


# Computing Protein Near–UV Circular Dichroism Spectra

Sarah Badri Jasim (M.Sci.)

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

School of Chemistry

**June 2019**

## Abstract

A fully quantitative theory connecting protein conformation and optical spectroscopy would facilitate deeper insights into biophysical and simulate studies of protein dynamics and folding. The regular, repeating, chiral nature of secondary structure elements leads to intense bands in the far-ultraviolet (UV). The near-UV bands are much weaker and have been challenging to compute theoretically. We report some advances in the accuracy of calculations in the near-UV by using the vibrational structure parameters in the circular dichroism (CD) spectra of proteins.

The near-UV CD region is sensitive to electronic transitions in aromatic amino acids. This can be used to study local structural changes arising from processes such as a ligand binding to a protein. The improvements of the accuracy of the calculated spectra have been assessed over a set of eight proteins. The near-UV CD spectra have been calculated with the new parameters describing the aromatic side chain transitions. An improvement has been noticed for some calculated spectra. The contribution of individual chromophores to the CD spectra has been calculated. A detailed analysis has been presented for the tyrosine chromophores and the tryptophan chromophores, separately in different chapters. Further investigations have been conducted to study the interactions which are most important in determining

the near-UV CD spectrum. An examination of the mechanisms has been made to understand the types of mechanisms which participate in generating the calculated spectra of the proteins in this thesis. The one-electron mechanism has the predominant role in generating the calculated spectra in the near-UV region. This mechanism appears mainly in all tyrosine and tryptophan chromophores. Some chromophores have shown coupling and mixing between their transitions through different mechanisms, i.e. the $\mu - m$ mechanism.

Exciton theory is discussed in detail and in the context of computing CD in the near-UV regions. We outline the importance and the influence of electrostatics of the protein environment on the electronic transitions in the chromophores responsible for CD signals in the near-UV. The accuracy of the computational approaches should allow more quantitative studies applying a combination of experimental data and modelling to a variety of interesting questions.

## Acknowledgement

Undertaking this PhD has been a truly life-changing experience for me, and it would not have been possible to do without the support and guidance that I received from many people.

First and foremost, I want to thank my supervisor Prof. Jonathan Hirst. It has been an honour to be his PhD student and I really appreciate all his contributions of time and ideas to make my PhD experience productive.

Besides my supervisor, I would like to thank Dr. David Rogers, Dr. Zhuo Li, and Dr. Hainam Do, who provided their precious explanations and discussions. Without their help it would not be possible to conduct this research.

I would like to thank Dr. Pritesh Tailor and Peter Morgan-Tansley for their support in solving problems I met on computers. I also want to thank former and current group members, for all their support. Thank you for all my friends around the School.

I have been fortunate to meet a lot of friends in Nottingham. I am grateful for their support, especially, Dr. Hadel, Dr. Haneen, Dr. Dhuha, Dr. Atheel, Dr. Maha and Dr. Siham.

I would like to express a sincere acknowledgement to the School of Chemistry, University of Nottingham for providing me the opportunity to complete my PhD study. I also want to thank our local cluster and university High Performance Computing facilities. I would like to express my great appreciation to the Iraqi Ministry of Higher Education and Scientific Research, and the Iraqi Cultural Attaché for the financial support during my study.

I would like to thank my colleagues in Wasit University. I am grateful to Dr. Subhi, Dr. Ahmed, Dr. Ghasan, Dr. Fatima and Dr. Wasan for being my grantors during my study. Special thanks goes to Dr. Hasanin, Dr. Sadiq, Dr. Haider, Dr. Alaa, Dr. Zaid, Dr. Jawad, Dr. Aqeel and Dr. Mohanad for being supportive always.

There are no words to express my heartfelt gratitude to my beloved parents, brothers, sisters, and all my relatives and friends in Iraq for their support. I have special people who passed away during my study. I will never see them again, but they will always live in my heart and mind. I am deeply thankful of my lovely husband (Wathiq) and my awesome son (Tariq) for their love, full support, and encouragement along the way and for being the source of my strength whenever I felt weak.

## Publications and Presentations

- Jasim, S.B., Li, Z., Guest, E.E. and Hirst, J.D. DichroCalc: improvements in computing protein circular dichroism spectroscopy in the near-ultraviolet. *J. Mol. Biol.* 2018, *430*, 2196-2202.

- Rogers, D.M., Jasim, S.B., Dyer, N.M., Auvray, F., Refregiers, M. and Hirst, J.D. Electronic circular dichroism spectroscopy of proteins. *Chem*, 2019, *5*, 2751–2774.

- July 2018. RSC TCG Conference, Sheffield. Poster presentation.

- July 2017. CD/LD Summer School, Warwick. Poster presentation.

- July 2018. Midlands Computational Chemistry Meeting. Nottingham. Oral presentation.

# List of contents


| | |
|---|---|
| Abstract | I |
| Acknowledgement | III |
| Publications and Presentations | V |
| List of contents | VI |
| List of figures | VIII |
| List of tables | XV |
| List of Abbreviations | XVIII |
| Chapter 1: Introduction | 1 |
| 1.1. Protein Structure | 1 |
| 1.2. Structural biology determination | 7 |
| 1.2.1 X-ray crystallography | 7 |
| 1.2.2 NMR techniques | 9 |
| 1.2.3 Optical spectroscopy | 10 |
| 1.2.3.1 Absorption and fluorescence in the ultra-violet | 12 |
| 1.2.3.2 Infra-red techniques | 14 |
| 1.2.3.3 Circular dichroism | 15 |
| 1.2.4 Software used for studying CD of proteins | 16 |
| 1.3. Chirality | 18 |
| 1.4. Circular Dichroism | 21 |
| 1.5. Outline of the thesis | 28 |
| 1.6. Aim of the work | 29 |
| Chapter 2: Methods | 31 |
| 2.1. Introduction | 31 |
| 2.2. The matrix method | 32 |
| 2.3. Mechanisms of interactions associated with different types of matrix element | 36 |
| 2.4. Electrostatic interactions from the ground electronic state | 39 |
| 2.5. Franck Condon Principle and vibrational parameters | 41 |
| Chapter 3: Analysis of the tyrosine chromophore in proteins | 47 |
| 3.1 Introduction | 47 |
| 3.2 Background on the proteins studied in this chapter | 49 |

## List of contents


| | |
|---|---|
| 3.2.1 Bovine pancreatic trypsin inhibitor | 49 |
| 3.2.2 Cardiotoxin | 50 |
| 3.2.3 Ribonuclease A | 51 |
| 3.3. Methods | 53 |
| 3.4. Results and discussion | 55 |
| 3.4.1. The role of Tyrosine residues in generating CD spectrum | 55 |
| 3.4.1.1 Bovine pancreatic trypsin inhibitor | 55 |
| 3.4.1.2. Cardiotoxin | 64 |
| 3.4.1.3. Ribonuclease A | 72 |
| 3.5. Influence of electrostatics | 80 |
| 3.6. Conclusion | 88 |
| Chapter 4: Analysis of the tryptophan chromophore in proteins | 90 |
| 4.1 Introduction | 90 |
| 4.2 Background on the proteins studied in this chapter | 92 |
| 4.2.1 Apolipophorin III | 92 |
| 4.2.2 Barnase | 93 |
| 4.2.3 Beta 2 microglobulin human | 95 |
| 4.2.4 Thioredoxin | 96 |
| 4.2.5 Hen egg white lysozyme | 97 |
| 4.3 Results and discussion | 98 |
| 4.3.1 The role of Trp residues in generating CD spectrum | 98 |
| 4.3.1.1 Apolipohorin III | 98 |
| 4.3.1.2 Barnase | 105 |
| 4.3.1.3 Beta 2 microglobulin human | 114 |
| 4.3.1.4 Thioredoxin | 122 |
| 4.3.1.5 Hen egg white lysozyme | 133 |
| 4.4. Influence of electrostatics | 138 |
| 4.5. Conclusion | 141 |
| Chapter 5: Conclusion and Future Work | 143 |
| References and Appendix | |

# List of Figures


| | |
|---|---|
| Figure 1.1: Structure of an amino acid | 2 |
| Figure 1.2: Structures of methyl derivatives, representing Phe, Tyr and Trp, respectively | 2 |
| Figure 1.3: A polypeptide | 3 |
| Figure 1.4: Main chain dihedral angles $\varphi$ and $\psi$ | 4 |
| Figure 1. 5: The regions of the electromagnetic spectrum | 11 |
| Figure 1. 6: The mirror image of a chiral atom represented in grey linked to four different substituents: $\alpha, \beta, \gamma$, and $\delta$ | 19 |
| Figure 1.7: Elliptically polarized light | 24 |
| Figure 2.1: Franck Condon principle for vibronic transitions | 42 |
| Figure 2. 2: The $L_a$ and $L_b$ electronic states for the aromatic amino acid: Tyr, Phe and Trp | 43 |
| Figure 3.1. Bovine pancreatic trypsin inhibitor (1PIT), a model from the family of the NMR structure. Tyrosine in green and phenylalanine in red. | 50 |
| Figure 3.2. Cardiotoxin (1CRF), a model from the family of the NMR structure. Tyrosine in green and phenylalanine in red. | 51 |
| Figure 3.3. Ribonuclease A (2AAS), a model from the family of the NMR structure. Tyrosine in green and phenylalanine in red. | 52 |
| Figure 3.4. Experimental near-UV CD spectrum (black solid line), and spectra calculated using the X-ray structure (blue solid line), calculated using an average over 20 frames of the NMR structure (red solid line) and calculated from individual frames of the NMR structure: frame 3 (red dotted line) and frame 17 (red triangles). | 56 |

## List of Figures


| | |
|---|---|
| Figure 3.5. Conformational diversity investigation. The black solid line is the experimental spectrum the red solid line is the average over 20 models and the red lines with triangles (upper and lower) represent the mean average ± standard deviation, respectively. | 57 |
| Figure 3.6: Calculated contributions for each Tyr at different wavelengths for frame 3 of bovine pancreatic trypsin inhibitor (1PIT). The upper panel shows calculated rotational strengths in (DBM) multiplied by 10000 for each residue at each wavelength. The lower panel shows the calculated spectrum. | 60 |
| Figure 3.7. Experimental near-UV CD spectrum (black solid line), and spectra calculated using the X-ray structure (blue solid line), calculated using an average over 12 frames of the NMR structure (red solid line) and calculated from individual frames of the NMR structure: frame 7 (red triangles) and frame 3 (red dotted line). | 65 |
| Figure 3.8. Conformational diversity investigation. The black solid line is the experimental spectrum, the red solid line is the average over 12 models and the red lines with triangles (upper and lower) represent the mean average ± standard deviation, respectively. | 66 |
| Figure 3.9. Calculated contributions for each Tyr at different wavelength for frame 3 of cardiotoxin (1CRF). The upper panel shows calculated rotational strengths in (DBM) multiplied by 1000 for each residue at each wavelength. The lower panel shows the calculated spectrum. | 69 |
| Figure 3.10. Experimental near-UV CD spectrum (black solid line), and spectra calculated using the X-ray structure (blue solid line), calculated using an average over 32 frames of the NMR structure (red solid line) and calculated from individual | 72 |

## List of Figures


| | |
|---|---|
| frames of the NMR structure: frame 1 (red triangles) and frame 12 (red dotted line). | |
| Figure 3.11. Conformational diversity investigation. The black solid line is the experimental, the red solid line is the average over 32 models and the red lines with triangles (upper and lower) represent the mean average ± standard deviation, respectively. | 74 |
| Figure 3.12. Calculated contributions for each Tyr at different wavelength for frame 12 of ribonuclease A (2AAS). The upper panel shows calculated rotational strengths in (DBM) multiplied by 100 for each residue at each wavelength. The lower panel shows the calculated spectrum. | 77 |
| Figure 3.13. Contour plane through $^{10}$Tyr in frame 3 from the family of the NMR structure of BPTI. The plane is coloured by calculated electrostatic potential of the local environment of the tyrosine residue. It is set to cut through $C\gamma, C\delta1$ and $C\delta2$ of the tyrosine residue. The colours from positive to negative: Blue (+20.0), green, yellow and red (-0.6); with points outside of the contour range (20.0 to -0.6 $k_BT/e_c$) not shown, i.e., black. | 81 |
| Figure 3.14. Contour plane through $^{21}$Tyr in frame 3 from the family of the NMR structure of BPTI. The plane is coloured by calculated electrostatic potential of the local environment of the tyrosine residue. It is set to cut through $C\gamma, C\delta1$ and $C\delta2$ of the tyrosine residue. The colours from positive to negative: Blue (+20.0), green, yellow and red (-6.0); with points outside of the contour range (20.0 to -0.6 $k_BT/e_c$) not shown, i.e., black. | 82 |
| Figure 3.15. Contour plane through $^{23}$Tyr in frame 3 from the family of the NMR structures of BPTI. The plane is coloured by calculated electrostatic potential of the local environment | 83 |

## List of Figures


| | |
|---|---|
| of the tyrosine residue. It is set to cut through $C\gamma, C\delta1$ and $C\delta2$ of the tyrosine residue. The colours from positive to negative: Blue (+20.0), green, yellow and red (-0.6); with points outside of the contour range (20.0 to -6.0 $k_BT/e_c$) not shown, i.e., black. | |
| Figure 3.16. Contour plane through $^{35}$Tyr in frame 3 from the family of the NMR structure of BPTI. The plane is coloured by calculated electrostatic potential of the local environment of the tyrosine residue. It is set to cut through $C\gamma, C\delta1$ and $C\delta2$ of the tyrosine residue. The colours from positive to negative: Blue (+20.0), green, yellow and red (-6.0); with points outside of the contour range (20.0 to -6.0 $k_BT/e_c$) not shown, i.e., black. | 84 |
| Figure 3.17. Superimposed image for frame 3 (in brown) and frame17 (in blue) of BPTI (1PIT) protein. Tyrosine in green and phenylalanine in red. | 86 |
| Figure 4.1. Apolipophorin III (1LS4), a model from the family of the NMR structure. Tryptophan in yellow and phenylalanine in red. | 93 |
| Figure 4.2. Barnase (1FW7), a model from the family of the NMR structure. Tryptophan in yellow, Tyrosine in green and phenylalanine in red. | 94 |
| Figure 4.3. β-2 micoglobulin (1JNJ), a model from the family of the NMR structure. Tryptophan in yellow, Tyrosine in green and phenylalanine in red. | 95 |
| Figure 4.4. Thioredoxin (1XOA), a model from the family of the NMR structure. Tryptophan in yellow, Tyrosine in green and phenylalanine in red. | 96 |

## List of Figures


| | |
|---|---|
| Figure 4.5. Hen egg white lysozyme (1E8L), a model from the family of the NMR structure. Tryptophan in yellow, Tyrosine in green and phenylalanine in red. | 97 |
| Figure 4.6. Experimental near-UV CD spectrum (black solid line), and spectra calculated using the X-ray structure (blue solid line), calculated using an average over 21 frames of the NMR structure (red solid line) and calculated from individual frames of the NMR structure: frame 5 (red triangles) and frame 15 (red dotted line). | 98 |
| Figure 4.7. Conformational diversity investigation. The black solid line is the experimental spectrum. The red solid line is the mean average computed from 21 NMR models and the red lines with triangles (upper and lower) represent the mean average ± standard deviation, respectively. | 100 |
| Figure 4.8. Experimental near-UV CD spectrum (black solid line), and spectra calculated using the X-ray structure (blue solid line), calculated using an average over 20 frames of the NMR structure (red solid line) and calculated from individual frames of the NMR structure: frame 8 (red triangles) and frame 20 (red dotted line). | 106 |
| Figure 4.9. Conformational diversity investigation. The black solid line is the experimental spectrum. The red solid line is the mean average computed from 20 NMR models and the red lines with triangles (upper and lower) represent the mean average ± standard deviation, respectively. | 107 |
| Figure 4.10. Calculated contributions to the intensity for each Trp at different λ for frame 20 of barnase (1FW7). The upper panel shows calculated rotational strengths for each residue at each wavelength were multiplied by 1000. The lower panel shows the calculated spectrum. | 110 |

## List of Figures


| | |
|---|---|
| Figure 4.11. Experimental near-UV CD spectrum (black solid line), and spectra calculated using the X-ray structure (blue solid line), calculated using an average over 20 frames of the NMR structure (red solid line) and calculated from individual frames of the NMR structure: frame 19 (dashed red line with triangles) and frame 16 (red dotted line). | 114 |
| Figure 4.12. Conformational diversity investigation. The black solid line is the experimental spectrum, the red solid line is the average over 20 models and the red lines with triangles (upper and lower) represent the mean average ± standard deviation, respectively. | 115 |
| Figure 4.13. Experimental near-UV CD spectrum (black solid line), and spectra calculated using the X-ray structure (blue solid line), calculated using an average over 20 frames of the NMR structure (red solid line) and calculated from individual frames of the NMR structure: frame 20 (dashed red line with triangles) and frame 5 (red dotted line). | 122 |
| Figure 4.14. Conformational diversity investigation. The black solid line is the experimental spectrum. The red solid line is the mean average computed from 20 NMR models and the red lines with triangles (upper and lower) represent the mean average ± standard deviation, respectively. | 124 |
| Figure 4.15. Calculated contributions to the intensity for each Trp at different λ for frame 20 of thioredoxin (1XOA). The upper panel shows calculated rotational strengths in (DBM) multiplied by 1000 for each residue at each wavelength. The lower panel shows the calculated spectrum. | 126 |
| Figure 4.16. Experimental near-UV CD spectrum (black solid line), and spectra calculated using the X-ray structure (blue solid line), calculated using an average over 50 frames of the NMR structure (red solid line) and calculated from individual | 131 |

## List of Figures


| | |
|---|---|
| frames of the NMR structure: frame 2 (red triangles) and frame 33 (red dotted line). | |
| Figure 4.17: Conformational diversity investigation. The black solid line is the experimental spectrum. The red solid line is the mean average computed from 50 NMR models and the red lines with triangles (upper and lower) represent the mean average ± standard deviation, respectively. | 132 |
| Figure 4.18. Calculated contributions to the intensity for each Trp at different λ for frame 33 of hen egg white lysozyme (1E8L). The upper panel shows calculated rotational strengths in (DBM) multiplied by 1000 for each residue at each wavelength. The lower panel shows the calculated spectrum. | 134 |
| Figure 4.19. Contour plane through $^{115}$Trp in frame 15 from the family of the NMR structure of Apolipophorin III (1LS4). The plane is coloured by calculated electrostatic potential of the local environment of the tryptophan residue. It is set to cut through $C\gamma, C\delta1$ and $C\delta2$ of the tryptophan residue. The colours from positive to negative: Blue (+20.0), green, yellow and red (-0.6); with points outside of the contour range (20.0 to -6.0 $k_BT/e_c$) not shown, i.e., black. | 139 |
| Figure 4.20. Contour plane through $^{130}$Trp in frame 15 from the family of the NMR structure of BPTI. The plane is coloured by calculated electrostatic potential of the local environment of the tryptophan residue. It is set to cut through $C\gamma, C\delta1$ and $C\delta2$ of the tryptophan residue. The colours from positive to negative: Blue (20.0), green, yellow and red (-6.0); with points outside of the contour range (20.0 to -6.0 $k_BT/e_c$) not shown, i.e., black. | 140 |

# List of Tables


| | |
|---|---|
| Table 2.1: Frank-Condon integrals and the transition energies | 44 |
| Table 3.1: The proteins studied in this chapter: name, PDB codes, length (i.e., number of residues in the X-ray structure) and the number of aromatic residues and disulfide bonds. | 49 |
| Table 3.2: Calculated interaction energies (shown only if greater than 1 $cm^{-1}$) of electronic transitions of aromatic side chains in frame 3 of the NMR structure. | 59 |
| Table 3.3: Wavelength and associated eigenvector coefficients of tyrosine side chain in frame 3 of the NMR structure. | 61 |
| Table 3.4: Wavelength and associated eigenvector coefficients of phenylalanine side chain in frame 3 of the NMR structure. | 61 |
| Table 3.5: The eight ($\nu$=1 to 8) tyrosine $^1L_b$ vibrational states, their wavelengths ($\lambda$) and their calculated contributions to the rotational strength ($R_\lambda$) in Debye Bohr Magneton (DBM) at that wavelength in the near-UV CD spectrum of frame 3 of the NMR structure. | 63 |
| Table 3.6: Calculated interaction energies (shown only if greater than 10 $cm^{-1}$) of electronic transitions of aromatic side chains in frame 3 of the NMR structure. | 67 |
| Table 3.7: Wavelength and associated eigenvector coefficients of tyrosine side chain in frame 3 of the NMR structures. | 70 |
| Table 3.8: The eight ($\nu$ =1 to 8) tyrosine $^1L_b$ vibrational states, their wavelengths ($\lambda$) and their calculated contributions to the rotational strength ($R_\lambda$) in Debye Bohr Magneton (DBM) at that wavelength in the near-UV CD spectrum of frame 3 of the NMR structures | 71 |
| Table 3.9: Calculated interaction energies (shown only if greater than 10 $cm^{-1}$) of electronic transitions of aromatic side chains in frame 3 of the NMR structures. | 75 |
| Table 3.10: Wavelength and associated eigenvector coefficients of tyrosine side chain in frame 12 of the NMR structures. | 78 |
| Table 3.11: Wavelength and associated eigenvector coefficients of phenylalanine side chain in frame 12 of the NMR structures. | 78 |
| Table 3.12: The eight ($\nu$ =1 to 8) tyrosine $^1L_b$ vibrational states, their wavelengths ($\lambda$) and their calculated contributions to the rotational strength ($R_\lambda$) in Debye Bohr | 79 |

## List of Tables


| | |
|---|---|
| Magneton (DBM) at that wavelength in the near-UV CD spectrum of frame 12 of the NMR structures. | |
| Table 3.13: Calculated distance between Tyr side chains and other aromatic side chains in frame 3 and frame 17 of 1pit of the NMR structure | 87 |
| Table 4.1: The proteins studied in this chapter: name, PDB codes, length (i.e., number of residues in the X-ray structure) and the number of aromatic residues and disulfide bonds. | 92 |
| Table 4.2: Calculated interaction energies (shown only if greater than 10 $cm^{-1}$) of electronic transitions of aromatic side chains in frame 15 of the NMR structure. | 101 |
| Table 4.3: Wavelength and associated eigenvector coefficients of tryptophan side chains transitions in frame 15 of the NMR structure. | 103 |
| Table 4.4: The seven ($\nu$ =1 to 7) $^1L_b$ and the nine ($\nu$ =1 to 9) tryptophan $^1L_a$ vibrational states, their wavelengths ($\lambda$) and their calculated contributions to the rotational strength ($R_\lambda$) in Debye Bohr Magneton (DBM) - shown only if greater than 1DBM- at that wavelength in the near-UV CD spectrum of frame 15 of the NMR. | 104 |
| Table 4.5: Calculated interaction energies (shown only if greater than 50 $cm^{-1}$) of electronic transitions of aromatic side chains in frame 20 of the NMR structure. | 108 |
| Table 4.6: Wavelength and associated eigenvector coefficients of tryptophan side chain transitions in frame 20 of the NMR structure. | 111 |
| Table 4.7: Wavelength and associated eigenvector coefficients of tyrosine side chain transitions in frame 20 of the NMR structure. | 111 |
| Table 4.8: Wavelength and associated eigenvector coefficients of phenylalanine side chain transitions in frame 20 of the NMR structure. | 112 |
| Table 4.9: The seven ($\nu$ =1 to 7) $^1L_b$ and the nine ($\nu$ =1 to 9) tryptophan $^1L_a$ vibrational states, their wavelengths ($\lambda$) and their calculated contributions to the rotational strength ($R_\lambda$) in Debye Bohr Magneton (DBM) - shown only if greater than 1DBM- at that wavelength in the near-UV CD spectrum of frame 20 of the NMR. | 113 |
| Table 4.10: Calculated interaction energies (shown only if greater than 10 $cm^{-1}$) of electronic transitions of aromatic side chains in frame 16 of the NMR structure. | 117 |
| Table 4.11: Wavelength and associated eigenvector coefficients of tryptophan side chain transitions in frame 16 of the NMR structure. | 118 |

## List of Tables


| | |
|---|---|
| Table 4.12: Wavelength and associated eigenvector coefficients of tyrosine side chain transitions in frame 16 of the NMR structure. | 119 |
| Table 4.13: Wavelength and associated eigenvector coefficients of phenylalanine side chain transitions in frame 16 of the NMR structure. | 119 |
| Table 4.14: The seven ($v$ =1 to 7) $^1L_b$ and the nine ($v$ =1 to 9) tryptophan $^1L_a$ vibrational states, their wavelengths ($\lambda$) and their calculated contributions to the rotational strength ($R_\lambda$) in Debye Bohr Magneton (DBM) at that wavelength in the near-UV CD spectrum of frame 16. | 121 |
| Table 4.15: Calculated interaction energies (shown only if greater than 15 $cm^{-1}$) of electronic transitions of aromatic side chains in frame 3 of the NMR structures. | 125 |
| Table 4.16: Wavelength and associated eigenvector coefficients of tryptophan side chain transitions in frame 5 of the NMR structure. | 127 |
| Table 4.17: Wavelength and associated eigenvector coefficients of phenylalanine side chain transitions in frame 5 of the NMR structure. | 128 |
| Table 4.18: The seven ($v$ =1 to 7) $^1L_b$ and the nine ($v$ =1 to 9) tryptophan $^1L_a$ vibrational states, their wavelengths ($\lambda$) and their calculated contributions to the rotational strength ($R_\lambda$) in Debye Bohr Magneton (DBM)- shown only if greater than 1DBM- at that wavelength in the near-UV CD spectrum of frame 5. | 129 |
| Table 4.19: Calculated interaction energies (shown only if greater than 100 $cm^{-1}$) of electronic transitions of aromatic side chains in frame 33 of the NMR structures. | 133 |
| Table 4.20: Wavelength and associated eigenvector coefficients of tryptophan side chain transitions in frame 33 of the NMR structure. | 135 |
| Table 4.21: Wavelength and associated eigenvector coefficients of phenylalanine side chain transitions in frame 33 of the NMR structure. | 135 |
| Table 4.22: The seven ($v$ =1 to 7) $^1L_b$ and the nine ($v$ =1 to 9) tryptophan $^1L_a$ vibrational states, their wavelengths ($\lambda$) and their calculated contributions to the rotational strength ($R_\lambda$) in Deby Bohr Magneton (DBM)- shown only if greater than 1DBM- at that wavelength in the near-UV CD spectrum of frame 33. | 137 |

## List of Abbreviations

| | |
|---|---|
| Ala | Alanine |
| Arg | Arginine |
| BPTI | Bovine pancreatic trypsin inhibitor |
| CASSCF | Complete active space self-consistent field |
| CD | Circular Dichroism |
| DBM | Debye Bohr Magneton |
| DFT | Density functional theory |
| DNA | Deoxyribonucleic acid |
| ECD | Electronic circular dichroism |
| Gln | Glutamine |
| HPLC | High-performance liquid chromatography |
| IR | Infrared |
| 3-MI | 3-Methylindole |
| MCD | Magnetic circular dichroism |
| MRE | Mean relative error |
| MD | Molecular dynamics |
| nm | nanometer |
| NMR | Nuclear Magnetic Resonance |
| ORD | Optical rotatory dispersion |
| PCDDB | Protein circular dichroism data bank |
| PDB | Protein Data Bank |
| Phe | Phenylalanine |
| Pro | Proline |
| QM | Quantum mechanical |
| RMSD | Root mean square deviation |
| Thr | Threonine |
| TRCD | Time resolved circular dichroism |
| Trp | Tryptophan |
| Tyr | Tyrosine |
| UV | Ultraviolet |
| VCD | Vibrational circular dichroism |

# Chapter 1
# Introduction

## 1.1. Protein Structure

Proteins are ubiquitous in all living organisms no matter whether they are simple or complex. The definition of proteins was presented at the same time as when Friedrich Wöhler in 1828 showed that the decomposition of ammonium cyanate yields urea. The new idea was the generation in living systems of organic compounds, such as urea, from simple inorganic chemicals. It opened the door to many investigations in biochemistry and the discovery of proteins.

Amino acids are the building blocks of proteins. There are 20 commonly and naturally occurring standard amino acids. 18 amino acids consist of carbon, hydrogen, nitrogen and oxygen, while cysteine and methionine also have sulfur. All of them are $\alpha$-amino acids with the general structure shown in Figure 1.1. The structure of an amino acid contains a chiral centre, the $\alpha$-carbon, bonded with a carboxyl group, an amino group, a single hydrogen atom and a side chain group R.

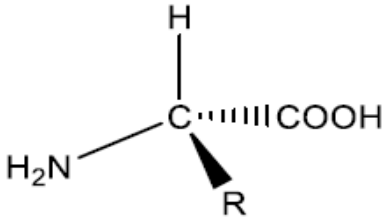



Figure 1.1: Structure of an amino acid

These vary in structure by the side chain R and fall into three general categories: hydrophobic (e.g. alanine), charged (e.g. lysine) and polar (e.g. serine). Glycine (Gly), which is not chiral, is usually considered by itself, because its side chain consists simply of a single hydrogen atom. The different properties of amino acids come from the R group.

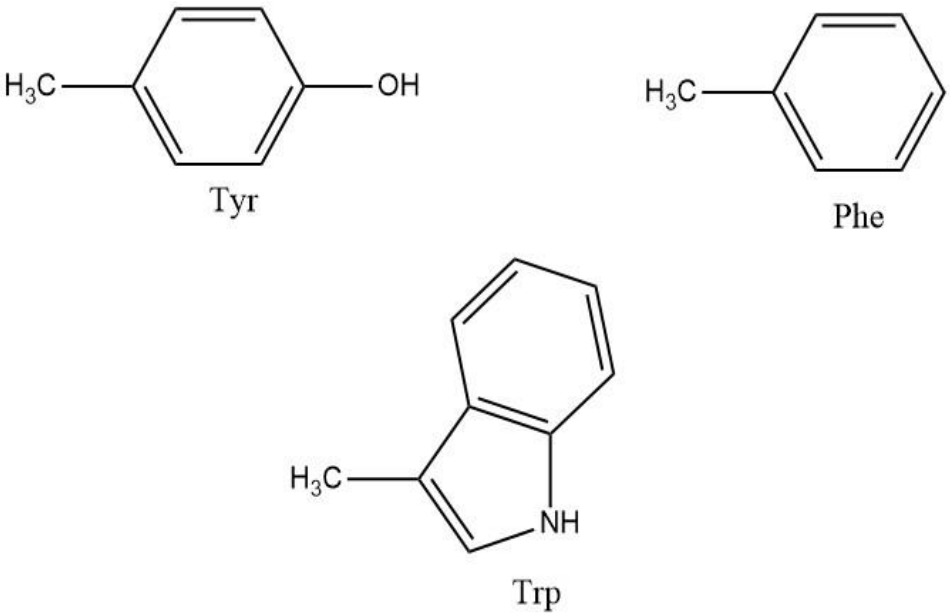



Figure 1.2: Structures of methyl derivatives, representing Phe, Tyr and Trp, respectively

The focus in this thesis is on aromatic side chains of phenylalanine (Phe), tyrosine (Tyr) and tryptophan (Trp). The methyl derivatives of these side chains are shown in Figure 1.2. In this thesis we do not consider the disulfide bond chromophore which is formed between two cysteine side chains. The amino acids are linked together by a peptide bond, resulting from a condensation reaction of the $\alpha$-carboxyl group of one amino acid, and the $\alpha$-amino group of another [1]. That will lead to the creation of a polymer known as a polypeptide. The backbone of a protein is shown in Figure 1.3 with the R group sidechains.

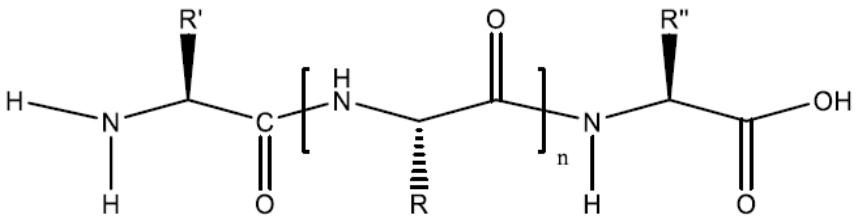



Figure 1.3: A polypeptide

The polypeptide backbone comprises single bonds, either C-C or C-N, which give flexibility and contribute to the variety of complex shapes of proteins. The shape or structure of a protein can be defined by the dihedral angles $\varphi$ and $\psi$, as shown in Figure 1.4, where $\Phi$ is the angle of rotation about the N-$C_\alpha$ bond and $\psi$ is the angle about the C-$C_\alpha$ bond. There is another dihedral angle known as $\omega$, which is $180^0$ for *trans*–peptides and $0^0$ for *cis*-peptides [2].

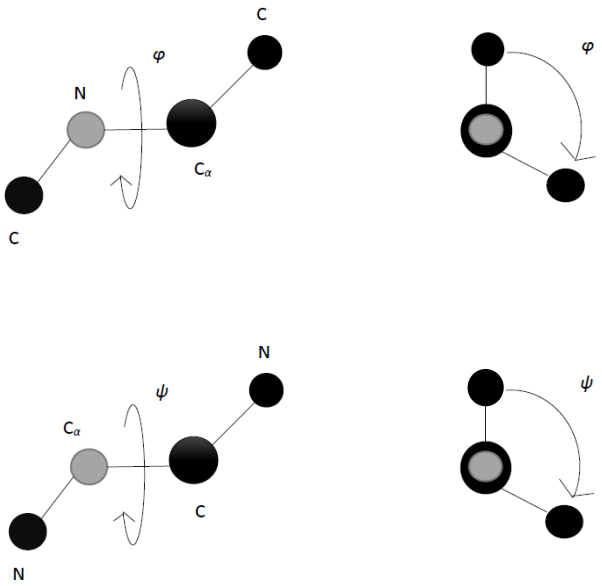



Figure 1.4: Main chain dihedral angles $\varphi$ and $\psi$

The two dihedral angles $\varphi$ and $\psi$ are the basis of the Ramachandran plot. In 1963, Ramachandran and his co-workers investigated the idea of a plot based on the dihedral angles $\varphi$ and $\psi$ [3]. Since that time, this plot has become one of the simplest methods to study the structures of proteins [4, 5].

Protein structure is exclusively determined by the sequence of amino acids along its length, and can be classified into a structural hierarchy, ranging from primary to quaternary. Firstly, the primary structure of a protein is the sequence of amino acids as prescribed by genomic information held by the sequence of nucleotides in the DNA. It is often represented by a series of a single-letter amino acid abbreviations. Secondly, the locally

ordered structure of the peptide backbone is defined as secondary structure. The most common secondary structural elements are $\alpha$-helices and $\beta$-sheets. Different types of amino acids tend to form different secondary structure elements. For example, methionine, alanine, leucine and glutamate and lysine are commonly found in $\alpha$-helices; large aromatic residues such as tyrosine, phenylalanine and tryptophan often occur in $\beta$-sheets. Some branched amino acids like, threonine, valine and isoleucine are found in $\beta$-sheets as well. Proline tends to break secondary structure, especially $\alpha$-helices due to its inability to form hydrogen bonds and because of steric hindrance due to its cyclised side chain. Proline will also break $\beta$-sheets as the allowed dihedral angles for proline lie outside of those generally found in $\beta$-sheets.

The tertiary structure of a protein is the three-dimensional structure of the folded polypeptide chain. There are several ways in which the amino acids—particularly the side chains—may interact with one another, and these interactions play a role in determining and maintaining the tertiary structure. Hydrogen bonding, important in maintaining the stability of secondary structure elements, also plays a role in stabilising the tertiary structure. Also important are disulfide bonds between sulfur containing cysteine residues, which can form covalent linkages

between two portions of the protein. The disulfide bonds bias the protein towards its native state and can aid in protein folding.

Hydrophobic interactions also play an important role in protein folding and formation of the tertiary structures. Residues with hydrophobic side-chains will tend to cluster together near the centre of the protein especially in aqueous solution [6]. The amino acids are classified in different groups. For example, hydrophobic groups do not favour contact with water, whereas polar groups and charged groups do. Phenylalanine is considered to be hydrophobic and it is usually buried inside the protein core. The aromatic residues tyrosine and tryptophan are often called amphipathic due to their ability to exhibit both polar and non-polar character. Being partially hydrophobic, tyrosine and tryptophan prefer to be buried in protein hydrophobic cores. The aromatic side chain can also mean that phenylalanine, tyrosine and tryptophan have an effect on stacking interactions with other aromatic side-chains [7].

For most proteins, the unique three-dimensional structures determine the biological functions. In order to form the three-dimensional structure, a protein undergoes a conformational change, known as folding. Finally, the quaternary structure describes the stable association of two or more chains of amino acids [1]. One example of a protein with quaternary structure is haemoglobin, which is made up of four haem-containing subunits; haemoglobin is heterotetrametric, comprising two $\alpha$ subunits and

two $\beta$ subunits. Not all proteins have quaternary structure, since they might be functional as monomers. Many proteins contain multiple structurally distinct regions which are known as domains. Each domain may have a specific function and can fold distinctly. Protein structure, specifically protein folding, has been the intense focus for many researchers over many decades. The biological properties of a protein depend on its three-dimensional structure and the way in which that structure can be influenced by chemical and physical factors, for example, pH, temperature, and ligands. Although the structures of over 100,000 proteins are now known, much less is known about the pathway of protein folding.

## 1.2. Structural biology determination

### 1.2.1 X-ray crystallography

X-ray crystallography is a technique used for determining the atomic and molecular structure of a crystal, in which the crystalline structure causes a beam of incident X-rays to diffract forming a distinctive pattern. By measuring the angles and intensities of these diffracted beams, a crystallographer can produce a three-dimensional picture of the density of electrons within the crystal. From this electron density, the mean positions of the atoms in the crystal can be determined, as well as their chemical bonds, their crystallographic disorder, and various other

information. Structure flexibility can be observed in very high-resolution electron density maps produced by X-ray crystallography, particularly when diffraction data is collected at room temperature instead of the traditional cryogenic temperature (typically near 100 K).

Production of good quality samples is a prerequisite to successful structural study by any method, and modern approaches to protein expression and purification used in structural biology laboratories help to facilitate this. X-ray crystallography requires milligram amounts of highly pure, monodisperse protein that remains soluble at high concentration. Structure analysis by X-ray crystallography is distinctive in its ability to determine atomic positions in a crystal lattice and hence yield a detailed and precise description of a protein's architecture. For this to be realized, however, large single protein crystals diffracting at high resolution are required. Unfortunately, this need is not always easily met due to, amongst other factors, the technical difficulties in crystallizing proteins, and, as a result, can delay the understanding of protein structure [8-11].

Meaningful models can be constructed from the high-resolution structures of individual subunits by macromolecular docking or rigid body refinement using X-ray crystallography tools. Some X-ray methods have been developed to generate probable

configurations of missing fragments in the incomplete high-resolution structures and they were successfully applied in practice [8].

### 1.2.2 NMR techniques

Nuclear magnetic resonance (NMR) is one of the most effective techniques applied since the first experimental observation of NMR in bulk matter more than 80 years ago. Its history has been punctuated by a series of revolutionary advances that have greatly expanded its horizons. Indeed, methodological and instrumental developments witnessed over the past decades have turned NMR into the most diverse spectroscopic tool currently available [12].

In addition to structures, NMR can yield information on the dynamics of various parts of the protein. This usually involves measuring relaxation times to determine order parameters, correlation times, and chemical exchange rates. NMR relaxation is a consequence of local fluctuating magnetic fields within a molecule. Local fluctuating magnetic fields are generated by molecular motions. In this way, measurements of relaxation times can provide information of motions within a molecule on the atomic level.

$^{1}H$ proton is used for protein structure determination. $^{1}H$ NMR spectra of proteins have been used to investigate secondary structure and tertiary structure. It is the origin of their unique value in the probing of atomic detail. The systems that can be studied range from simple organic molecules to complex biological macromolecules, because NMR techniques enable the resolution of distinct signals from even chemically identical groups when located in different local or global environments.

In NMR studies of protein dynamics, the $^{15}N$ isotope is the preferred nucleus to study because its relaxation times are relatively simple to relate to molecular motions. However, this requires isotope labelling of the protein and it is very expensive. $^{13}C$ is used commonly in NMR studies. NMR is used to study the flexibility of protein structure. 2D-NMR and Nuclear Overhauser Effects (NOEs) are methods used for protein structure determination.

### 1.2.3 Optical spectroscopy

Spectroscopy is the study of the interaction of electromagnetic radiation with matter. Detailed information regarding molecular structure and chemical properties can be obtained from the atomic and molecular spectra. Visible light was first considered as a form of electromagnetic waves by Maxwell in 1855. An electromagnetic

wave of any type includes an electric field component and magnetic field component [13].

The frequency range is divided into separate bands, and the electromagnetic waves within each frequency band are called by different names. Beginning at the low frequency (long wavelength) end of the spectrum these are: the radio waves, the microwaves, the terahertz waves, the infrared (IR), the visible light (vis), the ultraviolet (UV), the X-rays, and the gamma rays at the high frequency (short wavelength) end shown in Figure 1.5 [14].

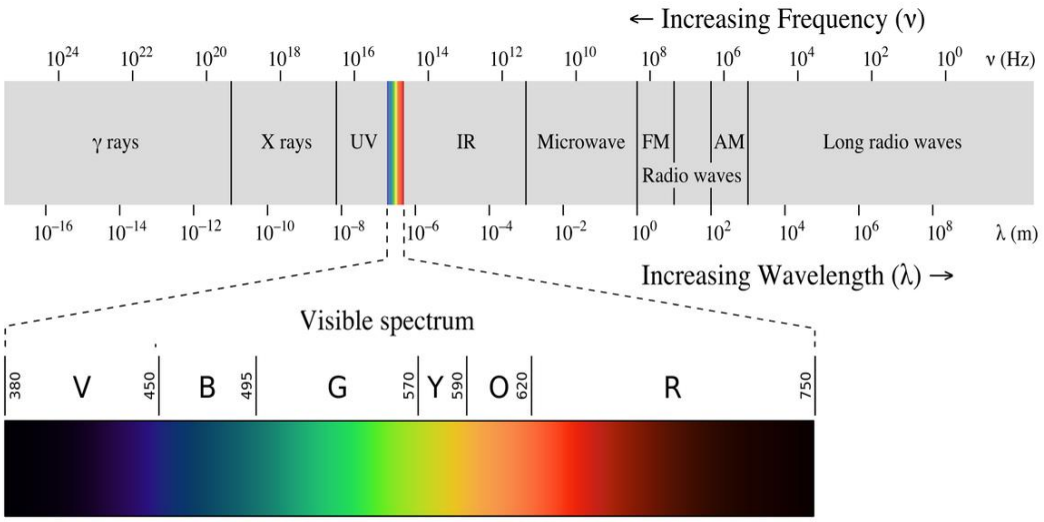



Figure 1.5: The regions of the electromagnetic spectrum taken from reference [14]

The electromagnetic waves in each of these bands have different characteristics, such as how they are produced, how they interact with matter, and their practical applications. Most of the frequency in a technique called spectroscopy can be used to physically separate waves different frequencies. This technique can produce a spectrum showing the constituent different frequencies.

#### 1.2.3.1 Absorption and fluorescence in the ultra-violet

Many studies have been carried out in the UV-vis region to investigate the structure of proteins and nucleic acids. For example, absorbance measurement is very sensitive and non-destructive, requiring only small amounts of material for analysis. It is widely used for measuring concentrations, for the detection of conformational changes and of ligand binding, and for following enzyme reactions [15].

UV absorption spectroscopies are powerful tools for the study of the reactivity of chemical and biological systems because they are highly sensitive, rapid, and simple. The concentrations of proteins or nucleic acids in solution can be easily and accurately determined by absorbance measurements. The absorbance is related to the intensity of the light before and after passage

through the protein solution, and the absorbance depends linearly on concentration, according to the Beer-Lambert law [14].

$$A = \log_{10}\left(\frac{I_0}{I}\right) = \varepsilon l c \qquad (1.1)$$

Equation (1.1) shows the relation between $A$ the absorbance and the initial and final intensities $I_0$ and $I$, and the role of the concentration $c$, the pathlength of the cell $l$, and $\varepsilon$ the molar absorptivity or molar absorption coefficient.

UV fluorescence occurs when a molecule absorbs energy causing an electron to be excited to a higher energy level, and as the electron returns to the original ground state, energy is lost as light or fluorescence. Thus, the excitation and emission wavelengths at which fluorescence occurs are characteristic to specific molecular structures. Fluorescence spectroscopy is an extremely sensitive multiparametric technique that provides simultaneous information about the conformation and size. Fluorescence signatures are observed with protein fragments and proteins of known composition and are caused by their tryptophan and tyrosine content. The fluorescence of aromatic amino acids in proteins depends on the specific sequence of amino acids and the presence of surrounding molecular groups [16]. Generally, buried tryptophans show different fluorescence to solvent-exposed tryptophans. Therefore, fluorescence is often used to follow folding.

### 1.2.3.2 Infra-red techniques

Infrared spectroscopy is another valuable technique used to characterize the structure of proteins, because of the high information content in its spectrum. A further advantage of this spectroscopy is a large application range from small soluble proteins to large membrane proteins. The absorption of infrared radiation excites vibrational transitions of molecules. This is generally the case when the dipole moment changes during the vibration. Since vibrational frequency and probability of absorption depend on the strength and polarity of the vibrating bonds, they are influenced by intra- and intermolecular effects. The approximate position of an infrared absorption band is determined by the vibrating masses and the type of bond (single, double, triple). Some bands are strongly associated with the presence of functional groups. The strength of absorption increases with increasing polarity of the vibrating bonds. In protein science, the effect of the environment on vibrational frequencies is often a sign of how proteins work [17].

Raman spectroscopy is another important technique in the infrared. In the IR absorption spectroscopy, the molecules absorb infrared radiation and pass to the excited vibrational state from the ground state and generate an IR absorption spectrum. Raman scattering is a two-photon process.

The molecule in the ground state gains energy from the incoming photon and transfers some energy to the excited state. This results in Raman scattering, which is measured in a conventional Raman experiment. Raman spectroscopy has found wide applications in physics, chemistry, biology and material science. For example, biological molecules including DNA, proteins, and carbohydrates can all be studied with Raman scattering spectroscopy [18].

### 1.2.3.3 Circular dichroism

One of the important spectroscopic techniques for studying proteins is circular dichroism (CD). It has become increasingly recognised as a valuable biophysical structural technique for addressing the relationship between the structure and the function of proteins. An advantage of the technique is that it is a relatively rapid means of obtaining secondary structural information and it can be the key method to monitor changes in structure brought about by ligand binding, or by modifications to the surrounding pH or temperature. The CD spectroscopy of proteins is a simple and quick way of evaluating their conformations for similarities and differences [19-21]. More details and information about CD will be introduced in section 1.4 in this

chapter. A CD signal will be observed if a component has a chiral centre.

### 1.2.4 Software used for studying CD of proteins

Different software and web interfaces are used to investigate large molecules. Most of them are well-established ways of studying proteins. Many factors affect the level of accuracy in different software, for example, including the solvent in the experimental structure determination of proteins and efficient computer modelling algorithms. In addition, the sequences of similar structures in Protein Data bank (PDB) [22] and the sequences without appropriate templates both have been considered as important factors. Development of more sophisticated and automated computer modelling approaches will dramatically enlarge the scope of modellable proteins in structural genomics projects. Different websites have been developed to make the predictions and the analysis of protein structures clear and accessible even for the non-expert user.

Analysis tools introduced include a wide range of structure prediction algorithms, for example, the CDtools software package for processing, calibrating and archiving CD data [23] and the Dichroweb analysis webserver for analysis of protein secondary structure from CD data [24]. The Protein Circular Dichroism Data

Bank (PCDDB) [25] is a public repository that archives CD spectral data and their associated experimental metadata. The PCDDB website includes links to other related resources for CD spectroscopy, including the DichroWeb analysis and CDtools [25]. 2Struc is a website that generates assignments for protein structures for up to eight different methods enabling easy analyses of similarities and differences between them [26]. Dichromatch provides the closest three matched spectra to the query spectrum produced by different methods [27]. ValiDichro runs a suit of analyses checks on a CD spectrum [28]. CAPITO, is a web server-based analysis tool for interpreting CD spectra. It allows the simultaneous evaluation of multiple datasets [29]. In 2016 new developments have been made to the PCDDB website to improve the use and the benefit [30]. PDB2CD is a website used to generate a CD spectrum from input protein atom structure coordinates [31]. Our in-house code DichroCalc web interface is used to calculate protein CD spectra with the matrix method [32]. Most of the other websites have statistical methods implemented in them such as, the PCDDB and the PDB2CD, while the DichroCalc web interface is based on first principles theory. Calculations of protein CD from first principles by using high level *ab initio* methods have been performed previously by our group [33-36].

## 1.3. Chirality

Chirality is a geometric property of various natural and non-natural objects that makes them non-identical and nonsuperposable with their own mirror images (Figure 1.6). Every amino acid (except glycine) can occur in two isomeric forms, because of the possibility of forming two different enantiomers (stereoisomers) around the central carbon atom. One enantiomer of a chiral amino acid is labelled D and the other L. They represent the direction of the rotation of the different groups linked to a chiral centre.

Many biologically active molecules are chiral, including the naturally occurring amino acids. In biological systems, most of these compounds are of the same chirality: most amino acids are levorotatory L and sugars are dextrorotatory D. Typical naturally occurring proteins, made of L amino acids, are known as *left-handed proteins*, whereas D amino acids produce *right-handed proteins*. D-amino acids are very rare in nature and have only been found in small peptides attached to bacterial cell walls.

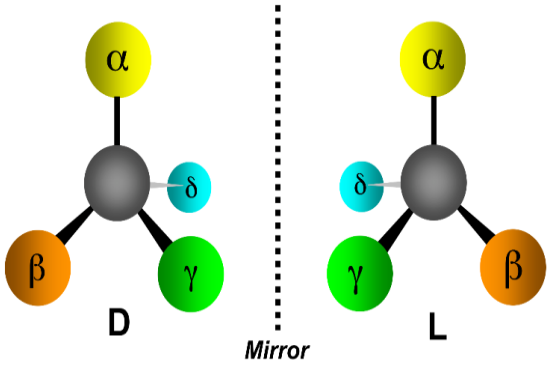



Figure 1.6: The mirror image of a chiral atom represented in grey linked to four different substituents: α, β, γ, and δ

This broad definition of chirality includes diverse macroscopic objects such as hands (the word chirality originated from the Greek word *cheir* meaning hand), animal organs, biological organisms, crystals and macromolecules as well as microscopic objects such as various molecules as chiral entities. Chirality is an important feature of the three-dimensional structures of peptides and proteins. The chirality of the secondary structure is key for protein CD. An effect will occur when a chromophore is chiral either intrinsically because of its structure or by being covalently linked to a chiral centre or by being placed in an asymmetric environment [37].

