickett, A. R. Pitt and D. R. Klug, *J. Phys. Chem. B*, 2014, **118**, 12855–12864.

110 H. Han and L. H. Hurley, *Trends Pharmacol. Sci.*, 2000, **21**, 136–42.

111 F. Fournier, R. Guo, E. M. Gardner, P. M. Donaldson, C. Loeffeld, I. R. Gould, K. R. Willison and D. R. Klug, *Acc. Chem. Res.*, 2009, **42**, 1322–31.

112 M. D. Morris and D. J. Wallan, *Anal. Chem.*, 1979, **51**, 182A–192A.

113 F. Fournier, R. Guo, E. M. Gardner, P. M. Donaldson, C. Loeffeld, I. R. Gould, K. R. Willison and D. R. Klug, *Acc. Chem. Res.*, 2009, **42**, 1322–1331.

114 R. Paschotta, Intracavity Frequency Doubling, *Encycl. laser Phys. Technol.* 1. ed October 2008, Wiley-VCH, ISBN 978-3-527-40828-3, 2008.

115 D. J. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenb, 2009.

116 M. Mohammadi, G. McMahon, L. Sun, C. Tang, P. Hirth, B. K. Yeh, S. R. Hubbard and J. Schlessinger, *Science (80-.).*, 1997, **276**, 955–60.

117 Spectra-Physics, *Tsunami User's Manual*, Spectra-Physics, Mountain View, California, 2002.

118 Spectra-Physics, *Millennia Pro s-Series User's Manual*, Spectra-Physics, Mountain View, California, 2006.

119 Spectra-Physics, *Spitfire Pro User Manual*, Spectra-Physics, Mountain View, California, 2008.

120 Spectra-Physics, *Empower User's Manual*, Spectra-Physics, Mountain View, California, 2005.

121 R. Paschotta, Regenerative Amplifiers, *Encycl. laser Phys. Technol.* 1. ed October 2008, Wiley-VCH, ISBN 978-3-527-40828-3, 2008.

122 C. Li, in *Nonlinear Optics: Principles and Applications*, Springer Singapore, Singapore, 2017, pp. 109–147.

123 M. Premaratne and G. Agrawal, in *Light Propagation in Gain Media: Optical Amplifiers*, Cambridge University Press, Cambridge, 2011, pp. 208–236.

124 Spectra-Physics, *OPA-800C Ultrafast Optical Parametric Amplifier*, 1999.

125 National Institute of Standards and Technology (NIST), *NIST Chem. Webb.*, 2017.

126 Edmund Optics, *Calcium Fluoride Opt.*, 2018.

127 L. M. Hanssen and C. Zhu, *Handb. Vib. Spectrosc.*, 2006, 1–10.

128 R. A. Nyquist, C. L. Putzig, M. A. Leugers, R. D. McLachlan and B. Thill, *Appl. Spectrosc.*, 1992, **46**, 981–987.

129 R. D. Deegan, O. Bakajin, T. F. Dupont, G. Huber, S. R. Nagel and T. A. Witten, *Nature*, 1997, **389**, 827–829.

130 L. Greenspan, *J. Res. Natl. Bur. Stand. Sect. A Phys. Chem.*, 1977, **81A**, 89–96.

131 I. Cordero, *Community eye Heal.*, 2010, **23**, 57.

132 J. K. Towns, *J. Chromatogr. A*, 1995, **705**, 115–127.

133 A. Bax, *Annu. Rev. Biochem.*, 1989, **58**, 223–256.

134 M. W. Parker, *J. Biol. Phys.*, 2003, **29**, 341–362.

135 K. Suguna, *Resonance*, 2015, **19**, 1093–1103.

136 M. Matringe, A. Sailland, B. Pelissier, A. Rolland and O. Zink, *Pest Manag. Sci.*, 2005, **61**, 269–276.

137 A. Santucci, G. Bernardini, D. Braconi, E. Petricci and F. Manetti, *J. Med. Chem.*, 2017, **60**, 4101–4125.

138 P. T. Wingfield, in *Current Protocols in Protein Science*, John Wiley & Sons, Inc., Hoboken, NJ, USA, 1995, vol. 00, p. A.3A.1-A.3A.4.

139 S. O. Ugwu and S. P. Apte, *Pharm. Technol.*, 2004, **28**, 86–109.

140 S. Lindstedt, E. Holme, E. A. Lock, O. Hjalmarson and B. Strandvik, *Lancet (London, England)*, 1992, **340**, 813–7.