The question of the origin of biomolecular chirality- Why life is based on L amino acids and D sugars and not their mirror images? – has been unresolved since Pasteur's hand resolution of tartaric acid in 1848 [38]. A characteristic hallmark of life is its chirality, especially in biological systems. Biology uses only one enantiomer

and not the other, with animals made of proteins based exclusively on L amino acids, coded for by DNA based exclusively on D-deoxyribose. However, some of the unnatural D amino acids and L sugars do occur with specific roles, as in bacterial cell walls, but not bacterial protein or nucleic acids and antibiotics [39]. These examples in fact illustrate the dominance of the L amino acids and the general rule of chirality. D amino acids are used in cell walls precisely because the enzymes of predators only digest L amino acids; furthermore, the genetic code does not code for these D amino acids, which are synthesized by enzymatic inversion of the L form [40].

Chirality is the most fundamental property underlying protein CD spectroscopy[2]. Optical rotatory dispersion (ORD) is another spectroscopy used successfully and specifically to study chiral molecules. They show variations, as a function of wavelength, of the optical rotation angle and ellipticity angle. Although CD and ORD give equivalent information, there are some points which differentiate the two techniques and offer particular advantages in certain cases [41].

One of the main differences between them is that ORD has a characteristic background effect. This effect arises from the shape of dispersion curves, where the observation of optical rotation at a given wavelength results from several contributions due to the

local environment effect based on position, sign and amplitude. These cases usually cannot be investigated in the absorption region because of various factors or instrumental limitations. As CD is an absorption phenomenon, the shape of the curve is such that its contribution is detectable only in the close vicinity of the absorption. Therefore, there are in practice no background effects, and it is possible to evaluate more accurately the rotational strength of individual electronic transitions [42]. The application of these two chiroptical techniques often depends on the type of instrument used, on the nature of compound examined and on the type of information required. Generally, CD measurements are preferred for the interpretation of vibrational fine structure and for low temperature and vapour phase measurements.

## 1.4. Circular Dichroism

The interaction of optical activity with molecular structure has led to a vast range of studies of chiroptical properties, from synthetic organic compounds to natural products, and encouraged the development of techniques, such as ORD and CD spectroscopy. CD spectroscopy on proteins [19-21] provides characteristic bands in the near- and far-ultraviolet (UV) region. The far–UV spectrum (190-250 nm) arises in large part from the secondary structure. This spectrum is affected by the backbone structure of a protein

while the near-UV bands (>250 nm) arise from aromatic side chains. The spectra in the near-UV which are the focus of this thesis, arise from the $\pi\pi^*$ electronic excitations on the aromatic rings of phenylalanine (Phe), tyrosine (Tyr) and tryptophan (Trp) which can reflect the conformation of aromatic side chains as well as their surrounding environment.

Each of the aromatic amino acids has a characteristic spectral profile, tryptophan, a peak close to 290 nm with fine structure between 290 and 305 nm; tyrosine, a peak between 275 and 282 nm - the fine structure at longer wavelengths may be obscured by that from tryptophan; phenylalanine, sharp fine structure between 255 and 270 nm [43]. Three factors tend to influence the intensities of aromatic CD bands: i) the rigidity of the protein, with the more highly mobile side chains having lower intensities; ii) interactions between aromatic amino acids, which are especially significant if the distance between them is less than 1 nm; iii) the number of aromatic amino acids.

The side chains of aromatic amino acids play an essential role in both protein structure and function. The aromatic structures tend to be buried in the interior of proteins as part of the hydrophobic core [44]. They also form key structures with themselves or with other amino acid side chains in the protein folding process. They are essential for the function of proteins, such as molecular

recognition and catalysis. Thus, biophysical measurements of aromatic amino acids can provide new insight into fundamental aspects of protein behaviour.

Over many years, CD spectroscopy has been used in numerous protein studies. One major strategy is the comparison of the CD spectra of site-directed mutants. Changes in the CD intensity have been used to study the contribution of individual aromatic residues using mutagenesis [45, 46]. Although CD is one of the most helpful spectroscopies for examining protein structure, other spectroscopies are used to achieve similar purposes [19-21]. However, CD spectroscopy has specific advantages, such as using lower concentrations and having no size limitations in contrast to NMR. Compared with X-ray crystallography, results have been gained without any of the problems in a structure that may occur during crystallization [21].

CD is the difference in absorption of left and right circularly polarized light. For chiral molecules, the extinction coefficients for the left and right circularly polarized light are different $\varepsilon_l \neq \varepsilon_r$ giving a differential molar absorption $\Delta\varepsilon$.

$$\Delta\varepsilon = \varepsilon_l - \varepsilon_r \qquad (1.2)$$

The units here are $\mathrm{cm}^{-1}.\mathrm{M}^{-1}$.

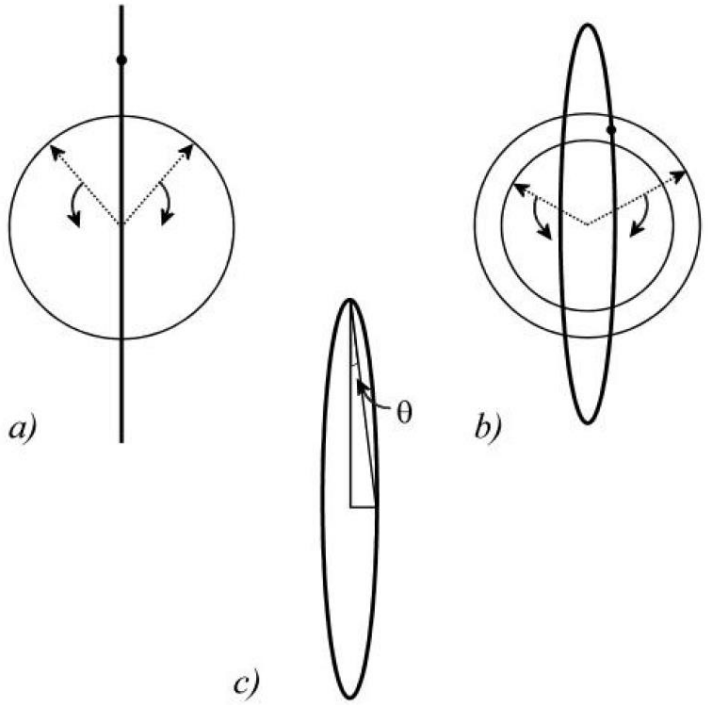



Figure 1.7: Elliptically polarized light. Redrawn from [41]

Figure 1.7 [41] shows the electric vector component of the beams looking down the direction of propagation of the light source. a) As the two beams pass through an achiral sample no CD signal is generated. b) Two circularly polarised beams when passing through a chiral sample will be unequally absorbed and the resultant light beam will transcribe an elliptically polarised path. c) The representation of ellipticity as a measure for CD signal.

In terms of ellipticity the mean residue ellipticity $[\theta]$ at wavelength λ is given by:

$$[\theta] = \frac{100\,\theta}{d.c} \qquad (1.3)$$

where $\theta$ is the observed ellipticity ($\mathrm{degrees}$), $d$ is the pathlength ($\mathrm{cm}$) and $c$ is the solute molarity (in units of mol $dm^{-3}$). The units of mean residue ellipticity are $\mathrm{deg.cm^2.dmol^{-1}}$, at any wavelength. The numerical relationship between the two quantities:

$$[\theta] = 3298\Delta\varepsilon \qquad (1.4)$$

The CD signal is related to the scalar product of the electric transition dipole moment $\vec{\mu}$ and magnetic transition dipole moment $\vec{m}$. Therefore, non-zero values of these two quantities during the excitation process of a chiral molecule leads to the CD signal [47]. Even though CD spectroscopy has become a useful tool to understand protein structure, and has different applications in studies of protein folding and unfolding, it still has intrinsic limitations; for example, it provides low structural resolution [2, 47]. Structure determination is essential to understand protein function, and to achieve this purpose different CD-based techniques have been used. Electronic CD (ECD) is the main technique that we focus on in this thesis with new parameters including new electronic vibrational levels of the chromophores which have been introduced by a previous PhD student in our group [48] and this research used these updated parameters. ECD arises from transitions that are electric dipole and magnetic dipole allowed. During the excitation of an electron from the ground state to a higher energy state, some change will occur in the electric

and magnetic fields. As a result, the net of these fields during the transition will produce the electric transition dipole moment $\vec{\mu}$ and the magnetic transition dipole moment $\vec{m}$. The rotational strength of a band in an ECD spectrum is analogous to the oscillator strength in electronic absorption spectroscopy [42]. More details about the rotational strength will be discussed in chapter two.

Vibrational CD (VCD) shows the differential response to left versus right circularly polarized light and it arises from simultaneous change in electric and magnetic dipole moments of the molecule due to the nuclear motion. VCD spectra are rich in information, but the instrumentation is not developed or as widely used as that for ECD. The rotational strength of VCD is less intense than ECD, which makes it a challenge to achieve the signal/noise ratio [20, 49 - 50].

Magnetic CD (MCD) is another CD-based technique used for studying bioinorganic systems [51]. It is used to interpret how many transitions are involved in one absorption envelope, because MCD signals with different signs may indicate different polarizations of transitions. Another type of CD which is used in some research is X-ray CD [52], which studies transitions localized near the core-excited atom; and nuclear spin-induced CD [53] can provide information about nuclear magnetization.

Time resolved CD spectroscopy was possible on millisecond timescale in the 1950s and 1960s when CD became more common than ORD with the development of photo elastic modulators [54]. The time regime was improved to nanosecond in 1985 when an ellipsometry approach was introduced for time-resolved CD (TRCD) measurements [55]. These techniques bring some new directions of using CD in structure determination. CD spectroscopy becomes even more powerful when it is coupled with other techniques such as high-performance liquid chromatography (HPLC) or stopped-flow apparatus. HPLC-CD was first introduced in 1980 to analyse acids in wines [56]. A combination of spectroscopy with stopped-flow devices can study fast kinetic events [57]; CD stopped-flow techniques have been used to study protein folding [58, 59].

The calculation of protein CD has been heavily studied in our group and improved through the use of more accurate calculations of chromophores from the first principles [34, 35, 48, 60-63]. Different aspects have been considered to improve the accuracy of the calculations. Recently the contributions of the transitions of aromatic side chains have been studied and incorporated into the CD calculations[60, 64, 65].

## 1.5. Outline of the thesis

The opening chapter has provided a general background of proteins and their structure and the different spectroscopic techniques that are commonly used to investigate protein structure. The origin of chirality and its relationship to study protein structure has been discussed to illustrate the role of optical activity in CD measurements. In addition, CD has been introduced with the principles and different applications of this technique.

Chapter two will be focusing on the theoretical and computational methods that have been used in this research, starting with a detailed discussion on the basic principles of the matrix method which is implemented in our in-house software DichroCalc [32]. The role of the Frank-Condon effect [48] on the CD calculations is incorporated by using the new parameters to include the vibrational electronic transitions in our calculations. The mechanisms leading to bands in the CD spectrum due to the aromatic residues in near-UV CD will be illustrated in detail and relevant work from the literature will be discussed. Calculations of electrostatic potentials surrounding a specific aromatic residue will be described.

Chapters three and four present investigations and analysis of the interactions between aromatic residues and the mechanisms that generate our calculated CD spectra. Two different groups of

proteins are considered. The first group will be discussed in chapter three when we discuss on tyrosine chromophores in proteins. The second group will show the analysis of tryptophan chromophores. A further electrostatic analysis for Tyr and Trp of some proteins from each group will be presented in each chapter. We investigate the effect of electrostatic environment surrounding these residues and how that contributes to the CD spectra.

Chapter five provides the conclusions drawn from this research. Ideas about how this research could be improved and progressed will be presented.

## 1.6. Aim of the work

This work forms a part of an ongoing effort to explore the relationship between structure and spectrum and the contributions of aromatic residues of some proteins in near-UV CD. In this thesis, we primarily focused on incorporating the vibrational structure of the excited states of the aromatic residues (Phe, Tyr and Trp) to improve the quality of our CD calculations. The conformational diversity is also touched on by computing protein CD from NMR structures. Consideration of the interactions of aromatic chromophores should give us a better understanding of the near-UV CD spectra. We investigate the mechanisms contributing to the calculated CD spectra and study the effect of the environment surrounding some interesting aromatic amino acids. A consideration of electrostatic analyses for the surrounding environment is useful to provide a further understanding of the structure-spectrum relationship.

# Chapter 2

# Methods

## 2.1. Introduction

The on-going developments in computing technologies since last century have revolutionised the use of different theoretical and computational methods to study protein structure and function. Some methods have an acceptable level of accuracy and agree well with the experimental data, enabling further in-depth analysis. This chapter aims to introduce and describe all methods used in this research, both computationally and theoretically. The matrix method is a successful formulation derived from the exciton model [66] to study large systems such as proteins. Including more vibration levels in our calculation helps to improve the calculated spectra. We will explain the mechanisms by which CD bands arise from the aromatic residues in the near-UV CD. The electrostatic interactions within proteins are studied by using one of the most widespread models to evaluate the electrostatic properties, the Poisson-Boltzmann equation [67].

## 2.2. The matrix method

The CD intensity or rotational strength for an excitation is analogous to the oscillator strength in traditional electronic absorption spectroscopy. The rotational strength ($R^{0k}$) for an excitation from the ground state 0 to an electronic excited state *k* is the imaginary part of the product of the electric transition dipole moment, $\vec{\mu}$, and the magnetic transition dipole moment, $\vec{m}$, given by the Rosenfeld equation [68]:

$$R^{0k} = \mathrm{Im}(\langle\psi^0|\vec{\mu}|\psi^k\rangle\langle\psi^k|\vec{m}|\psi^0\rangle) \qquad (2.1)$$

$\psi^0$ and $\psi^k$ represent the wave functions of the ground and excited state respectively and the CD spectrum can be calculated based on this equation. The wave functions for the ground and excited states can only be fully determined *ab initio* for relatively small compounds. For large systems, such as proteins, there are computational challenges. However, some approximate approaches are well established, for example, the matrix method [69-71]. The matrix method, which is based on exciton theory [66], is widely used for large molecular systems [72-76], particularly for proteins. It uses an improved version of an approach that was introduced by Tinoco in 1962 [77]. This method involves the diagonalization of a Hamiltonian matrix to solve an eigenvalue problem which comprises the interactions between chromophores. We calculate protein CD spectra with the matrix method as

implemented in our in-house code DichroCalc [32]. In the matrix method, the protein is treated as *M* independent chromophores, with a monomer wave function $\phi_{is}$ for each chromophoric group *i* and electronic state *s*. The protein's $k^{th}$ excited state wavefunction, $\psi^k$, is written as a linear combination of electronic configurations, $\Phi_{ia}$, in which only one chromophoric group *i*, is in an excited state *a* and the others are in the ground state, 0. Thus

$$\Phi_{ia} = \phi_{10} \ \ldots \ \phi_{ia} \ \ldots \ \phi_{j0} \ \ldots \ \phi_{M0} \qquad (2.2)$$

where $\phi_{ia}$ is the wave function of chromophore *i* after the transition $0 \rightarrow a$, and

$$\psi^k = \sum_i^M \sum_a^{n_i} c_{ia}^k \ \Phi_{ia} \qquad (2.3)$$

where $n_i$ is the number of excitations on chromophore *i* and $c_{ia}^k$ are the eigenvector coefficients. The eigenvector coefficients tell us about the nature of the transition of the protein, for example, whether it is localized and similar to the transition of the isolated chromophore (reflected in a value closer to one). If the transition is delocalized that means it is a linear combination of several chromophore transitions.

The wave function $\psi^k$ of electronic excited state *k* of protein, and its corresponding energy, can be calculated by solving the Schrödinger equation:

$$\hat{H}\psi^k = E^k\psi^k \qquad (2.4)$$

$$\hat{H} = \sum_{i=1}^{M} \hat{H}_i + \sum_{i=1}^{M-1} \sum_{j=i+1}^{M} \hat{V}_{ij} \qquad (2.5)$$

The Hamiltonian $\hat{H}$ of the system with *M* independent chromophores, is constructed as the sum of the local Hamiltonian $\hat{H}_i$ of the independent group *i* plus the interactions between the separated groups, which is denoted as intergroup potential $\hat{V}_{ij}$, while $\hat{V}_{ii}$ is the interaction between states on the same chromophore. In the Hamiltonian matrix below, a two-chromophore model is considered (denoted 1 and 2) with two transitions (denoted *a* and *b*) for each chromophore. Two vibrational levels (denoted $v_1$ and $v_2$) of the transition to the excited state *a* are shown in equation (2.6).

$$\hat{H} = \begin{pmatrix} E_{1a_v_1} & 0 & V_{1a_v_1;1b} & V_{1a_v_1;2a_v_1} & V_{1a_v_1;2a_v_2} & V_{1a_v_1;2b} \\ 0 & E_{1a_v_2} & V_{1a_v_2;1b} & V_{1a_v_2;2a_v_1} & V_{1a_v_2;2a_v_2} & V_{1a_v_2;1b} \\ V_{1a_v_1;1b} & V_{1a_v_2;1b} & E_{1b} & V_{1b;2a_v_1} & V_{1b;2a_v_2} & V_{1b;2b} \\ V_{1a_v_1;2a_v_1} & V_{1a_v_2;2a_v_1} & V_{1b;2a_v_1} & E_{2a_v_1} & 0 & V_{2a_v_1;2b} \\ V_{1a_v_1;2a_v_2} & V_{1a_v_2;2a_v_2} & V_{1b;2a_v_2} & 0 & E_{2a_v_2} & V_{2a_v_2;2b} \\ V_{1a_v_1;2b} & V_{1a_v_2;2b} & V_{1b;2b} & V_{2a_v_1;2b} & V_{2a_v_2;2b} & E_{2b} \end{pmatrix} \qquad (2.6)$$

The notation 1*a*_$v_1$ refers to the first vibrational level of transition *a* on chromophore 1. diagonal element *E* is the energy of a distinct transition of one chromophore, the off-diagonal element *V* is either the intergroup interaction or the static-field mixing of excitations of a single group, also known as the one-electron mixing, as described by Condon, Alter and Eyring [78]. The

interaction between different vibrational levels of a given electronic excited state in a chromophore is set to zero.

The monopole-monopole approximation is used, and we assume that the interactions between chromophoric groups are approximated as being solely electrostatic in nature. These interactions are calculated as:

$$V_{i0a;j0b} = \int\int \frac{\rho_{i0a}(r_i)\rho_{j0b}(r_j)}{4\pi\varepsilon_0 r_{ij}} dr_i dr_j \qquad (2.7)$$

where $\rho_{i0a}(r_i)$ and $\rho_{j0b}(r_j)$ are the transition electron densities of chromophores *i* and *j*, $\varepsilon_0$ is the vacuum permittivity, $r_i$ and $r_j$ are the coordinates of the excited electrons on chromphores i and j, respectively, and $r_{ij}$ is the distance between the chromophores.

The Coulomb interactions between monopoles account for the interaction between groups:

$$V_{i0a;j0b} = \sum_{s=1}^{N_s}\sum_{t=1}^{N_t} \frac{q_{s(i)}q_{t(j)}}{r_{st}} \qquad (2.8)$$

where $q_s$ and $q_t$ are the charges (i.e., the monopoles) on two chromophores *i* and *j*, and $r_{st}$ is their separation. $N_s$ and $N_t$ are the number of charges on the chromophores.

The matrix method involves the diagonalization of the Hamiltonian matrix given in equation (2.6), solving the eigenvalue problem. Diagonalization of the Hamiltonian matrix yields the energy of each transition (eigenvalues) and the expansion coefficients (eigenvectors).

The diagonalization process is effected by a unitary matrix *U*:

$$U^{-1} \cdot \hat{H} \cdot U = H_{diag} \quad (2.9)$$

The results of the diagonalization are used to calculate the electric and magnetic transition dipole moments of the interacting system from the initial dipole moments of the single groups [68].

$$\overrightarrow{\mu_i} = \sum_a U_{ai}\, \vec{\mu}_a^0 \quad (2.10)$$

$$\overrightarrow{m_i} = \sum_a U_{ai}\, \vec{m}_a^0 \quad (2.11)$$

In the matrix method, we consider the electrostatic potential of each electronic transition. Monopoles are fitted to reproduce the electrostatic potential. For small chromophores, the relevant electrostatic potentials can be calculated *ab initio* using, for example, the Complete Active Space Self-Consistent Field (CASSCF) method [79].

## 2.3. Mechanisms of interactions associated with different types of matrix element

The rotational strength can be generated by three mechanisms namely: i) *the one-electron mechanism* which includes both transitions with electric and magnetic transition dipoles within the same chromophore, which is also known as intra-chromophore mixing; ii) *the μ-μ mechanism* which is the coupling between electric transition dipole moments of two separate chromophores; and iii) the $\mu$-$m$ mechanism which is the coupling between the

electric and the magnetic transition dipole moments of two separate chromophores [80]. The last two mechanisms are called inter-chromophore mixing or coupled oscillator type.

A couple of studies have considered these mechanisms, to investigate the influence of conformational changes and the effects of the electrostatic interactions on the near-UV CD spectrum of TEM-1 $\beta$-lactamase [81-83]. Further research has studied the relationship between chiroptical properties and structural changes, with more detailed analysis of aromatic contributions of individual chromophore in TEM-1 $\beta$-lactamase by investigating these mechanisms which generate CD rotational strength [84-88].

Additionally, a computational study of the CD spectrum of human carbonic anhydrase was used to evaluate the effects of protein flexibility on the quality of the calculated spectra [89]. Therefore, we have seen that many studies focus on these three mechanisms to study aromatic side chain interactions and/or other properties in different proteins.

The aromatic side chains of a protein are planar and their location in the chiral environment of a peptide or protein leads to a CD signal arising from their electronic transitions [19]. Proteins where the aromatic side chains are isolated from one another exhibit CD signals which can be particularly useful to study conformational changes. Protein CD due to the interactions between side chains

can make some groups that do not contribute to the absorption spectrum become significant in the CD because they have transitions that are both magnetic and electric dipole allowed.

In our calculations the interactions of different transitions in the protein environment have been computed in the exciton matrix method calculation as purely electrostatic. The Coulombic interaction energy between two transition charge densities, each represented by a set of point charges as part of the parameter set, is described in detail elsewhere [48]. It depends on the charge densities and on their relative orientation and separation. In the case of the one-electron mechanism, there is a dependence on the surrounding ground state electronic density.

The Hamiltonian matrix in equation (2.6) was used to describe the different mechanisms that control the generation of the CD spectrum in our calculations. The *one-electron mechanism* or as it is also called the Condon, Alter, Eyring (CAE) [90] theory has a predominant role as examined in chapters 3 and 4. This corresponds two different transitions within the same chromophore, for example, $V_{1a-v_{1b;\,1b}}$ which represents the $(n\pi^*\,\pi\pi^*)$ transition of one chromophore in the Hamiltonian matrix as presented in equation (2.6).

## 2.4. Electrostatic interactions from the ground electronic state

Electrostatics influence many aspects of biochemical reactions and play a fundamental role in virtually all processes involving biomolecules in solution. One of the most fundamental approaches to study electrostatic properties is based on the Poisson-Boltzmann equation (PBE) [67, 91].

$$-\nabla.\varepsilon(\mathrm{x})\,\nabla\phi(x) + \bar{\mathrm{K}}^2(\mathrm{x})\,\sinh\phi(x) = \mathrm{f}(x) \qquad (2.12)$$

The above differential equation shows the relationship between the electrostatic potential $(\phi)$ and the dielectric properties of the solute and the solvent $(\varepsilon)$, the ionic strength of the solution and the accessibility of ions to the solute interior $(\bar{K}^2)$, and the distribution of solute atomic partial charges $(f)$. They are all functions of the position vector $(x)$. An additional approximation is often used in PBE, the linearized PBE, LPBE which assumes $\sinh\phi(x) \approx \phi(x)$.

Efficiency and accuracy are two major concerns in finding solutions of the PBE for applications in biophysics and chemistry. It is a useful equation in many settings, whether it be to understand physiological interfaces, polymer science, electron interactions in a semiconductor, or more. It aims to describe the distribution of the electric potential in solution in the direction normal to a charged surface. From the PBE many other equations

have been derived based on different assumptions. However, obtaining accurate solutions to the PBE for a large system such as a protein has a significant computational cost [92].

A continuum solvation model represents solvent as a continuous medium instead of individual explicit solvent molecules. It is often used in molecular dynamics simulations and in other applications of molecular mechanics and is sometimes termed implicit solvation. The method can be applied to estimate the free energy of solute-solvent interactions in structural and chemical processes, such as folding or different conformations of proteins, association of biological macromolecules with ligands, or transport of drugs across biological membranes. Although the implicit solvent model is useful, it is an approximate method with certain limitations and problems related to parameterization and treatment of ionization effects [93]. A continuum solvation model is implemented in the software APBS [91]. We use the APBS software to generate all the potential grid points surrounding the interesting aromatic amino acid in our calculations.

Defining the molecular surface area and the atomic charges which are usually taken from some force field enabled the investigation of the molecular-solvent boundary. Some evidence suggests that the solvent near proteins is mobile, indicating a high dielectric

constant [94]. The software APBS has different force field options and the CHARMM force field [95] has been chosen in our calculations. The potential grid points surrounding the aromatic amino acids are a special focus in this study. The electrostatic potentials are generated using the force field in APBS.

## 2.5. Franck Condon Principle and vibrational parameters

The Franck Condon principle explains the intensity of a vibrational progression. In an absorption spectrum, transitions from the electronic ground state to electronic excited state form a progression corresponding to the vibrational levels of the excited state. An electronic transition takes place much more rapidly than vibrational motion due to the slow motion of nuclei compared to the instantaneous motion of electrons because electrons are much lighter than nuclei. The electron configuration of the new state may result in a shift of the equilibrium position of the nuclei constituting the molecule.

The intensity distribution along the progression is proportional to the Franck Condon factor (FCFs), and the Franck Condon factor is equal to the overlap integral of vibrational wave functions of the ground and the excited electronic states. This is called the Franck Condon principle.

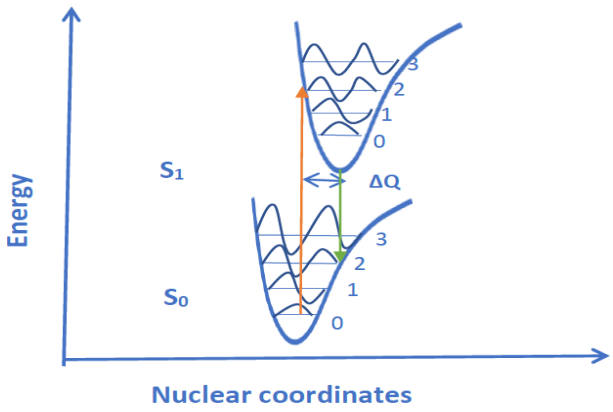



Figure 2.1: Franck Condon principle for vibronic transitions, where $S_0$ is the ground electronic state and $S_1$ is the first excited electronic state, ΔQ is the shift in nuclear coordinates

Figure 2.1 illustrates the Franck Condon principle for vibronic transitions in a molecule. This figure shows the shift in nuclear coordinates, labeled ΔQ, between the ground and the first excited state. The vibrational levels (0, 1, 2, 3) of the ground electronic state $S_0$ are shown in the above figure. When the necessary amount of energy is available, the molecule can undergo a transition to a vibrational level of the excited state $S_1$ when the photon is absorbed.

We used new parameters considering the vibrational transitions [48] of the phenylalanine (Phe), tyrosine (Tyr), and tryptophan (Trp) side chains. Each has four electronic states ($L_b$, $L_a$, $B_b$, $B_a$) as labeled by Platt [96]. They have been investigated previously in many studies [36, 60]. Figure 2.2 shows the ($L_b$, $L_a$) electronic states for the three aromatic amino acids. The importance of including

vibrational structure in CD calculations has been shown in different studies for various molecules, and specifically for the three aromatic residues that are the focus of this research [97-99]. The detailed features of near-UV CD spectra evident in experimental measurements are associated with vibrational structure of the aromatic chromophores.

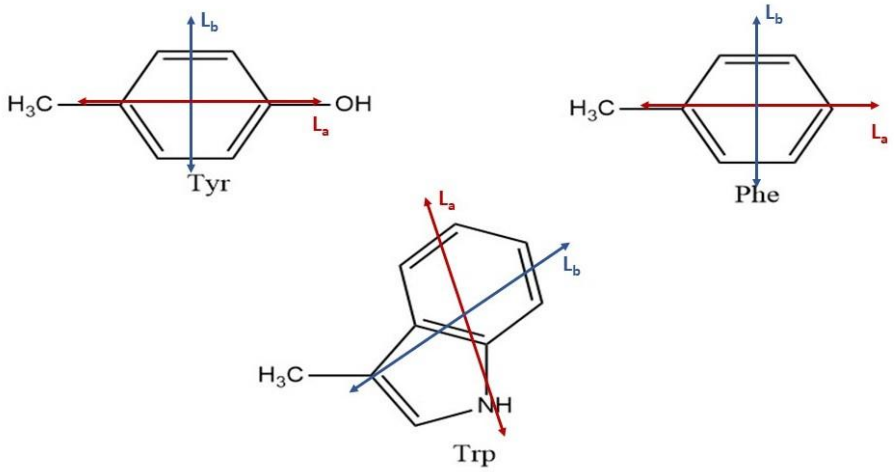



Figure 2.2: The $L_a$ and $L_b$ electronic states for the aromatic amino acid: Tyr, Phe and Trp

In our calculations the vibrational transitions are incorporated by modifying the electric transition dipole and the monopole charges considering the transition charge densities based on the normalized Frank-Condon overlap integral. The first six to nine vibrational transitions are considered in a vibrational progression, with the precise number determined from the threshold of 0.1 for the Frank-Condon integral. Interaction between the vibrational levels of the same chromophore was not included in the CD

calculations, as they are assumed to be orthogonal. The charge transfer between chromophores was not accounted for. We considered the $L_b$ and $L_a$ transitions for the near-UV CD calculations. Toluene was used as a representative of the aromatic side chain of phenylalanine.

Table 2.1:Frank-Condon integrals and the transition energies reprinted from [48]

| Toluene | | | p-cresol | | |
|---|---|---|---|---|---|
| '*vib*' level | Energy/cm$^{-1}$ | FC integral | '*vib*' level | Energy/cm$^{-1}$ | FC integral |
| 0 | 37310 | 0.59 | 0 | 34965 | 0.52 |
| $1\nu B_{CC}$ | 37830 | -0.43 | $1\nu B_{CC}$ | 35810 | 0.58 |
| $1\nu B_{CH}$ | 38288 | 0.49 | $1\nu B_{CH}$ | 36221 | 0.17 |
| $1\nu B_{CC}$/ $1\nu B_{CH}$ | 38808 | -0.35 | $2\nu B_{CC}$ | 36655 | 0.45 |
| $2\nu B_{CH}$ | 39266 | 0.26 | $1\nu B_{CH}$/ $1\nu B_{CC}$ | 37066 | 0.18 |
| $1\nu B_{CC}$/ $2\nu B_{CH}$ | 39786 | -0.18 | $3\nu B_{CC}$ | 37500 | 0.28 |
| | | | $1\nu B_{CH}$/ $2\nu B_{CC}$ | 37911 | 0.14 |
| | | | $4\nu B_{CC}$ | 38345 | 0.14 |
| 3-methylindole $^1L_b$ | | | 3-methylindole $^1L_a$ | | |
| '*vib*' level | Energy/cm$^{-1}$ | FC integral | '*vib*' level | Energy/cm$^{-1}$ | FC integral |
| 0 | 34424 | 0.62 | 0 | 35088 | 0.57 |
| $1\nu S_{CC}$ | 35199 | 0.47 | $1\nu S_{CC}$ | 35872 | 0.53 |
| $1\nu B_{CH}$ | 35381 | 0.44 | $1\nu B_{CH}$ | 36135 | 0.32 |
| $2\nu S_{CC}$ | 35974 | 0.23 | $2\nu S_{CC}$ | 36656 | 0.34 |
| $1\nu B_{CH}$/ $1\nu S_{CC}$ | 36156 | 0.33 | $1\nu S_{CC}$/ $1\nu B_{CH}$ | 36919 | 0.30 |
| $3\nu S_{CC}$ | 36749 | 0.09 | $3\nu S_{CC}$ | 37440 | 0.18 |
| $1\nu B_{CH}$/ $2\nu S_{CC}$ | 36931 | 0.17 | $1\nu S_{CC}$/ $2\nu B_{CH}$ | 37966 | 0.19 |
| | | | $4\nu S_{CC}$ | 38224 | 0.08 |
| | | | $1\nu S_{CC}$/ $3\nu B_{CH}$ | 39013 | 0.10 |

For phenylalanine only the Franck-Condon integrals of the $^1L_b$ first excited state were considered, since it is the only transition of

phenylalanine located in the near-UV region. *p*-Cresol was used to represent tyrosine. Its electronic ground and first singlet excited state have been extensively studied experimentally [100, 101] and theoretically [102, 103]

The Franck-Condon integrals of the $^1L_b$ first excited state of tyrosine were considered. 3-Methylindole was used as a model of tryptophan and we considered the two electronic excited states $^1L_b$ and $^1L_a$. The normalized Frank-Condon integrals and the transition energies are shown in Table 2.1 [48]. The table shows the transition energies and Franck-Condon integrals of transitions to different vibrational levels ('vib' level) of the $^1L_b$ states of toluene, p-cresol and 3-methylindole and the $^1L_a$ state of 3-methylindole. The vibrational level is denoted as M$v$N, where N is the assignment of the vibrational mode and M is the vibrational level of that mode. The C-C bend, C-H bend and C-C stretch mode are denoted as $B_{CC}$, $B_{CH}$ and $S_{CC}$, respectively.

We place the origin of the toluene $^1L_b$ transition at 268 nm to match the experimental spectra of phenylalanine, while for *p*-cresol the $^1L_b$ transition is located at 286 nm based on previous studies [100-103]. For 3-methylindole the origins of the $^1L_b$ and $^1L_a$ bands are placed at 290.5 and 285 nm, respectively [48]. At higher energies, the vibrational structure of the spectra becomes more complicated and overlapped, due to combination modes and

anharmonicity. Calculation of the CD spectra considering only the predominant modes represents a stepwise approach to increasing the complexity of the computational methodology.

For the calculations, we took the coordinates of the X-ray crystal structure of proteins from the Protein Data Bank. Later, to investigate the effect of different conformations on our calculations and the generated CD spectra we used the NMR models for the same proteins.

# Chapter 3

# Analysis of the tyrosine chromophore in proteins

## 3.1 Introduction

Aromatic amino acid chromophores play an essential role in understanding protein structure, stability, and function. CD in the near-UV is a widely utilized method to study structural changes in proteins because of its sensitivity[104]. Aromatic chromophores in the near-ultraviolet (near-UV) region (240 to 300 nm) are sensitive to environmental electrostatic changes, for example, the $L_a$ transition of tryptophan is very sensitive to its surroundings, due to the large dipole moment of this excited state [105, 106].

A combination of CD spectrum and computational approaches leads to a better understanding of the conformation of some proteins. All proteins we consider have a family of NMR structures, which allow us to investigate the effect of different interactions of the aromatic side chains. These appear to explain the sensitivity

of the calculated near-UV CD spectra to the precise conformation of the protein.

The proteins in this chapter have only phenylalanine and tyrosine chromophores. The analysis of this chapter focuses mainly on the tyrosine chromophores in proteins. Tyrosine has a predominant role in generating CD spectra in the near-UV, which is related to the effect of the OH substituent. This OH has a stronger effect than that of an alkyl group, and it shifts the bands to longer wavelengths [107]. Three proteins will be discussed in this chapter, for each protein the X-ray crystallographic structure and the NMR models from the PDB are used to generate the calculated CD spectrum.

Further analysis focuses on only two of the NMR structures. The selection of these two NMR models is based on statistical investigations, explained later in the methods section. The interactions and contributions will be investigated because of their important role in generating CD spectrum in near-UV. The results showed that phenylalanine residues have a weak contribution in this region [65]. Therefore, no results will be presented for any phenylalanine in this chapter. Table 3.1 shows the proteins that have been studied in this chapter.

Table 3.1: The proteins studied in this chapter: name, PDB codes, length (i.e., number of residues in the X-ray structure) and the number of aromatic residues and disulfide bonds.

| Protein | PDB entry | | Length | No. aromatic residue | | | |
|---|---|---|---|---|---|---|---|
| | X-ray | NMR | | Phe | Tyr | Trp | S-S |
| bovine pancreatic trypsin inhibitor | 5PTI | 1PIT | 58 | 4 | 4 | 0 | 3 |
| Cardiotoxin | 4OM4 | 1CRF | 60 | 2 | 3 | 0 | 4 |
| ribonuclease A | 1AFU | 2AAS | 124 | 3 | 6 | 0 | 4 |

## 3.2 Background on the proteins studied in this chapter

### 3.2.1 Bovine pancreatic trypsin inhibitor

Bovine pancreatic trypsin inhibitor (BPTI) is a small monomeric and single-chain globular polypeptide. It is derived from bovine lung tissue and is the classic member of the Kunitz family protease inhibitors. It has a chain of 58 residues that folds into a compact and stable tertiary structure. The high stability of BPTI comes from three disulfide bonds linking the six cysteine residues of the chain. It has four tyrosine and four phenylalanine residues, and many of these side chains are close. Such aromatic clusters are important since they lead to mixing of optical transitions, thus enhancing the CD signal. BPTI has been characterized by X-ray diffraction in three different crystal forms and these are available in the Protein Data Bank (PDB codes 4pti, [108] 5pti, [109] 6pti[110]).

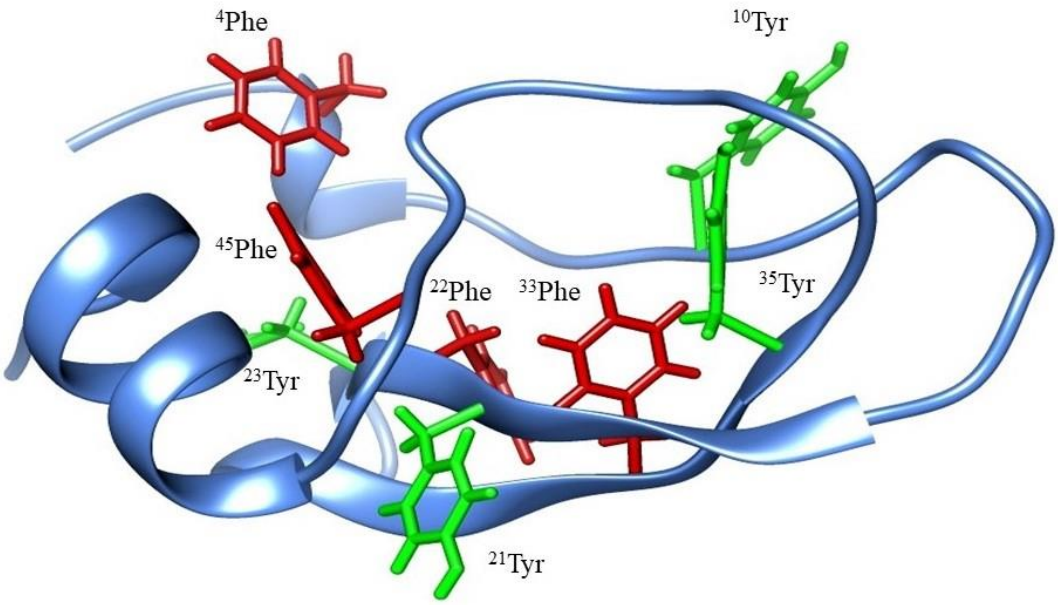



Figure 3.1. Bovine pancreatic trypsin inhibitor (1PIT), a model from the family of the NMR structure. Tyrosine in green and phenylalanine in red

This protein has a structure determined by NMR (PDB code 1pit [111]), comprising 20 models. The visualization of the structure of this protein and all the proteins in this thesis was done with Chimera [112] and is shown in Figure 3.1.

### 3.2.2 Cardiotoxin

Another protein used in this investigation is cardiotoxin. This protein belongs to a large family of snake venom toxins. It is a polymer protein with 60 residues, (PDB code 1CRF [113]).

It is derived from the venom of Taiwan cobra, naja atra. The structure determined by X-ray crystallography shows that this

protein has five chains with the same length of 60 residues (PDB code 4OM4 [114]). Only one chain is considered in the calculations. The 3D structure of this protein is highly stable because of four disulfide bonds. This protein has 12 models in the NMR structure. The structure is shown in Figure 3.2.

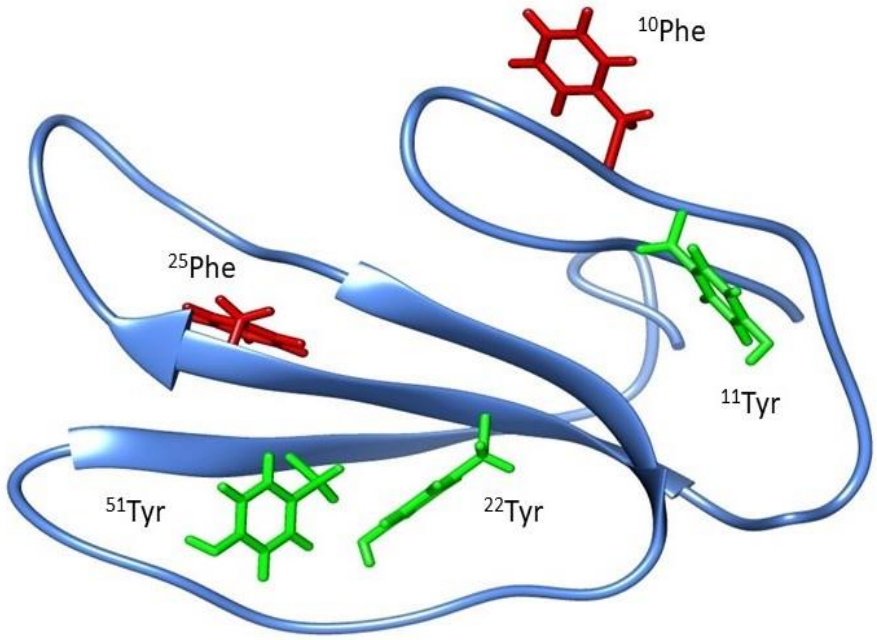



Figure 3.2. Cardiotoxin (1CRF), a model from the family of the NMR structure. Tyrosine in green and phenylalanine in red

### 3.2.3 Ribonuclease A

Ribonuclease A or RNase A is a widely studied protein. It is involved in the degradation of RNA. All organisms contain different types of this protein, showing that RNA degradation is a very ancient and important process. Also, it is involved in cleaning of

cellular RNA that is no longer required. It has a key role in the maturation of all RNA molecules, both messenger RNAs that carry genetic material for making proteins, and non-coding RNAs that function in varied cellular processes. In addition, active RNA degradation systems are a first defence against RNA viruses.

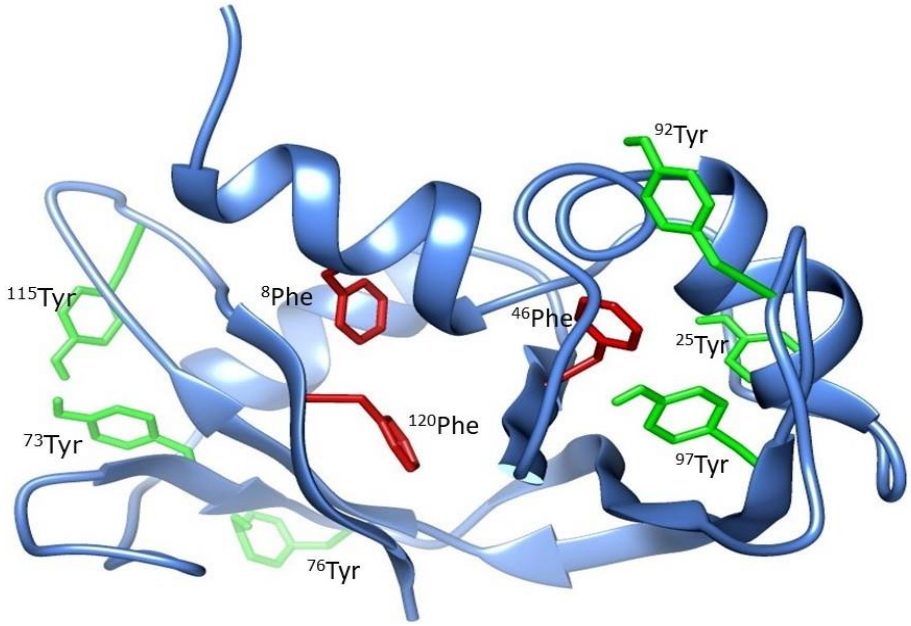



Figure 3.3. Ribonuclease A (2AAS), a model from the family of the NMR structure. Tyrosine in green and phenylalanine in red

It has two chains with a length of 124 residues and four disulfide bonds in each chain. Only one chain is considered in our calculations. The PDB code for the crystal structure is (1AFU [115]) and for NMR the code is (2AAS[116]), comprising 32 models. The structure is shown in Figure 3.3.

## 3.3. Methods

The matrix method [69, 70] is used, as implemented in our in-house software DichroCalc [32]. The parameters to construct the exciton Hamiltonian are from *ab initio* calculations of small molecules, which represent chromophores in proteins. The parameters consider, *N*-methyl acetamide (NMA) for studying the peptide bond while toluene, *p*-cresol and 3-methylindole are used to represent aromatic side chains of phenylalanine, tyrosine and tryptophan, respectively [36, 60].

In the parameter files, the excitations, the interactions between transitions and the permanent dipole moments are represented by monopoles, which have been modified to include electronic vibration structure based on the Franck Condon factor [48]. A line spectrum of rotational strength of each transition is generated and a band shape function [117] is added so that the calculated spectrum could be compared to the experimental spectrum. A Gaussian function has been applied to the calculated line spectra with a bandwidth of 4 nm.

Transitions of the peptide backbone ($n\pi^*$and $\pi\pi^*$) [118] and the $B_a$, $B_b$ and $L_a$ transitions of the aromatic side chains were included in the calculation, but no vibrational structure is considered for these transitions. For phenylalanine and tyrosine only $L_b$ electronic transitions states are considered, since their $L_a$ states are located at higher energies. For each protein only two models of the NMR

structures are chosen in our analysis. They were chosen based on the level of (dis)agreement with experiment as assessed by the root mean square deviation (RMSD) and mean relative error (MRE). These measures in Equations (3.1) and (3.2).

$$RMSD = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(r_i^x - r_i^y\right)^2} \qquad (3.1)$$

$$MRE = \sum_{i=1}^{N}\left|\frac{r_i^x - r_i^y}{r_i^y}\right| \qquad (3.2)$$

*N* is the total number of wavelengths considered, *i* is an integer index for the wavelength, $r_i^x$ is the intensity of the calculated CD spectrum at wavelength *i*, and $r_i^y$ is the intensity of the experimental spectrum at wavelength *i*.

To investigate the physical origin of variations of the calculated spectrum, the electrostatic potentials around the aromatic residues of interest in BPTI were calculated using the software APBS [119] (with the CHARMM force field). For a residue of interest, to compute the potential arising from its surrounding environment, the atomic charge and radius field parameters were set to zero in the APBS input (PQR file) for all atoms of that side chain. A Matlab script was used to interpolate the potentials to the atom centres on the residues of interest. A 2D plot of the electrostatic potential was generated using the Jmol visualisation software[120].

To examine the influence of the electrostatic environment and the structure on the predicted near-UV CD spectra, we choose BPTI out of the three proteins to discuss in detail in this chapter. To study the effect of the rest of the protein and assess its impact on the mechanism generating the calculated near-UV CD spectrum we consider the interpolated electrostatic potential at the atom centres of tyrosine chromophore in frame 3 of the NMR structures.

## 3.4. Results and discussion

### 3.4.1. The role of Tyr residues in generating CD spectrum

#### 3.4.1.1 Bovine pancreatic trypsin inhibitor

Most of the analysis of BPTI has been presented in our published work [65]. The near-UV CD spectra of BPTI calculated using the X-ray crystal structure (PDB code: 5PTI) and separately an average over the 20 NMR structures (PDB code: 1PIT) are compared to the experimental spectrum [121] in Figure. 3.4. The X-ray crystal structure leads to a computed spectrum that reproduces the overall shape and location of the peaks reasonably well, but the intensity across the wavelength range is noticeably underestimated.

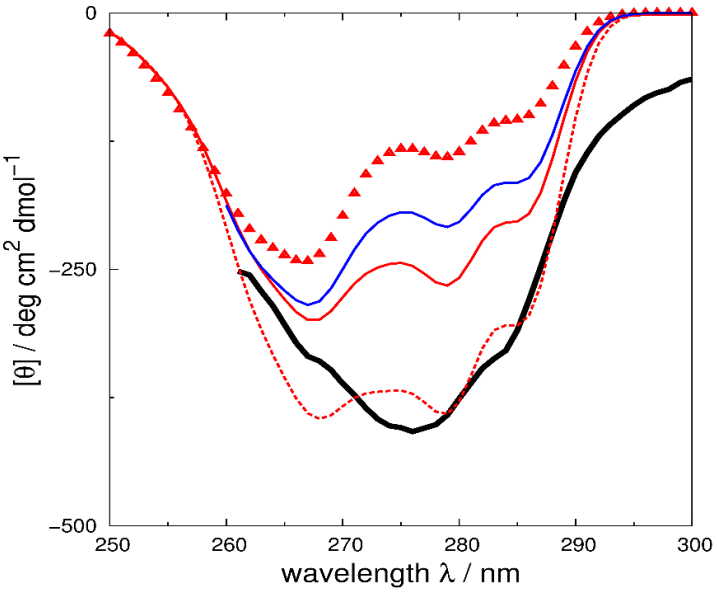



Figure 3.4. Experimental [121] near-UV CD spectrum (black solid line), and spectra calculated using the X-ray structure (blue solid line), calculated using an average over 20 frames of the NMR structure (red solid line) and calculated from individual frames of the NMR structure: frame 3 (red dotted line) and frame 17 (red triangles).

The spectrum calculated by averaging over the NMR structures is about 20% more intense and the RMSD with respect to the experimental spectrum is correspondingly reduced (from 128 to 91 deg cm$^2$ dmol$^{-1}$). The conformational diversity of BPTI in solution is reflected, at least in part, by the family of NMR structures. Figure 3.4 shows the CD spectrum calculated using specific NMR structures: frame 3 which gives the more intense spectrum (in closest agreement with experiment), and by way of contrast frame 17, which gives the calculated spectrum of weakest

intensity. The structural difference across these two frames is small, 0.9 Å all-atom RMSD and the RMSD for Tyr residues is 0.2 to 0.3 Å as calculated by the VMD software [122].

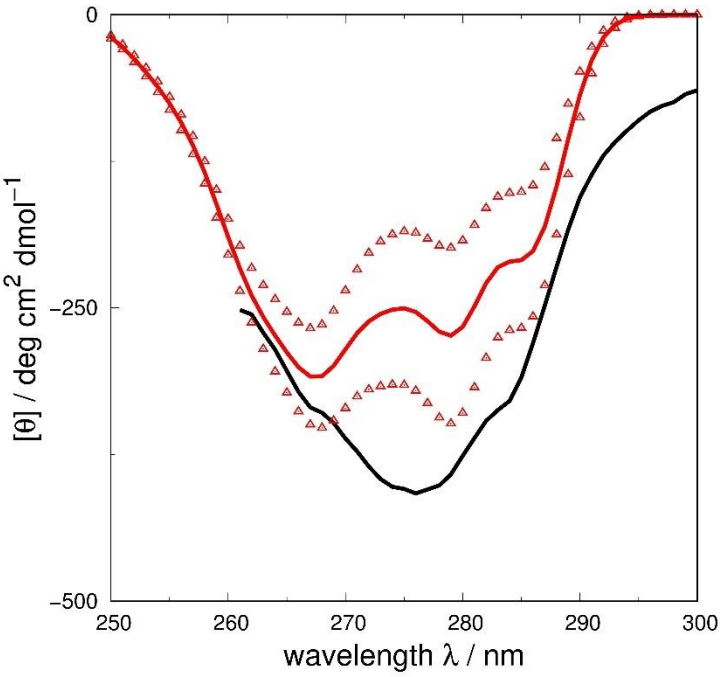



Figure 3.5. Conformational diversity investigation. The black solid line is the experimental spectrum [121], the red solid line is the average over 20 models and the red lines with triangles (upper and lower) represent the mean average ± standard deviation, respectively.

The level of sensitivity of the calculated CD spectra with respect to small variations in conformation is of interest. To investigate the influence of conformational diversity on the computed spectra, we calculated the standard deviation of the computed intensity at each wavelength over all the NMR structures as is shown in Figure

3.5. The spectrum computed from averaging over all the NMR frames is less intense than the experimental spectrum. However, one frame of the NMR structures frame-3 leads to a computed spectrum in excellent agreement (with an RMSD of 49 deg $cm^2$ $dmol^{-1}$) with experiment. The contributions of the phenylalanine and tyrosine chromophores are clearly well reproduced. The analysis will investigate only frame 3, which has the closest intensity and shape to the experimental spectrum.

The good agreement and the sensitivity to modest structural differences motivated us to investigate the detailed interactions that govern the spectrum. Table 3.2 shows the significant interaction energies for phenylalanine and tyrosine. The interaction is presented only if it is greater than $\pm 1$ $cm^{-1}$. It appears that only $v1$ and $v2$ electronic vibrational levels of $^1L_b$ states showed significant interactions. The zero values (0.0 Å) in the separation column correspond to the one-electron mixing of two transitions on the same chromophore. Coupling between L states and B states (the latter occurring in the far-UV) is weak.

The largest interaction energy is 84.9 $cm^{-1}$ and corresponds to the one-electron mixing of the $^1L_b$, $v2$ and $^1L_a$ transitions of $^{23}$Tyr. $v2$ denotes that the transition is to the second vibrational level of the excited state. The analogous transitions on $^{10}$Tyr interact via the same mechanism and with a very similar magnitude. Inter-

chromophore interactions are much weaker, and with the exception of the interaction between $^{22}$Phe and $^{33}$Phe (27.4 cm$^{-1}$), they are all less than 10 cm$^{-1}$; for example, the interaction between $^{22}$Phe $^{1}L_b$, $v1$ and $^{45}$Phe $^{1}L_a$ is $-1.4$ cm$^{-1}$.

Table 3.2: Calculated interaction energies (shown only if greater than 1 cm$^{-1}$) of electronic transitions of aromatic side chains in frame 3 of the NMR structure.

| Chromophore/Transitions | Chromophore/Transitions | Distance (Å) | Interaction Energy (cm$^{-1}$) |
|---|---|---|---|
| $^{4}$Phe-$^{1}L_{b1}$ | $^{45}$Phe-$^{1}L_{b1}$ | 9.6 | 1.3 |
| $^{22}$Phe-$^{1}L_{b1}$ | $^{33}$Phe-$^{1}L_{a}$ | 5.2 | 27.4 |
| $^{22}$Phe-$^{1}L_{b1}$ | $^{45}$Phe-$^{1}L_{a}$ | 7.8 | -1.4 |
| $^{22}$Phe-$^{1}L_{b1}$ | $^{23}$Tyr-$^{1}L_{b2}$ | 9.8 | -1.2 |
| $^{33}$Phe-$^{1}L_{a}$ | $^{23}$Tyr-$^{1}L_{a}$ | 8.8 | 1.9 |
| $^{33}$Phe-$^{1}L_{a}$ | $^{35}$Tyr-$^{1}L_{b2}$ | 8.9 | -5.4 |
| $^{45}$Phe-$^{1}L_{b1}$ | $^{21}$Tyr-$^{1}L_{a}$ | 8.6 | 3.5 |
| $^{10}$Tyr-$^{1}L_{b2}$ | $^{10}$Tyr-$^{1}L_{a}$ | 0.0 | 80.9 |
| $^{10}$Tyr-$^{1}L_{a}$ | $^{35}$Tyr-$^{1}L_{b2}$ | 8.9 | 2.8 |
| $^{21}$Tyr-$^{1}L_{b2}$ | $^{35}$Tyr-$^{1}L_{a}$ | 13.5 | 1.6 |
| $^{23}$Tyr-$^{1}L_{b2}$ | $^{23}$Tyr-$^{1}L_{a}$ | 0.0 | 84.9 |
| $^{23}$Tyr-$^{1}L_{a}$ | $^{35}$Tyr-$^{1}L_{a}$ | 19.8 | 1.6 |
| $^{35}$Tyr-$^{1}L_{b2}$ | $^{35}$Tyr-$^{1}L_{a}$ | 0.0 | 6.7 |

From the interaction energies in Table 3.2, it is clear that the one-electron mixing of the tyrosine $^{1}L_b$ states is an important factor in the appearance of the near-UV spectrum. The relative strong intensity of the calculated CD spectrum of frame 3 is highlighted as a point of investigation.

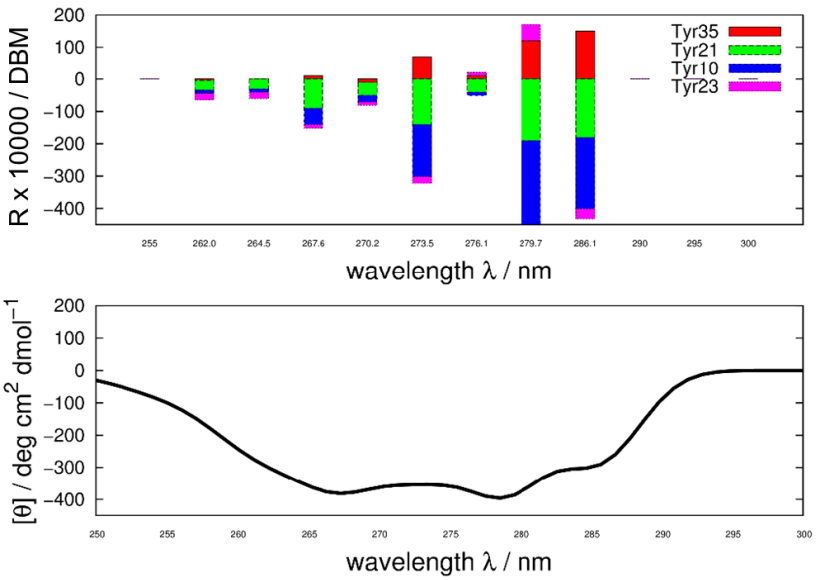



Figure 3.6: Calculated contributions for each Tyr at different $\lambda$ for frame 3 of bovine pancreatic trypsin inhibitor (1PIT). The upper panel shows calculated rotational strengths (in DBM) multiplied by 10000 for each residue at each wavelength. The lower panel shows the calculated spectrum.

The contributions of the Tyr residues are shown in Figure 3.6, along with the calculated CD spectrum for frame 3.

Here we can see that $^{10}$Tyr and $^{21}$Tyr make the biggest contributions to the overall calculated rotational strength for the CD spectrum of this structure. Both residues contribute towards the strong intensity negative peaks across the near-UV region. Figure 3.6 shows that $^{23}$Tyr makes a direct but very modest contribution to the combined negative rotational strength. $^{35}$Tyr makes a positive contribution.

It is clear that tyrosine residues are mainly responsible for generating the near-UV CD-spectrum for this protein.

Further investigations have been carried out to examine the role of each aromatic residue.

Table 3.3: Wavelength and associated eigenvector coefficients of tyrosine side chain transitions in frame 3 of the NMR structure.

| Eigenvector index | Values of Eigenvector coefficients | Participation ratio (W) | Residues/ Transitions | Wavelength (nm) |
|---|---|---|---|---|
| 1 | 0.98 | 1.0 | $^{35}Tyr_{Lb1}$ | 286.11 |
| 2 | 0.99 | 1.0 | $^{21}Tyr_{Lb1}$ | 286.08 |
| 3 | 0.99 | 1.0 | $^{10}Tyr_{Lb1}$ | 286.05 |
| 4 | 0.99 | 1.0 | $^{23}Tyr_{Lb1}$ | 286.02 |
| 5 | 0.98 | 1.0 | $^{35}Tyr_{Lb2}$ | 279.74 |
| 6 | 0.99 | 1.0 | $^{21}Tyr_{Lb2}$ | 279.71 |
| 7 | 0.99 | 1.0 | $^{10}Tyr_{Lb2}$ | 279.66 |
| 8 | 0.99 | 1.0 | $^{23}Tyr_{Lb2}$ | 279.63 |

Table 3.4: Wavelength and associated eigenvector coefficients of phenylalanine side chain in frame 3 of the NMR structure.

| Eigenvector index | Values of Eigenvector coefficients | Participation ratio (W) | Residues/ Transitions | Wavelength (nm) |
|---|---|---|---|---|
| 21 | -0.71 | 2.7 | $^{33}Phe_{Lb1}$ | 268.08 |
| 22 | -0.73 | 2.6 | $^{45}Phe_{Lb1}$ | 268.04 |
| 23 | -0.73 | 2.6 | $^{4}Phe_{Lb1}$ | 269.02 |
| 24 | 0.71 | 2.7 | $^{22}Phe_{Lb1}$ | 267.98 |
| 33 | 0.71 | 2.7 | $^{33}Phe_{Lb2}$ | 264.02 |
| 34 | -0.73 | 2.6 | $^{45}Phe_{Lb2}$ | 264.00 |
| 35 | -0.73 | 2.7 | $^{4}Phe_{Lb2}$ | 263.99 |
| 36 | 0.71 | 2.7 | $^{22}Phe_{Lb2}$ | 263.97 |

The wavelengths and the associated eigenvector coefficients for tyrosine and phenylalanine residues, is investigated and some of the data are presented in Tables 3.3 and 3.4, respectively.

From these tables we can see that the tyrosine transitions have eigenvector coefficients close to one with a participation ratio *(W*) also close to one. The participation factors represent the localisation of the transition on a specific chromophore and it is calculated based on the following equation [123]:

$$W = \frac{1}{\sum_{i=1}^{N} C_i^4} \qquad (3.3)$$

*N* is the number of transitions; *C* represents the values of the eigenvector coefficients. For phenylalanine almost all the eigenvector coefficients are close to 0.7 with a participation ratio of 2.6 which represents the delocalisation of the transitions. These data emphasise that the one electron mechanism appears clearly in the tyrosines in BPTI.

This is confirmed by a detailed examination of the intensity (i.e., rotational strength) of each transition in the calculated spectrum. The most intense transitions in this region arise from the tyrosine residues, with the phenylalanine transitions making a more modest contribution; the electric and magnetic transition dipole moments of the phenylalanine residues are substantially smaller than those for the tyrosine residues. The data in Table 3.5 are computed using frame 3 of the NMR structure.

Table 3.5: The eight ($\nu$ =1 to 8) tyrosine $^1L_b$ vibrational states, their wavelengths ($\lambda$) and their calculated contributions to the rotational strength ($R_\lambda$)in Debye Bohr Magneton (DBM) at that wavelength in the near-UV CD spectrum of frame 3 of the NMR structure.

| $\nu$ | $\lambda$/nm | $^{35}$Tyr | | $^{21}$Tyr | | $^{10}$Tyr | | $^{23}$Tyr | | $R_\lambda \times 1000$ / DBM |
|---|---|---|---|---|---|---|---|---|---|---|
| | | $\mu_1 m_1$ | $\cos\theta_1$ | $\mu_2 m_2$ | $\cos\theta_2$ | $\mu_3 m_3$ | $\cos\theta_3$ | $\mu_4 m_4$ | $\cos\theta_4$ | |
| 1 | 286.1 | 0.52 | 0.03 | 0.39 | -0.06 | 0.45 | -0.05 | 0.39 | -0.01 | -29 |
| 2 | 279.7 | 0.64 | 0.02 | 0.39 | -0.05 | 0.59 | -0.05 | 0.51 | -0.01 | -34 |
| 3 | 276.1 | 0.05 | 0.03 | 0.01 | -0.47 | 0.02 | -0.06 | 0.01 | -0.09 | -8 |
| 4 | 273.5 | 0.36 | 0.02 | 0.21 | -0.07 | 0.33 | -0.05 | 0.29 | -0.01 | -25 |
| 5 | 270.2 | 0.05 | 0.02 | 0.01 | -0.42 | 0.03 | -0.07 | 0.02 | -0.07 | -9 |
| 6 | 267.6 | 0.13 | 0.01 | 0.06 | -0.16 | 0.11 | -0.05 | 0.09 | -0.02 | -16 |
| 7 | 264.5 | 0.04 | 0.00 | 0.01 | -0.37 | 0.01 | -0.09 | 0.01 | -0.19 | -7 |
| 8 | 262.0 | 0.04 | -0.01 | 0.01 | -0.36 | 0.01 | -0.09 | 0.01 | -0.19 | -7 |

This protein has four tyrosines; thus, there are four contributions at each wavelength. The $\mu m$ value given in the table is not the scalar product, but simply the product of the magnitudes of the electric and magnetic transition dipole moments. In an achiral environment, the electric and magnetic transition dipole moments would be perpendicular ($\cos\theta = 0$). The rotational strength, $R = \mu m \cos\theta$, where $\mu$ is the magnitude of the electric transition dipole moment, $m$ is the magnitude of the magnetic transition dipole moment and $\theta$ is the angle between them. In the absence of any chirality, the electric and magnetic transition dipole moments of an excitation would be perpendicular, and the

rotational strength, which is directly proportional to the dot product of these two vectors, would be zero. Even small deviations from perpendicularity give rise to a non-zero intensity.

Table 3.5 shows for each tyrosine $^1L_b$ transition, the contribution made by the angle between the electric and magnetic transition dipole moments to the intensity. An angle less than 90° would lead to a positive intensity; an angle greater than 90° gives a negative rotational strength. The angles in Table 3.5 exhibit some variation and a delicate balance (of positive and negative contributions) across the transitions is responsible for the overall intensity. The predominance of the one-electron mechanism in the near-UV CD of BPTI leads to significant sensitivity of the spectrum to the precise nature of the local environment surrounding the tyrosine residues.

### 3.4.1.2. Cardiotoxin

The near-UV CD spectra of this protein calculated using the X-ray crystal structure (PDB code: 4OM4) and separately an average over the 20 NMR structures (PDB code:1CRF) in comparison to the experimental spectrum [45] are presented in Figure 3.7. The overall shape is not well reproduced by either X-ray crystal structure or the average of the NMR structures. Peaks are clearly missed, especially in the wavelength range between 250 and 270 nm. The

average CD spectrum computed from the NMR structure showed some positive intensities similar to the experimental spectrum. The variation depending on the structure emphasises the effect of conformation on the spectrum.

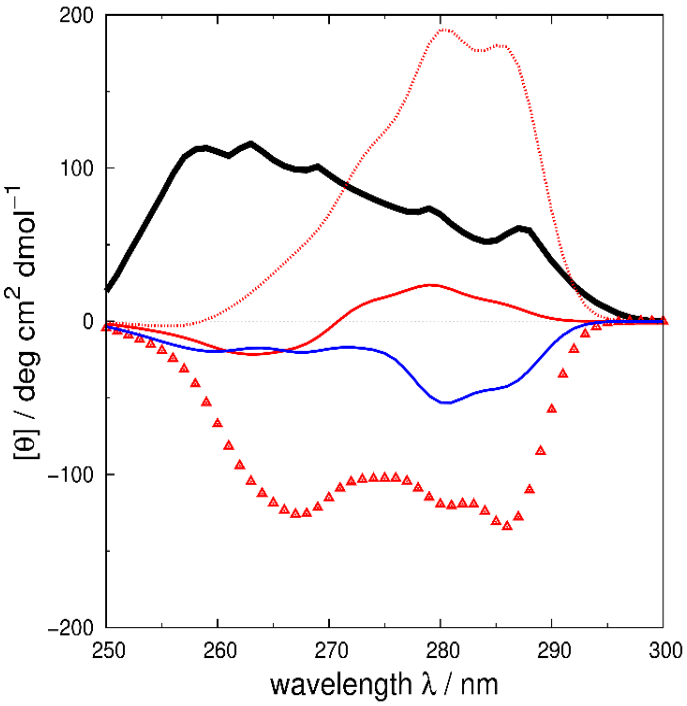



Figure 3.7. Experimental [45] near-UV CD spectrum (black solid line), and spectra calculated using the X-ray structure (blue solid line), calculated using an average over 12 frames of the NMR structure (red solid line) and calculated from individual frames of the NMR structure: frame 7 ( red triangles) and frame 3 (red dotted line).

Figure 3.7 shows the CD spectrum calculated using X-ray crystal structure and different NMR structures: frame 3 which gives the

more intense spectrum (in closest agreement with experiment), and by way of contrast frame 7, which gives the calculated spectrum of most negative intensity. The structural difference across these two frames is small, 1.6 Å all-atom RMSD and the RMSD for Tyr residues is 0.5 to 0.7 Å.

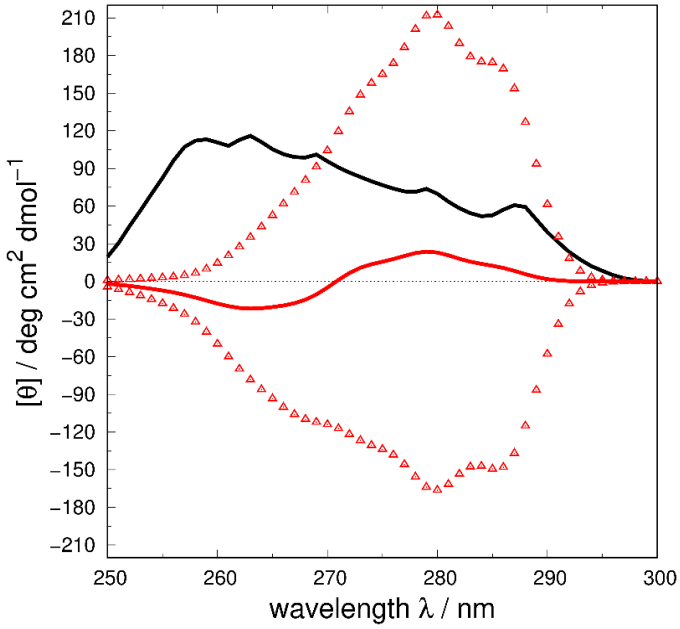



Figure 3.8. Conformational diversity investigation .The black solid line is the experimental spectrum [45], the red solid line is the average over 12 models and the red lines with triangles (upper and lower) represent the mean average ± standard deviation, respectively.

To investigate the influence of conformational diversity on the computed spectra, we calculated the standard deviation of the computed intensity at each wavelength over all the NMR

structures as shown in Figure 3.8. The spectrum computed from averaging over all the NMR frames is less intense than the experimental spectrum. However, one frame of the NMR structures frame 3 leads to computed spectra with an RMSD of 150 deg $cm^2$ $dmol^{-1}$ compared to experiment. However, this frame does not reproduce the overall shape or peaks positions well. Despite the limited agreement, we investigated the detailed interactions that govern the calculated spectrum. All the analysis was carried on frame 3 of the NMR structures.

Table 3.6: Calculated interaction energies (shown only if greater than 10 $cm^{-1}$) of electronic transitions of aromatic side chains in frame 3 of the NMR structure.

| Chromophore/Transitions | Chromophore/Transitions | Distance (Å) | Interaction Energy ($cm^{-1}$) |
|---|---|---|---|
| $^{22}$Tyr-$^{1}L_{b6}$ | $^{22}$Tyr-$^{1}L_{a}$ | 0.0 | 22.0 |
| $^{22}$Tyr-$^{1}L_{b4}$ | $^{22}$Tyr-$^{1}L_{a}$ | 0.0 | 35.3 |
| $^{22}$Tyr-$^{1}L_{b}$ | $^{22}$Tyr-$^{1}L_{a}$ | 0.0 | 40.8 |
| $^{22}$Tyr-$^{1}L_{b2}$ | $^{22}$Tyr-$^{1}L_{a}$ | 0.0 | 45.5 |
| $^{22}$Tyr-$^{1}L_{a}$ | $^{51}$Tyr-$^{1}L_{b}$ | 6.10 | 33.1 |
| $^{22}$Tyr-$^{1}L_{a}$ | $^{51}$Tyr-$^{1}L_{b6}$ | 6.10 | 17.9 |
| $^{22}$Tyr-$^{1}L_{a}$ | $^{51}$Tyr-$^{1}L_{b5}$ | 6.10 | 11.5 |
| $^{22}$Tyr-$^{1}L_{a}$ | $^{51}$Tyr-$^{1}L_{b4}$ | 6.10 | 28.7 |
| $^{22}$Tyr-$^{1}L_{a}$ | $^{51}$Tyr-$^{1}L_{b3}$ | 6.10 | 10.8 |
| $^{22}$Tyr-$^{1}L_{a}$ | $^{51}$Tyr-$^{1}L_{b2}$ | 6.10 | 37.0 |
| $^{51}$Tyr-$^{1}L_{b}$ | $^{51}$Tyr-$^{1}L_{a}$ | 0.0 | 24.0 |
| $^{51}$Tyr-$^{1}L_{b4}$ | $^{51}$Tyr-$^{1}L_{a}$ | 0.0 | 20.7 |
| $^{51}$Tyr-$^{1}L_{b2}$ | $^{51}$Tyr-$^{1}L_{a}$ | 0.0 | 26.7 |

Table 3.6 shows the significant interactions energies for tyrosine. The interaction is presented only if it is greater than $\pm 10$ cm$^{-1}$. The energies below this threshold are presented in Table A.1 in the Appendix. The following electronic vibrational levels $v1, v2, v3, v4, v5$ and $v6$ of the $^1L_b$ states showed significant interactions. The zero values (0.0 Å) in the separation column correspond to the one-electron mixing of two transitions on the same chromophore. The largest interaction energy is 45.5 cm$^{-1}$ and corresponds to the one-electron mixing of the $^1L_b$, $v2$ and $^1L_a$ transitions of $^{22}$Tyr $v2$ denotes that the transition is to the second vibrational level of the excited state.

It appears that mainly $^{22}$Tyr and $^{51}$Tyr have the strongest interactions. Interactions of other chromophores of this protein for example $^{25}$Phe and $^{11}$Tyr, are not shown because they are lower than the threshold that we have chosen for this protein. It is clear that the one-electron mixing of the tyrosine $^1L_b$ state and $^1L_a$ states are important factors in the appearance of the near-UV spectrum.

The contributions of the Tyr residues are shown in Figure 3.9. Here we can see that $^{51}$Tyr and $^{22}$Tyr make the biggest contributions to the calculated rotational strength. $^{51}$Tyr contributes negative intensity across the near-UV region, whereas $^{22}$Tyr has a positive contribution.

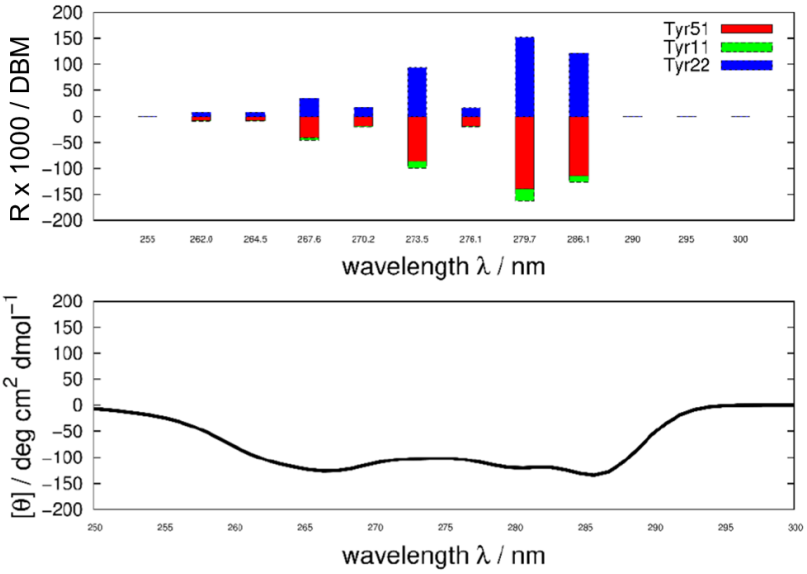



Figure 3.9. Calculated contributions for each Tyr at different wavelengths for frame 3 of cardiotoxin (1CRF). The upper panel shows calculated rotational strengths (in DBM) multiplied by 1000 for each residue at each wavelength. The lower panel shows the calculated spectrum.

$^{11}$Tyr makes weak contributions but it is important in leading to the calculated CD spectrum. The positive contribution of $^{22}$Tyr, reduces the contribution of the strong negative intensity CD spectrum of frame 3. It is clear that the near-UV CD-spectrum for this protein arises largely from the tyrosine residues. Further investigations have been carried to examine the role of each aromatic residue. The wavelengths (corresponding to the transition energies) and the associated eigenvectors coefficients for tyrosine residues are investigated and some of the data are presented in Table 3.7.

Table 3.7: Wavelength and associated eigenvector coefficients of tyrosine side chain in frame 3 of the NMR structures.

| Eigenvector index | Values of Eigenvector coefficients | Participation ratio (W) | Residues/ Transitions | Wavelength (nm) |
|---|---|---|---|---|
| 7 | 0.80 | 2.4 | $^{51}Tyr_{Lb6}$ | 276.1 |
| 8 | 0.99 | 1.0 | $^{11}Tyr_{Lb6}$ | 276.1 |
| 9 | 0.80 | 2.4 | $^{22}Tyr_{Lb6}$ | 276.1 |
| 10 | -0.80 | 2.4 | $^{51}Tyr_{Lb5}$ | 273.6 |
| 11 | 0.99 | 1.0 | $^{11}Tyr_{Lb5}$ | 273.5 |
| 12 | 0.80 | 2.4 | $^{22}Tyr_{Lb5}$ | 273.5 |
| 13 | 0.80 | 2.4 | $^{51}Tyr_{Lb4}$ | 270.2 |
| 14 | 0.99 | 1.0 | $^{11}Tyr_{Lb4}$ | 270.1 |
| 15 | -0.80 | 2.4 | $^{22}Tyr_{Lb4}$ | 270.1 |

The tyrosine residues have varying eigenvector coefficients. For example, most the values of $^{22}$Tyr and $^{51}$Tyr are close to 0.8 with a participation ratio of 2.4. $^{11}$Tyr the transitions show little mixing with eigenvectors and participation ratios of 1.0. Data are not shown for $^{10}$Phe and $^{25}$Phe because all transitions of these residues have eigenvector coefficients close to 1.0 and their participation ratios are 1.0. The transitions of phenylalanines are thus highly localised. A detailed examination of the intensity (i.e., rotational strength) of each transition in the calculated spectrum is shown in Table 3.8. The most intense transitions in this region arise from the tyrosine residues. The electric and magnetic transition dipole

moments of the phenylalanine residues are substantially smaller than those for the tyrosine residues, so their data are not shown.

Table 3.8: The eight ($v$ =1 to 8) tyrosine $^1L_b$ vibrational states, their wavelengths ($\lambda$) and their calculated contributions to the rotational strength ($R_\lambda$) in Debye Bohr Magneton (DBM) at that wavelength in the near-UV CD spectrum of frame 3 of the NMR structures

| $v$ | $\lambda$/nm | $^{51}$Tyr | | $^{11}$Tyr | | $^{22}$Tyr | | $R_\lambda$ ×1000/ DBM |
|---|---|---|---|---|---|---|---|---|
| | | $\mu_1 m_1$ | $\cos\theta_1$ | $\mu_2 m_2$ | $\cos\theta_2$ | $\mu_3 m_3$ | $\cos\theta_3$ | |
| 1 | 286.1 | 0.38 | -0.30 | 0.43 | -0.03 | 0.16 | 0.76 | -13 |
| 2 | 279.7 | 0.50 | -0.28 | 0.55 | -0.04 | 0.20 | 0.76 | -10 |
| 3 | 276.1 | 0.02 | -0.98 | 0.05 | -0.02 | 0.02 | 0.83 | -4 |
| 4 | 273.5 | 0.27 | -0.32 | 0.32 | -0.04 | 0.12 | 0.78 | -5 |
| 5 | 270.2 | 0.02 | -0.95 | 0.05 | -0.03 | 0.02 | 0.85 | -3 |
| 6 | 267.6 | 0.08 | -0.51 | 0.12 | -0.04 | 0.04 | 0.86 | -5 |
| 7 | 264.5 | 0.01 | -0.91 | 0.03 | -0.02 | 0.01 | 0.79 | -3 |
| 8 | 262.0 | 0.01 | -0.92 | 0.04 | -0.02 | 0.01 | 0.79 | -3 |

The data in Table 3.8 are computed using frame 3 of the NMR structure. This protein has three tyrosines; thus, there are three contributions at each wavelength. The strongest rotational strength values ($-13\times10^{-3}$ and $-10\times10^{-3}$) resulted from the $^1L_b$ transitions of the three tyrosines in the wavelength region between 279.2 and 286.1 nm. The contribution to the intensity is high in the tyrosines. The angle between the magnitudes of the electric and magnetic transition dipole moments are larger than 90°for $^{51}$Tyr and $^{11}$Tyr with a negative $\cos\theta$. $^{22}$Tyr has a positive

$cos\theta$, so we can conclude that the angle in this case is less than 90°.

### 3.4.1.3. Ribonuclease A

Figure 3.10 shows the calculated near-UV CD spectrum of RNase A computed by using the X-ray crystallography structure and the average spectrum from the NMR structures (PDB code: 1AFU and 2AAS), in comparison to the experimental spectrum [124].

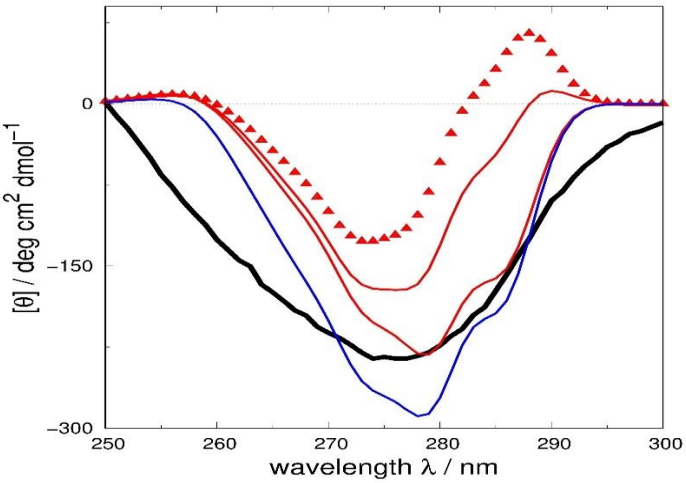



Figure 3.10. Experimental [124] near-UV CD spectrum (black solid line), and spectra calculated using the X-ray structure (blue solid line), calculated using an average over 32 frames of the NMR structure (red solid line) and calculated from individual frames of the NMR structure: frame 1 (red triangles) and frame 12 (red line).

There is modest quantitative (RMSD drops from 123 to 49 cm$^{-1}$) improvement with the new vibrational parameters. A detailed explanation about the improvement showed that a combination of

the NMR structures with the '*vib*' parameters better reproduced the experimental spectra than with the '*nonvib*' parameters. Calculation with the NMR structures predicted a less intense spectrum respect to the an experimental data. The X-ray crystal structure leads to a computed spectrum that reproduces the overall shape and location of the peaks reasonably well, but the calculated intensity in general is weaker than the intensity of experimental spectrum.

Figure 3.10 shows the CD spectrum calculated using specific NMR structures: frame 12 which gives the most intense spectrum (in closest agreement with experiment, but less intense than the spectrum generated from the X-ray crystal structure), and by way of contrast frame 1, which gives the calculated spectrum of weakest intensity and a positive peak in region 290 nm, which is not observed in the experimental spectrum. The structural difference across these two frames is small, 0.7 Å all-atom RMSD. To investigate the influence of conformational diversity on the computed spectra we calculated the standard deviation of the computed intensity at each wavelength over all the NMR structures as shown in Figure 3.11.

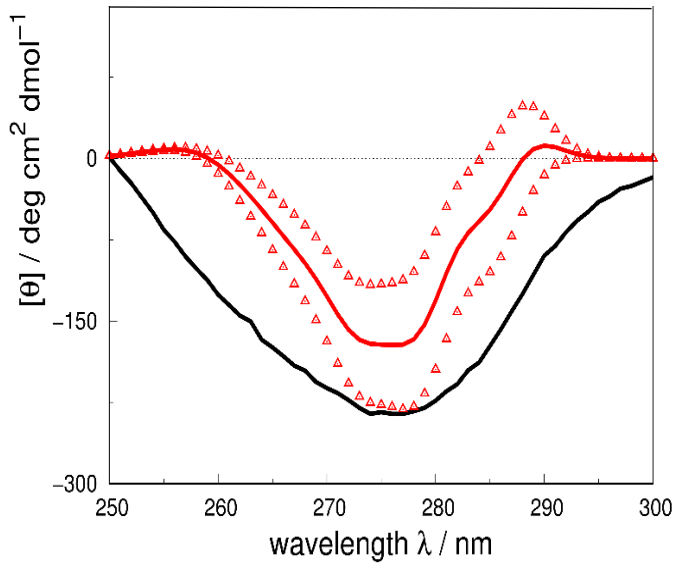



Figure 3.11. Conformational diversity investigation .The black solid line is the experimental [124], the red solid line is the average over 32 models and the red lines with triangles (upper and lower) represent the mean average ± standard deviation, respectively.

The computed spectrum from averaging over all the NMR frames is less intense than the experimental spectrum. However, one frame of the NMR structures frame 12 leads to a computed spectrum in close agreement (with an RMSD of 78 deg cm$^2$ dmol$^{-1}$) with experiment. The contributions of the phenylalanine and tyrosine chromophores have acceptable intensities. The focus in the analysis will be on frame 12.

Table 3.9: Calculated interaction energies (shown only if greater than 10 $cm^{-1}$) of electronic transitions of aromatic side chains in frame 12 of the NMR structures.

| Chromophore/Transitions | Chromophore/Transitions | Distance (Å) | Interaction Energy ($cm^{-1}$) |
|---|---|---|---|
| 8Phe-$^1L_b$ | 120Phe-$^1L_a$ | 6.3 | 11.9 |
| 8Phe-$^1L_b$ | 120Phe-$^1L_{b3}$ | 6.3 | -10.0 |
| 46Phe-$^1L_a$ | 25Tyr-$^1L_a$ | 5.8 | 16.1 |
| 25Tyr-$^1L_{b4}$ | 25Tyr-$^1L_a$ | 0.0 | 14.6 |
| 25Tyr-$^1L_b$ | 25Tyr-$^1L_a$ | 0.0 | 16.9 |
| 25Tyr-$^1L_{b2}$ | 25Tyr-$^1L_a$ | 0.0 | 18.9 |
| 73Tyr-$^1L_{b8}$ | 73Tyr-$^1L_a$ | 0.0 | 24.8 |
| 73Tyr-$^1L_{b7}$ | 73Tyr-$^1L_a$ | 0.0 | 24.8 |
| 73Tyr-$^1L_b$ | 73Tyr-$^1L_a$ | 0.0 | 92.0 |
| 73Tyr-$^1L_{b6}$ | 73Tyr-$^1L_a$ | 0.0 | 49.5 |
| 73Tyr-$^1L_{b5}$ | 73Tyr-$^1L_a$ | 0.0 | 31.8 |
| 73Tyr-$^1L_{b4}$ | 73Tyr-$^1L_a$ | 0.0 | 80.0 |
| 73Tyr-$^1L_{b3}$ | 73Tyr-$^1L_a$ | 0.0 | 30.0 |
| 73Tyr-$^1L_{b2}$ | 73Tyr-$^1L_a$ | 0.0 | 102.7 |
| 115Tyr-$^1L_{b8}$ | 115Tyr-$^1L_a$ | 0.0 | 10.3 |
| 115Tyr-$^1L_{b7}$ | 115Tyr-$^1L_a$ | 0.0 | 10.3 |
| 115Tyr-$^1L_{b6}$ | 115Tyr-$^1L_a$ | 0.0 | 20.5 |
| 115Tyr-$^1L_{b5}$ | 115Tyr-$^1L_a$ | 0.0 | 13.2 |
| 115Tyr-$^1L_{b4}$ | 115Tyr-$^1L_a$ | 0.0 | 33.0 |
| 115Tyr-$^1L_{b3}$ | 115Tyr-$^1L_a$ | 0.0 | 12.5 |
| 115Tyr-$^1L_{b2}$ | 115Tyr-$^1L_a$ | 0.0 | 42.5 |
| 115Tyr-$^1L_b$ | 115Tyr-$^1L_a$ | 0.0 | 38.0 |

Table 3.9 shows the significant interaction energies for phenylalanine and tyrosine. The interaction is presented only if it is greater than $\pm$10 $cm^{-1}$. The energies below this threshold are presented in Table A.2 in the Appendix. It appears that the eight electronic vibrational levels of $^1L_b$ states showed significant

interactions from $v1$ to $v8$. The zero values (0.0 Å) in the separation column corresponds to the one-electron mixing of two transitions on the same chromophore. Coupling between L states and B states (the latter occurring in the far-UV) is weak.

The largest interaction energy is 102.7 $cm^{-1}$ and corresponds to the one-electron mixing of the $^1L_b$, $v2$ and $^1L_a$ transitions of $^{73}$Tyr. The interactions related to two different phenylalanine chromophores are small (-10 and 11.9 $cm^{-1}$). $^{115}$Tyr and $^{25}$Tyr show weak interactions in comparison to the values of the interaction energies of $^{73}$Tyr. From this table we can see that the one electron mechanism of tyrosine has a predominant role in generating the near-UV CD spectrum.

Figure 3.12 shows the contributions of the Tyr residues, along with the calculated CD spectrum for frame 12. $^{73}$Tyr makes the biggest contributions to the CD spectrum of this structure. This residue contributes towards the strong negative peaks across the near-UV region. It shows that $^{115}$Tyr and $^{25}$Tyr make a direct contribution to the combined negative rotational strength, but with less effect. The strength of these contributions, however, is modest. For $^{76}$Tyr a positive contribution appears clearly at the wavelength 286.1 nm. This positive contribution could explain the shifting towards less intensity that appears in the calculated spectrum at the wavelength mentioned.

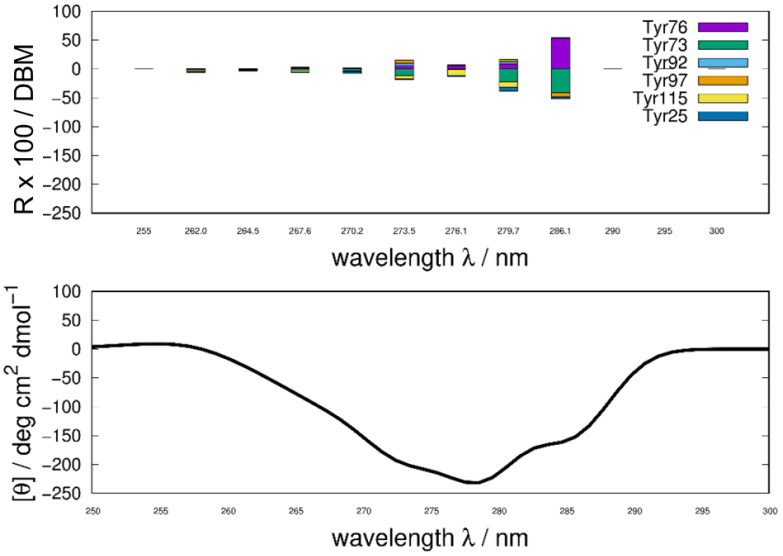



Figure 3.12. Calculated contributions for each Tyr at different wavelengths for frame 12 of ribonuclease A (2AAS). The upper panel shows calculated rotational strengths (in DBM) multiplied by 100 for each residue at each wavelength. The lower panel shows the calculated spectrum.

The wavelengths and the associated eigenvector coefficients for tyrosine and phenylalanine residues are investigated and some of the data are presented in Tables 3.10 and 3.11, respectively. These tables are for frame 12 of the NMR structures.

Table 3.10: Wavelength and associated eigenvector coefficients of tyrosine side chains in frame 12 of the NMR structures.

| Eigenvector index | Values of Eigenvector coefficients | Participation ratio (W) | Residues/ Transitions | Wavelength (nm) |
|---|---|---|---|---|
| 1 | 0.98 | 1.0 | $^{76}Tyr_{Lb8}$ | 286.1 |
| 2 | 0.94 | 1.0 | $^{73}Tyr_{Lb8}$ | 286.1 |
| 3 | 0.85 | 1.5 | $^{92}Tyr_{Lb8}$ | 286.0 |
| 4 | -0.85 | 1.5 | $^{97}Tyr_{Lb8}$ | 286.0 |
| 5 | -0.96 | 1.0 | $^{115}Tyr_{Lb8}$ | 286.0 |
| 6 | -0.99 | 1.0 | $^{25}Tyr_{Lb8}$ | 286.0 |
| 7 | -0.98 | 1.0 | $^{76}Tyr_{Lb7}$ | 279.7 |
| 8 | -0.95 | 1.0 | $^{73}Tyr_{Lb7}$ | 279.7 |
| 9 | -0.85 | 1.5 | $^{92}Tyr_{Lb7}$ | 279.6 |
| 10 | 0.85 | 1.5 | $^{97}Tyr_{Lb7}$ | 279.6 |
| 11 | 0.96 | 1.0 | $^{115}Tyr_{Lb7}$ | 279.6 |
| 12 | 0.99 | 1.0 | $^{25}Tyr_{Lb7}$ | 279.6 |

Table 3.11: Wavelength and associated eigenvector coefficients of phenylalanine side chain in frame 12 of the NMR structures.

| Eigenvectors | Values of Eigenvectors coefficients | Participation ratio (W) | Residues/ Transitions | Wavelength (nm) |
|---|---|---|---|---|
| 55 | 0.71 | 2.5 | $^{8}Phe_{Lb4}$ | 261.1 |
| 56 | 0.99 | 1.0 | $^{46}Phe_{Lb4}$ | 261.1 |
| 57 | 0.71 | 2.5 | $^{120}Phe_{Lb4}$ | 261.1 |
| 58 | 0.71 | 2.5 | $^{8}Phe_{Lb3}$ | 257.7 |
| 59 | -0.99 | 1.0 | $^{46}Phe_{Lb3}$ | 257.7 |
| 60 | 0.71 | 2.5 | $^{120}Phe_{Lb3}$ | 257.7 |
| 61 | 0.71 | 2.5 | $^{8}Phe_{Lb2}$ | 254.7 |
| 62 | -0.99 | 1.0 | $^{46}Phe_{Lb2}$ | 245.7 |
| 63 | 0.71 | 2.5 | $^{120}Phe_{Lb2}$ | 245.7 |

Most of the tyrosine residues have eigenvector coefficient values close to one and a participation ratio equal to 1.0, i.e., the transitions are localised on the individual chromophores. $^{92}$Tyr and $^{97}$Tyr showed eigenvector coefficients close to 0.85 with a participation ratio of 1.5, indicating weak delocalisation. Phenylalanine chromophores showed delocalisation, especially, $^{8}$Phe and $^{120}$Phe with a participation ratio of 2.5. However, $^{45}$Phe has a highly localised transition.

Table 3.12: The eight ($v$ =1 to 8) tyrosine $^{1}L_b$ vibrational states, their wavelengths ($\lambda$) and their calculated contributions to the rotational strength ($R_\lambda$) in Debye Bohr Magneton (DBM)at that wavelength in the near-UV CD spectrum of frame 12 of the NMR structures

| $v$ | $\lambda$/nm | $^{76}$Tyr | | $^{73}$Tyr | | $^{92}$Tyr | | $^{97}$Tyr | | $^{115}$Tyr | | $^{25}$Tyr | | $R_\lambda \times 1000$/ DBM |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $\mu_1 m_1$ | $\cos\theta_1$ | $\mu_2 m_2$ | $\cos\theta_2$ | $\mu_3 m_3$ | $\cos\theta_3$ | $\mu_4 m_4$ | $\cos\theta_4$ | $\mu_5 m_5$ | $\cos\theta_5$ | $\mu_6 m_6$ | $\cos\theta_6$ | |
| 1 | 286.1 | 0.91 | 0.01 | 1.85 | -0.45 | 0.75 | 0.08 | 0.50 | 0.001 | 0.63 | -0.14 | 0.18 | -0.42 | -32 |
| 2 | 279.7 | 1.13 | 0.02 | 2.22 | -0.04 | 0.92 | 0.08 | 0.61 | -0.002 | 0.81 | -0.12 | 0.23 | -0.40 | -44 |
| 3 | 276.1 | 0.05 | 0.12 | 0.21 | -0.25 | 0.11 | 0.06 | 0.08 | -0.002 | 0.08 | -0.19 | 0.02 | -0.44 | -7 |
| 4 | 273.5 | 0.63 | 0.02 | 1.31 | -0.04 | 0.58 | 0.08 | 0.39 | 0.004 | 0.51 | -0.12 | 0.14 | -0.39 | -33 |
| 5 | 270.2 | 0.06 | 0.07 | 0.23 | -0.02 | 0.12 | 0.06 | 0.09 | 0.005 | 0.09 | -0.17 | 0.02 | -0.42 | -8 |
| 6 | 267.6 | 0.21 | 0.004 | 0.53 | -0.03 | 0.26 | 0.07 | 0.18 | 0.007 | 0.21 | -0.13 | 0.06 | 0.04 | -16 |
| 7 | 264.5 | 0.02 | 0.13 | 0.15 | -0.02 | 0.09 | 0.06 | 0.06 | 0.008 | 0.06 | -0.17 | 0.01 | -0.39 | -6 |
| 8 | 262.0 | 0.02 | 0.11 | 0.15 | -0.02 | 0.09 | 0.05 | 0.06 | 0.01 | 0.06 | -0.16 | 0.01 | -0.38 | -6 |

A detailed examination is carried out on the rotational strength of each transition in the calculated spectrum. The most intense transitions in this region arise from the tyrosine residues, with the phenylalanine transitions making a more modest contribution; the electric and magnetic transition dipole moments of the phenylalanine residues are substantially smaller than those for the tyrosine residues.

The data in Table 3.12 are computed using frame 12 of the NMR structure. This protein has six tyrosines; thus, there are six contributions at each wavelength. The $\mu m$ values are clearly high for $^{73}$Tyr with high negative values of $\cos\theta$, so it shows noticeable contributions in the calculated spectrum. $^{25}$Tyr and $^{115}$Tyr have large negative values of $\cos\theta$ but they have low $\mu m$ so their contributions to the calculated CD spectrum are small. Both $^{76}$Tyr and $^{92}$Tyr have positive values of $\cos\theta$ indicating that the angle between the electric and magnetic transition dipoles are less than 90°. $^{97}$Tyr shows cos θ of almost zero, which indicated that this chromophore is isolated. No coupling or mixing appears in it, even it has $\mu m$ with a high value.

## 3.5. Influence of electrostatics

The aromatic chromophores in the near-UV region (240 to 300 nm) are sensitive to environmental electrostatic changes.

Investigating the surrounding environment for each aromatic residue in BPTI enabled a better understanding of the origin of the calculated CD spectrum.

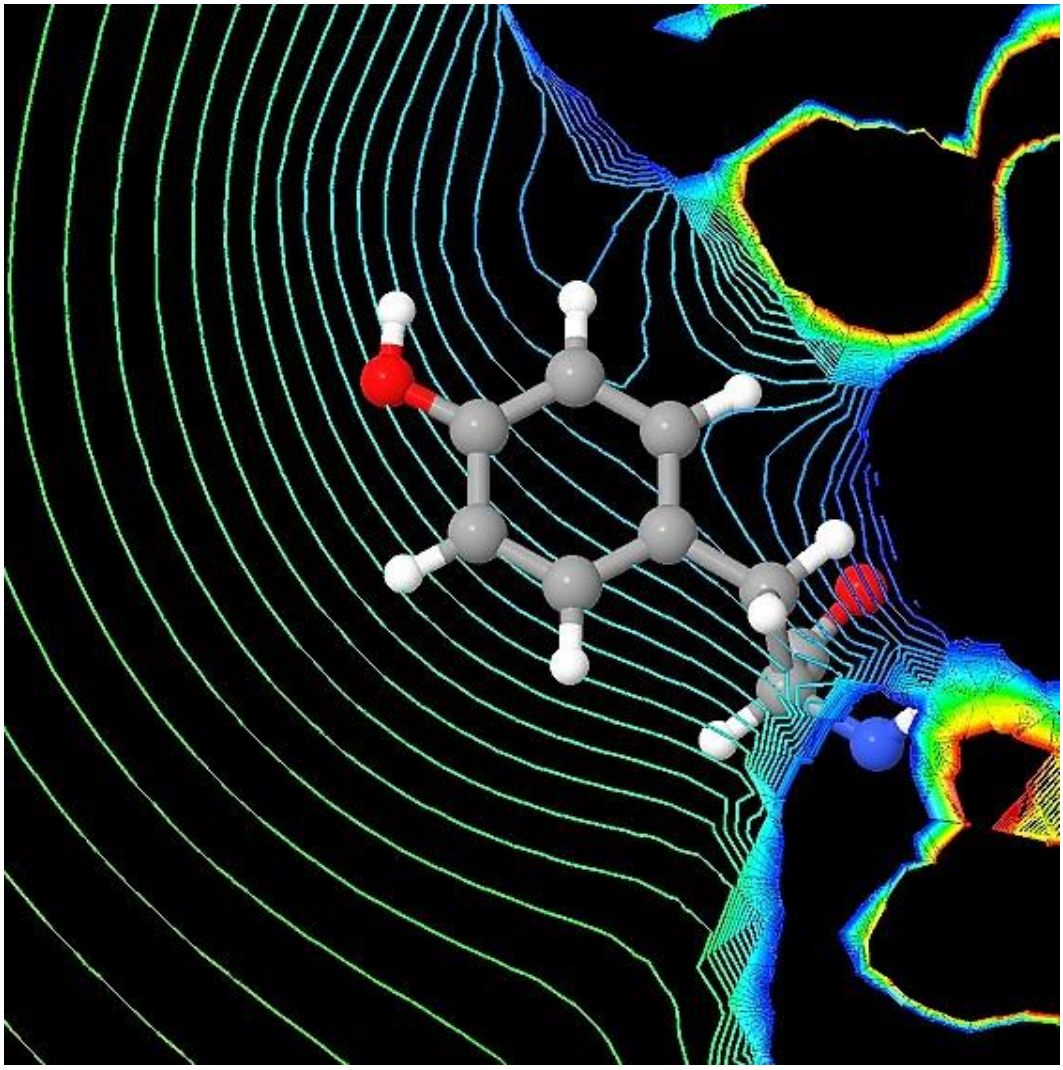


Figure 3.13. Contour plane through $^{10}$Tyr in frame 3 from the family of the NMR structure of BPTI. The plane is coloured by calculated electrostatic potential of the local environment of the tyrosine residue. It is set to cut through $C\gamma$, $C\delta1$ and $C\delta2$ of the tyrosine residue. The colours from positive to negative: blue (+20.0), green, yellow, and red (-6.0); with points outside of the contour range (20.0 to -6.0 $k_BT/e_c$) not shown, i.e., black.