141 A. Barth and C. Zscherp, *Q. Rev. Biophys.*, 2002, **35**, 369–430.

142 R. Guo, M. Miele, E. M. Gardner, F. Fournier, K. M. Kornau, I. R. Gould and D. R. Klug, *Faraday Discuss.*, 2011, **150**, 161–174.

143 A. Barth, *Biochim. Biophys. Acta*, 2007, **1767**, 1073–1101.

144 H. Ahrens, G. Lange, T. Müller, C. Rosinger, L. Willms and A. van Almsick, *Angew. Chem. Int. Ed. Engl.*, 2013, **52**, 9388–98.

145 A. Barth, *Prog. Biophys. Mol. Biol.*, 2000, **74**, 141–173.

146 K. Rogers, *Encycl. Br.*, 2018.

147 N. Yako, T. R. Young, J. M. Cottam Jones, C. A. Hutton, A. G. Wedd and Z. Xiao, *Metallomics*, 2017, **9**, 278–291.

148 K. . Viola, P. . Velasco, P. . Lacor and K. . Wilcox, 2018, 1–5.

149 W. B. Stine, K. N. Dahlgren, G. A. Krafft and M. J. LaDu, *J. Biol. Chem.*, 2003, **278**, 11612–11622.

150 S. D. Moran and M. T. Zanni, *J. Phys. Chem. Lett.*, 2014, 1984–1993.

151 Y. Su and P. T. Chang, *Brain Res.*, 2001, **893**, 287–91.

152 L. Gremer, D. Schölzel, C. Schenk, E. Reinartz, J. Labahn, R. B. G. Ravelli, M. Tusche, C. Lopez-iglesias, W. Hoyer, H. Heise, D. Willbold and G. F. Schröder, 2017, **358**, 116–119.

153 W. Xi and U. H. E. Hansmann, *J. Chem. Phys.*, 2018, **148**, 045103-1-045103-7.

154 B. Alies, A. Conte-Daban, S. Sayen, F. Collin, I. Kieffer, E. Guillon, P. Faller and C. Hureau, *Inorg. Chem.*, 2016, **55**, 10499–10509.

155 K. Brännström, T. Islam, L. Sandblad and A. Olofsson, *FEBS Lett.*, 2017, **591**, 1167–1175.

156 L. O. Tjernberg, J. Näslund, F. Lindqvist, J. Johansson, a R. Karlström, J. Thyberg, L. Terenius and C. Nordstedt, *J. Biol. Chem.*, 1996, **271**, 8545–8548.

157 Q. F. Ma, J. Hu, W. H. Wu, H. D. Liu, J. T. Du, Y. Fu, Y. W. Wu, P. Lei, Y. F. Zhao and Y. M. Li, *Biopolymers*, 2006, **83**, 20–31.

Appendix A: MATLAB codes

A.1 Smoothing code:

```
directory = dir;
mkdir('smoothed_data')



for i = 1:length(directory)
    if ~isdir(directory(i,1).name)
        disp(directory(i,1).name)
        data_raw = load(directory(i,1).name);
        data_smoothed = smooth2a(data_raw,1);
        cd smoothed_data
        dlmwrite([directory(i,1).name,'_smoothed'],data_smoothed)
        cd ..
    end
end
```

A.2 Background removal code:

```
directory = dir;
mkdir('noBG')

for i = 3:length(directory)
    disp(directory(i,1).name)
    data_raw = load(directory(i,1).name);

    bg=mean(mean(data_raw(10:20,85:95)))
    %bg=bg-zbg

    data_noBG = data_raw-bg;
    cd noBG
    dlmwrite([directory(i,1).name,'_noBG'],data_noBG)
    cd ..
end

%y values then x values (ie column, then row)
```

A.3 OPA energy normalisation code:

This code requires a matrix to be produced of the two OPA energy profiles, and saved as "normat". The code then calls this matrix, and divides each of the EVV 2DIR matrices by the OPA matrix normat.

```
directory = dir;
mkdir('data_opa_norm')
```

```
%normat=transpose(normat)

for i = 3:length(directory)
    disp(directory(i,1).name)
    data_raw = load(directory(i,1).name);
    data_opa_cor = data_raw./normat;
    cd data_opa_norm
    dlmwrite([directory(i,1).name,'_opa_norm'],data_opa_cor)
    cd ..
end
```