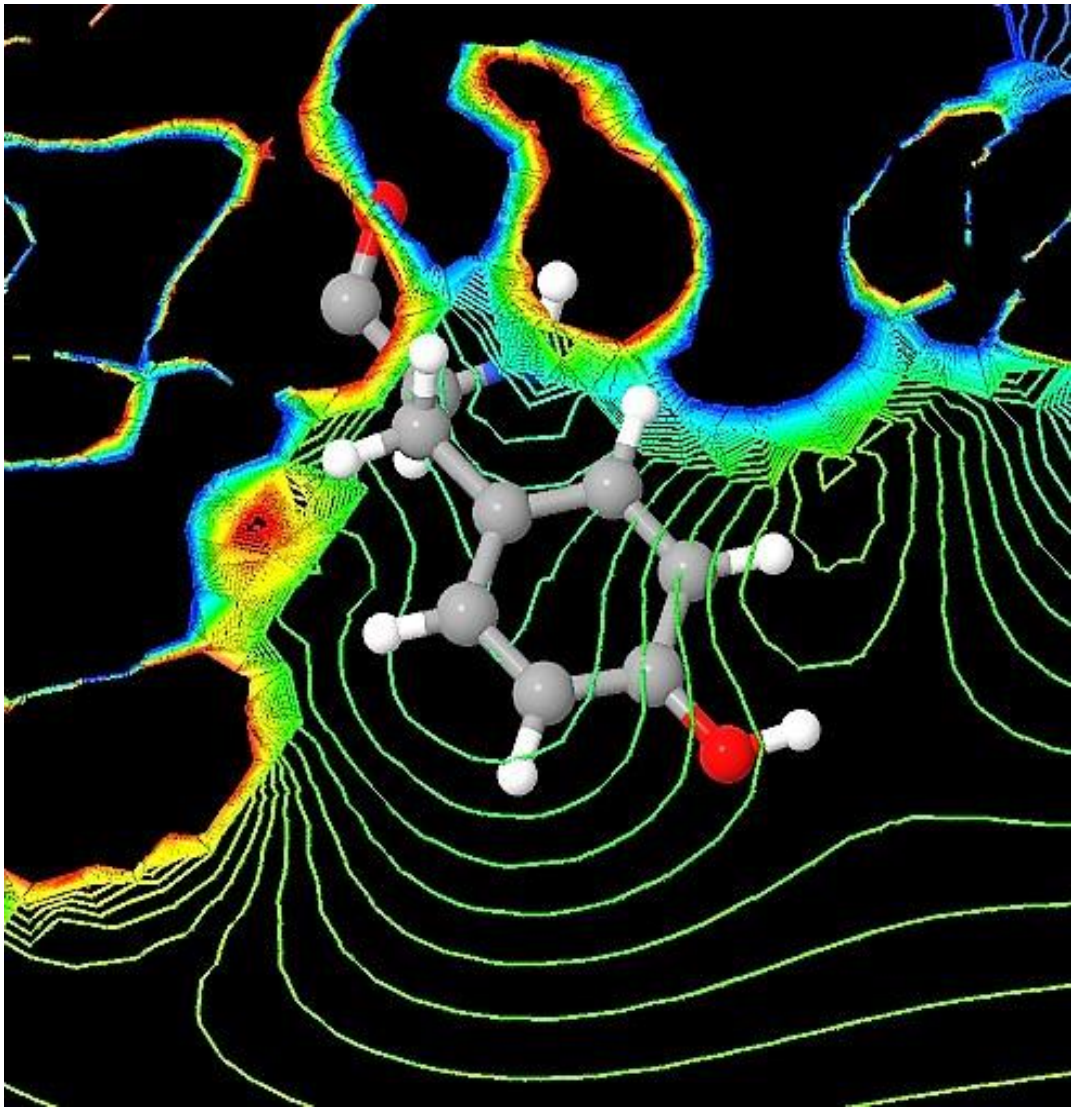


Figure 3.14. Contour plane through [21]Tyr in frame 3 from the family of the NMR structure of BPTI. The plane is coloured by calculated electrostatic potential of the local environment of the tyrosine residue. It is set to cut through $C\gamma$, $C\delta1$ and $C\delta2$ of the tyrosine residue. The colours from positive to negative: blue (+20.0), green, yellow, and red (-6.0); with points outside of the contour range (20.0 to -6.0 $k_BT/e_c$) not shown, i.e., black.

.

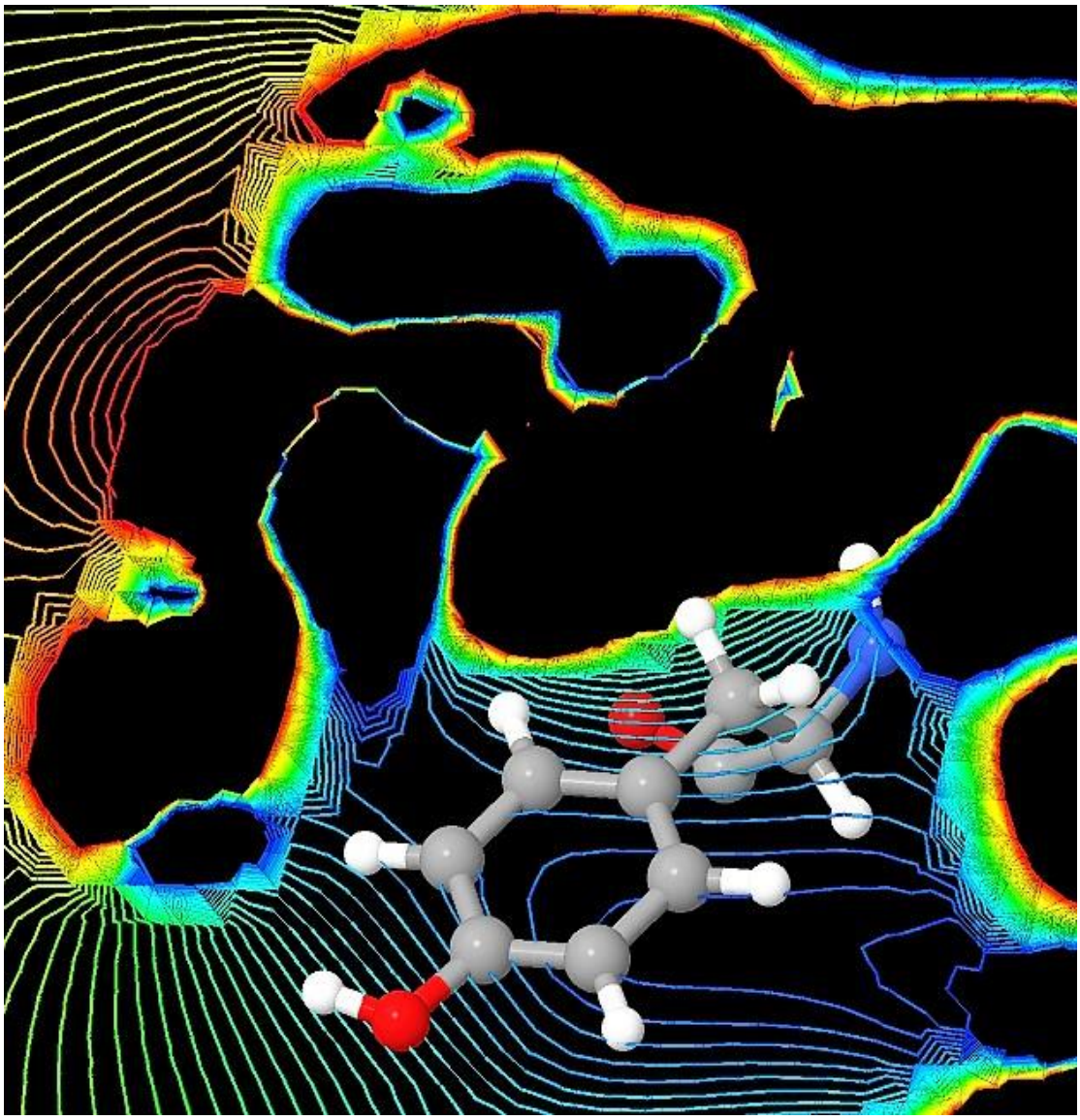


Figure 3.15. Contour plane through $^{23}$Tyr in frame 3 from the family of the NMR structures of BPTI. The plane is coloured by calculated electrostatic potential of the local environment of the tyrosine residue. It is set to cut through $C\gamma$, $C\delta1$ and $C\delta2$ of the tyrosine residue. The colours from positive to negative: blue (+20.0), green, yellow, and red (-6.0); with points outside of the contour range (20.0 to -6.0 $k_BT/e_c$) not shown, i.e., black.

.

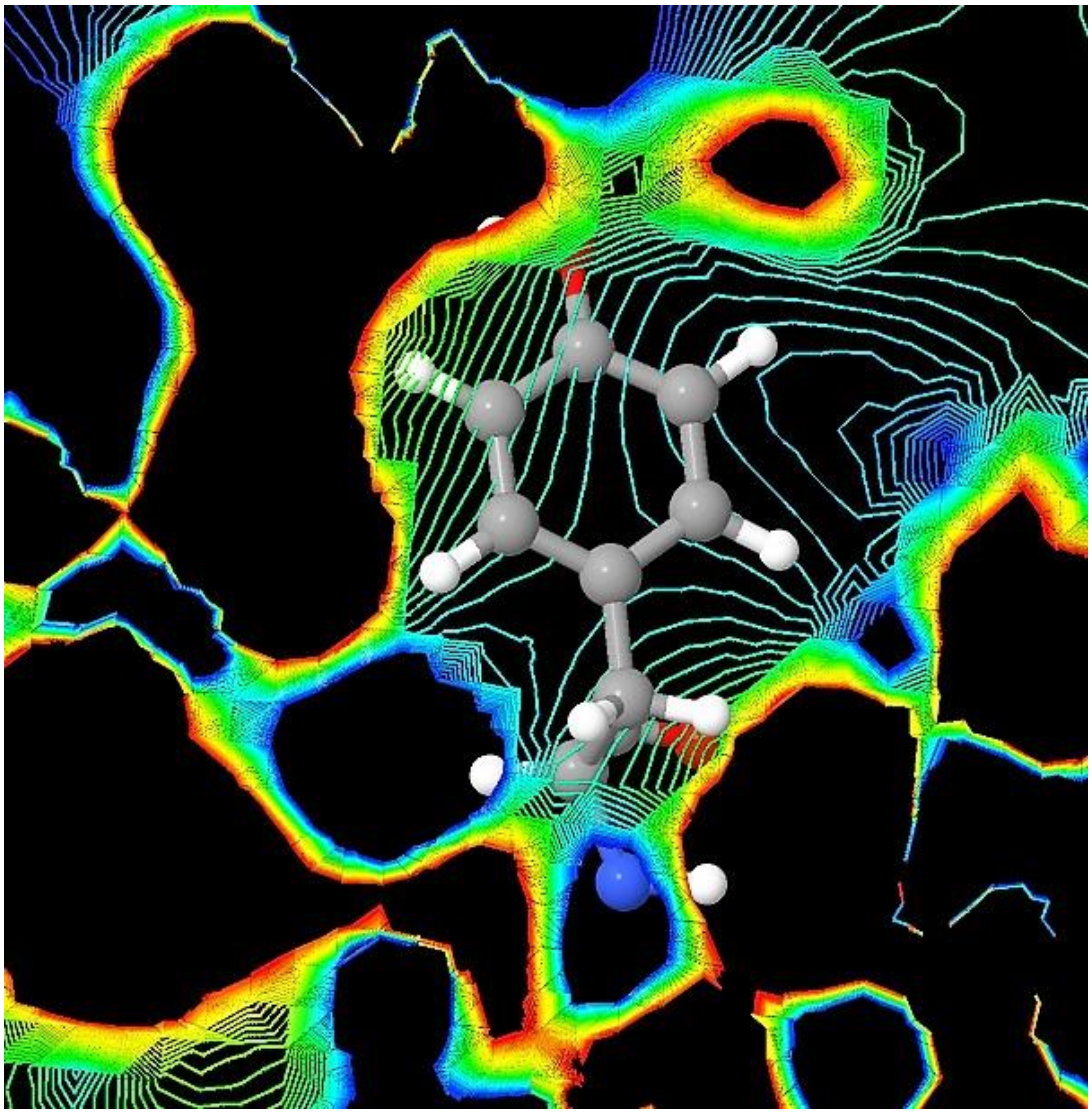


Figure 3.16. Contour plane through $^{35}$Tyr in frame 3 from the family of the NMR structure of BPTI. The plane is coloured by calculated electrostatic potential of the local environment of the tyrosine residue. It is set to cut through $C\gamma$, $C\delta1$ and $C\delta2$ of the tyrosine residue. The colours from positive to negative: blue (+20.0), green, yellow, and red (-6.0); with points outside of the contour range (20.0 to -6.0 $k_BT/e_c$) not shown, i.e., black.

A couple of studies [81, 82] have considered the three mechanisms outlined in chapter two in section 2.3, to investigate the influence of conformational changes and the effects of the electrostatic interactions on the near-UV CD spectrum.

We characterise the electrostatic potential generated by the rest of the protein and assess its impact on the mechanism generating the calculated CD spectrum. We compute electrostatic properties using the Poisson-Boltzmann equation [67, 91] as explained in chapter two.

Both $^{10}$Tyr and $^{21}$Tyr have an essential role in generating the CD spectrum. They are surrounded by a positive electrostatic potential as shown in Figure 3.13 and Figure 3.14. $^{23}$Tyr shows a noticeable contribution but is lower than those for $^{10}$Tyr and $^{21}$Tyr, it is also surrounded by a positive electrostatic potential (Figure 3.15). The case is similar for $^{35}$Tyr, which makes the smallest contribution and is surrounded by a positive electrostatic potential (Figure 3.16). The ranking of Tyr contributions to the CD spectra could help explain why the electrostatic environment and the local structure perturbs the electronic transitions on $^{10}$Tyr, $^{21}$Tyr and $^{23}$Tyr to a degree that would affect the spectra, as seen in the computed spectra for NMR frames 3.

For NMR frame 3, $^{10}$Tyr is located between $^{9}$Pro which is neutral and $^{11}$Thr which has a polar side chain and it is buried in the middle

of the protein. While $^{21}$Tyr is also buried inside the protein, it is located between $^{20}$Arg with a positive charge and $^{22}$Phe with a hydrophobic aromatic side chain. The interpolated electrostatic potentials evaluated at the eight side-chain atoms for $^{10}$Tyr and $^{21}$Tyr are all positive for NMR frame 3 of BPTI.

Further analysis has been performed on frame 3 and frame 17 of BPTI to investigate the conformation and the surrounding environment of Tyr residues. Figure 3.17 shows the superimposed structures of frame 3 and 17 of the NMR conformations. All Tyr and Phe residues are indicated in the figure.

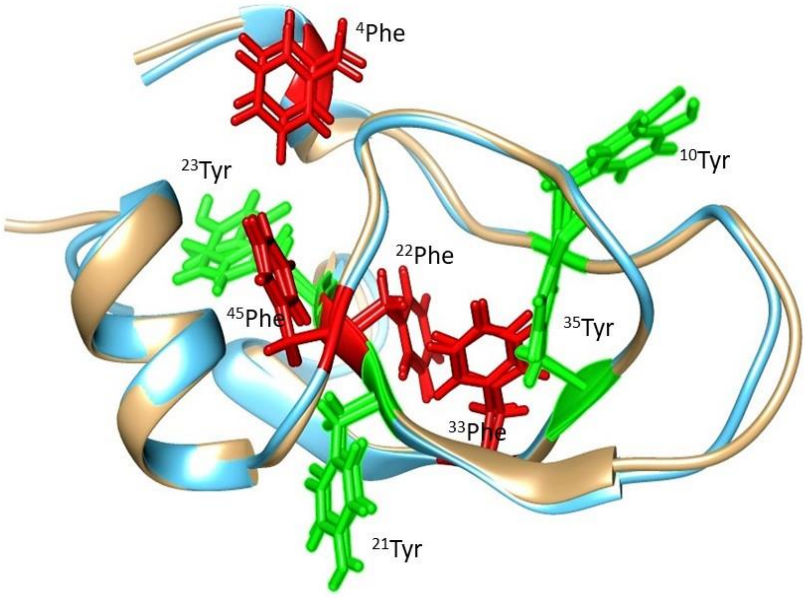



Figure 3.17. Superimposed image for frame 3 (in brown) and frame 17 (in blue) of BPTI (1PIT). Tyrosine in green and phenylalanine in red

Table 3.13: Calculated distance between Tyr side chains and other aromatic side chains in frame 3 and frame 17 of the NMR structure of BPTI.

| Chromophore1 | Chromophore2 | Distance (Å) Frame3 | Distance (Å) Frame17 |
|---|---|---|---|
| $^{10}$ Tyr | $^{21}$ Tyr | 13.0 | 13.0 |
| $^{10}$ Tyr | $^{23}$ Tyr | 12.9 | 13.5 |
| $^{10}$ Tyr | $^{35}$ Tyr | 7.8 | 7.8 |
| $^{10}$ Tyr | $^{4}$ Phe | 10.5 | 11.0 |
| $^{10}$ Tyr | $^{22}$ Phe | 8.8 | 9.2 |
| $^{10}$ Tyr | $^{33}$ Phe | 9.2 | 9.3 |
| $^{10}$ Tyr | $^{45}$ Phe | 13.2 | 13.7 |
| $^{21}$ Tyr | $^{23}$ Tyr | 8.1 | 8.3 |
| $^{21}$ Tyr | $^{35}$ Tyr | 9.0 | 9.0 |
| $^{21}$ Tyr | $^{4}$ Phe | 13.7 | 13.8 |
| $^{21}$ Tyr | $^{22}$ Phe | 5.5 | 5.3 |
| $^{21}$ Tyr | $^{33}$ Phe | 7.1 | 7.3 |
| $^{21}$ Tyr | $^{45}$ Phe | 4.4 | 4.4 |
| $^{23}$ Tyr | $^{35}$ Tyr | 13.8 | 14.2 |
| $^{23}$ Tyr | $^{4}$ Phe | 8.7 | 8.8 |
| $^{23}$ Tyr | $^{22}$ Phe | 5.4 | 5.6 |
| $^{23}$ Tyr | $^{33}$ Phe | 11.0 | 11.8 |
| $^{23}$ Tyr | $^{45}$ Phe | 7.9 | 8.2 |
| $^{35}$ Tyr | $^{4}$ Phe | 14.5 | 15.0 |
| $^{35}$ Tyr | $^{22}$ Phe | 8.5 | 8.8 |
| $^{35}$ Tyr | $^{33}$ Phe | 6.3 | 6.3 |
| $^{35}$ Tyr | $^{45}$ Phe | 10.2 | 10.3 |

This table shows the distances between each Tyr and the other aromatic side chains. They have been calculated for frame 3 and frame 17 of the NMR structures by using Chimera [112]. From Table

3.13 and Figure 3.17 we can say that there are no essential differences between frame 3 and frame 17. The similarity in the distances and the visualization structures both confirm that the Tyr residues have similar conformations and surrounding environments. This supports our focus on studying only one frame of the NMR structures for each protein.

## 3.6. Conclusion

In this chapter, the vibrational structure of electronic transitions of aromatic amino acids is used in our calculations of the near-UV CD spectra focusing on Tyr residues. We study the interactions and the mechanisms that generate the rotational strength by using new parameters incorporating the vibrational structure of electronic transitions. Detailed analysis has been presented for the tyrosine chromophores of the three proteins. Generating a spectrum by including the vibrational transitions shows an improvement compared to calculations neglecting the vibrational structure. Overall, the three experimental CD spectra are well reproduced in terms of the shape and the peaks, especially, in the case of BPTI. However, the calculated CD spectrum of cardiotoxin shows a less close similarity to the experimental spectrum.

Using different frames of the NMR structures in calculating a CD spectrum shows a noticeable improvement from the X-ray crystal

structure. In addition, using different frames of the NMR structures enables us to investigate the protein-structure and protein-dynamics relationships and their impact on the generated spectrum. The investigation of the interactions between residues helps to clarify the role of each chromophore in generating the calculated CD spectra. Also calculating contributions to the CD spectrum enables us to conclude that the *one-electron* mechanism has the main role in generating the spectrum of all the proteins examined in this chapter.

Further analysis of BPTI has characterised the electrostatic potentials generated by the rest of the protein. The investigation of its impact on the mechanisms generating the calculated CD spectrum in near-UV region emphasises the role and the influence of the electrostatic environment surrounding the most interesting aromatic residues of the BPTI protein. We will be developing the work further, for example, by generating parameters for disulfide bonds and by further consideration of the influence of solvent and conformational dynamics.

# Chapter 4

# Tryptophan chromophores in proteins

## 4.1 Introduction

The contribution of the aromatic side chains to the CD spectrum has an important role. The sign, the magnitude, and the wavelength are aspects of the aromatic CD bands that are necessary to be understood. It is therefore essential to investigate how aromatic side chains influence the appearance of a CD spectrum. Tryptophan residues have been studied to investigate the effects of these residues on the near-UV CD spectrum [99], both experimentally and theoretically [125, 126]. Tryptophan in proteins has a special importance because of the information that can be obtained from the CD spectrum. Tryptophan transitions can couple and develop an exciton chirality, which could be used to determine the relative stereochemistry of the chromophores. Among aromatic residues, the near-UV spectrum of Trp is most sensitive to its environment.

Tryptophan is sensitive to the polarity of its local environment. It has been used to probe the internal electrostatics of chromophores, which are vital to the structure and function of proteins [127]. The near-UV CD of Trp results from overlapping $^1L_a$ and $^1L_b$ transitions. It has advantages because both the $^1L_a$ and the $^1L_b$ bands are revealed simultaneously. However, the sign and intensity of those bands in the CD spectra vary depending on their surrounding environment [128].

Five proteins will be discussed in this chapter with a special focus on the tryptophan chromophores. For each protein the X-ray crystallographic structure and the NMR models from the PDB are used to generate the calculated CD spectrum. Further analysis focuses on only two of the NMR structures. The selection of these two NMR models is based on some statistical investigations, as mentioned previously in chapter three. The interactions and contributions will be investigated because of their important role in generating the CD spectrum in the near-UV. The method of this chapter is exactly as explained in chapter three section 3.3, so to avoid repetition we will not re-explain it in this chapter. Table 4.1 shows the proteins studied in this chapter.

Table 4.1: The proteins studied in this chapter: name, PDB codes, length (i.e., number of residues in the X-ray structure) and the number of aromatic residues and disulfide bonds.

| Protein | PDB entry | | Length | No. aromatic residue | | | |
|---|---|---|---|---|---|---|---|
| | X-ray | NMR | | Phe | Tyr | Trp | S-S |
| apolipophorin III | 1AEP | 1LS4 | 161 | 2 | 0 | 2 | 0 |
| Barnase | 1A2P | 1FW7 | 110 | 4 | 7 | 3 | 0 |
| $\beta$-2 microglobulin (human) | 1LDS | 1JNJ | 100 | 5 | 6 | 2 | 1 |
| Thioredoxin | 2TRX | 1XOA | 108 | 4 | 2 | 2 | 1 |
| hen egg white lysozyme | 1HF4 | 1E8L | 258 | 3 | 3 | 6 | 4 |

## 4.2 Background on the proteins studied in this chapter

### 4.2.1 Apolipophorin III

The apolipophorin III family of proteins are a family of exchangeable apolipoproteins, which play a critical and important roles in lipid transport and lipoprotein metabolism. The structure is shown in Figure 4.1. Apolipophorin III (apoLp-III) is a prototypical exchangeable apolipoprotein found in many insect species that functions in the transport of diacylglycerol. It is derived from locust migratoria. This protein does not have any disulfide bonds. It has been crystallographically characterized and it is available in the Protein Data Bank (PDB code 1AEP [129]). Additionally, this protein has a structure determined by NMR (PDB

code 1LS4 [130]), with 21 models. It has a single chain with length of 161 amino acid residues.

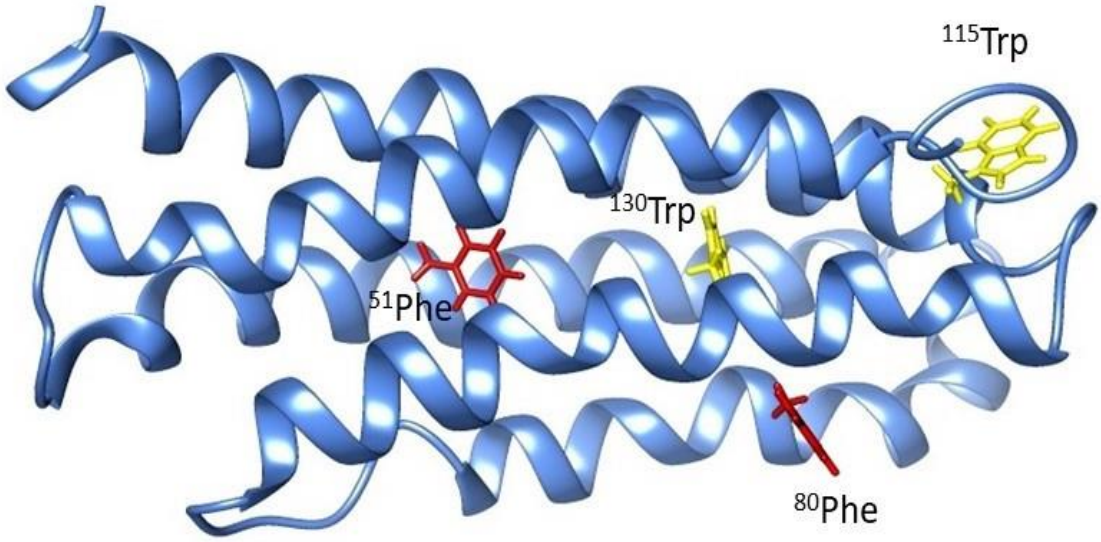



Figure 4.1. Apolipophorin III (1LS4), a model from the family of the NMR structure. Tryptophan in yellow and phenylalanine in red

### 4.2.2 Barnase

Barnase is a bacterial protein that consists of 110 amino acids and has ribonuclease activity. It is synthesized and secreted by the bacterium *Bacillus amyloliquefaciens*, but is lethal to the cell when expressed without its inhibitor barstar. The inhibitor binds to and occludes the ribonuclease active site, preventing barnase from damaging the RNA of the cell after it has been synthesized but before it has been secreted. It does not require divalent cations or

non-peptide components to fold. This simplicity, in combination with its reversible folding transition, means that barnase has been extensively studied in order to understand how proteins fold [131]. The X-ray crystallography code in the Protein Data Bank is (PDB:1A2P [132]) and there is an NMR family comprising 20 models (PDB:1FW7 [133]). Figure 4.2 shows the structure of this protein.

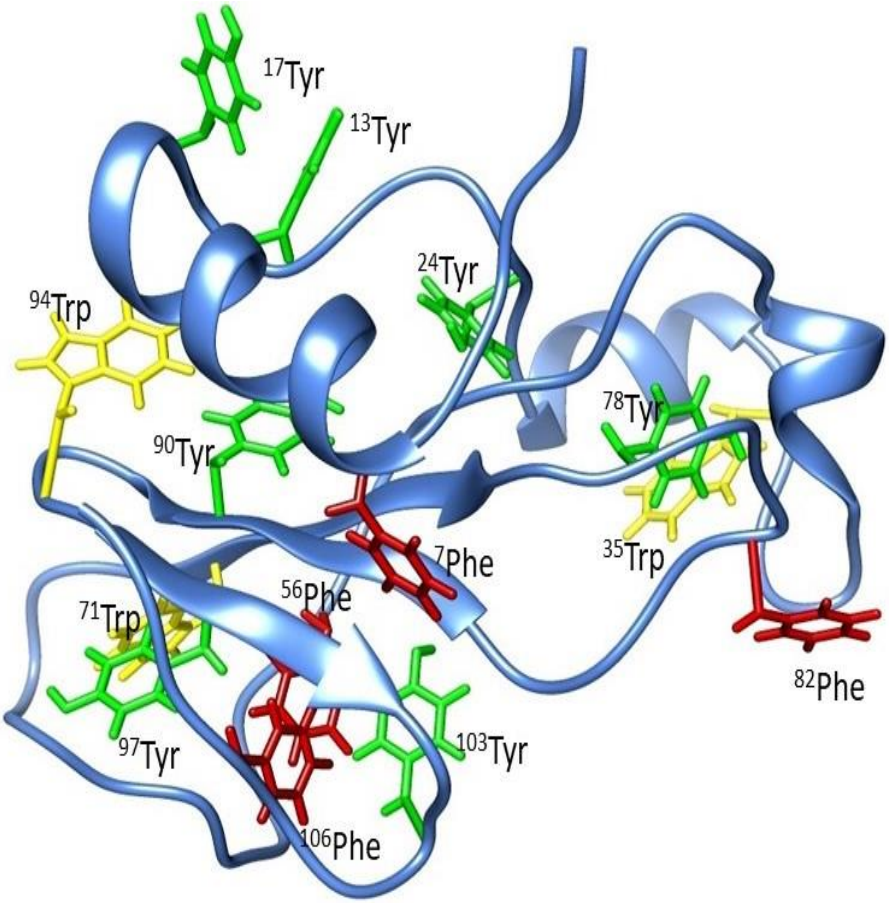



Figure 4.2. Barnase (1FW7), a model from the family of the NMR structure. Tryptophan in yellow, tyrosine in green and phenylalanine in red

### 4.2.3 Beta 2 microglobulin human

Beta 2 microglobulin ($\beta$-2 micoglobulin) is a protein associated with the outer membrane of many cells including lymphocytes. It is the small subunit of the MHC class I molecule. Association with beta 2-microglobulin is generally required for the transport of class I heavy chains from the endoplasmic reticulum to the cell surface. In the past years it has been recognized that B2M has a prominent beta structure which has been suggested to be important in the stabilization of the tertiary conformation of class I HLA heavy chains and other differentiation antigens with which it is associated on cell surfaces [134]. Figure 4.3 shows the structure of this protein. The X-ray crystallography code is (PDB:1LDs [135]) and the NMR code is (PDB:1JNJ [136]), the NMR family has 20 models.

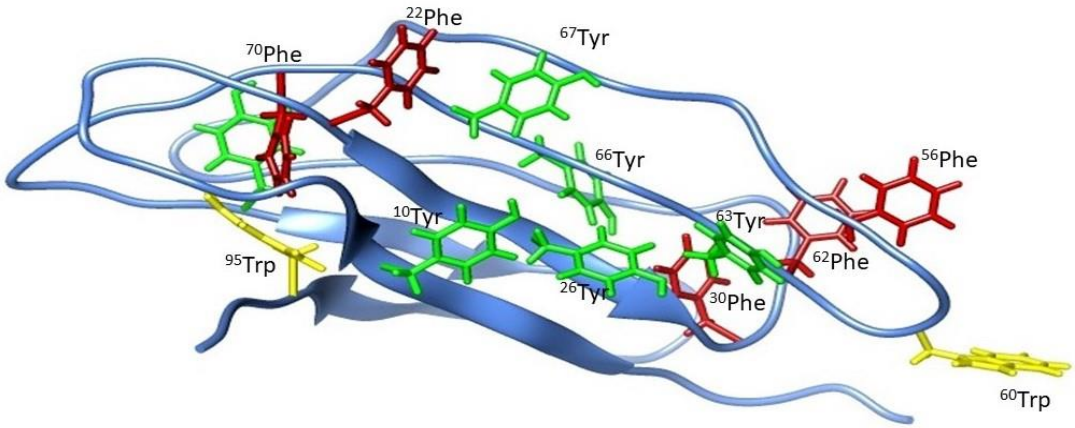



Figure 4.3. $\beta$-2 micoglobulin (1JNJ), a model from the family of the NMR structure. Tryptophan in yellow, Tyrosine in green and phenylalanine in red

### 4.2.4 Thioredoxin

Thioredoxin belongs to a class of small redox proteins present in all organisms. It plays a role in many important biological processes, including redox reactions. Figure 4.4 shows the structure of this protein.

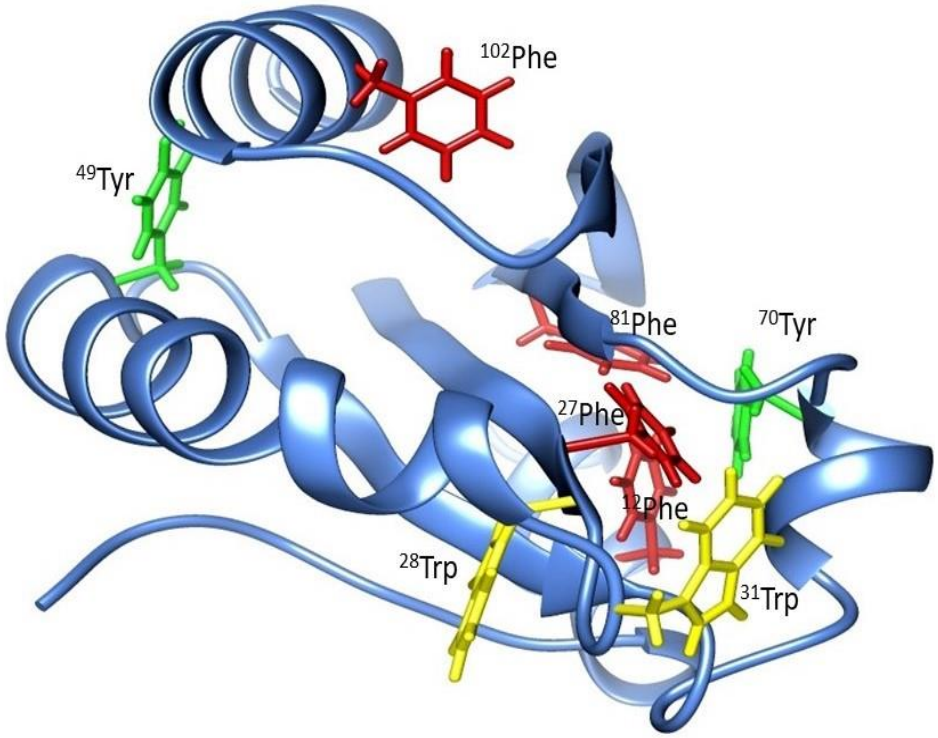



Figure 4.4. Thioredoxin (1XOA), a model from the family of the NMR structure. Tryptophan in yellow, tyrosine in green and phenylalanine in red

Thioredoxin plays a central role in humans and is increasingly linked to medicine through their response to reactive oxygen species. In plants, thioredoxins regulate a spectrum of critical functions, ranging from photosynthesis to growth, flowering and the development and germination of seeds. They have also recently been found to play a role in cell-to-cell communication.

The X-ray crystallography code in the Protein Data Bank is (PDB: 2TRX [137]) and the NMR family comprising 20 models is (PDB: 1XOA [138]). Thioredoxins act as antioxidants by facilitating the reduction of other proteins by cysteine disulfide exchange [139].

### 4.2.5 Hen egg white lysozyme

Lysozyme is a powerful enzyme of biological significance found in abundance in tears, saliva, and human milk. The nature of the amino acid sidechains in the lysozyme polypeptide sequence leads to regions of varying hydrophobicities and polarities of the enzyme structure. The code of X-ray crystallography is (PDB:1HF4 [140]) and the NMR code is (1E8L [141]) comprising 50 models.

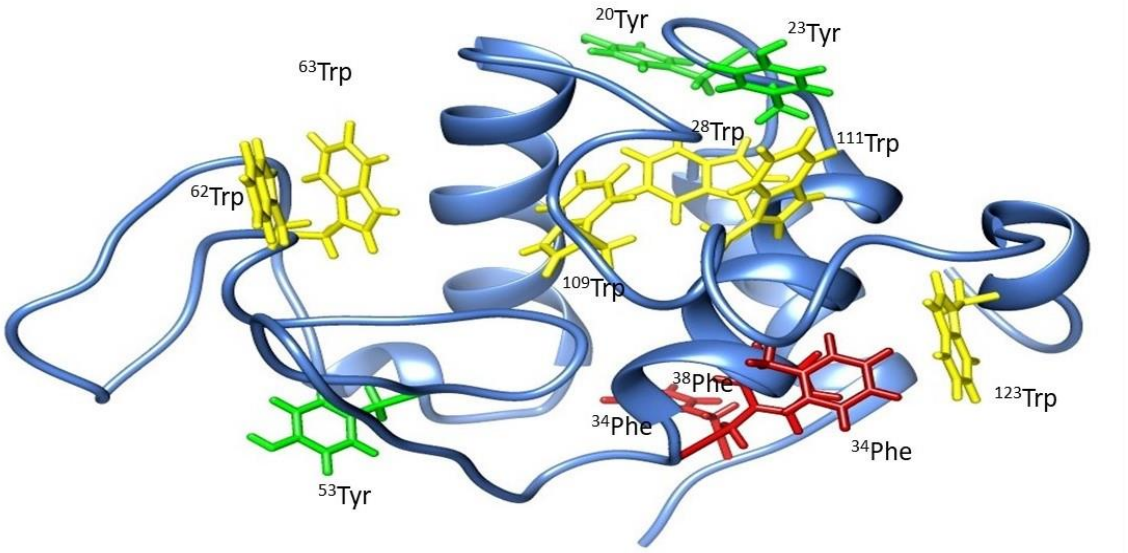



Figure 4.5. Hen egg white lysozyme (1E8L), a model from the family of the NMR structure. Tryptophan in yellow, tyrosine in green and phenylalanine in red

## 4.3 Results and discussion

### 4.3.1 The role of Trp residues in generating CD spectrum

#### 4.3.1.1 Apolipophorin III

The calculated near-UV CD spectrum of this protein was computed using the X-ray crystallography structure and from the average of the NMR structures (PDB ID: 1AEP and 1LS4). They are compared with the experimental spectrum [142] as shown in Figure 4.6.

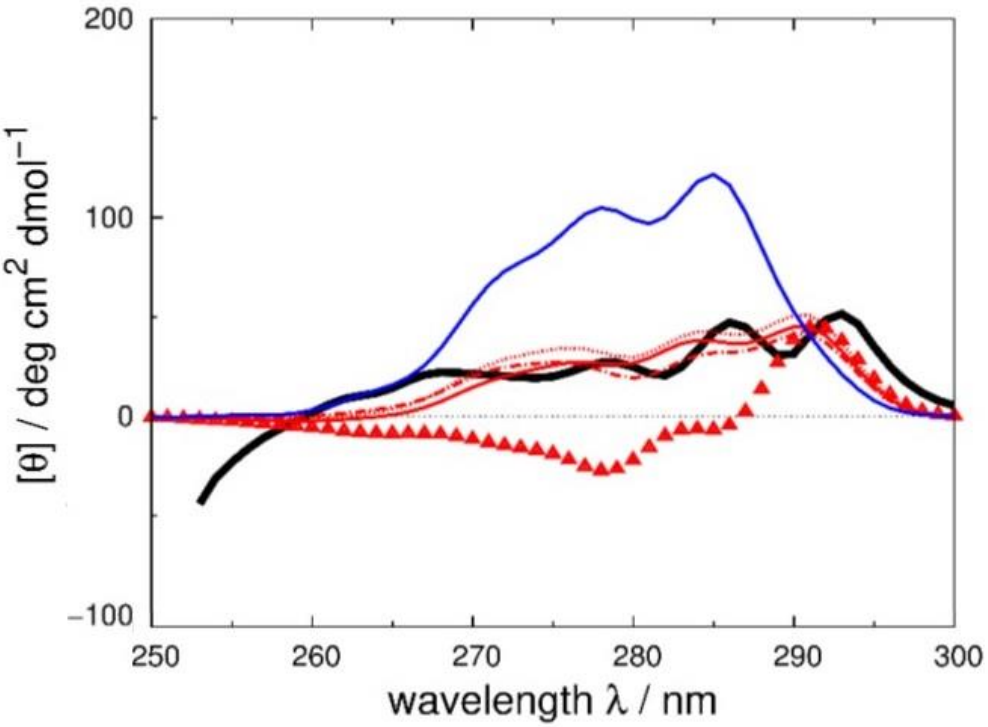



Figure 4.6. Experimental [142] near-UV CD spectrum (black solid line), and spectra calculated using the X-ray structure (blue solid line), calculated using an average over 21 frames of the NMR structure (red solid line) and calculated from individual frames of the NMR structure: frame 5 (red triangles) and frame 15 (red dotted line).

The X-ray crystal structure leads to a computed spectrum that reproduces the overall shape and location of the peaks reasonably well, but the intensity across the wavelength range is noticeably higher than the experimental spectrum. The conformational diversity of this protein in solution is reflected, at least in part, by the family of NMR structures.

Figure 4.6 shows the CD spectrum calculated using specific NMR structures: frame 15 which gives the more intense spectrum (in closest agreement with experiment), and by way of contrast frame 5, which gives the calculated spectrum of t intensity far from the experimental. Despite the good intensity and general shape of the spectrum by using frame 15, the peak positions are not well reproduced, especially in the region 280- 290 nm. The structural difference across these two frames is small, 0.9 Å all-atom RMSD and the RMSD of Trp residues is 0.3 to 0.4 Å.

To investigate the influence of conformational diversity on the computed spectra, we calculated the standard deviation of the computed intensity at each wavelength over all the NMR structures as shown in Figure 4.7. The spectrum computed from averaging over all the NMR frames is very similar to the experimental spectrum. Frame 15 leads to computed spectra in acceptable agreement (with an RMSD of 90 deg $cm^2$ $dmol^{-1}$) with experiment.

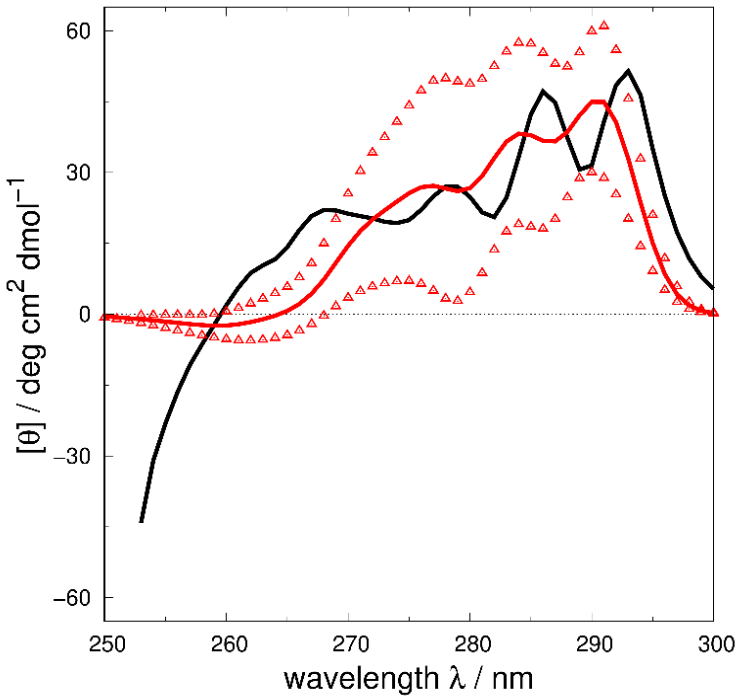



Figure 4.7. Conformational diversity investigation. The black solid line is the experimental spectrum [142]. The red solid line is the mean average computed from 21 NMR models and the red lines with triangles (upper and lower) represent the mean average ± standard deviation, respectively.

The contributions of the phenylalanine and tryptophan chromophores are generally well reproduced. The analysis focused on frame 15. Table 4.2 shows the significant interactions energies among chromophores, with interaction energies presented if greater than ±10 cm$^{-1}$. The energies below this threshold are presented in Table A.3 in the Appendix. Only tryptophan shows interactions above the 10 cm$^{-1}$ threshold, and $v1$ electronic vibrational levels of the $^1L_b$ states have a predominant role.

Different electronic vibrational levels of the $^1L_a$ states interact with $^1L_b$ states to give the relatively high interaction energies.

Table 4.2: Calculated interaction energies (shown only if greater than 10 $cm^{-1}$) of electronic transitions of aromatic side chains in frame 15 of the NMR structure.

| Chromophore/Transitions | Chromophore/Transitions | Distance (Å) | Interaction Energy ($cm^{-1}$) |
|---|---|---|---|
| $^{115}$Trp-$^1L_b$ | $^{115}$Trp-$^1L_{a7}$ | 0.0 | 17.7 |
| $^{115}$Trp-$^1L_b$ | $^{115}$Trp-$^1L_{a6}$ | 0.0 | 16.7 |
| $^{115}$Trp-$^1L_b$ | $^{115}$Trp-$^1L_{a5}$ | 0.0 | 27.9 |
| $^{115}$Trp-$^1L_b$ | $^{115}$Trp-$^1L_{a4}$ | 0.0 | 31.6 |
| $^{115}$Trp-$^1L_b$ | $^{115}$Trp-$^1L_{a3}$ | 0.0 | 29.8 |
| $^{115}$Trp-$^1L_b$ | $^{115}$Trp-$^1L_{a2}$ | 0.0 | 49.3 |
| $^{115}$Trp-$^1L_b$ | $^{115}$Trp-$^1L_a$ | 0.0 | 53.0 |
| $^{130}$Trp-$^1L_b$ | $^{130}$Trp-$^1L_{a9}$ | 0.0 | 17.3 |
| $^{130}$Trp-$^1L_b$ | $^{130}$Trp-$^1L_{a8}$ | 0.0 | 13.9 |
| $^{130}$Trp-$^1L_b$ | $^{130}$Trp-$^1L_{a7}$ | 0.0 | 33.0 |
| $^{130}$Trp-$^1L_b$ | $^{130}$Trp-$^1L_{a6}$ | 0.0 | 31.2 |
| $^{130}$Trp-$^1L_b$ | $^{130}$Trp-$^1L_{a5}$ | 0.0 | 52.0 |
| $^{130}$Trp-$^1L_b$ | $^{130}$Trp-$^1L_{a4}$ | 0.0 | 59.0 |
| $^{130}$Trp-$^1L_b$ | $^{130}$Trp-$^1L_{a3}$ | 0.0 | 55.5 |
| $^{130}$Trp-$^1L_b$ | $^{130}$Trp-$^1L_{a2}$ | 0.0 | 91.9 |
| $^{130}$Trp-$^1L_b$ | $^{130}$Trp-$^1L_a$ | 0.0 | 98.8 |

The zero values (0.0 Å) in the separation column correspond to the one-electron mixing of two transitions on the same chromophore. Coupling between L states and B states (the latter occurring in the far-UV) is weak. The largest interaction energy is 98.8 $cm^{-1}$ and corresponds to the one-electron mixing of the $^1L_b$,

$v1$ and the $^1L_a$ $v1$ transitions of $^{130}$Trp. There are very few interaction energies for phenylalanine, and these are not shown as they are smaller than $\pm$10 cm$^{-1}$. The interactions between different chromophores were very weak compared to the one-electron mixing.

The analogous transitions on $^{115}$Trp interact via the same mechanism and with a similar magnitude. Inter-chromophore interactions are much weaker. From the interaction energies in Table 4.2, it is clear that the one-electron mixing of the tryptophan $^1L_b$ states is an important factor in the appearance of the near-UV spectrum. The relatively strong intensity of the calculated CD spectrum of frame 15 is highlighted as a point of investigation.

The wavelengths and the associated eigenvector coefficients for tryptophan residue, are investigated to examine the nature of the transitions and the data for tryptophan are presented in Table 4.3. No table is presented for phenylalanine because all eigenvector coefficients are equal to one.

Further investigations have been carried to examine the role of each aromatic residue. Most tryptophan residues have an eigenvector coefficient close to one, and therefore a participation ratio equal to one. These values confirm the localisation of the transitions of tryptophan aromatic chromophores.

Table 4.3: Wavelength and associated eigenvector coefficients of tryptophan side chain transitions in frame 15 of the NMR structure.

| Eigenvector index | Values of Eigenvector coefficients | Participation ratio (W) | Residues/ Transitions | Wavelength (nm) |
|---|---|---|---|---|
| 1 | 0.98 | 1.0 | $^{130}Trp_{Lb7}$ | 291.4 |
| 2 | 0.99 | 1.0 | $^{115}Trp_{Lb7}$ | 290.6 |
| 3 | 0.86 | 2.0 | $^{130}Trp_{La9}$ | 285.7 |
| 4 | -0.96 | 1.0 | $^{115}Trp_{La9}$ | 285.2 |
| 5 | 0.90 | 1.0 | $^{130}Trp_{Lb6}$ | 283.4 |
| 6 | 0.97 | 1.0 | $^{115}Trp_{Lb6}$ | 283.3 |
| 7 | -0.95 | 1.0 | $^{130}Trp_{Lb5}$ | 282.7 |
| 8 | -0.98 | 1.0 | $^{115}Trp_{Lb5}$ | 282.6 |
| 9 | -0.87 | 2.0 | $^{130}Trp_{La8}$ | 278.7 |
| 10 | 0.94 | 1.0 | $^{115}Trp_{La8}$ | 278.7 |
| 11 | -0.98 | 1.0 | $^{130}Trp_{Lb4}$ | 277.9 |
| 12 | 0.93 | 1.0 | $^{115}Trp_{Lb4}$ | 277.8 |
| 13 | 0.77 | 2.5 | $^{130}Trp_{La7}$ | 276.9 |
| 14 | 0.89 | 2.0 | $^{115}Trp_{La7}$ | 276.8 |

A few cases showed coefficient factors close to 0.87 or 0.77 with a participation ratio equal to 2.0 and 2.5, respectively. This shows weak delocalisation of these transitions. Phenylalanine chromophores showed a high level of localisation of the transitions with a coefficient factor equal to one and the participation ratio equal to one.

Table 4.4: The seven ($v$ =1 to 7) $^1L_b$ and the nine ($v$ =1 to 9) tryptophan $^1L_a$ vibrational states, their wavelengths ($\lambda$) and their calculated contributions to the rotational strength ($R_\lambda$) in Debye Bohr Magneton (DBM) - shown only if greater than 1 DBM- at that wavelength in the near-UV CD spectrum of frame 15 of the NMR.

| $v$ | Residue transitions | $\lambda$/nm | $^{130}$Trp | | $^{115}$Trp | | $R_\lambda \times 1000$ / DBM |
|---|---|---|---|---|---|---|---|
| | | | $\mu_1 m_1$ | $\cos\theta_1$ | $\mu_2 m_2$ | $\cos\theta_2$ | |
| 1 | $^1L_b v1$ | 290.5 | 0.44 | 0.01 | 0.39 | 0.01 | 12 |
| 3 | $^1L_b v3$ | 282.6 | 0.25 | -0.009 | 0.20 | 0.01 | 2 |
| 5 | $^1L_b v5$ | 276.5 | 2.15 | -0.09 | 1.81 | 0.10 | 3 |
| 7 | $^1L_b v7$ | 270.7 | 0.42 | 0.01 | 0.38 | 0.01 | 3 |
| 8 | $^1L_a v1$ | 285.0 | 0.23 | -0.005 | 0.80 | 0.005 | 6 |
| 11 | $^1L_a v4$ | 272.8 | 0.55 | -0.01 | 0.39 | 0.01 | 2 |
| 13 | $^1L_a v6$ | 267.0 | 0.13 | -0.30 | 0.40 | 0.10 | 1 |
| 14 | $^1L_a v7$ | 263.4 | 0.20 | -0.30 | 0.40 | 0.16 | 1 |

A detailed examination of the rotational strength at each transition in the calculated spectrum is presented in Table 4.4. This protein has two tryptophans, the values of cos θ of $^{115}$Trp is smaller than the values of $^{130}$Trp, so there is less mixing and coupling of $^{115}$Trp chromophore. In addition, cos θ of $^{115}$Trp all are positive, which indicated that the angles between the electric and magnetic transition dipoles are less than 90°. Other residue transitions with small contributions, (i.e. less than $1 \times 10^{-3}$ DBM) are not shown in the table. The strongest rotational strength ($12 \times 10^{-3}$ and $6 \times 10^{-3}$ DBM) resulted from the first transition of the $^1L_b$ states and the first transition of the $^1L_a$ states, respectively. The surrounding

environment leads to a significant effect on the tryptophan residues.

#### 4.3.1.2 Barnase

Figure 4.8 shows the calculated near-UV CD spectrum of barnase computed by using the X-ray crystallography structure and the average spectrum from the NMR structures (PDB code: 1A2P and 1FW7), in comparison to the experimental spectrum [143]. The overall shape generally is well reproduced. The NMR structures predicted spectra of a higher intensity than the X-ray crystal structure in the wavelength region 280 and 290 nm. The X-ray crystal structure had some peaks shifted towards shorter wavelengths. The intensity from both the NMR and the X-ray structures across the wavelength range is lower than the experimental spectrum. The CD spectra in Figure 4.9 show two models of the family of the NMR structures. Frame 20 gives the most intense spectrum close to the experimental (with an RMSD of 70 deg $cm^2$ $dmol^{-1}$). Frame 8 shows weakest intensity and a negative peak in region 280 nm. The structural differences across these two frames is 0.6 Å all-atom RMSD and the RMSD of Trp residues is 0.2 to 0.3 Å. All the calculated spectra show a negative peak in region 290 nm, which is not found in the experimental spectrum.

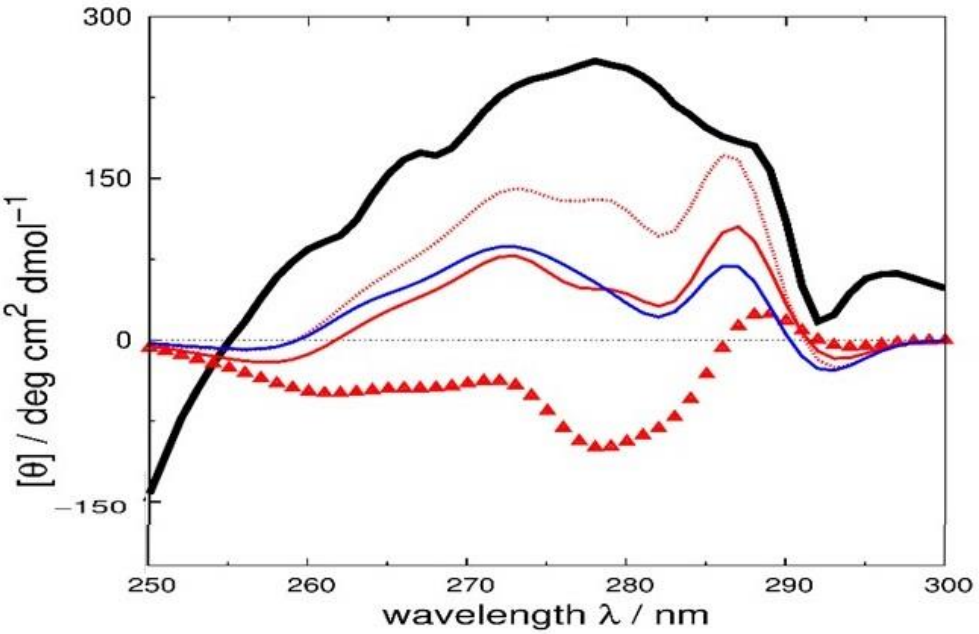



Figure 4.8. Experimental [143] near-UV CD spectrum (black solid line), and spectra calculated using the X-ray structure (blue solid line), calculated using an average over 20 frames of the NMR structure (red solid line) and calculated from individual frames of the NMR structure: frame 8 (dashed red line with triangles) and frame 20 (red dot-dashed line).

We calculated the standard deviation of the computed intensity at each wavelength over all the NMR frames, to investigate the influence of the conformational diversity on the computed spectra. Figure 4.9 shows that the computed spectrum from averaging over all the NMR frames is less intense than the experimental spectrum.

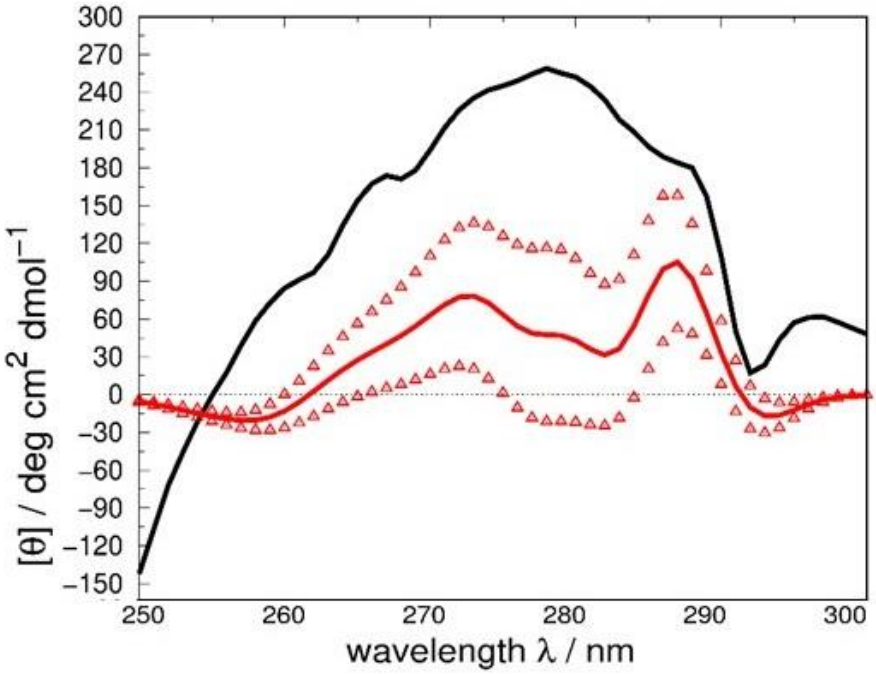



Figure 4.9. Conformational diversity investigation. The black solid line is the experimental spectrum [143]. The red solid line is the mean average computed from 20 NMR models and the red lines with triangles (upper and lower) represent the mean average ± standard deviation, respectively.

The general contributions of the tryptophan residues are well reproduced. The focus in the analysis of this protein will be on frame 20.

Table 4.5 shows the significant interaction energies, i.e. those greater than 50 $cm^{-1}$. Table A.4 in the Appendix shows the energies below this threshold. Most of the tryptophans and a few tyrosines show interactions above the 50 $cm^{-1}$ threshold. Different electronic vibrational levels of the $^1L_a$ states interact with $^1L_b$ states to give the relatively high interaction energies.

Table 4.5: Calculated interaction energies (shown only if greater than 50 $cm^{-1}$) of electronic transitions of aromatic side chains in frame 20 of the NMR structure.

| Chromophore/Transitions | Chromophore/Transitions | Distance (Å) | Interaction Energy ($cm^{-1}$) |
|---|---|---|---|
| $^{90}$Tyr-$^{1}L_{b4}$ | $^{90}$Tyr-$^{1}L_{a}$ | 0.0 | 61.2 |
| $^{90}$Tyr-$^{1}L_{b2}$ | $^{90}$Tyr-$^{1}L_{a}$ | 0.0 | 79.0 |
| $^{90}$Tyr-$^{1}L_{b}$ | $^{90}$Tyr-$^{1}L_{a}$ | 0.0 | 71.0 |
| $^{71}$Trp-$^{1}L_{b}$ | $^{71}$Trp-$^{1}L_{a}$ | 0.0 | 121.0 |
| $^{71}$Trp-$^{1}L_{b}$ | $^{71}$Trp-$^{1}L_{a5}$ | 0.0 | 63.7 |
| $^{71}$Trp-$^{1}L_{b}$ | $^{71}$Trp-$^{1}L_{a4}$ | 0.0 | 72.1 |
| $^{71}$Trp-$^{1}L_{b}$ | $^{71}$Trp-$^{1}L_{a3}$ | 0.0 | 67.9 |
| $^{71}$Trp-$^{1}L_{b}$ | $^{71}$Trp-$^{1}L_{a2}$ | 0.0 | 112.5 |
| $^{71}$Trp-$^{1}L_{b5}$ | $^{71}$Trp-$^{1}L_{a}$ | 0.0 | 64.4 |
| $^{71}$Trp-$^{1}L_{b3}$ | $^{71}$Trp-$^{1}L_{a}$ | 0.0 | 86.0 |
| $^{71}$Trp-$^{1}L_{b3}$ | $^{71}$Trp-$^{1}L_{a2}$ | 0.0 | 80.0 |
| $^{71}$Trp-$^{1}L_{b2}$ | $^{71}$Trp-$^{1}L_{a}$ | 0.0 | 91.7 |
| $^{71}$Trp-$^{1}L_{b2}$ | $^{71}$Trp-$^{1}L_{a2}$ | 0.0 | 85.3 |
| $^{94}$Trp-$^{1}L_{b5}$ | $^{94}$Trp-$^{1}L_{a}$ | 0.0 | 61.8 |
| $^{94}$Trp-$^{1}L_{b3}$ | $^{94}$Trp-$^{1}L_{a2}$ | 0.0 | 76.6 |
| $^{94}$Trp-$^{1}L_{b3}$ | $^{94}$Trp-$^{1}L_{a}$ | 0.0 | 82.4 |
| $^{94}$Trp-$^{1}L_{b2}$ | $^{94}$Trp-$^{1}L_{a2}$ | 0.0 | 82.0 |
| $^{94}$Trp-$^{1}L_{b2}$ | $^{94}$Trp-$^{1}L_{a}$ | 0.0 | 88.0 |
| $^{94}$Trp-$^{1}L_{b}$ | $^{94}$Trp-$^{1}L_{a5}$ | 0.0 | 61.1 |
| $^{94}$Trp-$^{1}L_{b}$ | $^{94}$Trp-$^{1}L_{a4}$ | 0.0 | 69.3 |
| $^{94}$Trp-$^{1}L_{b}$ | $^{94}$Trp-$^{1}L_{a3}$ | 0.0 | 65.2 |
| $^{94}$Trp-$^{1}L_{b}$ | $^{94}$Trp-$^{1}L_{a2}$ | 0.0 | 108.0 |
| $^{94}$Trp-$^{1}L_{b}$ | $^{94}$Trp-$^{1}L_{a}$ | 0.0 | 116.1 |

There are some other noticeable interaction energies, for example, $^{58}$Phe (18.8 $cm^{-1}$), $^{17}$Tyr (16.5 $cm^{-1}$) and $^{35}$Trp (25.4 $cm^{-1}$), but they are not shown because they are below the threshold.

The largest interaction energies are 121 $cm^{-1}$ and 116.1 $cm^{-1}$, corresponding to the one-electron mixing of the $\nu 1$ of the $^1L_b$ and the $\nu 1$ of the $^1L_a$ transitions of $^{71}$Trp and $^{94}$Trp, respectively. The interactions between transitions on different chromophores are not shown because they are below the threshold chosen for this protein. $^{90}$Tyr shows weak interactions in comparison to the interaction energies of the tryptophan chromophores. From this table we can see that the one electron mixing mechanism between the $^1L_b$ and the $^1L_a$ of the tryptophan has a predominant role in generating the near-UV CD spectrum.

The contributions of the Trp residues are shown in Figure 4.10, along with the calculated CD spectrum for frame 20. Here we can see that $^{94}$Trp and $^{71}$Trp make the biggest contributions to the overall calculated rotational strength of the CD spectrum of this frame from the NMR family. Both $^{94}$Trp and $^{71}$Trp contribute at some wavelength with positive intensities and at other wavelength show negative intensities. $^{71}$Trp residue contributes more towards the strong intensity positive peaks across the near-UV region. $^{35}$Trp makes a very modest contribution to the combined rotational strength. Mainly tryptophan residues are responsible for generating the near-UV CD spectrum.

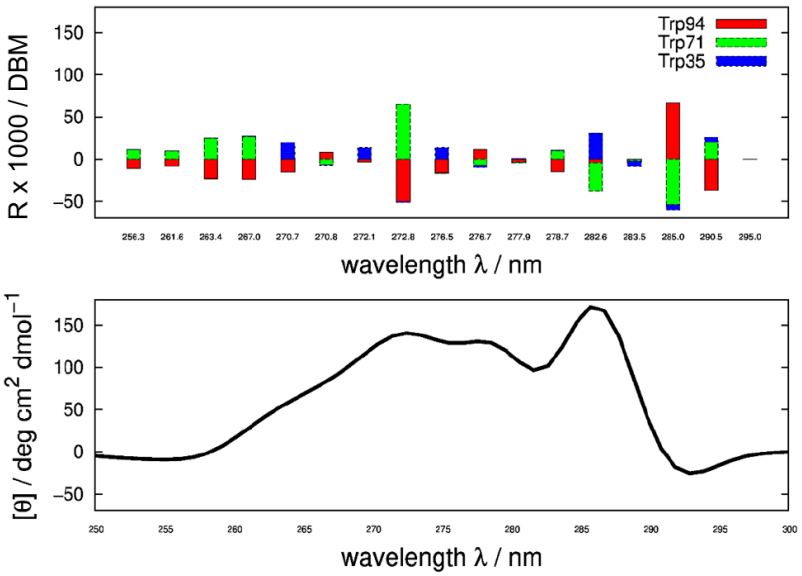



Figure 4.10. Calculated contributions to the intensity for each Trp at different $\lambda$ for frame 20 of barnase (1FW7). The upper panel shows calculated rotational strengths in (DBM) multiplied by 1000 for each residue at each wavelength. The lower panel shows the calculated spectrum.

Further investigations have been carried to examine the role of each aromatic residue. The wavelengths and the associated eigenvector coefficients for tryptophan, tyrosine and phenylalanine residues are investigated and the data are presented in Tables 4.6, 4.7 and 4.8, respectively.

Table 4.6: Wavelength and associated eigenvector coefficients of tryptophan side chain transitions in frame 20 of the NMR structure.

| Eigenvectors index | Values of Eigenvector coefficients | Participation ratio (W) | Residues/ Transitions | Wavelength (nm) |
|---|---|---|---|---|
| 27 | 0.72 | 2.5 | $^{71}Trp_{La8}$ | 278.7 |
| 28 | 0.72 | 2.5 | $^{94}Trp_{La8}$ | 278.7 |
| 29 | 0.92 | 1.0 | $^{35}Trp_{La8}$ | 278.7 |
| 30 | -0.95 | 1.0 | $^{35}Trp_{Lb4}$ | 277.9 |
| 31 | 0.90 | 1.0 | $^{94}Trp_{Lb4}$ | 277.8 |
| 32 | -0.90 | 1.0 | $^{71}Trp_{Lb4}$ | 277.8 |
| 33 | -0.69 | 2.5 | $^{71}Trp_{La7}$ | 276.9 |
| 34 | 0.67 | 2.5 | $^{94}Trp_{La7}$ | 276.8 |
| 35 | 0.82 | 1.8 | $^{35}Trp_{La7}$ | 276.8 |
| 36 | -0.81 | 1.8 | $^{35}Trp_{Lb3}$ | 276.4 |
| 37 | -0.68 | 2.5 | $^{94}Trp_{Lb3}$ | 276.3 |
| 38 | 0.68 | 2.5 | $^{71}Trp_{Lb3}$ | 276.3 |

Table 4.7: Wavelength and associated eigenvector coefficients of tyrosine side chain transitions in frame 20 of the NMR structure.

| Eigenvectors index | Values of Eigenvector coefficients | Participation ratio (W) | Residues/ Transitions | Wavelength (nm) |
|---|---|---|---|---|
| 4 | 0.99 | 1.0 | $^{78}Tyr_{Lb8}$ | 286.1 |
| 5 | 0.99 | 1.0 | $^{90}Tyr_{Lb8}$ | 286.1 |
| 6 | 0.97 | 1.0 | $^{13}Tyr_{Lb8}$ | 286.1 |
| 7 | 0.94 | 1.0 | $^{17}Tyr_{Lb8}$ | 286.1 |
| 10 | 0.90 | 1.0 | $^{24}Tyr_{Lb8}$ | 286.1 |
| 20 | -0.99 | 1.0 | $^{78}Tyr_{Lb7}$ | 279.7 |
| 21 | 0.99 | 1.0 | $^{90}Tyr_{Lb7}$ | 279.7 |
| 22 | 0.97 | 1.0 | $^{13}Tyr_{Lb7}$ | 279.7 |
| 23 | 0.95 | 1.0 | $^{17}Tyr_{Lb7}$ | 279.7 |
| 24 | 0.95 | 1.0 | $^{97}Tyr_{Lb7}$ | 279.7 |
| 25 | 0.88 | 1.3 | $^{103}Tyr_{Lb7}$ | 279.7 |
| 26 | 0.89 | 1.3 | $^{24}Tyr_{Lb7}$ | 279.7 |

Table 4.8: Wavelength and associated eigenvector coefficients of phenylalanine side chain transitions in frame 20 of the NMR structure.

| Eigenvectors index | Values of Eigenvector coefficients | Participation ratio (W) | Residues/ Transitions | Wavelength (nm) |
|---|---|---|---|---|
| 72 | -0.85 | 1.8 | $^{106}Phe_{Lb6}$ | 268.0 |
| 73 | -0.99 | 1.0 | $^{82}Phe_{Lb6}$ | 268.0 |
| 74 | 0.99 | 1.0 | $^{7}Phe_{Lb6}$ | 268.0 |
| 75 | -0.85 | 1.8 | $^{56}Phe_{Lb6}$ | 268.0 |
| 93 | 0.86 | 1.8 | $^{106}Phe_{Lb5}$ | 264.0 |
| 94 | 1.0 | 1.0 | $^{82}Phe_{Lb5}$ | 264.0 |
| 95 | 0.99 | 1.0 | $^{7}Phe_{Lb5}$ | 264.0 |
| 96 | 0.86 | 1.8 | $^{56}Phe_{Lb5}$ | 264.0 |

The tryptophan eigenvector coefficients in Table 4.6 show various values. $^{94}$Trp and $^{71}$Trp have eigenvector coefficients of 0.75 with a participation ratio of 2.5. This shows delocalisation of the transitions for those two tryptophan residues. $^{35}$Trp has eigenvector coefficients of 0.92 with a participation ratio of one, but in some cases, it shows less localisation of the transition with a participation ratio of 1.8 and eigenvector coefficients close to 0.8. Phe eigenvector coefficients in Table 4.8 have similar values to $^{35}$Trp. The tyrosine eigenvector coefficients in Table 4.6 are close to one, and therefore the participation ratio is close to one. This means these transitions are localised on the individual residues.

Table 4.9: The seven ($v$ =1 to 7) $^1L_b$ and the nine ($v$ =1 to 9) tryptophan $^1L_a$ vibrational states, their wavelengths ($\lambda$) and their calculated contributions to the rotational strength ($R_\lambda$) in Debye Bohr magneton (DBM) - shown only if greater than 1 DBM- at that wavelength in the near-UV CD spectrum of frame 20.

| $v$ | Residue transitions | $\lambda$/nm | $^{94}$Trp | | $^{71}$Trp | | $^{35}$Trp | | $R_\lambda$ ×1000 / DBM |
|---|---|---|---|---|---|---|---|---|---|
| | | | $\mu_1 m_1$ | $\cos\theta_1$ | $\mu_2 m_2$ | $\cos\theta_2$ | $\mu_3 m_3$ | $\cos\theta_3$ | |
| 1 | $^1L_b v1$ | 290.5 | 0.43 | 0.29 | 0.73 | -0.15 | 0.50 | 0.03 | 22 |
| 2 | $^1L_b v2$ | 283.5 | 0.04 | -0.005 | 0.17 | 0.07 | 0.18 | 0.05 | 4 |
| 3 | $^1L_b v3$ | 282.6 | 0.02 | 0.0002 | 0.25 | -0.02 | 0.14 | 0.09 | 9 |
| 4 | $^1L_b v4$ | 277.9 | 0.09 | 0.05 | 0.10 | 0.005 | 0.06 | -0.09 | 1 |
| 5 | $^1L_b v5$ | 276.5 | 0.26 | -0.006 | 0.33 | 0.50 | 0.33 | 0.45 | 1.5 |
| 9 | $^1L_a v2$ | 278.7 | 0.52 | -0.06 | 0.91 | 0.02 | 0.87 | -0.001 | 3 |
| 10 | $^1L_a v3$ | 276.7 | 0.09 | -0.002 | 0.21 | 0.23 | 0.21 | 0.15 | 2 |
| 11 | $^1L_a v4$ | 272.8 | 0.25 | -0.22 | 0.41 | 0.01 | 0.46 | 0.11 | 16 |
| 12 | $^1L_a v5$ | 270.8 | 0.17 | -0.007 | 0.12 | 0.05 | 0.16 | 0.09 | 1 |

A detailed examination of the rotational strength at each transition in the calculated spectrum of frame 20 from the NMR family is presented in Table 4.9. This protein has three tryptophans, so there are three contributions at each wavelength. The most intense transitions in the near-UV region arise from the tryptophan residues. Tyrosines and phenylalanines make modest contributions. The values of cos θ of $^{71}$Trp is the highest among the tryptophan residues.

### 4.3.1.3 Beta 2 microglobulin human

The near-UV CD spectra of this protein calculated using the X-ray crystal structure (PDB code: 1LDS) and separately an average over the 20 NMR structures (PDB code:1JNJ) in comparison to the experimental spectrum [144] is presented in Figure 4.11. The overall shape is not well reproduced by either X-ray crystal structure or the average of the NMR structures. Peaks are clearly missed, especially, in the wavelength range between 270 and 290 nm.

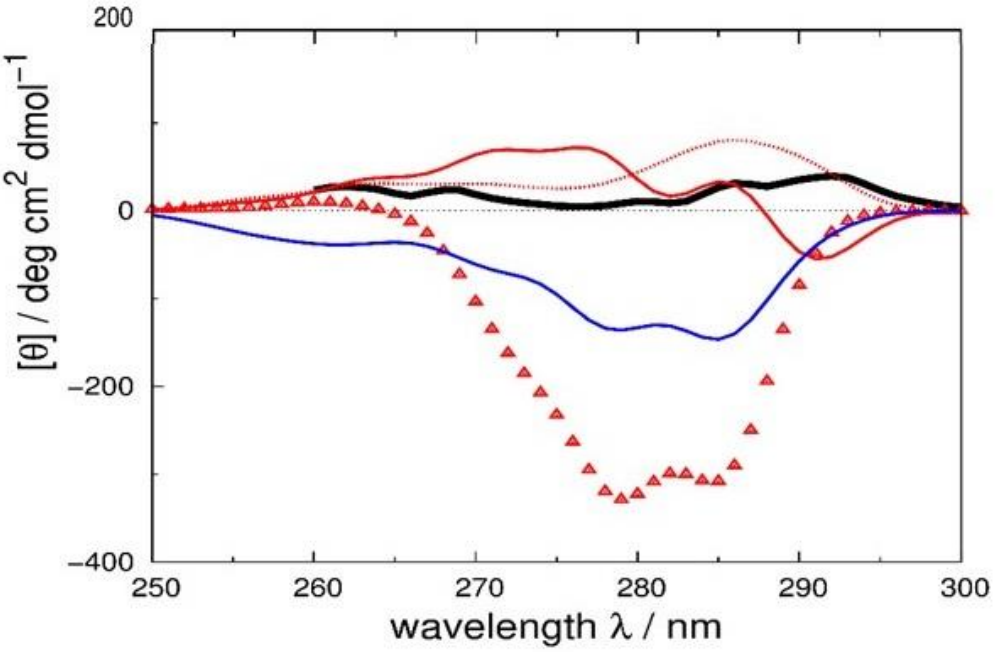



Figure 4.11. Experimental [144] near-UV CD spectrum (black solid line), and spectra calculated using the X-ray structure (blue solid line), calculated using an average over 20 frames of the NMR structure (red solid line) and calculated from individual frames of the NMR structure: frame 19 (dashed red line with triangles) and frame 16 (red dotted line).

The average CD spectrum computed from the NMR structure showed some positive intensities similar to the experimental spectrum. The variation depending on the structure emphasises the effect of conformational changes on the spectrum. Figure 4.11 shows the CD spectrum calculated using specific NMR structures: frame 16 which gives a spectrum in closest agreement with experiment (RMSD 40 deg $cm^2$ $dmol^{-1}$), and by way of contrast frame 19, which gives the calculated spectrum of intensity far from the experimental. The structural difference across these two frames is small, 0.9 Å all-atom RMSD.

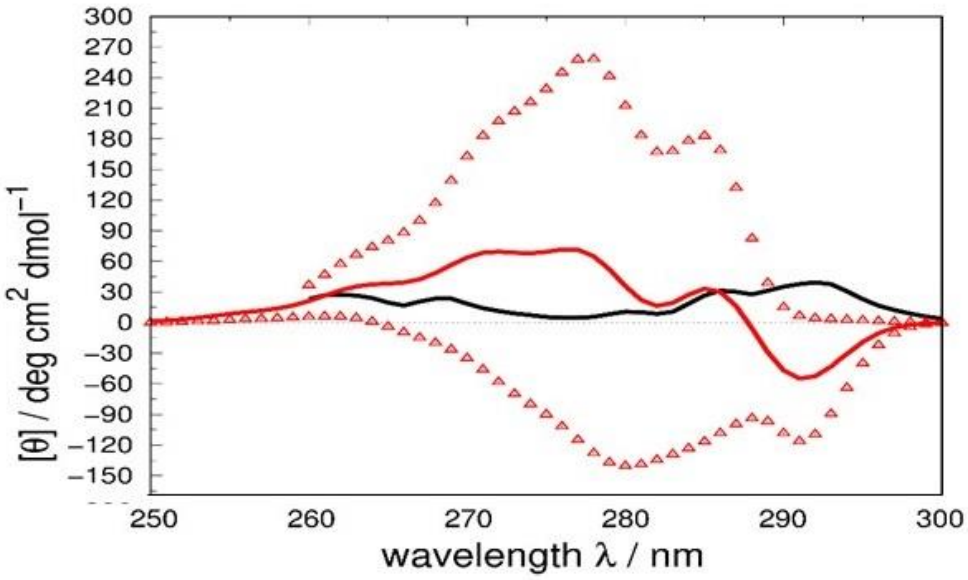



Figure 4.12. Conformational diversity investigation. The black solid line is the experimental spectrum [144], the red solid line is the average over 20 models and the red lines with triangles (upper and lower) represent the mean average ± standard deviation, respectively.

To investigate the influence of conformational diversity on the computed spectra, we calculated the standard deviation of the computed intensity at each wavelength over all the NMR structures as shown in Figure 4.12. The spectrum computed from averaging over all the NMR frames is more intense than the experimental spectrum in the wavelength region 270 to 280 nm. Some peaks appear with negative intensity between 285 and 290 nm, which differs from the experimental spectrum for the same region. We investigated the detailed interactions that govern the calculated spectrum. All the analysis was carried on frame 16 of the NMR structures.

Different types of interactions appear in Table 4.10. Some are intra- and other are inter chromophore mixing. Therefore, there are two mechanisms have a role in generating the CD spectrum. The most significant interactions are from the tryptophan residues. The largest interaction energy is -40.1 $cm^{-1}$ corresponding to the one-electron mixing of the $^{1}L_{b}$, $v5$ and the $^{1}L_{a}$, $v5$ transitions of $^{95}$Trp $v5$ denotes that the transition is to the fifth vibrational level of the excited state. $^{60}$Trp also shows one of the large interaction energies of -36.8 $cm^{-1}$ between the $^{1}L_{b}$, $v7$ and the $^{1}L_{a}$, $v5$ transitions.

Table 4.10: Calculated interaction energies (shown only if greater than 10 $cm^{-1}$) of electronic transitions of aromatic side chains in frame 16 of the NMR structure.

| Chromophore/Transitions | Chromophore/Transitions | Distance (Å) | Interaction Energy ($cm^{-1}$) |
|---|---|---|---|
| $^{70}$Phe-$^{1}L_{a}$ | $^{78}$Tyr-$^{1}L_{b4}$ | 4.5 | 15.1 |
| $^{70}$Phe-$^{1}L_{a}$ | $^{78}$Tyr-$^{1}L_{b2}$ | 4.5 | 19.5 |
| $^{70}$Phe-$^{1}L_{a}$ | $^{78}$Tyr-$^{1}L_{b}$ | 4.5 | 17.5 |
| $^{66}$Tyr-$^{1}L_{b6}$ | $^{66}$Tyr-$^{1}L_{a}$ | 0.0 | 16.6 |
| $^{66}$Tyr-$^{1}L_{b4}$ | $^{66}$Tyr-$^{1}L_{a}$ | 0.0 | 26.7 |
| $^{66}$Tyr-$^{1}L_{b2}$ | $^{66}$Tyr-$^{1}L_{a}$ | 0.0 | 34.4 |
| $^{66}$Tyr-$^{1}L_{b}$ | $^{66}$Tyr-$^{1}L_{a}$ | 0.0 | 30.8 |
| $^{60}$Trp-$^{1}L_{b}$ | $^{60}$Trp-$^{1}L_{a8}$ | 0.0 | -15.4 |
| $^{60}$Trp-$^{1}L_{b7}$ | $^{60}$Trp-$^{1}L_{a5}$ | 0.0 | -36.8 |
| $^{60}$Trp-$^{1}L_{b6}$ | $^{60}$Trp-$^{1}L_{a}$ | 0.0 | -29.7 |
| $^{60}$Trp-$^{1}L_{b6}$ | $^{60}$Trp-$^{1}L_{a2}$ | 0.0 | -24.1 |
| $^{95}$Trp-$^{1}L_{b5}$ | $^{95}$Trp-$^{1}L_{a5}$ | 0.0 | -40.1 |
| $^{95}$Trp-$^{1}L_{b5}$ | $^{95}$Trp-$^{1}L_{a3}$ | 0.0 | -39.7 |

Some tyrosine residues have large interaction energies. In paticular, $^{66}$Tyr shows the one-electron mixing clearly. The mixing between two transitions of different chromophores is found between the $^{1}L_{a}$ of $^{70}$Phe and different transitions of the $^{1}L_{b}$ of $^{78}$Tyr. One more interaction of significance occurs between the $^{1}L_{b}$ of $^{78}$Tyr and different transitions of the $^{1}L_{b3}$ of $^{95}$Trp . The residues are separated by 8.6 Å. Other tyrosines and phenylalanines have some interactions, but they are not shown because their energies

are below the threshold (10 $cm^{-1}$). The energies below this threshold are presented in Table A.5 in the Appendix.

Further investigations have been carried to examine the role of each aromatic residue. The wavelengths and the associated eigenvector coefficients for tryptophan, tyrosine and phenylalanine residues are presented in Tables 4.11, 4.12 and 4.13, respectively.

Table 4.11: Wavelength and associated eigenvector coefficients of tryptophan side chain transitions in frame 16 of the NMR structure.

| Eigenvectors index | Values of Eigenvector coefficients | Participation ratio (W) | Residues/ Transitions | Wavelength (nm) |
|---|---|---|---|---|
| 1 | -0.99 | 1.0 | $^{95}Trp_{Lb7}$ | 291.4 |
| 2 | 0.99 | 1.0 | $^{60}Trp_{Lb7}$ | 290.6 |
| 9 | 0.95 | 1.0 | $^{60}Trp_{La9}$ | 285.7 |
| 10 | 0.99 | 1.0 | $^{95}Trp_{La9}$ | 285.2 |
| 11 | -0.98 | 1.0 | $^{95}Trp_{Lb6}$ | 283.4 |
| 12 | -0.96 | 1.0 | $^{60}Trp_{Lb6}$ | 283.3 |
| 13 | 0.99 | 1.0 | $^{95}Trp_{Lb5}$ | 282.7 |
| 14 | 0.98 | 1.0 | $^{60}Trp_{Lb5}$ | 282.6 |
| 21 | 0.99 | 1.0 | $^{95}Trp_{La8}$ | 278.7 |
| 22 | -0.96 | 1.0 | $^{60}Trp_{La8}$ | 278.7 |
| 23 | -0.99 | 1.0 | $^{95}Trp_{Lb4}$ | 277.9 |
| 24 | -0.97 | 1.0 | $^{60}Trp_{Lb4}$ | 277.8 |

Table 4.12: Wavelength and associated eigenvector coefficients of tyrosine side chain transitions in frame 16 of the NMR structure.

| Eigenvectors index | Values of Eigenvector coefficients | Participation ratio (W) | Residues/ Transitions | Wavelength (nm) |
|---|---|---|---|---|
| 3 | 0.80 | 2.3 | $^{10}Tyr_{Lb8}$ | 286.1 |
| 4 | 0.99 | 1.0 | $^{78}Tyr_{Lb8}$ | 286.1 |
| 5 | -0.98 | 1.0 | $^{63}Tyr_{Lb8}$ | 286.1 |
| 6 | -0.80 | 2.3 | $^{66}Tyr_{Lb8}$ | 286.1 |
| 7 | 0.80 | 2.3 | $^{67}Tyr_{Lb8}$ | 286.1 |
| 8 | 0.80 | 2.3 | $^{26}Tyr_{Lb8}$ | 286.1 |
| 15 | -0.80 | 2.3 | $^{10}Tyr_{Lb7}$ | 279.7 |
| 16 | 0.99 | 1.0 | $^{78}Tyr_{Lb7}$ | 279.7 |
| 17 | -0.98 | 1.0 | $^{63}Tyr_{Lb7}$ | 279.7 |
| 18 | -0.80 | 2.3 | $^{66}Tyr_{Lb7}$ | 279.7 |
| 19 | -0.70 | 2.5 | $^{67}Tyr_{Lb7}$ | 279.7 |
| 20 | -0.70 | 2.5 | $^{26}Tyr_{Lb7}$ | 279.7 |

Table 4.13: Wavelength and associated eigenvector coefficients of phenylalanine side chain transitions in frame 16 of the NMR structure.

| Eigenvectors index | Values of Eigenvector coefficients | Participation ratio (W) | Residues/ Transitions | Wavelength (nm) |
|---|---|---|---|---|
| 55 | -0.70 | 2.5 | $^{30}Phe_{Lb6}$ | 268.0 |
| 56 | -0.97 | 1.0 | $^{62}Phe_{Lb6}$ | 268.0 |
| 57 | 0.94 | 1.0 | $^{22}Phe_{Lb6}$ | 268.0 |
| 58 | 0.95 | 1.0 | $^{70}Phe_{Lb6}$ | 268.0 |
| 59 | 0.73 | 2.5 | $^{56}Phe_{Lb6}$ | 268.0 |
| 74 | 0.71 | 2.5 | $^{30}Phe_{Lb5}$ | 264.0 |
| 75 | -0.98 | 1.0 | $^{62}Phe_{Lb5}$ | 264.0 |
| 76 | -0.94 | 1.0 | $^{22}Phe_{Lb5}$ | 264.0 |
| 77 | 0.98 | 1.0 | $^{70}Phe_{Lb5}$ | 264.0 |
| 78 | -0.73 | 2.5 | $^{56}Phe_{Lb5}$ | 264.0 |

All the tryptophan eigenvector coefficients in Table 4.11 are one, so their participation ratios are one, i.e. the transitions are localised on the individual chromophores. For tyrosines and phenylalanines the eigenvector coefficients vary between 0.99 and 0.7 with participation ratios of 1.0 and 2.5, respectively. The participation ratio when it is less than one represents a delocalised transition for a particular chromophore.

A detailed examination is carried on the rotational strength of each transition in the calculated spectrum. The most intense transitions in this region arise from the tryptophan residues, with the tyrosine and phenylalanine transitions making a more modest contribution. The electric and magnetic transition dipole moments of the tyrosine and phenylalanine residues are substantially smaller than those for the tryptophan residues.

The data in Table 4.14 are computed using frame 16 of the NMR structure. This protein has two tryptophan residues; thus, there are two contributions at each wavelength.

The $\mu m$ values are clearly high for $^{95}$Trp and $^{60}$Trp with a variety between negative and positive values of $\cos\theta$, so it shows noticeable contributions in the calculated spectrum.

Table 4.14: The seven ($v$ =1 to 7) $^1L_b$ and the nine ($v$ =1 to 9) tryptophan $^1L_a$ vibrational states, their wavelengths ($\lambda$) and their calculated contributions to the rotational strength ($R_\lambda$) in Debye Bohr magneton (DBM ) at that wavelength in the near-UV CD spectrum of frame 16.

| $v$ | Residue transitions | $\lambda$/nm | $^{95}$Trp | | $^{60}$Trp | | $R_\lambda \times 1000$ / DBM |
|---|---|---|---|---|---|---|---|
| | | | $\mu_1 m_1$ | $\cos\theta_1$ | $\mu_2 m_2$ | $\cos\theta_2$ | |
| 1 | $^1L_b\nu 1$ | 290.5 | 0.13 | -0.06 | 1.24 | 0.01 | 6 |
| 2 | $^1L_b\nu 2$ | 283.5 | 0.08 | -0.01 | 0.51 | 0.002 | 0.06 |
| 3 | $^1L_b\nu 3$ | 282.6 | 0.04 | -0.05 | 0.55 | 0.007 | 1 |
| 4 | $^1L_b\nu 4$ | 277.9 | 0.02 | -0.01 | 0.16 | -0.001 | -0.7 |
| 5 | $^1L_b\nu 5$ | 276.5 | 0.10 | 0.02 | 0.01 | -0.01 | -1 |
| 6 | $^1L_b\nu 6$ | 272.1 | 0.30 | -0.1 | 0.22 | 0.007 | 0.2 |
| 7 | $^1L_b\nu 7$ | 270.7 | 0.12 | 0.01 | 0.02 | -0.01 | -1 |
| 8 | $^1L_a\nu 1$ | 285.0 | 1.38 | 0.01 | 0.99 | -0.01 | 8 |
| 9 | $^1L_a\nu 2$ | 278.7 | 0.84 | 0.0006 | 0.99 | 0.002 | 2 |
| 10 | $^1L_a\nu 3$ | 276.7 | 0.34 | 0.02 | 0.22 | -0.01 | 4 |
| 11 | $^1L_a\nu 4$ | 272.8 | 0.34 | 0.007 | 0.43 | 0.003 | 3 |
| 12 | $^1L_a\nu 5$ | 270.8 | 0.30 | 0.05 | 0.22 | -0.05 | 5 |
| 13 | $^1L_a\nu 6$ | 267.0 | 0.09 | 0.01 | 0.12 | 0.01 | 2 |
| 14 | $^1L_a\nu 7$ | 263.4 | 0.10 | 0.01 | 0.14 | 0.01 | 2 |
| 15 | $^1L_a\nu 8$ | 261.6 | 0.01 | 0.07 | 0.02 | 0.03 | 1 |
| 16 | $^1L_a\nu 9$ | 256.3 | 0.02 | 0.03 | 0.04 | 0.02 | 2 |

Both $^{95}$Trp and $^{60}$Trp have some positive values of $\cos\theta$ indicating that the angle between the electric and magnetic transition dipoles are less than 90°. However, there are a few negative values of

cos θ which indicated that the angle is more than 90° in this case. No coupling or mixing appears, even if $\mu m$ has a high value.

#### 4.3.1.4 Thioredoxin

Figure 4.13 shows the calculated near-UV CD spectrum of thioredoxin computed by using the X-ray crystallography structure and the average spectrum from the NMR structures (PDB code: 2TRX and 1XOA), in comparison to the experimental spectrum [145].

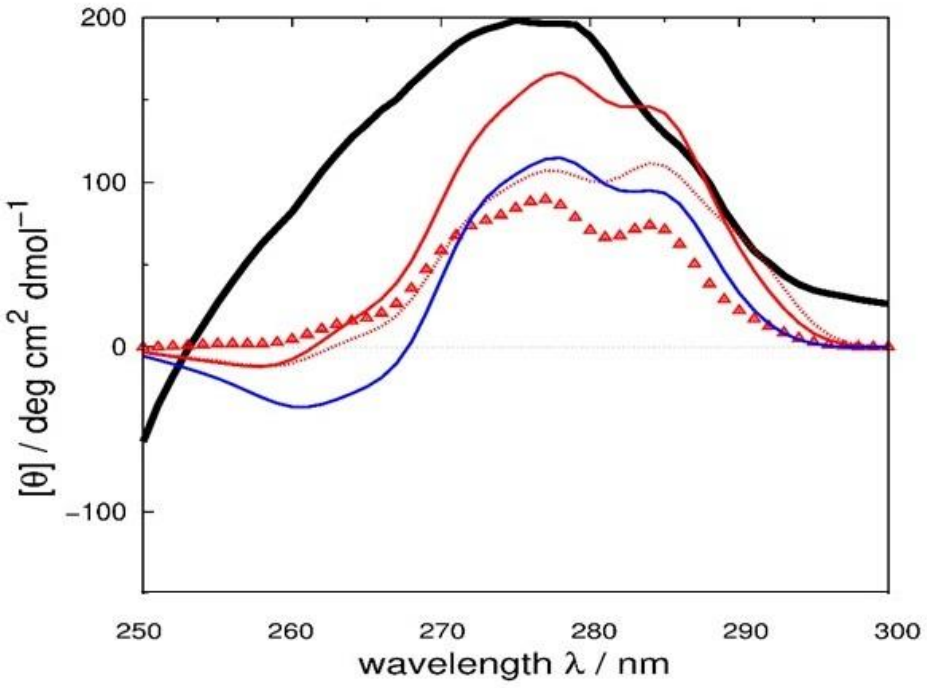



Figure 4.13. Experimental [145] near-UV CD spectrum (black solid line), and spectra calculated using the X-ray structure (blue solid line), calculated using an average over 20 frames of the NMR structure (red solid line) and calculated from individual frames of the NMR structure: frame 20 (dashed red triangles) and frame 5 (red dotted line).

The overall shape generally is well reproduced. The NMR average structures predict a spectrum of a similar intensity to that from the X-ray crystal structure. The computed intensity from both the NMR structures and the X-ray across the wavelength range is lower than the experimental spectrum. The region between 260 - 270 nm is not well reproduced in the calculated spectra and the intensity is weaker than the experimental spectrum. The CD spectra in Figure 4.13 show two models of the family of the NMR structures. Frame 5 gives a spectrum close to the experiment (with an RMSD of 39.5 deg $cm^2$ $dmol^{-1}$), but it is less intense than the experimental spectrum. Frame 20 shows weakest intensity. The structural differences across these two frames is 0.4 Å all-atom RMSD.

To investigate the influence of conformational diversity on the computed spectra, we calculated the standard deviation of the computed intensity at each wavelength over all the NMR structures (Figure 4.14). The spectrum computed from averaging over all the NMR frames is less intense than the experimental spectrum. The contributions of the aromatic chromophores are clearly well reproduced. The focus in the analysis will be on frame 5 of the NMR family.

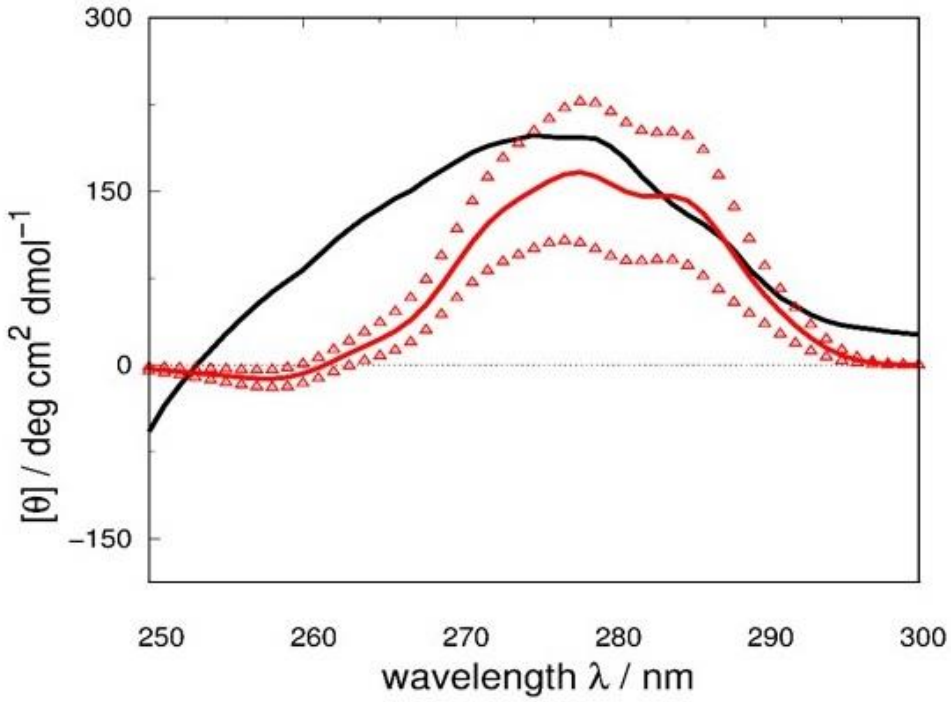



Figure 4.14. Conformational diversity investigation. The black solid line is the experimental spectrum [145]. The red solid line is the mean average computed from 20 NMR models and the red lines with triangles (upper and lower) represent the mean average ± standard deviation, respectively.

Table 4.15 shows the main interactions in the table related to the one-electron mixing of the tryptophan residues. The largest interaction energy is 53.0 $cm^{-1}$ corresponding to the mixing of the $^{1}L_{b}$, $\nu 1$ and the $^{1}L_{a}$, $\nu 1$ of $^{31}$Trp. Generally, the interaction energies of $^{31}$Trp are the highest, and there is no noticeable interaction energy involving $^{28}$Trp. $^{49}$Tyr also shows the one-electron mixing with large interaction energy values (27.2, 24.4 and 21.1 $cm^{-1}$).

Table 4.15: Calculated interaction energies (shown only if greater than 15 $cm^{-1}$) of electronic transitions of aromatic side chains in frame 5 of the NMR structures.

| Chromophore/Transitions | Chromophore/Transitions | Distance (Å) | Interaction Energy ($cm^{-1}$) |
|---|---|---|---|
| $^{12}$Phe-$^{1}L_{b2}$ | $^{81}$Phe-$^{1}L_{a}$ | 7.7 | 19.0 |
| $^{12}$Phe-$^{1}L_{a}$ | $^{81}$Phe-$^{1}L_{b2}$ | 7.7 | 17.5 |
| $^{12}$Phe-$^{1}L_{b4}$ | $^{81}$Phe-$^{1}L_{a}$ | 7.7 | 15.5 |
| $^{12}$Phe-$^{1}L_{b2}$ | $^{70}$Tyr-$^{1}L_{a}$ | 4.9 | 16.2 |
| $^{49}$Tyr-$^{1}L_{b}$ | $^{49}$Tyr-$^{1}L_{a}$ | 0.0 | 24.4 |
| $^{49}$Tyr-$^{1}L_{b4}$ | $^{49}$Tyr-$^{1}L_{a}$ | 0.0 | 21.1 |
| $^{49}$Tyr-$^{1}L_{b2}$ | $^{49}$Tyr-$^{1}L_{a}$ | 0.0 | 27.2 |
| $^{31}$Trp -$^{1}L_{b}$ | $^{31}$Trp -$^{1}L_{a}$ | 0.0 | 28.2 |
| $^{31}$Trp -$^{1}L_{b5}$ | $^{31}$Trp -$^{1}L_{a}$ | 0.0 | 26.2 |
| $^{31}$Trp -$^{1}L_{b5}$ | $^{31}$Trp -$^{1}L_{a5}$ | 0.0 | 27.9 |
| $^{31}$Trp -$^{1}L_{b}$ | $^{31}$Trp -$^{1}L_{a}$ | 0.0 | 53.0 |
| $^{31}$Trp -$^{1}L_{b}$ | $^{31}$Trp -$^{1}L_{a4}$ | 0.0 | 31.6 |
| $^{31}$Trp -$^{1}L_{b}$ | $^{31}$Trp -$^{1}L_{a3}$ | 0.0 | 29.7 |
| $^{31}$Trp -$^{1}L_{b}$ | $^{31}$Trp -$^{1}L_{a2}$ | 0.0 | 49.3 |
| $^{31}$Trp -$^{1}L_{b3}$ | $^{31}$Trp -$^{1}L_{a}$ | 0. 0 | 37.6 |
| $^{31}$Trp -$^{1}L_{b3}$ | $^{31}$Trp- $^{1}L_{a4}$ | 0.0 | 22.4 |
| $^{31}$Trp -$^{1}L_{b3}$ | $^{31}$Trp -$^{1}L_{a3}$ | 0.0 | 21.1 |
| $^{31}$Trp -$^{1}L_{b3}$ | $^{31}$Trp -$^{1}L_{a2}$ | 0.0 | 35.0 |
| $^{31}$Trp -$^{1}L_{b2}$ | $^{31}$Trp -$^{1}L_{a5}$ | 0.0 | 21.1 |
| $^{31}$Trp -$^{1}L_{b2}$ | $^{31}$Trp -$^{1}L_{a}$ | 0.0 | 40.2 |
| $^{31}$Trp -$^{1}L_{b2}$ | $^{31}$Trp -$^{1}L_{a4}$ | 0.0 | 23.9 |
| $^{31}$Trp -$^{1}L_{b2}$ | $^{31}$Trp -$^{1}L_{a3}$ | 0.0 | 22.5 |
| $^{31}$Trp -$^{1}L_{b2}$ | $^{31}$Trp-$^{1}L_{a2}$ | 0.0 | 37.4 |

Energies below the above threshold are presented in Table A.6 in the Appendix.

The mixing of the transitions of different chromophores appears between different phenylalanine residues. Different transitions of the $^{1}L_{a}$ and the $^{1}L_{b}$ of $^{12}$Phe chromophore are mixed with transitions of $^{81}$Phe, the two chromophores are separated by 7.7 Å, while the distance between $^{12}$Phe and $^{70}$Tyr is 4.9 Å. In both cases, the short distance contributes to the strength of the interaction.

The negative contributions for both $^{31}$Trp and $^{28}$Trp at different wavelengths have an effect on the positive contributions at the same wavelengths. This could explain the change in the computed intensity that appears in the CD spectrum.

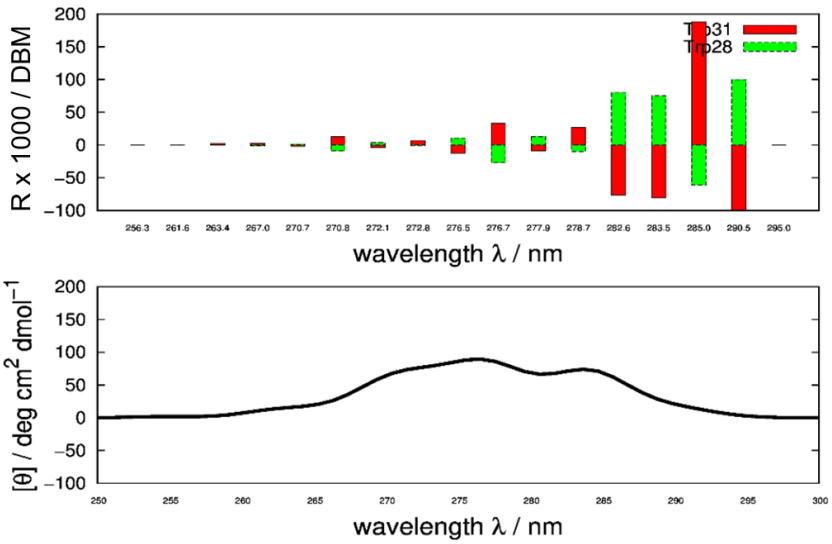



Figure 4.15. Calculated contributions to the intensity for each Trp at different $\lambda$ for frame 20 of thioredoxin (1XOA). The upper panel shows calculated rotational strengths in (DBM) multiplied by 1000 for each residue at each wavelength. The lower panel shows the calculated spectrum.

Further investigations have been carried to examine the role of each aromatic residue. The wavelengths, the associated eigenvector coefficients and the participation ratio for the tryptophan and phenylalanine residues of frame 5 of the NMR structure of thioredoxin are presented in Tables 4.16 and 4.17, respectively. No table is presented for Tyr because the focus of this chapter is on Trp. The tryptophans have eigenvector coefficients that vary from 0.99 with a participation ratio of one, to 0.7 with a participation ratio of 2.5. The latter shows the delocalised nature of the transition.

Table 4.16: Wavelength and associated eigenvector coefficients of tryptophan side chain transitions in frame 5 of the NMR structure.

| Eigenvector index | Values of Eigenvector coefficients | Participation ratio (W) | Residues/ Transitions | Wavelength (nm) |
|---|---|---|---|---|
| 1 | 0.99 | 1.0 | $^{31}Trp_{Lb7}$ | 291.4 |
| 2 | -0.99 | 1.0 | $^{28}Trp_{Lb7}$ | 290.6 |
| 5 | -0.96 | 1.0 | $^{31}Trp_{La9}$ | 285.7 |
| 6 | -0.99 | 1.0 | $^{28}Trp_{La9}$ | 285.2 |
| 7 | -0.87 | 2.0 | $^{28}Trp_{Lb6}$ | 283.4 |
| 8 | -0.87 | 2.0 | $^{31}Trp_{Lb6}$ | 283.3 |
| 9 | 0.71 | 2.5 | $^{28}Trp_{Lb5}$ | 282.7 |
| 10 | 0.71 | 2.5 | $^{31}Trp_{Lb5}$ | 282.6 |
| 15 | 0.95 | 1.0 | $^{28}Trp_{Lb4}$ | 277.9 |
| 16 | -0.95 | 1.0 | $^{31}Trp_{Lb4}$ | 277.8 |
| 17 | 0.7 | 2.5 | $^{31}Trp_{La7}$ | 276.9 |
| 18 | -0.7 | 2.5 | $^{28}Trp_{La7}$ | 276.8 |

Table 4.17: Wavelength and associated eigenvector coefficients of phenylalanine side chain transitions in frame 5 of the NMR structure.

| Eigenvectors index | Values of Eigenvector coefficients | Participation ratio (W) | Residues/ Transitions | Wavelength (nm) |
|---|---|---|---|---|
| 55 | -0.70 | 2.5 | $^{81}Phe_{Lb4}$ | 261.6 |
| 56 | -0.97 | 1.0 | $^{12}Phe_{Lb4}$ | 261.6 |
| 57 | 0.93 | 1.0 | $^{100}Phe_{Lb4}$ | 261.1 |
| 58 | -0.75 | 2.5 | $^{27}Phe_{Lb4}$ | 261.1 |
| 59 | -0.70 | 2.5 | $^{81}Phe_{Lb3}$ | 257.7 |
| 60 | -0.92 | 1.0 | $^{12}Phe_{Lb3}$ | 257.7 |
| 61 | 0.99 | 1.0 | $^{100}Phe_{Lb3}$ | 257.7 |
| 62 | 0.76 | 2.5 | $^{27}Phe_{Lb3}$ | 257.7 |

Some Trp transitions have eigenvector coefficients of 0.87 with a participation ratio of 2.0. Therefore, the behaviour of tryptophan residues varies, and this reflects the variation in the interaction energies presented in Table 4.15.

All the tyrosine residues have eigenvector coefficients of one, so their participation ratio is one. This represents a high localisation of the transitions of tyrosine residues. The role of tyrosine in generating the calculated CD spectrum for this protein is strong, but the focus in the analysis in this chapter is mainly on the role of tryptophan residues.

Phenylalanine residues have some eigenvector coefficients close to one, and others are 0.7 with participation ratio 2.5, especially

[12]Phe which has transition which couple via the $\mu$-$m$ mechanism with other chromophores as shown in Table 4.15.

A detailed examination is carried out on the rotational strength of each transition in the calculated spectrum. The most intense transitions in this region arise from the tryptophan residues, with the tyrosine and phenylalanine transitions making a more modest contribution.

Table 4.18: The seven ($v$ =1 to 7) $^1L_b$ and the nine ($v$ =1 to 9) tryptophan $^1L_a$ vibrational states, their wavelengths ($\lambda$) and their calculated contributions to the rotational strength ($R_\lambda$) in Debye Bohr Magneton (DBM) -shown only if it is greater than 1 DBM- at that wavelength ) in the near-UV CD spectrum of frame 5.

| $v$ | Residue transitions | $\lambda$/nm | [31]Trp | | [28]Trp | | $R_\lambda \times 1000$ / DBM |
|---|---|---|---|---|---|---|---|
| | | | $\mu_1 m_1$ | $\cos\theta_1$ | $\mu_2 m_2$ | $\cos\theta_2$ | |
| 1 | $^1L_b v1$ | 290.5 | 1.05 | -0.01 | 0.45 | 0.06 | 10 |
| 2 | $^1L_b v2$ | 283.5 | 0.16 | -0.05 | 0.79 | 0.01 | 1 |
| 3 | $^1L_b v3$ | 282.6 | 0.06 | -0.52 | 0.68 | 0.05 | 4 |
| 5 | $^1L_b v5$ | 276.5 | 0.08 | 0.009 | 0.68 | 0.0004 | 1 |
| 8 | $^1L_a v1$ | 285.0 | 1.92 | 0.01 | 0.95 | -0.01 | 18 |
| 9 | $^1L_a v2$ | 278.7 | 0.81 | -0.01 | 1.74 | 0.01 | 14 |
| 10 | $^1L_a v3$ | 276.7 | 0.42 | 0.01 | 0.26 | -0.01 | 4 |
| 11 | $^1L_a v4$ | 272.8 | 0.32 | -0.006 | 0.78 | 0.01 | 6 |
| 12 | $^1L_a v5$ | 270.8 | 0.46 | 0.02 | 0.22 | -0.02 | 4 |
| 13 | $^1L_a v6$ | 267.0 | 0.10 | 0.11 | 0.04 | -0.20 | 2 |
| 14 | $^1L_a v7$ | 263.4 | 0.09 | 0.002 | 0.25 | 0.007 | 2 |

The data in Table 4.18 are computed using frame 5 of the NMR structure. This protein has two tryptophan residues; thus, there are two contributions at each wavelength. The biggest contributions ($10 \times 10^{-3}$ and $18 \times 10^{-3}$ DBM) correspond to first transitions of the $^{1}L_a$ and the $^{1}L_b$ of $^{31}$Trp. This chromophore shows noticeable contributions in the calculated spectrum and has a predominant role. The $\mu m$ values are high for $^{28}$Trp, but costheta is small with positive and negative values. This indicates that this chromophore is isolated.

#### 4.3.1.5 Hen egg white lysozyme

Figure 4.16 shows the calculated spectra using the X-ray crystal structure (PDB code: 1HF4) and an average over the 50 NMR structures (PDB code:1E8L) in comparison to the experimental spectrum [146] . The overall shape is better reproduced by the X-ray crystal structure than the average of the NMR structures. Peaks are clearly missed, especially, in the wavelength range between 280 and 295 nm. The average CD spectrum computed from the NMR structure showed intensities similar to the experimental spectrum. The variation depending on the structure emphasises the effect of conformational changes on the spectrum.

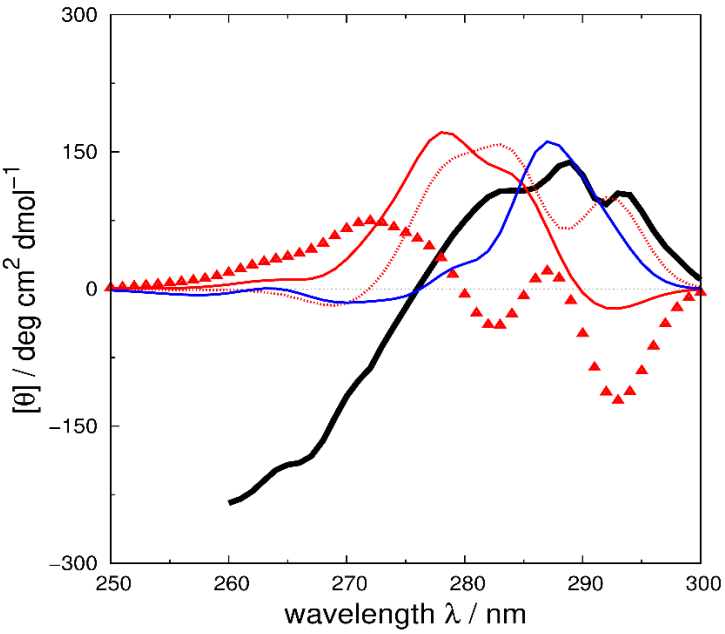



Figure 4.16. Experimental [146] near-UV CD spectrum (black solid line), and spectra calculated using the X-ray structure (blue solid line), calculated using an average over 50 frames of the NMR structure (red solid line) and calculated from individual frames of the NMR structure: frame 2 ( red triangles) and frame 33 (red dotted line).

Figure 4.16 shows the CD spectrum calculated using specific NMR structures: frame 33 which gives the more intense spectrum in closest agreement to the experimental spectrum with RMSD 75 deg cm$^2$ dmol$^{-1}$, and by way of contrast frame 2, which gives the calculated spectrum of weakest intensity. The difference between the two NMR frames is small, 0.5 Å all-atom RMSD.

To investigate the influence of conformational diversity on the computed spectra, we calculated the standard deviation of the computed intensity at each wavelength over all the NMR

structures (Figure 4.17). The spectrum computed from averaging over all the NMR frames is less intense than the experimental spectrum. The contributions of the aromatic chromophores are poorly reproduced at some wavelengths. The focus in the analysis will be on frame 33.

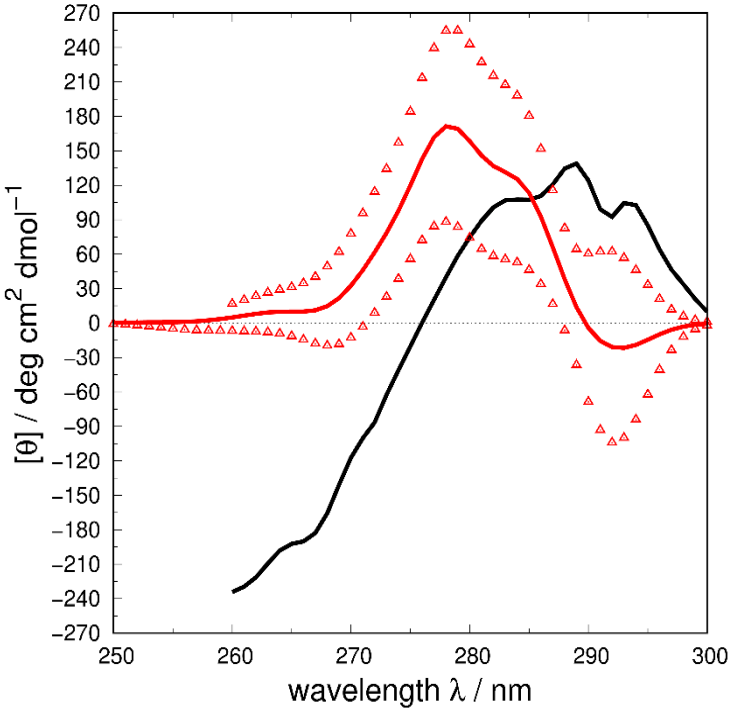



Figure 4.17: Conformational diversity investigation. The black solid line is the experimental spectrum [146]. The red solid line is the mean average computed from 50 NMR models and the red lines with triangles (upper and lower) represent the mean average ± standard deviation, respectively.

Table 4.19: Calculated interaction energies (shown if greater than 100 $cm^{-1}$) of electronic transitions of aromatic side chains in frame 33 of the NMR structure.

| Chromophore/Transitions | Chromophore/Transitions | Distance (Å) | Interaction Energy ($cm^{-1}$) |
|---|---|---|---|
| $^{28}$Trp -$^{1}L_b$ | $^{28}$Trp-$^{1}L_a$ | 0.0 | 129.8 |
| $^{28}$Trp -$^{1}L_b$ | $^{28}$Trp -$^{1}L_{a2}$ | 0.0 | 120.0 |
| $^{63}$Trp-$^{1}L_b$ | $^{63}$Trp -$^{1}L_a$ | 0.0 | 254.7 |
| $^{63}$Trp -$^{1}L_b$ | $^{63}$Trp $^{1}L_{a4}$ | 0.0 | 152.0 |
| $^{63}$Trp -$^{1}L_{b3}$ | $^{63}$Trp -$^{1}L_a$ | 0.0 | 180.8 |
| $^{63}$Trp -$^{1}L_{b3}$ | $^{63}$Trp -$^{1}L_{a2}$ | 0.0 | 168.1 |
| $^{63}$Trp -$^{1}L_{b2}$ | $^{63}$Trp -$^{1}L_{a2}$ | 0.0 | 179.5 |
| $^{123}$Trp -$^{1}L_{b3}$ | $^{123}$Trp-$^{1}L_{a2}$ | 0.0 | 102.3 |
| $^{123}$Trp -$^{1}L_{b3}$ | $^{123}$Trp -$^{1}L_a$ | 0.0 | 110.0 |
| $^{123}$Trp -$^{1}L_{b3}$ | $^{123}$Trp -$^{1}L_{a3}$ | 0.0 | 109.3 |
| $^{123}$Trp -$^{1}L_{b2}$ | $^{123}$Trp -$^{1}L_{a2}$ | 0.0 | 117.5 |
| $^{123}$Trp -$^{1}L_{b2}$ | $^{123}$Trp -$^{1}L_{a2}$ | 0.0 | 110.0 |
| $^{123}$Trp -$^{1}L_{b2}$ | $^{123}$Trp -$^{1}L_a$ | 0.0 | 118.4 |
| $^{123}$Trp -$^{1}L_b$ | $^{123}$Trp -$^{1}L_{a2}$ | 0.0 | 144.0 |
| $^{123}$Trp -$^{1}L_b$ | $^{123}$Trp -$^{1}L_a$ | 0.0 | 155.0 |

Table 4.19 shows the interaction energies of frame 33 of the NMR family. The energies below the threshold are presented in Table A.7 in the Appendix. The predominant interactions are the one-electron mixing of the tryptophan transitions. The largest interaction energy is 254.7 $cm^{-1}$ correspond the mixing of $^{1}L_b$, $\nu 1$ and $^{1}L_a$, $\nu 1$ of $^{63}$Trp. Generally, all the interaction energies are high. Most of the interactions energies of $^{123}$Trp corresponds to

the $^{1}L_b$, $\nu1, \nu2$ and $\nu3$ and the $^{1}L_a$, $\nu1, \nu2, \nu3$ and $\nu4$ of the vibrational states. The $^{28}$Trp interactions correspond to the $^{1}L_b$, $\nu1, \nu2$ and $\nu3$ and the $^{1}L_a$, $\nu1$ and $\nu2$ of the vibrational states.

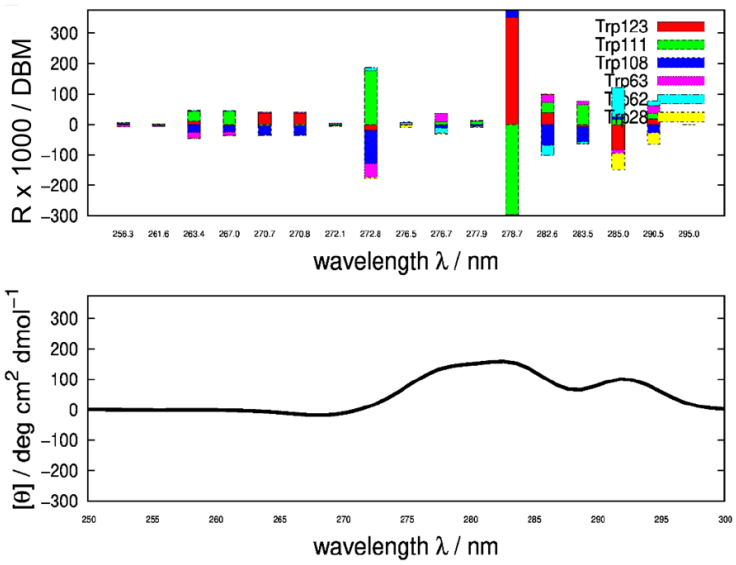



Figure 4.18. Calculated contributions to the intensity for each Trp at different $\lambda$ for frame 33 of hen egg white lysozyme (1E8L). The upper panel shows calculated rotational strengths in (DBM) multiplied by 1000 for each residue at each wavelength. The lower panel shows the calculated spectrum.

The contributions of the tryptophan residues of frame 33 of the NMR structure are shown in Figure 4.18. $^{111}$Trp and $^{123}$Trp make the biggest contributions to the calculated rotational strength. $^{111}$Trp contributes a negative intensity at 278.7 of the near-UV, whereas $^{123}$Trp has a positive contribution at the same wavelength. $^{108}$Trp makes small contributions but these are important in leading to the calculated CD spectrum, $^{62}$Trp makes a negative contribution.

Table 4.20: Wavelength and associated eigenvector coefficients of tryptophan side chain transitions in frame 33 of the NMR structure.

| Eigenvectors index | Eigenvector coefficients | Participation ratio (W) | Residues/ Transitions | Wavelength (nm) |
|---|---|---|---|---|
| 1 | 0.90 | 1.0 | $^{63}Trp_{Lb7}$ | 292.0 |
| 2 | -0.96 | 1.0 | $^{123}Trp_{Lb7}$ | 291.1 |
| 3 | 0.96 | 1.0 | $^{28}Trp_{Lb7}$ | 290.9 |
| 4 | 0.91 | 1.0 | $^{108}Trp_{Lb7}$ | 290.6 |
| 5 | 0.93 | 1.0 | $^{111}Trp_{Lb7}$ | 290.5 |
| 6 | -0.99 | 1.0 | $^{62}Trp_{Lb7}$ | 290.5 |
| 11 | 0.81 | 2.2 | $^{123}Trp_{La9}$ | 285.5 |
| 12 | 0.85 | 2.0 | $^{28}Trp_{La9}$ | 285.4 |
| 13 | 0.90 | 1.0 | $^{108}Trp_{La9}$ | 285.1 |
| 14 | -0.90 | 1.0 | $^{111}Trp_{La9}$ | 285.1 |
| 15 | -0.96 | 1.0 | $^{62}Trp_{La9}$ | 285.0 |

Table 4.21: Wavelength and associated eigenvector coefficients of phenylalanine side chain transitions in frame 33 of the NMR structure.

| Eigenvectors index | Eigenvector coefficients | Participation ratio (W) | Residues/ Transitions | Wavelength (nm) |
|---|---|---|---|---|
| 103 | -0.84 | 2.1 | $^{3}Phe_{Lb5}$ | 264 |
| 104 | -0.99 | 1.0 | $^{34}Phe_{Lb5}$ | 264 |
| 105 | 0.84 | 2.1 | $^{38}Phe_{Lb5}$ | 264 |
| 121 | 0.84 | 2.1 | $^{83}Phe_{Lb4}$ | 261.2 |
| 122 | -0.99 | 1.0 | $^{34}Phe_{Lb4}$ | 261.2 |
| 123 | -0.84 | 2.1 | $^{3}Phe_{Lb4}$ | 261.2 |
| 124 | 0.84 | 2.1 | $^{38}Phe_{Lb3}$ | 257.7 |
| 125 | 0.99 | 1.0 | $^{34}Phe_{Lb3}$ | 257.7 |
| 126 | -0.84 | 2.1 | $^{3}Phe_{Lb3}$ | 257.7 |

The wavelengths and the associated eigenvector coefficients for the tryptophan and the phenylalanine residues are investigated and some of the data are presented in Tables 4.20 and 4.21, respectively. No table has been presented for the tyrosine residues, because all the tyrosine eigenvector coefficients are close to one, i.e., they show a high localisation of the transitions of a specific chromophore. Most of the tryptophan eigenvector coefficients in Table 4.20 are one, so their participation ratio is one, i.e. the transitions are localised on the individual chromophores. Only a few transitions of $^{123}$Trp and $^{28}$Trp have eigenvector coefficients equal to 0.81 with a participation ratio of 2.0. For phenylalanines the eigenvector coefficients vary between 0.99 and 0.7 with participation ratio of 1.0 and 2.5, respectively. The most intense transitions in this region arise from the tryptophan residues, with the tyrosine and phenylalanine transitions making a more modest contribution. The data in Table 4.22 are computed using frame 33 of the NMR structure. This protein has six tryptophan residues; thus, there are six contributions at each wavelength. The $\mu m$ values are clearly high for all tryptophan residues with high values of $\cos\theta$, so they have the most predominant role in generating the near-UV CD calculated spectrum. The values of the $\cos\theta$ vary between negative and positive.

Table 4.22: The seven ($v$ =1 to 7) $^1L_b$ and the nine ($v$ =1 to 9) tryptophan $^1L_a$ vibrational states, their wavelengths ($\lambda$) and their calculated contributions to the rotational strength ($R_\lambda$) in Debye Bohr Magneton (DBM) -shown if greater than 1 DBM- at that wavelength in the near-UV CD spectrum of frame 33.

| $v$ | Residue transitions | $\lambda$/nm | $^{123}$Trp | $^{111}$Trp | $^{108}$Trp | $^{63}$Trp | $^{62}$Trp62 | $^{28}$Trp | $R_\lambda \times$ 1000 / DBM |
|---|---|---|---|---|---|---|---|---|---|
| | | | $\mu_1 m_1$ $\cos\theta_1$ | $\mu_2 m_2$ $\cos\theta_2$ | $\mu_3 m_3$ $\cos\theta_3$ | $\mu_4 m_4$ $\cos\theta_4$ | $\mu_5 m_5$ $\cos\theta_5$ | $\mu_6 m_6$ $\cos\theta_6$ | |
| 1 | $^1L_b v1$ | 290.5 | 0.94<br>0.02 | 0.90<br>0.02 | 0.46<br>-0.06 | 0.12<br>0.22 | 0.30<br>0.05 | 0.74<br>-0.05 | 22 |
| 2 | $^1L_b v2$ | 283.5 | 0.32<br>-0.02 | 0.13<br>0.51 | 0.14<br>-0.36 | 0.12<br>0.09 | 0.31<br>-0.02 | 0.13<br>-0.003 | 14 |
| 3 | $^1L_b v3$ | 282.6 | 0.04<br>0.97 | 0.50<br>0.07 | 0.08<br>-0.85 | 0.42<br>0.06 | 0.13<br>-0.25 | 0.22<br>0.01 | 8 |
| 8 | $^1L_a v1$ | 285.0 | 0.34<br>-0.24 | 0.86<br>0.02 | 0.79<br>0.01 | 0.37<br>-0.03 | 1.18<br>0.08 | 0.69<br>-0.08 | 13 |
| 9 | $^1L_a v2$ | 278.7 | 1.6<br>0.22 | 0.90<br>-0.33 | 0.70<br>0.24 | 0.20<br>-0.45 | 0.29<br>0.02 | -0.10<br>0.97 | 50 |
| 10 | $^1L_a v3$ | 276.7 | 0.03<br>-0.05 | 0.15<br>0.06 | 0.09<br>-0.10 | 0.16<br>0.17 | 0.14<br>-0.13 | -0.005<br>0.56 | 4 |

The positive values of cos θ indicate that the angle between the electric and magnetic transition dipoles are less than 90°. The negative values of cos θ indicate that the angle between the electric and magnetic transition dipoles is greater than 90°. Table

4.22 shows that all the tryptophan residues have shown a coupling or mixing in their transitions.

## 4.4. Influence of electrostatics

In this section the analysis of the electrostatic is presented for only one protein from the five that discussed in this chapter. $^{115}$Trp and $^{130}$Trp of frames 15 of the NMR structure of Apolipophorin III are investigated. APBS is used to generate the potentials and investigate the surrounding environment within the mentioned tryptophans.

We have investigated the surrounding environment for tryptophan residues. The $^{1}L_{a}$ transition of Trp is very sensitive to its surroundings, due to the large dipole moment of this electronic excited state [105,106]. The data of the electrostatic analysis will be discussed only for frame 15 of the NMR structure of Apolipophorin III. Further analysis to the electrostatic environment will provide a better understanding of the role of the tryptophan residues in the origin of the calculated spectrum in the near-UV region.

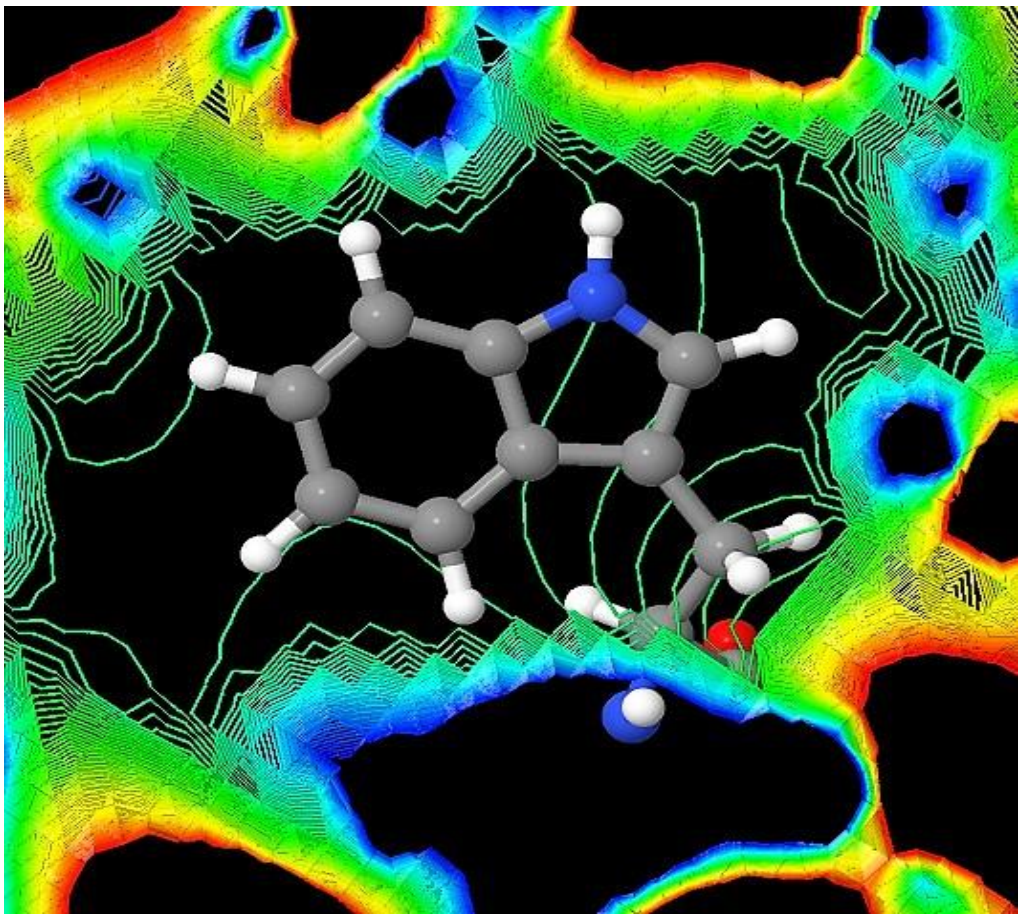


Figure 4.19. Contour plane through $^{115}$Trp in frame 15 from the family of the NMR structure of Apolipophorin III (1LS4). The plane is coloured by calculated electrostatic potential of the local environment of the tryptophan residue. It is set to cut through $C\gamma$, $C\delta1$ and $C\delta2$ of the tryptophan residue. The colours from positive to negative: blue (+20.0), green, yellow, and red (-6.0); with points outside of the contour range (20.0 to -6.0 $k_BT/e_c$) not shown, i.e., black.

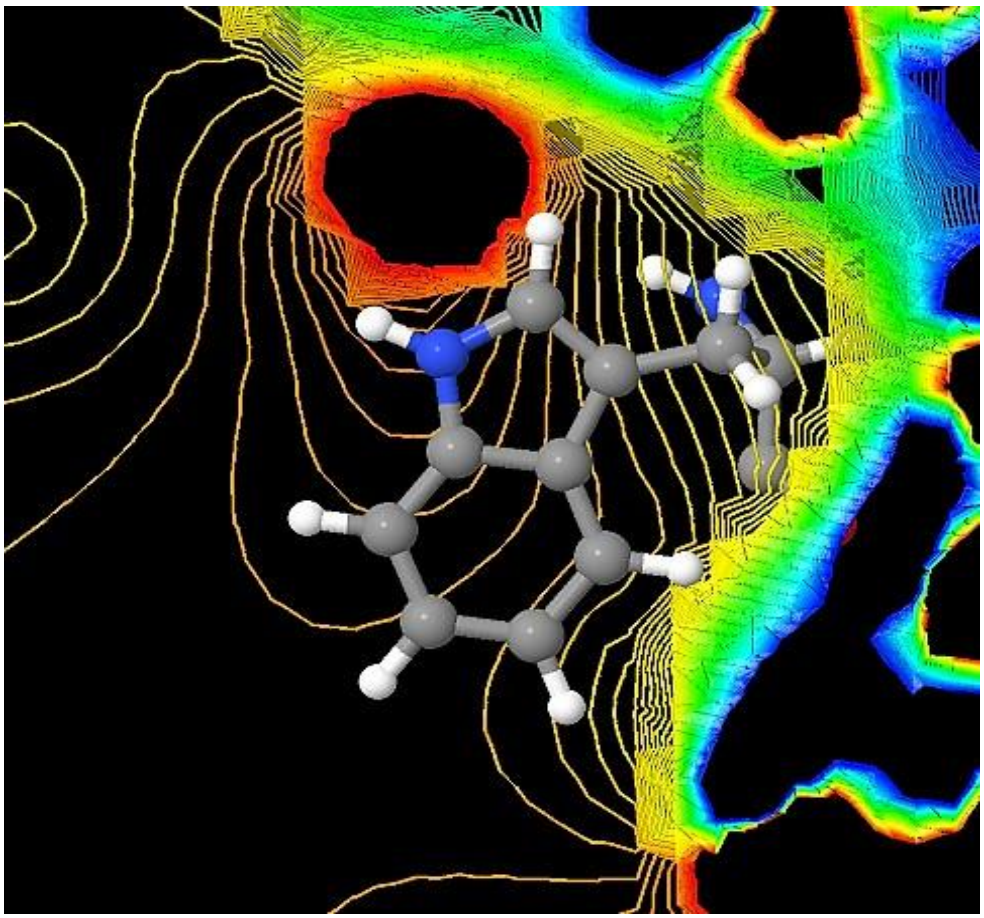


Figure 4.20. Contour plane through $^{130}$Trp in frame 15 from the family of the NMR structure of BPTI. The plane is coloured by calculated electrostatic potential of the local environment of the tryptophan residue. It is set to cut through $C\gamma$, $C\delta1$ and $C\delta2$ of the tryptophan residue. The colours from positive to negative: blue (+20.0), green, yellow, and red (-6.0); with points outside of the contour range (20.0 to -6.0 $k_BT/e_c$) not shown, i.e., black.

Figure 4.19 shows that the local environment surrounding $^{115}$Trp is positive. This residue is buried in the core of the protein and it is located between $^{114}$Gln and $^{116}$Gln. The Gln residue is a polar amino acid, so this may affect the potentials of the $^{115}$Trp. Figure 4.20 shows the potential around $^{130}$Trp. Mostly it expressed a

negative potential. $^{130}$Trp is also buried inside the protein and located between $^{129}$Ala and $^{131}$Ala. Ala is a neutral and nonpolar residue. This could help explain why the electrostatic environment and the local structure perturbs the electronic transitions on both residues to a degree that would affect the spectra, as seen in the computed spectra for frame 15 of the NMR structure of Apolipophorin III.

## 4.5. Conclusion

From the analysis on the five proteins in this chapter, we can conclude that using the vibrational structure of electronic transitions of aromatic amino acids in our calculations of the near-UV CD spectra shows an improvement over spectra calculated without considering the vibrational structure. Investigation and detailed analysis have been done for the tryptophan chromophores of the five proteins, using an extension of the matrix method approach to include new parameterizations of chromophores based on the Franck-Condon integrals of toluene, *p*-cresol and 3-methyl-indole [48].

Overall, the five calculated CD spectra reproduced some of the experimental band shapes and peaks, especially, in the case of apolipophorin III. However, the calculated CD spectrum of barnase and hen egg white lysozyme show weak similarities to

the experimental spectrum. Using different frames of the NMR structures in calculating the CD spectra shows a noticeable improvement from the X-ray crystal structure. In addition, using different frames of the NMR structures enable us to investigate the protein-structure and protein-dynamics relationships and their impact on the generated spectrum.

The investigation of the interactions between residues clarifies the role of each chromophore in generating the calculated CD spectra. Also calculate their contributions of generating CD spectrum enable us to indicate that the one-electron mechanism has the main role in generating the spectrum of all the proteins that examined in this chapter.

Further analysis of Apolipophorin III has characterised the electrostatic potentials generated by the rest of the protein. The investigation of its impact on the mechanisms generating the calculated CD spectrum in near-UV region emphasizes the role and the influence of the electrostatic environment surrounding the most interesting aromatic residues of the Apolipophorin III protein.

# Chapter 5

# Conclusion and Future Work

Proteins are an important class of biomolecules, studied in different scientific subjects, studying their structure-spectrum relationship is essential to understand their function and folding. CD spectroscopy is a useful technique used to study proteins, and it is used in revealing information about conformational changes of proteins. CD spectra in the near-UV region provide information on the aromatic side chains present in a protein due to their electronic transitions. However, the band intensity is highly sensitive to the environments of these aromatic side chains. Therefore, spectra can vary considerably for different conformations of the same protein.

In this thesis we have calculated CD spectra for a set of proteins in the near-UV region. Exciton theory is used to compute the spectra as it is implemented in our DichroCalc web interface, extending the matrix method approach to include new

parameterizations of chromophores based on the Franck-Condon integrals of toluene, *p*-cresol and 3-Methyl-indole [48]. We report some advances in the accuracy of calculations in the near-UV region by using the vibrational structure parameters in the electronic CD spectra of proteins. The new parameters are included, and an improvement in peak positions and intensity has been noticed for some calculated spectra. Considering the vibrational features of the chromophores provided some further insight into the physical origins of the nature of protein CD spectra. However, the difficulty in interpretation of the signal remains, as the CD spectrum is a summation of signals arising from individual chromophores and is affected by the conformational changes.

The reproduction of the calculated CD spectrum in the near-UV regions provides the opportunity to study the underlying mechanisms of interactions between individual aromatic chromophores which generate the spectrum. Two different types of mechanisms have been found the *one-electron* mechanism and the $\mu - m$ mechanism. Mainly the predominant role in generating CD spectra corresponds to the *one- electron* mechanism. Further investigations of each aromatic chromophore have been undertaken by calculating the contributions of each residue to

generate the rotational strength at each wavelength. These contributions emphasise the role of the *one-electron* mechanism in generating a CD spectrum.

Considering the NMR models in our calculations enable us to explore the influence of protein conformational flexibility and changes. Including the NMR models in our calculations shows the importance of the conformation diversity of proteins in the CD calculation. Most of the NMR models in this thesis shows a better reproducing of the calculated spectrum comparing to the X-ray crystallography. In addition, using different frames of the NMR structures enable us to investigate the protein-structure and protein-dynamics relationships and their impact on the generated spectrum.

Chapter three focuses on investigating the electronic properties of tyrosine depend on a combination of specific interactions between the chromophore and its surroundings. Better understanding of the effects of these interactions would make the spectroscopy of tyrosine a more valuable intrinsic probe of protein structure. Tyrosine residues for the three proteins in chapter three show an effective and predominant role in generating the calculated CD spectrum. The analysis has supported the generally held belief

that $^{1}L_{b}$ has the strong role in tyrosine spectrum in the near-UV region. Overall, the three protein calculated CD spectra are well reproduced in terms of the shape and the peaks, especially, in the case of bovine pancreatic trypsin inhibitor. However, the calculated CD spectrum of cardiotoxin shows a weak similarity to the experimental spectrum.

Chapter four focuses mainly on tryptophan chromophores, although some tyrosine residues- and rarely phenylalanine residues- make an important contribution to the calculated spectrum. For example, Beta-2-microglobulin shows significant interaction energies resulted from the coupling and the mixing of the transitions of $^{70}$Phe, $^{66}$Tyr and $^{78}$Tyr. The largest interaction energy mostly corresponds to the one-electron mixing of the $^{1}L_{b}$ and the $^{1}L_{a}$, transitions of the tryptophan chromophores. The *one-electron* mechanism of the tryptophan residues has the main role of generating the calculated spectra of the proteins investigated in chapter four.

Electrostatic interactions are vital to the structure and function of proteins. The investigation of the surrounding environment for aromatic residues is important. APBS is an established piece of software used to generate the electrostatic potentials of

chromophores to study their effect on the calculated CD spectrum. Therefore, further analysis of the tyrosines $^{10}$Tyr, $^{21}$Tyr, $^{23}$Tyr and $^{35}$Tyr of BPTI and the tryptothans $^{130}$Trp and $^{115}$Trp of Apolipophorin III have characterised the electrostatic potentials generated by the rest of the protein. The investigation of the impact of the aromatic residues for the above proteins is presented in detail.

The location of the aromatic residues whether they are buried inside the core of protein or they locate on the surface also has an impact on the electrostatic potentials of each residue. In addition, the location of the tyrosine and the tryptophan chromophores is affected by the neighbours residues whether they are polar or neutral. The positively charged residues near the benzene end of the Trp ring will shift the transitions to longer wavelengths (produce a redshift) [127]. The effects of the aromatic residues on the mechanisms that generate the calculated CD spectrum in near-UV region, emphasises the role and the influence of the electrostatic environment surrounding the aromatic residues of the BPTI and Apolipophorin III proteins.

In the future we will be developing the work further, for example, by generating parameters for disulfide bonds, which is not

considered in our current calculations. Disulfide bonds play an important role for the near-UV spectra, so including them will enable a better analysis and understanding to the calculated spectrum. Further improvement will be gained by consideration of the influence of solvent. Its effects are crucial in *ab initio* calculations. All the calculations in this thesis have been done in the gas phase for all the small molecules representing the chromophore. It would be worthwhile to consider solvent effects. The consideration of disulphide bond and the solvent will make a noticeable improvement in the calculations. The better the calculation can aid the interpretation of the experimental data, the more that CD spectroscopy will be used in protein studies.

A combination between CD calculations and Molecular Dynamics (MD) will enable more exploration of the influence of the protein conformational flexibility on the mechanisms of generation of rotational strength and chromophore interactions. In addition, the implementation of MD simulation snapshots will be useful because we will consider the effect of solvent. Therefore, the investigation of the potentials of the aromatic residues and the impact of the electrostatic environment will be more accurate.

Considering the role of hydrogen bonds is also important to investigate and understand the protein structure. Different secondary structures of proteins are connected by hydrogen bonds between peptide bonds in different patterns. The aromatic side chains of tyrosine and tryptophan can also form hydrogen bonds with other residues or water molecules. The hydrogen bonds have a role in hydrophilic and hydrophobic features for the side chains.

The presence of certain regions of hydrophilicity and hydrophobicity is a driving force in determining protein structure when folding. The varying polarities of the side chains influence the locations of residues in the structure [147]. These interactions will shift the CD band on a case by case basis. Considering these effects would help better assign the CD signals and further improve our understanding of the influence of protein structure.

# References


1. N. Darby and T. Creighton, *Protein structure* Oxford, UK: IRL Press at Oxford University Press, **1993**.
2. U. Langel, B. F. Cravatt, A. Graslund, N. Von Heijne, M. Zorko, T. Land and S. Niessen, *Introduction to peptides and proteins*, CRC Press, **2009**.
3. G. T. Ramachandran and V. Sasisekharan, in *Conformation of polypeptides and proteins*, Elsevier, **1968**, vol. 23, pp. 283-437.
4. G. J. Kleywegt and T. A. Jones. Phi/psi-chology: Ramachandran revisited, *Structure.* **1996**, 4, 1395-1400.
5. B. K. Ho, A. Thomas and R. Brasseur. Revisiting the Ramachandran plot: Hard-sphere repulsion, electrostatics, and H-bonding in the α-helix, *Prot. Sci.* **2003**, 12, 2508-2522.
6. K. A. Dill. Theory for the folding and stability of globular proteins, *Biochemistry*. **1985**, 24, 1501-1509.
7. K. A. Dill, S. Bromberg, K. Yue, H. S. Chan, K. M. Ftebig, D. P. Yee and P. Thomas. Principles of protein folding—a perspective from simple exact models, *Prot. Sci.* **1995**, 4, 561-602.
8. J. Drenth, *Principles of protein X-ray crystallography*, Springer Science & Business Media, **2007**.
9. F. Schotte, M. Lim, T. A. Jackson, A. V. Smirnov, J. Soman, J. S. Olson, G. N. Phillips, M. Wulff and P. A. Anfinrud. Watching a protein as it functions with 150-ps time-resolved x-ray crystallography, *Science.* **2003**, 300, 1944-1947.
10. R. A. Engh and R. Huber. Accurate bond and angle parameters for X-ray protein structure refinement, *Acta Crystallogr., Sect. A*. **1991**, 47, 392-400.
11. H. N. Chapman, P. Fromme, A. Barty, T. A. White, R. A. Kirian, A. Aquila, M. S. Hunter, J. Schulz, D. P. DePonte and U. Weierstall. Femtosecond X-ray protein nanocrystallography, *Nature.* **2011**, 470, 73-77.
12. K. Wüthrich. Nobel Lecture: NMR studies of structure and function of biological macromolecules, *Biosci. Rep.* **2003**, 23, 119-168.
13. C. Raymond, *Basic Principles of Spectroscopy*, McGraw Hill Kogakusha Limited, **1971**.
14. D. A. Skoog, D. M. West, F. J. Holler and S. Crouch, *Fundamentals of analytical chemistry*, Nelson Education, **2013**.
15. C. R. Cantor and P. R. Schimmel, *Biophysical chemistry: Part III: the behavior of biological macromolecules*, Macmillan, **1980**.
16. S. Determann, J. M. Lobbes, R. Reuter and J. Rullkötter. Ultraviolet fluorescence excitation and emission spectroscopy of marine algae and bacteria, *Mar. Chem.* **1998**, 62, 137-156.
17. A. Barth. Infrared spectroscopy of proteins, *Biochimica et Biophysica Acta - Bioenergetics.* **2007**, 1767, 1073-1101.

18. Z. Wen. Raman spectroscopy of protein pharmaceuticals, *J. Pharm. Sci.* .**2007**, 96, 2861-2878.

19. B. Nordén, A. Rodger and T. Dafforn, *Linear dichroism and circular dichroism*, The Royal Society of Chemistry, **2010**.

20. G. D. Fasman, in *Differentiation between transmembrane helices and peripheral helices by the deconvolution of circular dichroism spectra of membrane proteins*, Springer, **1996**, pp. 381-412.

21. S. M. Kelly and N. C. Price. The use of circular dichroism in the investigation of protein structure and function, *Curr. Prot. Pept. Sci.* **2000**, 1, 349-384.

22. H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov and P. E. Bourne. The protein data bank, *Nucleic Acids Res.* **2000**, 28, 235-242.

23. J. Lees, B. Smith, F. Wien, A. Miles and B. Wallace. CDtool—an integrated software package for circular dichroism spectroscopic data processing, analysis, and archiving, *Anal. Biochem.* **2004**, 332, 285-289.

24. L. Whitmore and B. A. Wallace. Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases, *Biopolymers*. **2008**, 89, 392-400.

25. L. Whitmore, B. Woollett, A. J. Miles, D. Klose, R. W. Janes and B. A. Wallace. PCDDB: the Protein Circular Dichroism Data Bank, a repository for circular dichroism spectral and metadata, *Nucleic Acids Res.* **2010**, 39, D480-D486.

26. D. Klose, B. A. Wallace and R. W. Janes. 2Struc: the secondary structure server, *Bioinformatics.* **2010**, 26, 2624-2625.

27. D. Klose, B. A. Wallace and R. W. Janes. DichroMatch: a website for similarity searching of circular dichroism spectra, *Nucleic Acids Res.* **2012**, 40, W547-W552.

28. B. Woollett, L. Whitmore, R. W. Janes and B. A. Wallace. ValiDichro: a website for validating and quality control of protein circular dichroism spectra, *Nucleic Acids Res.* **2013**, 41, W417-W421.

29. C. Wiedemann, P. Bellstedt and M. Görlach. CAPITO—a web server-based analysis and plotting tool for circular dichroism data, *Bioinformatics.* **2013**, 29, 1750-1757.

30. L. Whitmore, A. J. Miles, L. Mavridis, R. W. Janes and B. A. Wallace. PCDDB: new developments at the protein circular dichroism data bank, *Nucleic Acids Res.* **2016**, 45, D303-D307.

31. L. Mavridis and R. W. Janes. PDB2CD: a web-based application for the generation of circular dichroism spectra from protein atomic coordinates, *Bioinformatics.* **2016**, 33, 56-63.

32. B. M. Bulheller and J. D. Hirst. DichroCalc—circular and linear dichroism online, *Bioinformatics*. **2009**, 25, 539-540.

33. N. A. Besley, M.-J. Brienne and J. D. Hirst. Electronic structure of a rigid cyclic diamide, *J. Phys. Chem. B*. **2000**, 104, 12371-12377.

34. J. D. Hirst, D. M. Hirst and C. L. Brooks. Ab initio calculations of the excited states of formamide, *J. Phys. Chem.* **1996**, 100, 13487-13491.

35. J. D. Hirst and B. J. Persson. Ab initio calculations of the vibrational and electronic spectra of diketopiperazine, *J. Phys. Chem.* **1998**, 102, 7519-7524.

36. D. M. Rogers and J. D. Hirst. First-principles calculations of protein circular dichroism in the near ultraviolet, *Biochemistry.* **2004**, 43, 11092-11102.

37. G. H. Wagnière, *On chirality and the universal asymmetry: reflections on image and mirror image*, John Wiley & Sons, **2008**.

38. J. Lough and I. W. Wainer, *Chirality in the Natural & Applied Sciences*, Blackwell Publishing Limited, Oxford, **2002**.

39. A. Meister, *Biochemistry of the amino acids*, Elsevier, **2012**.

40. G. Spach and A. Brack, *Structure, Dynamics and Evolution of Biological Macromolecules* C. Helene. Reidel, **1983**.

41. B. A. Wallace and R. W. Janes, *Modern techniques for circular dichroism and synchrotron radiation circular dichroism spectroscopy*. IOS press. **2009.**

42. C. Djerassi, Proceedings Applications of optical rotatory dispersion and circular dichroism in stereochemistry, **1964**.

43. E. H. Strickland and S. Beychok. Aromatic contributions to circular dichroism spectra of protein, *Crit. Rev. Biochem.* **1974**, 2, 113-175.

44. S. Burley and G. A. Petsko. Aromatic-aromatic interaction: a mechanism of protein structure stabilization, *Science.* **1985**, 229, 23-28.

45. A. Y. M. Woody and R. W. Woody. Individual tyrosine side-chain contributions to circular dichroism of ribonuclease, *Biopolymers.* **2003**, 72, 500-513.

46. B. S. Frank, D. Vardar, D. A. Buckley and C. J. McKnight. The role of aromatic residues in the hydrophobic core of the villin headpiece subdomain, *Prot. Sci.* **2002**, 11, 680-687.

47. P. Crabbé, *ORD and CD in Chemistry and Biochemistry*, Elsevier, **2012**.

48. Z. Li and J. D. Hirst. Quantitative first principles calculations of protein circular dichroism in the near-ultraviolet, *Chem. Sci.* **2017**, 8, 4318-4333.

49. T. A. Keiderling. Protein and peptide secondary structure and conformational determination with vibrational circular dichroism, *Curr. Opini. Chem. Bio.* **2002**, 6, 682-688.

50. L. M. Thompson, A. l. Lasoroski, P. M. Champion, J. T. Sage, M. J. Frisch, J. J. van Thor and M. J. Bearpark. Analytical harmonic vibrational frequencies for the green fluorescent protein computed with ONIOM: chromophore mode character and its response to environment, *J. Chem. Theory Comput.* **2014**, 10, 751-766.

51. D. M. Dooley and J. H. Dawson. Bioinorganic applications of magnetic circular dichroism spectroscopy: copper, rare-earth ions, cobalt and non-heme iron systems, *Coord. Chem. Rev.* **1984**, 60, 1-66.

52. Y. Zhang, J. R. Rouxel, J. Autschbach, N. Govind and S. Mukamel. X-ray circular dichroism signals: a unique probe of local molecular chirality, *Chem. Sci.* **2017**, 8, 5969-5978.

53. P. Štěpánek, S. Coriani, D. Sundholm, V. A. Ovchinnikov and J. Vaara. Relation between molecular electronic structure and nuclear spin-induced circular dichroism, *Sci. Rep.* **2017**, 7, 46617.

54. D. Kliger and J. Lewis, in *Fast Time-Resolved Circular Dichroism Measurements*, **2000**, pp. 243-259.

55. J. Lewis, R. Tilton, C. Einterz, S. Milder, I. Kuntz and D. Kliger. New technique for measuring circular dichroism changes on a nanosecond time scale. Application to (carbonmonoxy) myoglobin and (carbonmonoxy) hemoglobin, *J. Phys. Chem.* **1985**, 89, 289-294.

56. A. F. Drake, J. M. Grould and S. F. Mason. Simultaneous monitoring of ligh-absorption and optical activity in the liquid chromatography of chiral substance, *J. Chromatogr. A.* **1980**, 202, 239-245.

57. M. D. Crespo, E. R. Simpson and M. S. Searle. Population of on-pathway intermediates in the folding of ubiquitin, *J. Mol. Biol.***2006**, 360, 1053-1066.

58. S. M. Kelly, T. J. Jess, N. Price and Proteomics. How to study proteins by circular dichroism, *Biochim. Biophys. Acta, Proteins Proteomics.* **2005**, 1751, 119-139.

59. P. Jemth, C. M. Johnson, S. Gianni and A. R. Fersht. Demonstration by burst-phase analysis of a robust folding intermediate in the FF domain, *Protein Eng., Des. Sel.* **2008**, 21, 207-214.

60. D. M. Rogers and J. D. Hirst. Ab initio study of aromatic side chains of amino acids in gas phase and solution, *J. Phys. Chem. A*. **2003**, 107, 11191-11200.

61. N. A. Besley and J. D. Hirst. Ab initio study of the effect of solvation on the electronic spectra of formamide and N-methylacetamide, *J. Phys. Chem. A.* **1998**, 102, 10791-10797.

62. M. T. Oakley, B. M. Bulheller and J. D. Hirst. First-principles calculations of protein circular dichroism in the far-ultraviolet and beyond, *Chirality.* **2006**, 18, 340-347.

63. Z. Li, D. Robinson and J. D. Hirst. Vibronic structure in the far-UV electronic circular dichroism spectra of proteins, *Faraday Discuss*. **2015**, 177, 329-344.

64. S. Bhattacharjee, G. Tóth, S. Lovas and J. D. Hirst. Influence of tyrosine on the electronic circular dichroism of helical peptides, *J. Phys. Chem. B.* **2003**, 107, 8682-8688.

65. S. B. Jasim, Z. Li, E. E. Guest and J. D. Hirst. DichroCalc: improvements in computing protein circular dichroism spectroscopy in the near-ultraviolet, *J. Mol. Biol.* **2018**, 430, 2196-2202.

66. A. Davydov, *Theory of molecular excitons*, Springer, **2013**.

67. M. E. Davis and J. A. McCammon. Electrostatics in biomolecular structure and dynamics, *Chem. Rev.* **1990**, 90, 509-521.

68. v. L. Rosenfeld. Quantenmechanische Theorie der natürlichen optischen Aktivität von Flüssigkeiten und Gasen, *Zeitschrift für Physik.* **1929**, 52, 161-174.

69. P. M. Bayley, E. B. Nielsen and J. A. Schellman. Rotatory properties of molecules containing two peptide groups: theory, *J. Phys. Chem.* **1969**, 73, 228-243.

70. R. W. Woody and I. Tinoco Jr. Optical rotation of oriented helices. III. Calculation of the rotatory dispersion and circular dichroism of the alpha-and 310-helix, *J. Chem. Phys.* **1967**, 46, 4927-4945.

71. R. W. Woody. Improved calculation of the nπ* rotational strength in polypeptides, *J. Chem. Phys.* **1968**, 49, 4797-4806.

72. S. Matile, N. Berova, K. Nakanishi, J. Fleischhauer and R. W. Woody. Structural studies by exciton coupled circular dichroism over a large distance: porphyrin derivatives of steroids, dimeric steroids, and brevetoxin B, *J. Am. Chem. Soc.* **1996**, 118, 5198-5206.

73. J. Gawroński, M. Brzostowska, K. Kacprzak, H. Kołbon and P. Skowronek. Chirality of aromatic bis-imides from their circular dichroism spectra, *Chirality*. **2000**, 12, 263-268.

74. M. T. Oakley, G. Guichard and J. D. Hirst. Calculations on the electronic excited states of ureas and oligoureas, *J. Phys. Chem. B*. **2007**, 111, 3274-3279.

75. D. Nachtigallová, P. Hobza and H.-H. Ritze. Electronic splitting in the excited states of DNA base homodimers and-trimers: an evaluation of short-range and Coulombic interactions, *Phys. Chem. Chem. Phys.* **2008**, 10, 5689-5697.

76. K. Tsubaki. Synthesis and properties of the chiral oligonaphthalenes, *Org. Biomol. Chem.* **2007**, 5, 2179-2188.

77. I. Tinoco Jr. Theoretical aspects of optical activity part two: polymers, *Adv. Chem. Phys.* **1962**, 4, 113-160.

78. E. Condon, W. Altar and H. Eyring. One-electron rotatory power, *J. Chem. Phys.* **1937**, 5, 753-775.

79. B. O. Roos, in *Book The multiconfigurational (MC) self-consistent field (SCF) theory*, Springer, **1992**, pp. 177-254.

80. J. Schellman. Symmetry rules for optical rotation, *Accounts of Chemical Research*. **1968**, 1, 144-151.

81. C. Christov, A. Kantardjiev, T. Karabencheva and F. Tielens. Mechanisms of generation of the rotational strengths in TEM-1 β-lactamase. Part II: theoretical study of the effects of the electrostatic interactions in the near-UV, *Chem. Phys. Lett*. **2004**, 400, 524-530.

82. C. Christov and T. Karabencheva. Mechanisms of generation of rotational strengths in TEM-1 β-lactamase. Part I: theoretical analysis of the influence of conformational changes in the near-UV, *Chem. Phys. Lett.* **2004**, 396, 282-287.

83. T. Karabencheva and C. Christov. Comparative theoretical study of the mechanisms of generation of rotational strengths in the near-UV in β-lactamases from class A, *Chem. Phys. Lett.* **2004**, 398, 511-516.

84. C. Christov, S. Gabriel, B. Atanasov and J. Fleischhauer. Calculation of the CD spectrum of class A β-lactamase from Escherichia coli (TEM-1), *Zeitsch. Natur. A*. **2001**, 56, 757-760.

85. C. Christov and T. Karabencheva. Computational insight into protein circular dichroism: detailed analysis of contributions of individual chromophores in TEM-1 β-lactamase, *Theor. Chem. Acco*. **2011**, 128, 25-37.

86. C. Christov, T. Karabencheva and A. Lodola. Aromatic interactions and rotational strengths within protein environment: An electronic structural study on β-lactamases from class A, *Chem. Phys. Lett.* **2008**, 456, 89-95.

87. C. Z. Christov, T. G. Karabencheva and A. Lodola. Relationship between chiroptical properties, structural changes and interactions in enzymes: A computational study on β-lactamases from class A, *Comp. Biol. Chem*. **2008**, 32, 167-175.

88. T. Karabencheva, R. Donev, K. Balali-Mood and C. Christov. Individual contributions of the aromatic chromophores to the near-UV Circular Dichroism in class A β-lactamases: A comparative computational analysis, *Biophys. Chem.* **2010**, 151, 39-45.

89. T. G. Karabencheva-Christova, U. Carlsson, K. Balali-Mood, G. W. Black and C. Z. Christov. Conformational effects on the circular dichroism of Human Carbonic Anhydrase II: a multilevel computational study, *PLoS One.* **2013**, 8, e56874.

90. J. A. Schellman. Symmetry rules for optical rotation, *J. Chem. Phys.* **1966**, 44, 55-63.

91. N. A. Baker, D. Sept, S. Joseph, M. J. Holst and J. A. McCammon. Electrostatics of nanosystems: application to microtubules and the ribosome, *Proc. Natl. Acad. Sci. U.S.A*. **2001**, 98, 10037-10041.

92. S. R. Edinger, C. Cortis, P. S. Shenkin and R. A. Friesner. Solvation free energies of peptides: Comparison of approximate continuum solvation models with accurate solution of the Poisson− Boltzmann equation, *J. Phys. Chem. B.* **1997**, 101, 1190-1197.

93. R. Zhou. Free energy landscape of protein folding in water: explicit vs. implicit solvent, *Proteins: Struct., Funct., Bioinf.* **2003**, 53, 148-161.

94. K. A. Sharp and B. Honig. Electrostatic interactions in macromolecules: theory and applications, *Annu. Rev. Biophys. Biophys. Chem.* **1990**, 19, 301-332.

95. B. R. Brooks, C. L. Brooks III, A. D. Mackerell Jr, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels and S. Boresch. CHARMM: the biomolecular simulation program, *J. Comput. Chem.* **2009**, 30, 1545-1614.

96. J. R. Platt. Classification of spectra of cata-condensed hydrocarbons, *J. Chem. Phys.* **1949**, 17, 484-495.

97. J. Horwitz, E. H. Strickland and C. Billups. Analysis of vibrational structure in the near-ultraviolet circular dichroism and absorption spectra of phenylalanine and its derivatives, *J. Am. Chem. Soc.* **1969**, 91, 184-190.

98. J. Horwitz, E. H. Strickland and C. Billups. Analysis of the vibrational structure in the near-ultraviolet circular dichroism and absorption spectra of tyrosine derivatives and ribonuclease-A at 77. deg. K, *J. Am. Chem. Soc.* **1970**, 92, 2119-2129.

99. E. H. Strickland, J. Horwitz and C. Billups. Fine structure in the near-ultraviolet circular dichroism and absorption spectra of tryptophan derivatives and chymotrypsinogen A at 77 K, *Biochemistry.* **1969**, 8, 3205-3213.

100. S. A. Asher, M. Ludwig and C. R. Johnson. UV resonance Raman excitation profiles of the aromatic amino acids, *J. Am. Chem. Soc*. **1986**, 108, 3186-3197.

101. G. Myszkiewicz, W. L. Meerts, C. Ratzer and M. Schmitt. The structure of 4-methylphenol and its water cluster revealed by rotationally resolved UV spectroscopy using a genetic algorithm approach, *J. Chem. Phys.* **2005**, 123, 044304.

102. V. Balachandran, M. Murugan, A. Nataraj, M. Karnan and G. Ilango. Comparative vibrational spectroscopic studies, HOMO–LUMO, NBO analyses and thermodynamic functions of p-cresol and 2-methyl-p-cresol based on DFT calculations, *Spectrochimica Acta Part A: Molecular Biomolecular Spectroscopy.* **2014**, 132, 538-549.

103. J. F. Gaff and S. Franzen. Excited-state geometry method for calculation of the absolute resonance Raman cross sections of the aromatic amino acids, *J. Phys. Chem. A*. **2009**, 113, 5414-5422.

104. R. W. Woody, in *Theory of circular dichroism of proteins*, Springer, **1996**, pp. 25-67.

105. G. Kurapkat, P. Krüger, A. Wollmer, J. Fleischhauer, B. Kramer, E. Zobel, A. Koslowski, H. Botterweck and R. W. Woody. Calculations of the CD spectrum of bovine pancreatic ribonuclease, *Biopolymers*. **1997**, 41, 267-287.

106. P.-S. Song and W. E. Kurtin. Photochemistry of the model phototropic system involving flavines and indoles. III. A spectroscopic study of the polarized luminescence of indoles, *J. Am. Chem. Soc.***1969**, 91, 4892-4906.

107. R. Woody, in *Book Aromatic Side-Chain Contributions to Protein Circular Dchroism*, eds. Uversky V and P. E. Hauppauge, NY: Nova Science Publishers, Inc. **2007**, pp. 291-344.

108. J. Deisenhofer and W. Steigemann. Crystallographic refinement of the structure of bovine pancreatic trypsin inhibitor at l. 5 Å resolution, *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem.* **1975**, 31, 238-250.

109. A. Wlodawer, J. Walter, R. Huber and L. Sjölin. Structure of bovine pancreatic trypsin inhibitor: Results of joint neutron and X-ray refinement of crystal form II, *J. Mol. Biol.* **1984**, 180, 301-329.

110. A. Wlodawer, J. Nachman, G. L. Gilliland, W. Gallagher and C. Woodward. Structure of form III crystals of bovine pancreatic trypsin inhibitor, *J. Mol. Biol.* **1987**, 198, 469-480.

111. K. D. Berndt, P. Güntert, L. P. Orbons and K. Wüthrich. Determination of a high-quality nuclear magnetic resonance solution structure of the bovine pancreatic trypsin inhibitor and comparison with three crystal structures, *J. Mol. Biol.* **1992**, 227, 757-775.

112. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin. UCSF Chimera—a visualization system for exploratory research and analysis, *J. Comput. Chem.* **2004**, 25, 1605-1612.

113. R. Bhaskaran, C.-C. Huang, Y.-C. Tsai, G. Jayaraman, D.-K. Chang and C. Yu. Cardiotoxin II from Taiwan cobra venom, Naja naja atra. Structure in solution and comparison among homologous cardiotoxins, *J. Biol. Chem.* **1994**, 269, 23500-23508.

114. S. Lee, C. Lin, C. Wang, P. Wu, H. Huang, C. Chang and W. Wu. Endocytotic routes of cobra cardiotoxins depend on spatial distribution of positively charged and hydrophobic domains to target distinct types of sulfated glycoconjugates on cell surface, *J. Biol. Chem.* **2014**, 289, 20170-20181.

115. D. D. Leonidas, R. Shapiro, L. I. Irons, N. Russo and K. R. Acharya. Crystal structures of ribonuclease A complexes with 5 ‘-diphosphoadenosine 3 ‘-phosphate and 5 ‘-diphosphoadenosine 2 ‘-phosphate at 1.7 Å resolution, *Biochemistry*. **1997**, 36, 5578-5588.

116. J. Santoro, C. González, M. Bruix, J. L. Neira, J. L. Nieto, J. Herranz and M. Rico. High-resolution three-dimensional structure of ribonuclease A in solution by nuclear magnetic resonance spectroscopy, *J. Mol. Biol.* **1993**, 229, 722-734.

117. P. Bayley. The analysis of circular dichroism of biomolecules, *Prog. Biophys. Mol. Biol.* **1973**, 27, 1-76.

118. N. A. Besley and J. D. Hirst. Theoretical studies toward quantitative protein circular dichroism calculations, *J. Am. Chem. Soc.* **1999**, 121, 9636-9644.

119. E. Jurrus, D. Engel, K. Star, K. Monson, J. Brandi, L. E. Felberg, D. H. Brookes, L. Wilson, J. Chen and K. Liles. Improvements to the APBS biomolecular solvation software suite, *Prot. Sci.* **2018**, 27, 112-128.

120. Jmol: an open-source Java viewer for chemical structures in 3D. http://www.jmol.org/.

121. N. Sreerama, M. C. Manning, M. E. Powers, J.-X. Zhang, D. P. Goldenberg and R. W. Woody. Tyrosine, phenylalanine, and disulfide contributions to the circular dichroism of proteins: circular dichroism spectra of wild-type and mutant bovine pancreatic trypsin inhibitor, *Biochemistry.* **1999**, 38, 10814-10822.

122. W. Humphrey, A. Dalke and K. Schulten. VMD: visual molecular dynamics, *J. Mol. Graphics*. **1996**, 14, 33-38.

123. D. J. Thouless. Electrons in disordered systems and the theory of localization, *Phys. Rep.* **1974**, 13, 93-142.

124. M. N. Pflumm and S. Beychok. Optical Activity of Cystine-containing Proteins II. Circular Dichroism Spectra of Pancreatic Ribonuclease A, Ribonuclease S, and Ribonuclease S-Protein, *J. Biol. Chem.* **1969**, 244, 3973-3981.

125. L. J. Andrews and L. S. Forster. Protein difference spectra. Effect of solvent and charge on tryptophan, *Biochemistry*. **1972**, 11, 1875-1879.

126. L. Serrano-Andrés and B. O. Roos. Theoretical study of the absorption and emission spectra of indole in the gas phase and in a solvent, *J. Am. Chem. Soc.* **1996**, 118, 185-195.

127. J. T. Vivian and P. R. Callis. Mechanisms of tryptophan fluorescence shifts in proteins, *Biophys. J.* **2001**, 80, 2093-2109.

128. O. K. Gasymov, A. R. Abduragimov and B. J. Glasgow. Probing tertiary structure of proteins using single Trp mutations with circular dichroism at low temperature, *J. Phys. Chem. B*. **2014**, 118, 986-995.

129. D. R. Breiter, M. R. Kanost, M. M. Benning, G. Wesenberg, J. H. Law, M. A. Wells, I. Rayment and H. M. Holden. Molecular structure of an apolipoprotein determined at 2.5-. ANG. resolution, *Biochemistry*. **1991**, 30, 603-608.

130. D. Fan, Y. Zheng, D. Yang and J. Wang. NMR solution structure and dynamics of an exchangeable apolipoprotein, Locusta migratoria apolipophorin III, *J. Biol. Chem.* **2003**, 278, 21212-21220.

131. G. Schreiber and A. R. Fersht. Interaction of barnase with its polypeptide inhibitor barstar studied by protein engineering, *Biochemistry*. **1993**, 32, 5145-5150.

132. C. Martin, V. Richard, M. Salem, R. Hartley and Y. Mauguen. Refinement and structural analysis of barnase at 1.5 Å resolution, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1999**, 55, 386-398.

133. M. Y. Reibarkh, L. I. Vasilieva, A. A. Schulga, M. P. Kirpichnikov and A. S. Arseniev. Refined solution structure and backbone dynamics of 15N-labeled Barnase studied by NMR, *J. Biomol. NMR*. **2003**, 15, 271-285.

134. P. D. Gorevic, T. Casey, W. J. Stone, C. DiRaimondo, F. Prelli and B. Frangione. Beta-2 microglobulin is an amyloidogenic protein in man, *J. Clin. Invest.* **1985**, 76, 2425-2429.

135. C. H. Trinh, D. P. Smith, A. P. Kalverda, S. E. Phillips and S. E. Radford. Crystal structure of monomeric human β-2-microglobulin reveals clues to its amyloidogenic properties, *Proc. Natl. Acad. Sci U.S.A.* **2002**, 99, 9771-9776.

136. G. Verdone, A. Corazza, P. Viglino, F. Pettirossi, S. Giorgetti, P. Mangione, A. Andreola, M. Stoppini, V. Bellotti and G. Esposito. The solution structure of human β2-microglobulin reveals the prodromes of its amyloid transition, *Prot. Sci.* **2002**, 11, 487-499.

137. S. K. Katti, D. M. LeMaster and H. Eklund. Crystal structure of thioredoxin from Escherichia coli at 1.68 Å resolution, *J. Mol. Biol.* **1990**, 212, 167-184.

138. M.-F. Jeng, A. P. Campbell, T. Begley, A. Holmgren, D. A. Case, P. E. Wright and H. J. Dyson. High-resolution solution structures of oxidized and reduced Escherichia coli thioredoxin, *Structure*. **1994**, 2, 853-868.

139. A. Holmgren. Thioredoxin and glutaredoxin systems, *J. Biol. Chem.* **1989**, 264, 13963-13966.

140. M. Vaney, I. Broutin, P. Retailleau, A. Douangamath, S. Lafont, C. Hamiaux, T. Prangé, A. Ducruix and M. Riès-Kautt. Structural effects of monovalent anions on polymorphic lysozyme crystals, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2001**, 57, 929-940.

141. H. Schwalbe, S. B. Grimshaw, A. Spencer, M. Buck, J. Boyd, C. M. Dobson, C. Redfield and L. J. Smith. A refined solution structure of hen lysozyme determined using residual dipolar coupling data, *Prot. Sci.* **2001**, 10, 677-688.

142. P. M. Weers, C. M. Kay, K. Oikawa, M. Wientzek, D. J. Van der Horst and R. O. Ryan. Factors affecting the stability and conformation of Locusta migratoria apolipophorin III, *Biochemistry.* **1994**, 33, 3617-3624.

143. M. Bycroft, S. Ludvigsen, A. R. Fersht and F. M. Poulsen. Determination of the three-dimensional solution structure of barnase using nuclear magnetic resonance spectroscopy, *Biochemistry.* **1991**, 30, 8697-8701.

144. E. M. Brown and M. L. Groves. Effect of temperature on the circular dichroism spectra of β2-microglobulins, *FEBS letters.* **1985**, 184, 36-39.

145. R. E. Georgescu, E. H. Braswell, D. Zhu and M. L. Tasayco. Energetics of Assembling An Artificial Heterodimer with An Alpha/Beta Motif: Cleaved versus Uncleaved Escherichia Coli Thioredoxin., *Biochemistry*. **1999**, 38, 13355-13366.

146. M. Buck, S. E. Radford and C. M. Dobson. A partially folded state of hen egg white lysozyme in trifluoroethanol: structural characterization and implications for protein folding, *Biochemistry.* **1993**, 32, 669-678.

147. D. C. Phillips. The hen egg-white lysozyme molecule, *Proc. Natl. Acad. Sci. U. S. A.* **1967**, 483-495.

# Appendix

The interaction energies which have not been presented in chapter three and chapter four are presented here in this appendix. And all the tables their threshold is $\pm 1$ $cm^{-1}$.

Table A.1: Calculated interaction energies of electronic transitions of aromatic side chains in frame 3 of (PDB:1CRF) of the NMR structure

| Chromophore/Transitions | Chromophore/Transitions | Distance (Å) | Interaction Energy ($cm^{-1}$) |
|---|---|---|---|
| $^{11}$Tyr-$^{1}L_{b8}$ | $^{11}$Tyr-$^{1}L_{a}$ | 0 | -7.5 |
| $^{11}$Tyr-$^{1}L_{b7}$ | $^{11}$Tyr-$^{1}L_{a}$ | 0 | -7.5 |
| $^{11}$Tyr-$^{1}L_{b5}$ | $^{11}$Tyr-$^{1}L_{a}$ | 0 | -9.6 |
| $^{11}$Tyr-$^{1}L_{b3}$ | $^{11}$Tyr-$^{1}L_{a}$ | 0 | -9.1 |
| $^{11}$Tyr-$^{1}L_{b2}$ | $^{22}$Tyr-$^{1}L_{a}$ | 16.11 | -1.1 |
| $^{22}$Tyr-$^{1}L_{b8}$ | $^{22}$Tyr-$^{1}L_{a}$ | 0 | 10.9 |
| $^{22}$Tyr-$^{1}L_{b7}$ | $^{22}$Tyr-$^{1}L_{a}$ | 0 | 10.9 |
| $^{22}$Tyr-$^{1}L_{b6}$ | $^{22}$Tyr-$^{1}L_{a}$ | 0 | 21.9 |
| $^{22}$Tyr-$^{1}L_{b5}$ | $^{22}$Tyr-$^{1}L_{a}$ | 0 | 14.1 |
| $^{22}$Tyr-$^{1}L_{b4}$ | $^{22}$Tyr-$^{1}L_{a}$ | 0 | 35.3 |
| $^{22}$Tyr-$^{1}L_{b}$ | $^{22}$Tyr-$^{1}L_{a}$ | 0 | 40.8 |
| $^{22}$Tyr-$^{1}L_{b3}$ | $^{22}$Tyr-$^{1}L_{a}$ | 0 | 13.3 |
| $^{22}$Tyr-$^{1}L_{b2}$ | $^{22}$Tyr-$^{1}L_{a}$ | 0 | 45.5 |
| $^{22}$Tyr-$^{1}L_{b8}$ | $^{51}$Tyr-$^{1}L_{b1}$ | 6.1 | 1.1 |
| $^{22}$Tyr-$^{1}L_{b8}$ | $^{51}$Tyr-$^{1}L_{b4}$ | 6.1 | 1.0 |
| $^{22}$Tyr-$^{1}L_{b8}$ | $^{51}$Tyr-$^{1}L_{b2}$ | 6.1 | 1.3 |
| $^{22}$Tyr-$^{1}L_{b8}$ | $^{51}$Tyr-$^{1}L_{a}$ | 6.1 | -3.5 |
| $^{22}$Tyr-$^{1}L_{b7}$ | $^{51}$Tyr-$^{1}L_{b}$ | 6.1 | 1.1 |
| $^{22}$Tyr-$^{1}L_{b7}$ | $^{51}$Tyr-$^{1}L_{b4}$ | 6.1 | 1.0 |
| $^{22}$Tyr-$^{1}L_{b7}$ | $^{51}$Tyr-$^{1}L_{b2}$ | 6.1 | 1.3 |
| $^{22}$Tyr-$^{1}L_{b7}$ | $^{51}$Tyr-$^{1}L_{a}$ | 6.1 | -3.5 |

| | | | |
|---|---|---|---|
| $^{22}$Tyr-$^{1}$L$_{b6}$ | $^{51}$Tyr-$^{1}$L$_{b1}$ | 6.1 | 2.3 |
| $^{22}$Tyr-$^{1}$L$_{b6}$ | $^{51}$Tyr-$^{1}$L$_{b6}$ | 6.1 | 1.2 |
| $^{22}$Tyr-$^{1}$L$_{b6}$ | $^{51}$Tyr-$^{1}$L$_{b4}$ | 6.1 | 2.0 |
| $^{22}$Tyr-$^{1}$L$_{b6}$ | $^{51}$Tyr-$^{1}$L$_{b2}$ | 6.1 | 2.6 |
| $^{22}$Tyr-$^{1}$L$_{b6}$ | $^{51}$Tyr-$^{1}$L$_{a}$ | 6.1 | -7.1 |
| $^{22}$Tyr-$^{1}$L$_{b5}$ | $^{51}$Tyr-$^{1}$L$_{b}$ | 6.1 | 1.5 |
| $^{22}$Tyr-$^{1}$L$_{b5}$ | $^{51}$Tyr-$^{1}$L$_{b4}$ | 6.1 | 1.3 |
| $^{22}$Tyr-$^{1}$L$_{b5}$ | $^{51}$Tyr-$^{1}$L$_{b2}$ | 6.1 | 1.7 |
| $^{22}$Tyr-$^{1}$L$_{b5}$ | $^{51}$Tyr-$^{1}$L$_{a}$ | 6.1 | -4.5 |
| $^{22}$Tyr-$^{1}$L$_{b4}$ | $^{51}$Tyr-$^{1}$L$_{b8}$ | 6.1 | 1.0 |
| $^{22}$Tyr-$^{1}$L$_{b4}$ | $^{51}$Tyr-$^{1}$L$_{b}$ | 6.1 | 3.8 |
| $^{22}$Tyr-$^{1}$L$_{b4}$ | $^{51}$Tyr-$^{1}$L$_{b7}$ | 6.1 | 1.0 |
| $^{22}$Tyr-$^{1}$L$_{b4}$ | $^{51}$Tyr-$^{1}$L$_{b6}$ | 6.1 | 2.0 |
| $^{22}$Tyr-$^{1}$L$_{b4}$ | $^{51}$Tyr-$^{1}$L$_{b5}$ | 6.1 | 1.3 |
| $^{22}$Tyr-$^{1}$L$_{b4}$ | $^{51}$Tyr-$^{1}$L$_{b4}$ | 6.1 | 3.3 |
| $^{22}$Tyr-$^{1}$L$_{b4}$ | $^{51}$Tyr-$^{1}$L$_{b3}$ | 6.1 | 1.2 |
| $^{22}$Tyr-$^{1}$L$_{b4}$ | $^{51}$Tyr-$^{1}$L$_{b2}$ | 6.1 | 4.2 |
| $^{22}$Tyr-$^{1}$L$_{b}$ | $^{51}$Tyr-$^{1}$L$_{b8}$ | 6.1 | 1.1 |
| $^{22}$Tyr-$^{1}$L$_{b}$ | $^{51}$Tyr-$^{1}$L$_{b}$ | 6.1 | 4.4 |
| $^{22}$Tyr-$^{1}$L$_{b}$ | $^{51}$Tyr-$^{1}$L$_{b7}$ | 6.1 | 1.1 |
| $^{22}$Tyr-$^{1}$L$_{b}$ | $^{51}$Tyr-$^{1}$L$_{b6}$ | 6.1 | 2.3 |
| $^{22}$Tyr-$^{1}$L$_{b}$ | $^{51}$Tyr-$^{1}$L$_{b5}$ | 6.1 | 1.5 |
| $^{22}$Tyr-$^{1}$L$_{b}$ | $^{51}$Tyr-$^{1}$L$_{b4}$ | 6.1 | 3.8 |
| $^{22}$Tyr-$^{1}$L$_{b}$ | $^{51}$Tyr-$^{1}$L$_{b3}$ | 6.1 | 1.4 |
| $^{22}$Tyr-$^{1}$L$_{b}$ | $^{51}$Tyr-$^{1}$L$_{b2}$ | 6.1 | 4.9 |
| $^{22}$Tyr-$^{1}$L$_{b3}$ | $^{51}$Tyr-$^{1}$L$_{b}$ | 6.1 | 1.4 |
| $^{22}$Tyr-$^{1}$L$_{b3}$ | $^{51}$Tyr-$^{1}$L$_{b4}$ | 6.1 | 1.2 |
| $^{22}$Tyr-$^{1}$L$_{b3}$ | $^{51}$Tyr-$^{1}$L$_{b2}$ | 6.1 | 1.6 |
| $^{22}$Tyr-$^{1}$L$_{b3}$ | $^{51}$Tyr-$^{1}$L$_{a}$ | 6.1 | -4.3 |
| $^{22}$Tyr-$^{1}$L$_{b2}$ | $^{51}$Tyr-$^{1}$L$_{b8}$ | 6.1 | 1.3 |
| $^{22}$Tyr-$^{1}$L$_{b2}$ | $^{51}$Tyr-$^{1}$L$_{b}$ | 6.1 | 4.9 |
| $^{22}$Tyr-$^{1}$L$_{b2}$ | $^{51}$Tyr-$^{1}$L$_{b7}$ | 6.1 | 1.3 |
| $^{22}$Tyr-$^{1}$L$_{b2}$ | $^{51}$Tyr-$^{1}$L$_{b6}$ | 6.1 | 2.6 |
| $^{22}$Tyr-$^{1}$L$_{b2}$ | $^{51}$Tyr-$^{1}$L$_{b5}$ | 6.1 | 1.7 |
| $^{22}$Tyr-$^{1}$L$_{b2}$ | $^{51}$Tyr-$^{1}$L$_{b4}$ | 6.1 | 4.2 |
| $^{22}$Tyr-$^{1}$L$_{b2}$ | $^{51}$Tyr-$^{1}$L$_{b3}$ | 6.1 | 1.6 |

| | | | |
|---|---|---|---|
| $^{22}$Tyr-$^{1}L_{b2}$ | $^{51}$Tyr-$^{1}L_{b2}$ | 6.1 | 5.5 |
| $^{22}$Tyr-$^{1}L_{a}$ | $^{51}$Tyr-$^{1}L_{b8}$ | 6.1 | 8.9 |
| $^{22}$Tyr-$^{1}L_{a}$ | $^{51}$Tyr-$^{1}L_{b}$ | 6.1 | 33.1 |
| $^{22}$Tyr-$^{1}L_{a}$ | $^{51}$Tyr-$^{1}L_{b7}$ | 6.1 | 8.9 |
| $^{22}$Tyr-$^{1}L_{a}$ | $^{51}$Tyr-$^{1}L_{b6}$ | 6.1 | 17.8 |
| $^{22}$Tyr-$^{1}L_{a}$ | $^{51}$Tyr-$^{1}L_{b5}$ | 6.1 | 11.4 |
| $^{22}$Tyr-$^{1}L_{a}$ | $^{51}$Tyr-$^{1}L_{b4}$ | 6.1 | 28.7 |
| $^{22}$Tyr-$^{1}L_{a}$ | $^{51}$Tyr-$^{1}L_{b3}$ | 6.1 | 10.8 |
| $^{22}$Tyr-$^{1}L_{a}$ | $^{51}$Tyr-$^{1}L_{b2}$ | 6.1 | 36.9 |
| $^{51}$Tyr-$^{1}L_{b8}$ | $^{51}$Tyr-$^{1}L_{a}$ | 0 | 6.4 |
| $^{51}$Tyr-$^{1}L_{b}$ | $^{51}$Tyr-$^{1}L_{a}$ | 0 | 24.0 |
| $^{51}$Tyr-$^{1}L_{b7}$ | $^{51}$Tyr-$^{1}L_{a}$ | 0 | 6.4 |
| $^{51}$Tyr-$^{1}L_{b6}$ | $^{51}$Tyr-$^{1}L_{a}$ | 0 | 12.9 |
| $^{51}$Tyr-$^{1}L_{b5}$ | $^{51}$Tyr-$^{1}L_{a}$ | 0 | 8.3 |
| $^{51}$Tyr-$^{1}L_{b4}$ | $^{51}$Tyr-$^{1}L_{a}$ | 0 | 20.7 |
| $^{51}$Tyr-$^{1}L_{b3}$ | $^{51}$Tyr-$^{1}L_{a}$ | 0 | 7.8 |
| $^{51}$Tyr-$^{1}L_{b2}$ | $^{51}$Tyr-$^{1}L_{a}$ | 0 | 26.7 |

Table A.2: Calculated interaction energies of electronic transitions of aromatic side chains in frame 12 of (PDB:2AAS) of the NMR structure

| Chromophore/Transitions | Chromophore/Transitions | Distance (Å) | Interaction Energy ($cm^{-1}$) |
|---|---|---|---|
| $^{8}$Phe-$^{1}L_{b}$ | $^{120}$Phe-$^{1}L_{a}$ | 6.28 | 11.8 |
| $^{8}$ Phe-$^{1}L_{b3}$ | $^{120}$ Phe-$^{1}L_{a}$ | 6.28 | 9.8 |
| $^{8}$ Phe-$^{1}L_{a}$ | $^{120}$ Phe-$^{1}L_{b2}$ | 6.28 | 8.7 |
| $^{8}$ Phe-$^{1}L_{a}$ | $^{120}$ Phe-$^{1}L_{b4}$ | 6.28 | 7.1 |
| $^{8}$Phe-$^{1}L_{b5}$ | $^{120}$ Phe-$^{1}L_{a}$ | 6.28 | 5.2 |
| $^{8}$ Phe-$^{1}L_{a}$ | $^{120}$ Phe-$^{1}L_{b6}$ | 6.28 | 3.6 |
| $^{8}$ Phe-$^{1}L_{b}$ | $^{120}$ Phe-$^{1}L_{b2}$ | 6.28 | 1.1 |
| $^{8}$ Phe-$^{1}L_{b2}$ | $^{120}$ Phe-$^{1}L_{b}$ | 6.28 | 1.1 |
| $^{8}$ Phe-$^{1}L_{b3}$ | $^{120}$ Phe-$^{1}L_{b3}$ | 6.28 | -1.0 |

| | | | |
|---|---|---|---|
| $^{8}$ Phe-$^{1}L_{b}$ | $^{120}$ Phe-$^{1}L_{b3}$ | 6.28 | -1.2 |
| $^{8}$ Phe-$^{1}L_{b3}$ | $^{120}$ Phe-$^{1}L_{b}$ | 6.28 | -1.2 |
| $^{8}$ Phe-$^{1}L_{b}$ | $^{120}$ Phe-$^{1}L_{b}$ | 6.28 | -1.5 |
| $^{8}$ Phe-$^{1}L_{b6}$ | $^{120}$ Phe-$^{1}L_{a}$ | 6.28 | -3.6 |
| $^{8}$ Phe-$^{1}L_{a}$ | $^{120}$ Phe-$^{1}L_{b5}$ | 6.28 | -5.2 |
| $^{8}$ Phe-$^{1}L_{b4}$ | $^{120}$ Phe-$^{1}L_{a}$ | 6.28 | -7.0 |
| $^{8}$ Phe-$^{1}L_{b2}$ | $^{120}$ Phe-$^{1}L_{a}$ | 6.28 | -8.6 |
| $^{8}$ Phe-$^{1}L_{a}$ | $^{120}$ Phe-$^{1}L_{b3}$ | 6.28 | -9.9 |
| $^{46}$ Phe-$^{1}L_{a}$ | $^{25}$ Tyr-$^{1}L_{b8}$ | 5.83 | -1.2 |
| $^{46}$ Phe-$^{1}L_{a}$ | $^{25}$ Tyr-$^{1}L_{b7}$ | 5.83 | -1.2 |
| $^{46}$ Phe-$^{1}L_{a}$ | $^{25}$ Tyr-$^{1}L_{b6}$ | 5.83 | -2.5 |
| $^{46}$ Phe-$^{1}L_{a}$ | $^{25}$ Tyr-$^{1}L_{b5}$ | 5.83 | -1.6 |
| $^{46}$ Phe-$^{1}L_{a}$ | $^{25}$ Tyr-$^{1}L_{b4}$ | 5.83 | -4.1 |
| $^{46}$ Phe-$^{1}L_{a}$ | $^{25}$ Tyr-$^{1}L_{b3}$ | 5.83 | -1.5 |
| $^{46}$ Phe-$^{1}L_{a}$ | $^{25}$ Tyr-$^{1}L_{b}$ | 5.83 | -4.7 |
| $^{46}$ Phe-$^{1}L_{a}$ | $^{25}$ Tyr-$^{1}L_{b2}$ | 5.83 | -5.3 |
| $^{73}$ Tyr-$^{1}L_{b8}$ | $^{73}$ Tyr-$^{1}L_{a}$ | 0 | 24.7 |
| $^{73}$ Tyr-$^{1}L_{b7}$ | $^{73}$ Tyr-$^{1}L_{a}$ | 0 | 24.7 |
| $^{73}$ Tyr-$^{1}L_{b}$ | $^{73}$ Tyr-$^{1}L_{a}$ | 0 | 91.9 |
| $^{73}$ Tyr-$^{1}L_{b6}$ | $^{73}$ Tyr-$^{1}L_{a}$ | 0 | 49.5 |
| $^{73}$ Tyr-$^{1}L_{b5}$ | $^{73}$ Tyr-$^{1}L_{a}$ | 0 | 31.8 |
| $^{73}$ Tyr-$^{1}L_{b4}$ | $^{73}$ Tyr-$^{1}L_{a}$ | 0 | 79.6 |
| $^{73}$ Tyr-$^{1}L_{b3}$ | $^{73}$ Tyr-$^{1}L_{a}$ | 0 | 30.0 |
| $^{73}$ Tyr-$^{1}L_{b2}$ | $^{73}$ Tyr-$^{1}L_{a}$ | 0 | 102.6 |
| $^{73}$ Tyr-$^{1}L_{b8}$ | $^{76}$ Tyr-$^{1}L_{a}$ | 10.56 | 1.4 |
| $^{73}$ Tyr-$^{1}L_{b7}$ | $^{76}$ Tyr-$^{1}L_{a}$ | 10.56 | 1.4 |
| $^{73}$ Tyr-$^{1}L_{b}$ | $^{76}$ Tyr-$^{1}L_{a}$ | 10.56 | 5.3 |
| $^{73}$ Tyr-$^{1}L_{b6}$ | $^{76}$ Tyr-$^{1}L_{a}$ | 10.56 | 2.8 |
| $^{73}$ Tyr-$^{1}L_{b5}$ | $^{76}$ Tyr-$^{1}L_{a}$ | 10.56 | 1.8 |
| $^{73}$ Tyr-$^{1}L_{b4}$ | $^{76}$ Tyr-$^{1}L_{a}$ | 10.56 | 4.6 |
| $^{73}$ Tyr-$^{1}L_{b3}$ | $^{76}$ Tyr-$^{1}L_{a}$ | 10.56 | 1.7 |
| $^{73}$ Tyr-$^{1}L_{b2}$ | $^{76}$ Tyr-$^{1}L_{a}$ | 10.56 | 5.9 |

| | | | |
|---|---|---|---|
| $^{73}$ Tyr-$^{1}L_{b8}$ | $^{115}$ Tyr-$^{1}L_{a}$ | 4.05 | -5.3 |
| $^{73}$ Tyr-$^{1}L_{b7}$ | $^{115}$ Tyr-$^{1}L_{a}$ | 4.05 | -5.3 |
| $^{73}$ Tyr-$^{1}L_{b}$ | $^{115}$ Tyr-$^{1}L_{b6}$ | 4.05 | -1. |
| $^{73}$ Tyr-$^{1}L_{b}$ | $^{115}$ Tyr-$^{1}L_{b4}$ | 4.05 | -2.2 |
| $^{73}$ Tyr-$^{1}L_{b}$ | $^{115}$ Tyr-$^{1}L_{b2}$ | 4.05 | -2.8 |
| $^{73}$ Tyr-$^{1}L_{b}$ | $^{115}$ Tyr-$^{1}L_{b}$ | 4.05 | -2.5 |
| $^{73}$ Tyr-$^{1}L_{b6}$ | $^{115}$ Tyr-$^{1}L_{b4}$ | 4.05 | -1.2 |
| $^{73}$ Tyr-$^{1}L_{b6}$ | $^{115}$ Tyr-$^{1}L_{b2}$ | 4.05 | -1.5 |
| $^{73}$ Tyr-$^{1}L_{b6}$ | $^{115}$Tyr-$^{1}L_{b}$ | 4.05 | -1.3 |
| $^{73}$ Tyr-$^{1}L_{b5}$ | $^{115}$ Tyr-$^{1}L_{b2}$ | 4.05 | -1.0 |
| $^{73}$ Tyr-$^{1}L_{b5}$ | $^{115}$ Tyr-$^{1}L_{a}$ | 4.05 | -6.8 |
| $^{73}$ Tyr-$^{1}L_{b4}$ | $^{115}$ Tyr-$^{1}L_{b6}$ | 4.05 | -1.2 |
| $^{73}$ Tyr-$^{1}L_{b4}$ | $^{115}$ Tyr-$^{1}L_{b4}$ | 4.05 | -1.9 |
| $^{73}$ Tyr-$^{1}L_{b4}$ | $^{115}$ Tyr-$^{1}L_{b2}$ | 4.05 | -2.5 |
| $^{73}$ Tyr-$^{1}L_{b4}$ | $^{115}$ Tyr-$^{1}L_{b}$ | 4.05 | -2.2 |
| $^{73}$ Tyr-$^{1}L_{b3}$ | $^{115}$ Tyr-$^{1}L_{a}$ | 4.05 | -6.4 |
| $^{73}$ Tyr-$^{1}L_{b2}$ | $^{115}$ Tyr-$^{1}L_{b6}$ | 4.05 | -1.5 |
| $^{73}$ Tyr-$^{1}L_{b2}$ | $^{115}$ Tyr-$^{1}L_{b5}$ | 4.05 | -1.0 |
| $^{73}$ Tyr-$^{1}L_{b2}$ | $^{115}$ Tyr-$^{1}L_{b4}$ | 4.05 | -2.5 |
| $^{73}$ Tyr-$^{1}L_{b2}$ | $^{115}$ Tyr-$^{1}L_{b2}$ | 4.05 | -3.2 |
| $^{73}$ Tyr-$^{1}L_{b2}$ | $^{115}$ Tyr-$^{1}L_{b}$ | 4.05 | -2.8 |
| $^{73}$ Tyr-$^{1}L_{a}$ | $^{115}$ Tyr-$^{1}L_{b8}$ | 4.05 | -3.5 |
| $^{73}$ Tyr-$^{1}L_{a}$ | $^{115}$ Tyr-$^{1}L_{b7}$ | 4.05 | -3.5 |
| $^{73}$ Tyr-$^{1}L_{a}$ | $^{115}$ Tyr-$^{1}L_{b6}$ | 4.05 | -7.1 |
| $^{73}$ Tyr-$^{1}L_{a}$ | $^{115}$ Tyr-$^{1}L_{b5}$ | 4.05 | -4.6 |
| $^{73}$ Tyr-$^{1}L_{a}$ | $^{115}$ Tyr-$^{1}L_{b3}$ | 4.05 | -4.3 |
| $^{76}$ Tyr-$^{1}L_{a}$ | $^{115}$ Tyr-$^{1}L_{b4}$ | 13.23 | 1.2 |
| $^{76}$ Tyr-$^{1}L_{a}$ | $^{115}$ Tyr-$^{1}L_{b2}$ | 13.23 | 1.6 |
| $^{76}$ Tyr-$^{1}L_{a}$ | $^{115}$ Tyr-$^{1}L_{b}$ | 13.23 | 1.4 |
| $^{92}$ Tyr-$^{1}L_{b8}$ | $^{92}$ Tyr-$^{1}L_{a}$ | 0 | -2.9 |
| $^{92}$ Tyr-$^{1}L_{b7}$ | $^{92}$ Tyr-$^{1}L_{a}$ | 0 | -2.9 |
| $^{92}$ Tyr-$^{1}L_{b6}$ | $^{92}$ Tyr-$^{1}L_{a}$ | 0 | -5. |

| | | | |
|---|---|---|---|
| $^{92}$ Tyr-$^{1}L_{b5}$ | $^{92}$ Tyr-$^{1}L_{a}$ | 0 | -3.7 |
| $^{92}$ Tyr-$^{1}L_{b4}$ | $^{92}$ Tyr-$^{1}L_{a}$ | 0 | -9.4 |
| $^{92}$ Tyr-$^{1}L_{b3}$ | $^{92}$ Tyr-$^{1}L_{a}$ | 0 | -3.5 |
| $^{92}$ Tyr-$^{1}L_{b}$ | $^{97}$ Tyr-$^{1}L_{a}$ | 12.29 | -2.1 |
| $^{92}$ Tyr-$^{1}L_{b6}$ | $^{97}$ Tyr-$^{1}L_{a}$ | 12.29 | -1.1 |
| $^{92}$ Tyr-$^{1}L_{b4}$ | $^{97}$ Tyr-$^{1}L_{a}$ | 12.29 | -1.8 |
| $^{92}$ Tyr-$^{1}L_{b2}$ | $^{97}$ Tyr-$^{1}L_{a}$ | 12.29 | -2.4 |
| $^{92}$ Tyr-$^{1}L_{a}$ | $^{97}$ Tyr-$^{1}L_{b}$ | 12.29 | 1.2 |
| $^{92}$ Tyr-$^{1}L_{a}$ | $^{97}$ Tyr-$^{1}L_{b4}$ | 12.29 | 1.1 |
| $^{92}$ Tyr-$^{1}L_{a}$ | $^{97}$ Tyr-$^{1}L_{b2}$ | 12.29 | 1.4 |
| $^{97}$ Tyr-$^{1}L_{b}$ | $^{97}$ Tyr-$^{1}L_{a}$ | 0 | -6.2 |
| $^{97}$ Tyr-$^{1}L_{b8}$ | $^{97}$ Tyr-$^{1}L_{a}$ | 0 | -1.6 |
| $^{97}$ Tyr-$^{1}L_{b7}$ | $^{97}$ Tyr-$^{1}L_{a}$ | 0 | -1.6 |
| $^{97}$ Tyr-$^{1}L_{b6}$ | $^{97}$ Tyr-$^{1}L_{a}$ | 0 | -3.3 |
| $^{97}$ Tyr-$^{1}L_{b5}$ | $^{97}$ Tyr-$^{1}L_{a}$ | 0 | -2.1 |
| $^{97}$ Tyr-$^{1}L_{b4}$ | $^{97}$ Tyr-$^{1}L_{a}$ | 0 | -5.3 |
| $^{97}$ Tyr-$^{1}L_{b3}$ | $^{97}$ Tyr-$^{1}L_{a}$ | 0 | -2.0 |
| $^{97}$ Tyr-$^{1}L_{b2}$ | $^{97}$ Tyr-$^{1}L_{a}$ | 0 | -6.9 |
| $^{115}$ Tyr-$^{1}L_{b8}$ | $^{115}$ Tyr-$^{1}L_{a}$ | 0 | 10.2 |
| $^{115}$ Tyr-$^{1}L_{b7}$ | $^{115}$ Tyr-$^{1}L_{a}$ | 0 | 10.2 |
| $^{115}$ Tyr-$^{1}L_{b6}$ | $^{115}$ Tyr-$^{1}L_{a}$ | 0 | 20.5 |
| $^{115}$ Tyr-$^{1}L_{b5}$ | $^{115}$ Tyr-$^{1}L_{a}$ | 0 | 13.1 |
| $^{115}$ Tyr-$^{1}L_{b4}$ | $^{115}$ Tyr-$^{1}L_{a}$ | 0 | 32.9 |
| $^{115}$ Tyr-$^{1}L_{b3}$ | $^{115}$ Tyr-$^{1}L_{a}$ | 0 | 12.4 |
| $^{115}$ Tyr-$^{1}L_{b2}$ | $^{115}$ Tyr-$^{1}L_{a}$ | 0 | 42.4 |
| $^{115}$ Tyr-$^{1}L_{b}$ | $^{115}$ Tyr-$^{1}L_{a}$ | 0 | 38.0 |

Table A.3: Calculated interaction energies of electronic transitions of aromatic side chains in frame 15 of (PDB:1LS4) of the NMR structure

| Chromophore/Transitions | Chromophore/Transitions | Distance (Å) | Interaction Energy ($cm^{-1}$) |
|---|---|---|---|
| $^{115}$Trp-$^{1}L_{b7}$ | $^{115}$Trp-$^{1}L_{a9}$ | 0 | 2.5 |
| $^{115}$ Trp-$^{1}L_{b7}$ | $^{115}$ Trp-$^{1}L_{a8}$ | 0 | 2.0 |
| $^{115}$ Trp-$^{1}L_{b7}$ | $^{115}$Trp-$^{1}L_{a7}$ | 0 | 4.8 |
| $^{115}$ Trp-$^{1}L_{b7}$ | $^{115}$ Trp-$^{1}L_{a6}$ | 0 | 4.5 |
| $^{115}$ Trp-$^{1}L_{b7}$ | $^{115}$Trp-$^{1}L_{a5}$ | 0 | 7.6 |
| $^{115}$ Trp-$^{1}L_{b7}$ | $^{115}$ Trp-$^{1}L_{a4}$ | 0 | 8.6 |
| $^{115}$ Trp-$^{1}L_{b7}$ | $^{115}$ Trp-$^{1}L_{a3}$ | 0 | 8.1 |
| $^{115}$ Trp-$^{1}L_{b7}$ | $^{115}$Trp-$^{1}L_{a2}$ | 0 | 13.5 |
| $^{115}$ Trp-$^{1}L_{b7}$ | $^{115}$Trp-$^{1}L_{a}$ | 0 | 14.5 |
| $^{115}$ Trp-$^{1}L_{b6}$ | $^{115}$ Trp-$^{1}L_{a9}$ | 0 | 1.3 |
| $^{115}$ Trp-$^{1}L_{b6}$ | $^{115}$ Trp-$^{1}L_{a8}$ | 0 | 1.0 |
| $^{115}$ Trp-$^{1}L_{b6}$ | $^{115}$ Trp-$^{1}L_{a7}$ | 0 | 2.5 |
| $^{115}$ Trp-$^{1}L_{b6}$ | $^{115}$ Trp-$^{1}L_{a6}$ | 0 | 2.4 |
| $^{115}$ Trp-$^{1}L_{b6}$ | $^{115}$ Trp-$^{1}L_{a5}$ | 0 | 4.0 |
| $^{115}$ Trp-$^{1}L_{b6}$ | $^{115}$ Trp-$^{1}L_{a4}$ | 0 | 4.5 |
| $^{115}$ Trp-$^{1}L_{b6}$ | $^{115}$ Trp-$^{1}L_{a3}$ | 0 | 4.3 |
| $^{115}$ Trp-$^{1}L_{b6}$ | $^{115}$ Trp-$^{1}L_{a2}$ | 0 | 7.1 |
| $^{115}$ Trp-$^{1}L_{b6}$ | $^{115}$ Trp-$^{1}L_{a}$ | 0 | 7.6 |
| $^{115}$ Trp-$^{1}L_{b5}$ | $^{115}$ Trp-$^{1}L_{a9}$ | 0 | 4.9 |
| $^{115}$ Trp-$^{1}L_{b5}$ | $^{115}$ Trp-$^{1}L_{a8}$ | 0 | 3.9 |
| $^{115}$ Trp-$^{1}L_{b5}$ | $^{115}$ Trp-$^{1}L_{a7}$ | 0 | 9.4 |
| $^{115}$ Trp-$^{1}L_{b5}$ | $^{115}$ Trp-$^{1}L_{a6}$ | 0 | 8.9 |
| $^{115}$ Trp-$^{1}L_{b5}$ | $^{115}$ Trp-$^{1}L_{a5}$ | 0 | 14.8 |
| $^{115}$ Trp-$^{1}L_{b5}$ | $^{115}$ Trp-$^{1}L_{a4}$ | 0 | 16.8 |
| $^{115}$ Trp-$^{1}L_{b5}$ | $^{115}$ Trp-$^{1}L_{a3}$ | 0 | 15.8 |
| $^{115}$ Trp-$^{1}L_{b5}$ | $^{115}$ Trp-$^{1}L_{a2}$ | 0 | 26.2 |
| $^{115}$ Trp-$^{1}L_{b5}$ | $^{115}$ Trp-$^{1}L_{a}$ | 0 | 28.2 |

| | | | |
|---|---|---|---|
| $^{115}$Trp-$^{1}$L$_{b4}$ | $^{115}$ Trp-$^{1}$L$_{a9}$ | 0 | 3.4 |
| $^{115}$ Trp-$^{1}$L$_{b4}$ | $^{115}$ Trp-$^{1}$L$_{a8}$ | 0 | 2.7 |
| $^{115}$ Trp-$^{1}$L$_{b4}$ | $^{115}$ Trp-$^{1}$L$_{a7}$ | 0 | 6.5 |
| $^{115}$ Trp-$^{1}$L$_{b4}$ | $^{115}$ Trp-$^{1}$L$_{a6}$ | 0 | 6.2 |
| $^{115}$ Trp-$^{1}$L$_{b4}$ | $^{115}$ Trp-$^{1}$L$_{a5}$ | 0 | 10.3 |
| $^{115}$ Trp-$^{1}$L$_{b4}$ | $^{115}$ Trp-$^{1}$L$_{a4}$ | 0 | 11.7 |
| $^{115}$ Trp-$^{1}$L$_{b4}$ | $^{115}$ Trp-$^{1}$L$_{a3}$ | 0 | 11.0 |
| $^{115}$ Trp-$^{1}$L$_{b4}$ | $^{115}$ Trp-$^{1}$L$_{a2}$ | 0 | 18.2 |
| $^{115}$ Trp-$^{1}$L$_{b4}$ | $^{115}$ Trp-$^{1}$L$_{a}$ | 0 | 19.6 |
| $^{115}$ Trp-$^{1}$L$_{b3}$ | $^{115}$ Trp-$^{1}$L$_{a9}$ | 0 | 6.5 |
| $^{115}$ Trp-$^{1}$L$_{b3}$ | $^{115}$ Trp-$^{1}$L$_{a8}$ | 0 | 5.2 |
| $^{115}$ Trp-$^{1}$L$_{b3}$ | $^{115}$ Trp-$^{1}$L$_{a7}$ | 0 | 12.5 |
| $^{115}$ Trp-$^{1}$L$_{b3}$ | $^{115}$ Trp-$^{1}$L$_{a6}$ | 0 | 11.8 |
| $^{115}$ Trp-$^{1}$L$_{b3}$ | $^{115}$ Trp-$^{1}$L$_{a5}$ | 0 | 19.7 |
| $^{115}$ Trp-$^{1}$L$_{b3}$ | $^{115}$ Trp-$^{1}$L$_{a4}$ | 0 | 22.4 |
| $^{115}$ Trp-$^{1}$L$_{b3}$ | $^{115}$ Trp-$^{1}$L$_{a3}$ | 0 | 21.1 |
| $^{115}$ Trp-$^{1}$L$_{b3}$ | $^{115}$ Trp-$^{1}$L$_{a2}$ | 0 | 34.9 |
| $^{115}$ Trp-$^{1}$L$_{b3}$ | $^{115}$ Trp-$^{1}$L$_{a}$ | 0 | 37.6 |
| $^{115}$ Trp-$^{1}$L$_{b2}$ | $^{115}$ Trp-$^{1}$L$_{a9}$ | 0 | 7.0 |
| $^{115}$ Trp-$^{1}$L$_{b2}$ | $^{115}$ Trp-$^{1}$L$_{a8}$ | 0 | 5.6 |
| $^{115}$ Trp-$^{1}$L$_{b2}$ | $^{115}$ Trp-$^{1}$L$_{a7}$ | 0 | 13.3 |
| $^{115}$ Trp-$^{1}$L$_{b2}$ | $^{115}$ Trp-$^{1}$L$_{a6}$ | 0 | 12.6 |
| $^{115}$ Trp-$^{1}$L$_{b2}$ | $^{115}$ Trp-$^{1}$L$_{a5}$ | 0 | 21.1 |
| $^{115}$ Trp-$^{1}$L$_{b2}$ | $^{115}$ Trp-$^{1}$L$_{a4}$ | 0 | 23.9 |
| $^{115}$ Trp-$^{1}$L$_{b2}$ | $^{115}$ Trp-$^{1}$L$_{a3}$ | 0 | 22.5 |
| $^{115}$ Trp-$^{1}$L$_{b2}$ | $^{115}$ Trp-$^{1}$L$_{a2}$ | 0 | 37.3 |
| $^{115}$ Trp-$^{1}$L$_{b2}$ | $^{115}$ Trp-$^{1}$L$_{a}$ | 0 | 40.1 |
| $^{115}$ Trp-$^{1}$L$_{b}$ | $^{115}$ Trp-$^{1}$L$_{a9}$ | 0 | 9.2 |
| $^{115}$ Trp-$^{1}$L$_{b}$ | $^{115}$ Trp-$^{1}$L$_{a8}$ | 0 | 7.4 |
| $^{115}$ Trp-$^{1}$L$_{b}$ | $^{115}$ Trp-$^{1}$L$_{a7}$ | 0 | 17.6 |
| $^{115}$ Trp-$^{1}$L$_{b}$ | $^{115}$ Trp-$^{1}$L$_{a6}$ | 0 | 16.7 |
| $^{115}$ Trp-$^{1}$L$_{b}$ | $^{115}$ Trp-$^{1}$L$_{a5}$ | 0 | 27.8 |

| | | | |
|---|---|---|---|
| $^{115}$ Trp-$^{1}L_{b}$ | $^{115}$ Trp-$^{1}L_{a4}$ | 0 | 31.6 |
| $^{115}$ Trp-$^{1}L_{b}$ | $^{115}$Trp-$^{1}L_{a3}$ | 0 | 29.7 |
| $^{115}$ Trp-$^{1}L_{b}$ | $^{115}$ Trp-$^{1}L_{a2}$ | 0 | 49.2 |
| $^{115}$ Trp-$^{1}L_{b}$ | $^{115}$ Trp-$^{1}L_{a}$ | 0 | 53.0 |
| $^{130}$ Trp-$^{1}L_{b7}$ | $^{130}$ Trp-$^{1}L_{a9}$ | 0 | 4.7 |
| $^{130}$ Trp-$^{1}L_{b7}$ | $^{130}$ Trp-$^{1}L_{a8}$ | 0 | 3.8 |
| $^{130}$Trp-$^{1}L_{b7}$ | $^{130}$ Trp-$^{1}L_{a7}$ | 0 | 9.0 |
| $^{130}$ Trp-$^{1}L_{b7}$ | $^{130}$ Trp-$^{1}L_{a6}$ | 0 | 8.5 |
| $^{130}$ Trp-$^{1}L_{b7}$ | $^{130}$ Trp-$^{1}L_{a5}$ | 0 | 14.2 |
| $^{130}$Trp-$^{1}L_{b7}$ | $^{130}$ Trp-$^{1}L_{a4}$ | 0 | 16.1 |
| $^{130}$ Trp-$^{1}L_{b7}$ | $^{130}$ Trp-$^{1}L_{a3}$ | 0 | 15.2 |
| $^{130}$ Trp-$^{1}L_{b7}$ | $^{130}$ Trp-$^{1}L_{a2}$ | 0 | 25.1 |
| $^{130}$ Trp-$^{1}L_{b7}$ | $^{130}$ Trp-$^{1}L_{a}$ | 0 | 27.0 |
| $^{130}$ Trp-$^{1}L_{b6}$ | $^{130}$ Trp-$^{1}L_{a9}$ | 0 | 2.5 |
| $^{130}$ Trp-$^{1}L_{b6}$ | $^{130}$ Trp-$^{1}L_{a8}$ | 0 | 2.0 |
| $^{130}$ Trp-$^{1}L_{b6}$ | $^{130}$ Trp-$^{1}L_{a7}$ | 0 | 4.7 |
| $^{130}$ Trp-$^{1}L_{b6}$ | $^{130}$ Trp-$^{1}L_{a6}$ | 0 | 4.5 |
| $^{130}$ Trp-$^{1}L_{b6}$ | $^{130}$ Trp-$^{1}L_{a5}$ | 0 | 7.5 |
| $^{130}$ Trp-$^{1}L_{b6}$ | $^{130}$ Trp-$^{1}L_{a4}$ | 0 | 8.5 |
| $^{130}$ Trp-$^{1}L_{b6}$ | $^{130}$ Trp-$^{1}L_{a3}$ | 0 | 8.0 |
| $^{130}$ Trp-$^{1}L_{b6}$ | $^{130}$ Trp-$^{1}L_{a2}$ | 0 | 13.3 |
| $^{130}$ Trp-$^{1}L_{b6}$ | $^{130}$Trp-$^{1}L_{a}$ | 0 | 14.3 |
| $^{130}$ Trp-$^{1}L_{b5}$ | $^{130}$ Trp-$^{1}L_{a9}$ | 0 | 9.2 |
| $^{130}$ Trp-$^{1}L_{b5}$ | $^{130}$ Trp-$^{1}L_{a8}$ | 0 | 7.3 |
| $^{130}$ Trp-$^{1}L_{b5}$ | $^{130}$ Trp-$^{1}L_{a7}$ | 0 | 17.5 |
| $^{130}$ Trp-$^{1}L_{b5}$ | $^{130}$ Trp-$^{1}L_{a6}$ | 0 | 16.6 |
| $^{130}$ Trp-$^{1}L_{b5}$ | $^{130}$ Trp-$^{1}L_{a5}$ | 0 | 27.6 |
| $^{130}$ Trp-$^{1}L_{b5}$ | $^{130}$ Trp-$^{1}L_{a4}$ | 0 | 31.3 |
| $^{130}$ Trp-$^{1}L_{b5}$ | $^{130}$ Trp-$^{1}L_{a3}$ | 0 | 29.5 |
| $^{130}$ Trp-$^{1}L_{b5}$ | $^{130}$ Trp-$^{1}L_{a2}$ | 0 | 48.8 |
| $^{130}$ Trp-$^{1}L_{b5}$ | $^{130}$Trp-$^{1}L_{a}$ | 0 | 52.5 |
| $^{130}$ Trp-$^{1}L_{b4}$ | $^{130}$ Trp-$^{1}L_{a9}$ | 0 | 6.4 |

| | | | |
|---|---|---|---|
| $^{130}$ Trp-$^{1}L_{b4}$ | $^{130}$ Trp-$^{1}L_{a8}$ | 0 | 5.1 |
| $^{130}$ Trp-$^{1}L_{b4}$ | $^{130}$ Trp-$^{1}L_{a7}$ | 0 | 12.2 |
| $^{130}$ Trp-$^{1}L_{b4}$ | $^{130}$ Trp-$^{1}L_{a6}$ | 0 | 11.5 |
| $^{130}$Trp-$^{1}L_{b4}$ | $^{130}$ Trp-$^{1}L_{a5}$ | 0 | 19.2 |
| $^{130}$ Trp-$^{1}L_{b4}$ | $^{130}$ Trp-$^{1}L_{a4}$ | 0 | 21.8 |
| $^{130}$ Trp-$^{1}L_{b4}$ | $^{130}$ Trp-$^{1}L_{a3}$ | 0 | 20.5 |
| $^{130}$Trp-$^{1}L_{b4}$ | $^{130}$ Trp-$^{1}L_{a2}$ | 0 | 34.0 |
| $^{130}$Trp-$^{1}L_{b4}$ | $^{130}$ Trp-$^{1}L_{a}$ | 0 | 36.6 |
| $^{130}$ Trp-$^{1}L_{b3}$ | $^{130}$ Trp-$^{1}L_{a9}$ | 0 | 12.3 |
| $^{130}$ Trp-$^{1}L_{b3}$ | $^{130}$ Trp-$^{1}L_{a8}$ | 0 | 9.8 |
| $^{130}$ Trp-$^{1}L_{b3}$ | $^{130}$ Trp-$^{1}L_{a7}$ | 0 | 23.3 |
| $^{130}$ Trp-$^{1}L_{b3}$ | $^{130}$ Trp-$^{1}L_{a6}$ | 0 | 22.1 |
| $^{130}$ Trp-$^{1}L_{b3}$ | $^{130}$ Trp-$^{1}L_{a5}$ | 0 | 36.9 |
| $^{130}$ Trp-$^{1}L_{b3}$ | $^{130}$ Trp-$^{1}L_{a4}$ | 0 | 41.8 |
| $^{130}$ Trp-$^{1}L_{b3}$ | $^{130}$ Trp-$^{1}L_{a3}$ | 0 | 39.3 |
| $^{130}$ Trp-$^{1}L_{b3}$ | $^{130}$ Trp-$^{1}L_{a2}$ | 0 | 65.1 |
| $^{130}$ Trp-$^{1}L_{b3}$ | $^{130}$ Trp-$^{1}L_{a}$ | 0 | 70.1 |
| $^{130}$ Trp-$^{1}L_{b2}$ | $^{130}$ Trp-$^{1}L_{a9}$ | 0 | 13.1 |
| $^{130}$Trp-$^{1}L_{b2}$ | $^{130}$ Trp-$^{1}L_{a8}$ | 0 | 10.5 |
| $^{130}$ Trp-$^{1}L_{b2}$ | $^{130}$ Trp-$^{1}L_{a7}$ | 0 | 24.9 |
| $^{130}$ Trp-$^{1}L_{b2}$ | $^{130}$ Trp-$^{1}L_{a6}$ | 0 | 23.6 |
| $^{130}$ Trp-$^{1}L_{b2}$ | $^{130}$ Trp-$^{1}L_{a5}$ | 0 | 39.4 |
| $^{130}$ Trp-$^{1}L_{b2}$ | $^{130}$ Trp-$^{1}L_{a4}$ | 0 | 44.6 |
| $^{130}$ Trp-$^{1}L_{b2}$ | $^{130}$Trp-$^{1}L_{a3}$ | 0 | 42.0 |
| $^{130}$ Trp-$^{1}L_{b2}$ | $^{130}$ Trp-$^{1}L_{a2}$ | 0 | 69.6 |
| $^{130}$ Trp-$^{1}L_{b2}$ | $^{130}$Trp-$^{1}L_{a}$ | 0 | 74.8 |
| $^{130}$ Trp-$^{1}L_{b}$ | $^{130}$ Trp-$^{1}L_{a9}$ | 0 | 17.3 |
| $^{130}$ Trp-$^{1}L_{b}$ | $^{130}$ Trp-$^{1}L_{a8}$ | 0 | 13.8 |
| $^{130}$ Trp-$^{1}L_{b}$ | $^{130}$ Trp-$^{1}L_{a7}$ | 0 | 32.9 |
| $^{130}$ Trp-$^{1}L_{b}$ | $^{130}$Trp-$^{1}L_{a6}$ | 0 | 31.1 |
| $^{130}$ Trp-$^{1}L_{b}$ | $^{130}$ Trp-$^{1}L_{a5}$ | 0 | 51.9 |
| $^{130}$ Trp-$^{1}L_{b}$ | $^{130}$Trp-$^{1}L_{a4}$ | 0 | 58.9 |

| | | | |
|---|---|---|---|
| $^{130}$ Trp-$^{1}L_b$ | $^{130}$ Trp-$^{1}L_{a3}$ | 0 | 55.4 |
| $^{130}$ Trp-$^{1}L_b$ | $^{130}$ Trp-$^{1}L_{a2}$ | 0 | 91.8 |
| $^{130}$ Trp-$^{1}L_b$ | $^{130}$Trp-$^{1}L_a$ | 0 | 98.7 |

Table A.4: Calculated interaction energies of electronic transitions of aromatic side chains in frame 20 of (PDB:1FW7) of the NMR structure

| Chromophore/Transitions | Chromophore/Transitions | Distance (Å) | Interaction Energy ($cm^{-1}$) |
|---|---|---|---|
| $^{56}$Phe-$^{1}L_a$ | $^{82}$Phe-$^{1}L_b$ | 8.07 | -1.3 |
| $^{56}$ Phe-$^{1}L_{b6}$ | $^{90}$ Tyr-$^{1}L_a$ | 9.22 | -2.0 |
| $^{56}$ Phe-$^{1}L_{b5}$ | $^{90}$ Tyr-$^{1}L_a$ | 9.22 | 13.0 |
| $^{56}$ Phe-$^{1}L_{b4}$ | $^{90}$ Tyr-$^{1}L_a$ | 9.22 | -40.0 |
| $^{56}$ Phe-$^{1}L_b$ | $^{90}$ Tyr-$^{1}L_a$ | 9.22 | 6.8 |
| $^{56}$ Phe-$^{1}L_{b3}$ | $^{90}$ Tyr-$^{1}L_a$ | 9.22 | 5.1 |
| $^{56}$ Phe-$^{1}L_{b2}$ | $^{90}$ Tyr-$^{1}L_a$ | 9.22 | -24.9 |
| $^{56}$ Phe-$^{1}L_a$ | $^{90}$ Tyr-$^{1}L_{b8}$ | 9.22 | -18.7 |
| $^{56}$ Phe-$^{1}L_a$ | $^{90}$ Tyr-$^{1}L_{b7}$ | 9.22 | -11.7 |
| $^{56}$ Phe-$^{1}L_a$ | $^{90}$ Tyr-$^{1}L_{b6}$ | 9.22 | -3.2 |
| $^{56}$ Phe-$^{1}L_a$ | $^{90}$ Tyr-$^{1}L_{b5}$ | 9.22 | -2.0 |
| $^{56}$Phe-$^{1}L_a$ | $^{90}$ Tyr-$^{1}L_{b4}$ | 9.22 | -5.2 |
| $^{56}$ Phe-$^{1}L_a$ | $^{90}$ Tyr-$^{1}L_{b3}$ | 9.22 | -9.2 |
| $^{56}$Phe-$^{1}L_a$ | $^{90}$ Tyr-$^{1}L_{b2}$ | 9.22 | -6.7 |
| $^{56}$Phe-$^{1}L_a$ | $^{90}$Tyr-$^{1}L_b$ | 9.22 | -15.0 |
| $^{56}$ Phe-$^{1}L_a$ | $^{56}$ Phe-$^{1}L_b$ | 0 | 4.0 |
| $^{56}$ Phe-$^{1}L_a$ | $^{56}$ Phe-$^{1}L_{b3}$ | 0 | 3.3 |
| $^{56}$ Phe-$^{1}L_{b2}$ | $^{56}$ Phe-$^{1}L_a$ | 0 | 3.8 |
| $^{56}$ Phe-$^{1}L_{b4}$ | $^{56}$ Phe-$^{1}L_a$ | 0 | 1.9 |
| $^{56}$ Phe-$^{1}L_a$ | $^{56}$ Phe-$^{1}L_{b5}$ | 0 | 1.5 |
| $^{56}$ Phe-$^{1}L_{b6}$ | $^{56}$ Phe-$^{1}L_a$ | 0 | 3.0 |
| $^{56}$ Phe-$^{1}L_a$ | $^{56}$ Phe-$^{1}L_{b6}$ | 0 | -1.1 |

| | | | |
|---|---|---|---|
| $^{56}$ Phe-$^{1}L_{b5}$ | $^{56}$ Phe-$^{1}L_{a}$ | 0 | -1.8 |
| $^{56}$ Phe-$^{1}L_{a}$ | $^{56}$ Phe-$^{1}L_{b4}$ | 0 | -2.1 |
| $^{56}$Phe-$^{1}L_{a}$ | $^{56}$ Phe-$^{1}L_{b2}$ | 0 | -2.6 |
| $^{56}$ Phe-$^{1}L_{b3}$ | $^{56}$ Phe-$^{1}L_{a}$ | 0 | -2.7 |
| $^{56}$ Phe-$^{1}L_{b}$ | $^{56}$ Phe-$^{1}L_{a}$ | 0 | -3.7 |
| $^{56}$ Phe-$^{1}L_{b}$ | $^{78}$ Tyr-$^{1}L_{b2}$ | 8.91 | -15.0 |
| $^{56}$ Phe-$^{1}L_{b}$ | $^{78}$ Tyr-$^{1}L_{a}$ | 8.91 | 10.0 |
| $^{56}$ Phe-$^{1}L_{a}$ | $^{78}$Tyr-$^{1}L_{b2}$ | 8.91 | 7.1 |
| $^{56}$ Phe-$^{1}L_{a}$ | $^{78}$ Tyr-$^{1}L_{b}$ | 8.91 | 1.0 |
| $^{56}$ Phe-$^{1}L_{b}$ | $^{71}$ Trp-$^{1}L_{a}$ | 8.96 | -16.5 |
| $^{56}$ Phe-$^{1}L_{b}$ | $^{78}$ Tyr-$^{1}L_{a}$ | 8.91 | -1.1 |
| $^{82}$ Phe-$^{1}L_{a}$ | $^{78}$Tyr-$^{1}L_{b8}$ | 5.21 | 24.7 |
| $^{82}$Phe-$^{1}L_{a}$ | $^{78}$Tyr-$^{1}L_{b7}$ | 5.21 | 4.7 |
| $^{82}$Phe-$^{1}L_{a}$ | $^{78}$ Tyr-$^{1}L_{b6}$ | 5.21 | 39.4 |
| $^{82}$Phe-$^{1}L_{a}$ | $^{78}$ Tyr-$^{1}L_{b5}$ | 5.21 | 6.0 |
| $^{82}$ Phe-$^{1}L_{a}$ | $^{78}$ Tyr-$^{1}L_{b4}$ | 5.21 | 15.1 |
| $^{82}$Phe-$^{1}L_{a}$ | $^{78}$Tyr-$^{1}L_{b3}$ | 5.21 | 5.7 |
| $^{82}$ Phe-$^{1}L_{a}$ | $^{78}$ Tyr-$^{1}L_{b2}$ | 5.21 | 19.5 |
| $^{82}$ Phe-$^{1}L_{a}$ | $^{78}$ Tyr-$^{1}L_{b}$ | 5.21 | 17.5 |
| $^{82}$ Phe-$^{1}L_{b3}$ | $^{94}$ Trp-$^{1}L_{a3}$ | 6.67 | -33.7 |
| $^{82}$ Phe-$^{1}L_{b3}$ | $^{94}$ Trp-$^{1}L_{a2}$ | 6.67 | -6.1 |
| $^{82}$ Phe-$^{1}L_{b3}$ | $^{94}$ Trp-$^{1}L_{a}$ | 6.67 | -13.9 |
| $^{82}$ Phe-$^{1}L_{b2}$ | $^{94}$ Trp-$^{1}L_{b2}$ | 6.67 | -10.0 |
| $^{82}$ Phe-$^{1}L_{b2}$ | $^{94}$ Trp-$^{1}L_{b}$ | 6.67 | -1.4 |
| $^{82}$ Phe-$^{1}L_{b2}$ | $^{94}$ Trp-$^{1}L_{a9}$ | 6.67 | 29.0 |
| $^{82}$ Phe-$^{1}L_{b2}$ | $^{94}$ Trp-$^{1}L_{a7}$ | 6.67 | 10.8 |
| $^{82}$ Phe-$^{1}L_{b2}$ | $^{94}$ Trp-$^{1}L_{a6}$ | 6.67 | 45.8 |
| $^{82}$ Phe-$^{1}L_{b2}$ | $^{94}$ Trp-$^{1}L_{a5}$ | 6.67 | 3.1 |
| $^{82}$ Phe-$^{1}L_{b2}$ | $^{94}$ Trp-$^{1}L_{a4}$ | 6.67 | 43.3 |
| $^{82}$ Phe-$^{1}L_{b2}$ | $^{94}$ Trp-$^{1}L_{a3}$ | 6.67 | 3.9 |
| $^{82}$ Phe-$^{1}L_{b2}$ | $^{94}$ Trp-$^{1}L_{a2}$ | 6.67 | 35.2 |
| $^{82}$ Phe-$^{1}L_{b2}$ | $^{94}$ Trp-$^{1}L_{a}$ | 6.67 | 44.7 |

| | | | |
|---|---|---|---|
| $^{82}$ Phe-$^{1}L_{b}$ | $^{94}$ Trp-$^{1}L_{b5}$ | 6.67 | 1.0 |
| $^{82}$ Phe-$^{1}L_{b}$ | $^{94}$ Trp-$^{1}L_{b3}$ | 6.67 | 11.3 |
| $^{82}$ Phe-$^{1}L_{b}$ | $^{94}$ Trp-$^{1}L_{b2}$ | 6.67 | 1.5 |
| $^{82}$ Phe-$^{1}L_{b}$ | $^{94}$ Trp-$^{1}L_{b}$ | 6.67 | 21.1 |
| $^{82}$ Phe-$^{1}L_{b}$ | $^{94}$ Trp-$^{1}L_{a9}$ | 6.67 | -1.9 |
| $^{82}$ Phe-$^{1}L_{b}$ | $^{94}$ Trp-$^{1}L_{a8}$ | 6.67 | -1.1 |
| $^{82}$ Phe-$^{1}L_{b}$ | $^{94}$ Trp-$^{1}L_{a7}$ | 6.67 | -32.1 |
| $^{82}$ Phe-$^{1}L_{b}$ | $^{94}$ Trp-$^{1}L_{a6}$ | 6.67 | -12.0 |
| $^{82}$ Phe-$^{1}L_{b}$ | $^{94}$ Trp-$^{1}L_{a5}$ | 6.67 | -4.5 |
| $^{82}$ Phe-$^{1}L_{b}$ | $^{94}$ Trp-$^{1}L_{a4}$ | 6.67 | -27.7 |
| $^{82}$ Phe-$^{1}L_{b}$ | $^{94}$ Trp-$^{1}L_{a3}$ | 6.67 | -4.5 |
| $^{82}$ Phe-$^{1}L_{b}$ | $^{94}$ Trp-$^{1}L_{a2}$ | 6.67 | -7.4 |
| $^{82}$ Phe-$^{1}L_{b}$ | $^{94}$ Trp-$^{1}L_{a}$ | 6.67 | -8.0 |
| $^{82}$ Phe-$^{1}L_{a}$ | $^{94}$Trp-$^{1}L_{b7}$ | 6.67 | -5.1 |
| $^{82}$ Phe-$^{1}L_{a}$ | $^{94}$Trp-$^{1}L_{b6}$ | 6.67 | -2.7 |
| $^{82}$Phe-$^{1}L_{a}$ | $^{94}$ Trp-$^{1}L_{b4}$ | 6.67 | -16.9 |
| $^{13}$ Tyr-$^{1}L_{b8}$ | $^{13}$ Tyr-$^{1}L_{a}$ | 0 | -3.3 |
| $^{13}$ Tyr-$^{1}L_{b7}$ | $^{13}$ Tyr-$^{1}L_{a}$ | 0 | -3.3 |
| $^{13}$ Tyr-$^{1}L_{b6}$ | $^{13}$ Tyr-$^{1}L_{a}$ | 0 | -19.7 |
| $^{13}$ Tyr-$^{1}L_{b5}$ | $^{13}$ Tyr-$^{1}L_{a}$ | 0 | -4.3 |
| $^{13}$ Tyr-$^{1}L_{b3}$ | $^{13}$ Tyr-$^{1}L_{a}$ | 0 | -4.1 |
| $^{13}$ Tyr-$^{1}L_{b8}$ | $^{24}$ Tyr-$^{1}L_{b2}$ | 6.81 | -41.3 |
| $^{13}$ Tyr-$^{1}L_{b7}$ | $^{24}$ Tyr-$^{1}L_{b}$ | 6.81 | -11.1 |
| $^{13}$ Tyr-$^{1}L_{b7}$ | $^{24}$ Tyr-$^{1}L_{b2}$ | 6.81 | -1.5 |
| $^{13}$ Tyr-$^{1}L_{b6}$ | $^{24}$ Tyr-$^{1}L_{b}$ | 6.81 | -21.1 |
| $^{13}$ Tyr-$^{1}L_{b6}$ | $^{24}$ Tyr-$^{1}L_{b2}$ | 6.81 | -2.4 |
| $^{13}$ Tyr-$^{1}L_{b5}$ | $^{24}$ Tyr-$^{1}L_{b4}$ | 6.81 | -1.5 |
| $^{13}$ Tyr-$^{1}L_{b5}$ | $^{24}$ Tyr-$^{1}L_{b}$ | 6.81 | -41.4 |
| $^{13}$ Tyr-$^{1}L_{b5}$ | $^{24}$ Tyr-$^{1}L_{b2}$ | 6.81 | -10.5 |
| $^{13}$ Tyr-$^{1}L_{b4}$ | $^{24}$ Tyr-$^{1}L_{b5}$ | 6.81 | -15.5 |
| $^{13}$ Tyr-$^{1}L_{b4}$ | $^{24}$ Tyr-$^{1}L_{b4}$ | 6.81 | -3.0 |
| $^{13}$ Tyr-$^{1}L_{b4}$ | $^{24}$ Tyr-$^{1}L_{b}$ | 6.81 | -3.5 |

| | | | |
|---|---|---|---|
| $^{13}$ Tyr-$^{1}L_{b4}$ | $^{24}$ Tyr-$^{1}L_{b2}$ | 6.81 | -33.9 |
| $^{13}$ Tyr-$^{1}L_{b4}$ | $^{24}$ Tyr-$^{1}L_{a}$ | 6.81 | 2.2 |
| $^{13}$ Tyr-$^{1}L_{b3}$ | $^{24}$ Tyr-$^{1}L_{b4}$ | 6.81 | -19.1 |
| $^{13}$ Tyr-$^{1}L_{b3}$ | $^{24}$ Tyr-$^{1}L_{b}$ | 6.81 | -41.3 |
| $^{13}$ Tyr-$^{1}L_{b2}$ | $^{24}$ Tyr-$^{1}L_{b6}$ | 6.81 | -2.4 |
| $^{13}$ Tyr-$^{1}L_{b2}$ | $^{24}$ Tyr-$^{1}L_{b5}$ | 6.81 | -1.5 |
| $^{13}$ Tyr-$^{1}L_{b2}$ | $^{24}$ Tyr-$^{1}L_{b4}$ | 6.81 | -48.9 |
| $^{13}$ Tyr-$^{1}L_{b2}$ | $^{24}$ Tyr-$^{1}L_{b}$ | 6.81 | -15.5 |
| $^{13}$ Tyr-$^{1}L_{b2}$ | $^{24}$ Tyr-$^{1}L_{b3}$ | 6.81 | -1.4 |
| $^{13}$ Tyr-$^{1}L_{b2}$ | $^{24}$ Tyr-$^{1}L_{b2}$ | 6.81 | -25.0 |
| $^{13}$ Tyr-$^{1}L_{b2}$ | $^{24}$ Tyr-$^{1}L_{a}$ | 6.81 | 2.9 |
| $^{13}$ Tyr-$^{1}L_{b}$ | $^{24}$ Tyr-$^{1}L_{b8}$ | 6.81 | -10.9 |
| $^{13}$ Tyr-$^{1}L_{b}$ | $^{24}$ Tyr-$^{1}L_{b6}$ | 6.81 | -24.1 |
| $^{13}$ Tyr-$^{1}L_{b}$ | $^{24}$ Tyr-$^{1}L_{b5}$ | 6.81 | -18.0 |
| $^{13}$ Tyr-$^{1}L_{b}$ | $^{24}$ Tyr-$^{1}L_{b4}$ | 6.81 | -31.0 |
| $^{13}$ Tyr-$^{1}L_{b}$ | $^{24}$ Tyr-$^{1}L_{b}$ | 6.81 | -44.0 |
| $^{13}$ Tyr-$^{1}L_{b}$ | $^{24}$ Tyr-$^{1}L_{b3}$ | 6.81 | -11.3 |
| $^{13}$ Tyr-$^{1}L_{b}$ | $^{24}$ Tyr-$^{1}L_{b2}$ | 6.81 | -4.2 |
| $^{13}$ Tyr-$^{1}L_{b}$ | $^{24}$ Tyr-$^{1}L_{a}$ | 6.81 | 32.6 |
| $^{13}$ Tyr-$^{1}L_{a}$ | $^{24}$ Tyr-$^{1}L_{b6}$ | 6.81 | 1.0 |
| $^{13}$ Tyr-$^{1}L_{a}$ | $^{24}$ Tyr-$^{1}L_{b4}$ | 6.81 | 10.6 |
| $^{13}$ Tyr-$^{1}L_{a}$ | $^{24}$ Tyr-$^{1}L_{b}$ | 6.81 | 41.8 |
| $^{13}$ Tyr-$^{1}L_{a}$ | $^{24}$ Tyr-$^{1}L_{b2}$ | 6.81 | 2.1 |
| $^{13}$ Tyr-$^{1}L_{b2}$ | $^{78}$ Tyr-$^{1}L_{a}$ | 13.05 | 1.0 |
| $^{13}$ Tyr-$^{1}L_{a}$ | $^{78}$ Tyr-$^{1}L_{b4}$ | 13.05 | -21.2 |
| $^{13}$ Tyr-$^{1}L_{a}$ | $^{78}$ Tyr-$^{1}L_{b2}$ | 13.05 | -10.7 |
| $^{13}$ Tyr-$^{1}L_{a}$ | $^{78}$ Tyr-$^{1}L_{b}$ | 13.05 | -1.4 |
| $^{13}$ Tyr-$^{1}L_{a}$ | $^{103}$ Tyr-$^{1}L_{b8}$ | 9.11 | 1.2 |
| $^{13}$ Tyr-$^{1}L_{a}$ | $^{103}$ Tyr-$^{1}L_{b7}$ | 9.11 | 1.2 |
| $^{13}$ Tyr-$^{1}L_{a}$ | $^{103}$ Tyr-$^{1}L_{b6}$ | 9.11 | 2.5 |
| $^{13}$ Tyr-$^{1}L_{a}$ | $^{103}$ Tyr-$^{1}L_{b5}$ | 9.11 | 10.9 |
| $^{13}$ Tyr-$^{1}L_{a}$ | $^{103}$ Tyr-$^{1}L_{b4}$ | 9.11 | 40.1 |

| | | | |
|---|---|---|---|
| $^{13}$ Tyr-$^{1}L_{a}$ | $^{103}$ Tyr-$^{1}L_{b3}$ | 9.11 | 1.5 |
| $^{13}$ Tyr-$^{1}L_{a}$ | $^{103}$ Tyr-$^{1}L_{b2}$ | 9.11 | 45.3 |
| $^{13}$ Tyr-$^{1}L_{a}$ | $^{103}$ Tyr-$^{1}L_{b}$ | 9.11 | 24.8 |
| $^{13}$ Tyr-$^{1}L_{b2}$ | $^{94}$ Trp-$^{1}L_{b2}$ | 13.57 | -1.8 |
| $^{13}$ Tyr-$^{1}L_{b2}$ | $^{94}$ Trp-$^{1}L_{b}$ | 13.57 | -2.5 |
| $^{13}$ Tyr-$^{1}L_{b2}$ | $^{94}$ Trp-$^{1}L_{a2}$ | 13.57 | -1.0 |
| $^{13}$ Tyr-$^{1}L_{b2}$ | $^{94}$ Trp-$^{1}L_{a}$ | 13.57 | -14.1 |
| $^{13}$ Tyr-$^{1}L_{b}$ | $^{94}$ Trp-$^{1}L_{b5}$ | 13.57 | -1.1 |
| $^{13}$ Tyr-$^{1}L_{b}$ | $^{94}$ Trp-$^{1}L_{b3}$ | 13.57 | -14.5 |
| $^{13}$ Tyr-$^{1}L_{b}$ | $^{94}$ Trp-$^{1}L_{b2}$ | 13.57 | -48.6 |
| $^{13}$ Tyr-$^{1}L_{b}$ | $^{94}$ Trp-$^{1}L_{b}$ | 13.57 | -27.2 |
| $^{13}$ Tyr-$^{1}L_{b}$ | $^{94}$ Trp-$^{1}L_{a}$ | 13.57 | -10.0 |
| $^{13}$ Tyr-$^{1}L_{a}$ | $^{94}$ Trp-$^{1}L_{b7}$ | 13.57 | -1.0 |
| $^{13}$ Tyr-$^{1}L_{a}$ | $^{94}$ Trp-$^{1}L_{b5}$ | 13.57 | -29.0 |
| $^{13}$ Tyr-$^{1}L_{a}$ | $^{94}$Trp-$^{1}L_{b4}$ | 13.57 | -1.4 |
| $^{13}$ Tyr-$^{1}L_{a}$ | $^{94}$ Trp-$^{1}L_{b3}$ | 13.57 | -29.7 |
| $^{13}$ Tyr-$^{1}L_{a}$ | $^{94}$ Trp-$^{1}L_{b2}$ | 13.57 | -20.9 |
| $^{13}$ Tyr-$^{1}L_{a}$ | $^{94}$ Trp-$^{1}L_{b}$ | 13.57 | -3.8 |
| $^{24}$ Tyr-$^{1}L_{b6}$ | $^{24}$ Tyr-$^{1}L_{a}$ | 0 | 15. 9 |
| $^{24}$ Tyr-$^{1}L_{b5}$ | $^{24}$ Tyr-$^{1}L_{a}$ | 0 | 1.2 |
| $^{24}$ Tyr-$^{1}L_{b4}$ | $^{24}$ Tyr-$^{1}L_{a}$ | 0 | 38.0 |
| $^{24}$ Tyr-$^{1}L_{b}$ | $^{24}$ Tyr-$^{1}L_{a}$ | 0 | 13.5 |
| $^{24}$ Tyr-$^{1}L_{b3}$ | $^{24}$ Tyr-$^{1}L_{a}$ | 0 | 11.1 |
| $^{24}$ Tyr-$^{1}L_{b2}$ | $^{24}$ Tyr-$^{1}L_{a}$ | 0 | 43.0 |
| $^{24}$ Tyr-$^{1}L_{b8}$ | $^{78}$ Tyr-$^{1}L_{a}$ | 7.38 | 1.0 |
| $^{24}$ Tyr-$^{1}L_{b7}$ | $^{78}$ Tyr-$^{1}L_{a}$ | 7.38 | 31.0 |
| $^{24}$ Tyr-$^{1}L_{b6}$ | $^{78}$ Tyr-$^{1}L_{a}$ | 7.38 | 25.0 |
| $^{24}$ Tyr-$^{1}L_{b5}$ | $^{78}$ Tyr-$^{1}L_{a}$ | 7.38 | 1.5 |
| $^{24}$ Tyr-$^{1}L_{b4}$ | $^{78}$ Tyr-$^{1}L_{a}$ | 7.38 | 3.2 |
| $^{24}$ Tyr-$^{1}L_{b}$ | $^{78}$ Tyr-$^{1}L_{a}$ | 7.38 | 38.7 |
| $^{24}$ Tyr-$^{1}L_{b3}$ | $^{78}$ Tyr-$^{1}L_{a}$ | 7.38 | 41.2 |
| $^{24}$ Tyr-$^{1}L_{b2}$ | $^{78}$ Tyr-$^{1}L_{a}$ | 7.38 | 24.2 |

| | | | |
|---|---|---|---|
| $^{24}$ Tyr-$^{1}L_{a}$ | $^{78}$ Tyr-$^{1}L_{b6}$ | 7.38 | -1.3 |
| $^{24}$ Tyr-$^{1}L_{a}$ | $^{78}$ Tyr-$^{1}L_{b4}$ | 7.38 | -22.1 |
| $^{24}$ Tyr-$^{1}L_{a}$ | $^{78}$ Tyr-$^{1}L_{b2}$ | 7.38 | -35.7 |
| $^{24}$ Tyr-$^{1}L_{a}$ | $^{78}$ Tyr-$^{1}L_{b}$ | 7.38 | -20.4 |
| $^{24}$ Tyr-$^{1}L_{b6}$ | $^{90}$ Tyr-$^{1}L_{a}$ | 12.07 | -12.1 |
| $^{24}$ Tyr-$^{1}L_{b4}$ | $^{90}$ Tyr-$^{1}L_{a}$ | 12.07 | -1.8 |
| $^{24}$ Tyr-$^{1}L_{b}$ | $^{90}$ Tyr-$^{1}L_{a}$ | 12.07 | -42.1 |
| $^{24}$ Tyr-$^{1}L_{b2}$ | $^{90}$ Tyr-$^{1}L_{a}$ | 12.07 | -2.4 |
| $^{24}$ Tyr-$^{1}L_{a}$ | $^{90}$ Tyr-$^{1}L_{b6}$ | 12.07 | -16.8 |
| $^{24}$ Tyr-$^{1}L_{a}$ | $^{90}$ Tyr-$^{1}L_{b5}$ | 12.07 | -1.1 |
| $^{24}$ Tyr-$^{1}L_{a}$ | $^{90}$ Tyr-$^{1}L_{b4}$ | 12.07 | -28.9 |
| $^{24}$ Tyr-$^{1}L_{a}$ | $^{90}$ Tyr-$^{1}L_{b3}$ | 12.07 | -1.1 |
| $^{24}$ Tyr-$^{1}L_{a}$ | $^{90}$ Tyr-$^{1}L_{b2}$ | 12.07 | -3.8 |
| $^{24}$ Tyr-$^{1}L_{a}$ | $^{90}$ Tyr-$^{1}L_{b}$ | 12.07 | -32.4 |
| $^{24}$ Tyr-$^{1}L_{b}$ | $^{103}$ Tyr-$^{1}L_{b2}$ | 13.84 | 11.1 |
| $^{24}$ Tyr-$^{1}L_{b}$ | $^{103}$ Tyr-$^{1}L_{b}$ | 13.84 | 35.0 |
| $^{24}$ Tyr-$^{1}L_{b2}$ | $^{103}$ Tyr-$^{1}L_{b2}$ | 13.84 | 17.2 |
| $^{24}$ Tyr-$^{1}L_{b2}$ | $^{103}$ Tyr-$^{1}L_{b}$ | 13.84 | 10.1 |
| $^{24}$ Tyr-$^{1}L_{a}$ | $^{94}$ Trp-$^{1}L_{b5}$ | 16.02 | -1.0 |
| $^{24}$ Tyr-$^{1}L_{a}$ | $^{94}$ Trp-$^{1}L_{b3}$ | 16.02 | -1.4 |
| $^{24}$ Tyr-$^{1}L_{a}$ | $^{94}$ Trp-$^{1}L_{b2}$ | 16.02 | -12.5 |
| $^{24}$ Tyr-$^{1}L_{a}$ | $^{94}$ Trp-$^{1}L_{b}$ | 16.02 | -2.0 |
| $^{78}$ Tyr-$^{1}L_{b8}$ | $^{78}$ Tyr-$^{1}L_{a}$ | 0 | -35.9 |
| $^{78}$ Tyr-$^{1}L_{b7}$ | $^{78}$ Tyr-$^{1}L_{a}$ | 0 | -5.9 |
| $^{78}$ Tyr-$^{1}L_{b5}$ | $^{78}$ Tyr-$^{1}L_{a}$ | 0 | -47.6 |
| $^{78}$ Tyr-$^{1}L_{b3}$ | $^{78}$ Tyr-$^{1}L_{a}$ | 0 | -27.1 |
| $^{78}$ Tyr-$^{1}L_{b4}$ | $^{90}$ Tyr-$^{1}L_{a}$ | 12.44 | 10.1 |
| $^{78}$ Tyr-$^{1}L_{b2}$ | $^{90}$ Tyr-$^{1}L_{a}$ | 12.44 | 1.5 |
| $^{78}$ Tyr-$^{1}L_{b}$ | $^{90}$ Tyr-$^{1}L_{a}$ | 12.44 | 1.3 |
| $^{78}$ Tyr-$^{1}L_{a}$ | $^{90}$Tyr-$^{1}L_{b4}$ | 12.44 | -12.2 |
| $^{78}$ TYR-1La | $^{90}$ Tyr-$^{1}L_{b2}$ | 12.44 | -14.5 |
| $^{78}$ TYR-1La | $^{90}$ Tyr-$^{1}L_{b}$ | 12.44 | -1.4 |

| | | | |
|---|---|---|---|
| $^{78}$ Tyr-$^{1}L_{b4}$ | $^{103}$ Tyr-$^{1}L_{a}$ | 13.75 | -1.0 |
| $^{78}$ Tyr-$^{1}L_{b2}$ | $^{103}$ Tyr-$^{1}L_{a}$ | 13.75 | -1.5 |
| $^{78}$ Tyr-$^{1}L_{b}$ | $^{103}$ Tyr-$^{1}L_{a}$ | 13.75 | -41.1 |
| $^{78}$ Tyr-$^{1}L_{b4}$ | $^{71}$ Trp-$^{1}L_{a}$ | 14.65 | 13.0 |
| $^{78}$ Tyr-$^{1}L_{b2}$ | $^{71}$ Trp-$^{1}L_{a}$ | 14.65 | 1.5 |
| $^{78}$ Tyr-$^{1}L_{b2}$ | $^{71}$ Trp-$^{1}L_{a2}$ | 14.65 | 11.2 |
| $^{78}$ Tyr-$^{1}L_{b}$ | $^{71}$ Trp-$^{1}L_{a}$ | 14.65 | 1.1 |
| $^{78}$ Tyr-$^{1}L_{b}$ | $^{71}$ Trp-$^{1}L_{a2}$ | 14.65 | 1.0 |
| $^{90}$ Tyr-$^{1}L_{b8}$ | $^{90}$ Tyr-$^{1}L_{a}$ | 0 | 48.5 |
| $^{90}$ Tyr-$^{1}L_{b7}$ | $^{90}$ Tyr-$^{1}L_{a}$ | 0 | 22.5 |
| $^{90}$ Tyr-$^{1}L_{b6}$ | $^{90}$ Tyr-$^{1}L_{a}$ | 0 | 16.6 |
| $^{90}$ Tyr-$^{1}L_{b5}$ | $^{90}$ Tyr-$^{1}L_{a}$ | 0 | 13.6 |
| $^{90}$ Tyr-$^{1}L_{b4}$ | $^{90}$ Tyr-$^{1}L_{a}$ | 0 | 24.8 |
| $^{90}$ Tyr-$^{1}L_{b3}$ | $^{90}$ Tyr-$^{1}L_{a}$ | 0 | 13.0 |
| $^{90}$ Tyr-$^{1}L_{b2}$ | $^{90}$ Tyr-$^{1}L_{a}$ | 0 | 34.4 |
| $^{90}$ Tyr-$^{1}L_{b}$ | $^{90}$ Tyr-$^{1}L_{a}$ | 0 | 16.8 |
| $^{90}$ Tyr-$^{1}L_{b4}$ | $^{103}$ Tyr-$^{1}L_{b2}$ | 10.49 | 10.2 |
| $^{90}$ Tyr-$^{1}L_{b4}$ | $^{103}$ Tyr-$^{1}L_{b}$ | 10.49 | 1.1 |
| $^{90}$ Tyr-$^{1}L_{b4}$ | $^{103}$ Tyr-$^{1}L_{a}$ | 10.49 | -1.0 |
| $^{90}$ Tyr-$^{1}L_{b2}$ | $^{103}$ Tyr-$^{1}L_{b4}$ | 10.49 | 21.2 |
| $^{90}$ Tyr-$^{1}L_{b2}$ | $^{103}$ Tyr-$^{1}L_{b2}$ | 10.49 | 1.6 |
| $^{90}$ Tyr-$^{1}L_{b2}$ | $^{103}$ Tyr-$^{1}L_{b}$ | 10.49 | 14.4 |
| $^{90}$ Tyr-$^{1}L_{b2}$ | $^{103}$ Tyr-$^{1}L_{a}$ | 10.49 | -1.3 |
| $^{90}$ Tyr-$^{1}L_{b}$ | $^{103}$ Tyr-$^{1}L_{b4}$ | 10.49 | 1.1 |
| $^{90}$ Tyr-$^{1}L_{b}$ | $^{103}$ Tyr-$^{1}L_{b2}$ | 10.49 | 16.4 |
| $^{90}$ Tyr-$^{1}L_{b}$ | $^{103}$ Tyr-$^{1}L_{b}$ | 10.49 | 1.2 |
| $^{90}$ Tyr-$^{1}L_{b}$ | $^{103}$ Tyr-$^{1}L_{a}$ | 10.49 | -12.1 |
| $^{103}$ Tyr-$^{1}L_{b4}$ | $^{103}$ Tyr-$^{1}L_{a}$ | 0 | 10.1 |
| $^{103}$ Tyr-$^{1}L_{b2}$ | $^{103}$ Tyr-$^{1}L_{a}$ | 0 | 21.2 |
| $^{103}$ Tyr-$^{1}L_{b}$ | $^{103}$ Tyr-$^{1}L_{a}$ | 0 | 31.3 |
| $^{78}$ Tyr-$^{1}L_{b8}$ | $^{94}$ Trp-$^{1}L_{a2}$ | 9.16 | -1.2 |
| $^{78}$ Tyr-$^{1}L_{b8}$ | $^{94}$ Trp-$^{1}L_{a}$ | 9.16 | -41.3 |

| | | | |
|---|---|---|---|
| $^{78}$ Tyr-$^{1}L_{b7}$ | $^{94}$ Trp-$^{1}L_{a2}$ | 9.16 | -1.2 |
| $^{78}$ Tyr-$^{1}L_{b7}$ | $^{94}$ Trp-$^{1}L_{a}$ | 9.16 | -1.3 |
| $^{78}$ Tyr-$^{1}L_{b6}$ | $^{94}$ Trp-$^{1}L_{a5}$ | 9.16 | -1.4 |
| $^{78}$ Tyr-$^{1}L_{b6}$ | $^{94}$ Trp-$^{1}L_{a4}$ | 9.16 | -1.6 |
| $^{78}$ Tyr-$^{1}L_{b