A.4 Normalisation code:

```
directory = dir;
mkdir('normalised_data')


for i = 1:length(directory)
    if ~isdir(directory(i,1).name)
         disp(directory(i,1).name)
         data_raw = load(directory(i,1).name);
         particular_value = data_raw([167],[47]);
         data_normalised = data_raw/particular_value;
         cd normalised_data
         dlmwrite([directory(i,1).name,'_normalised'],data_normalised)
         cd ..
    end
end

%type [y],[x] ie [column],[row]
```

A.5 Lower-frequency range “stripe” removal code

```
%% MATLAB CODE TO GENERATE AND REMOVE THE STRIPE PROFILE FROM EXTENDED
SPECTRUM 2DIR DATA.
% STRIPE PROFILE SPECTRUM IS GENERATED FROM A 1D PROFILE CONSISTING OF
THE
% AVERAGE OF n_average ADJACENT PIXELS AT THE LEFT HAND EDGE OF THE 2D
SPECTRUM

%% Get list of files in folder [must be prefixed with DOVE]
DOVE_file_list = dir(fullfile(pwd, 'DOVE*'));
DOVE_fnames={DOVE_file_list.name};
n_datasets=numel(DOVE_fnames);

mkdir('stripe_sub')

%% import datasets into cell array
for k=1:n_datasets
    DOVE_data{k} = csvread(DOVE_fnames{k})
end
```

```
[Nrows,Ncolumns]= size(DOVE_data{1,1});      %find size of datasets
stripe_matrix=zeros(Nrows,Ncolumns);         %create empty matrix to
hold standard deviations
stripe_profile=zeros(Nrows,1);               %create column vector for
stripe profile
x_j=zeros(Ncolumns,1);                       %create column vector to
temporarily hold all x values for each spectrum
n_average=5;                                 %specify number of pixels
to average over

%% Loop over N datasets, extract 1D stripe profiles and propagate into
a 2D spectrum for subtraction
for i=1:n_datasets
    for j=1:length(stripe_profile)
        x_j=DOVE_data{1,i}(j,1:end); %temp 1D vector for X values
        stripe_profile(j,1)=mean(x_j(1:n_average)); %calculate average
of specified number of pixels [defined as n_average above]
    end
    for j=1:Ncolumns
        stripe_matrix(1:end,j)=stripe_profile(1:end); % create a 2d
matrix where the stripe profile is propagated across the whole x axis
of the spectrum
    end

    %% Create stripe-subtracted spectra
    stripe_sub=DOVE_data{1,i}-stripe_matrix;
    stripe_div=DOVE_data{1,i}./stripe_matrix;

    %%OPTIONAL: display graphs of raw data, generated profile, and
    %%subtracted or divided profiles
%     figure('Name','stripe_sub','NumberTitle','off');
%     surface(stripe_sub)
%     title('stripe sub')
%     figure('Name','strip_matrix','NumberTitle','off');
%     surface(stripe_matrix)
%     title('stripe matrix')
%     figure('Name','Raw Data','NumberTitle','off');
%     surface(DOVE_data{1,i})
%     title('Raw DATA')
%     figure('Name','divide by stripe','NumberTitle','off');
%     surface(stripe_div)
%     title('stripe divide')
    cd stripe_sub
    dlmwrite([DOVE_file_list(i,1).name, '_stripesub'], stripe_sub);
    cd ..
end
```

Appendix B: *Arabidopsis thaliana* HPPD sequence

MHHHHHHFSKFVRKNPKSDKFKVKRFHHIEFWCGDATNVARRFSWGLGMRFSAKSDLSTGNM
VHASYLLTSGDLRFLFTAPYSPSLSAGEIKPTTTASIPSFDHGSCRSFFSSHGLGVRAVAIEVEDAES
AFSISVANGAIPSSPPIVLNEAVTIAEVKLYGDVVLRYVSYKAEDTEKSEFLPGFERVEDASSFPLDY
GIRRLDHAVGNVPELGPALTYVAGFTGFHQFAEFTADDVGTAESGLNSAVLASNDEMVLLPINEP
VHGTKRKSQIQTYLEHNEGAGLQHLALMSEDIFRTLREMRKRSSIGGFDFMPSPPPTYYQNLKKR
VGDVLSDDQIKECEELGILVDRDDQGTLLQIFTKPLGDRPTIFIEIIQRVGCMMKDEEGKAYQSGG
CGGFGKGNFSELFKSIEEYEKTLEAKQLVG

Appendix C: Amyloid beta 1-40 peptide sequence

DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVV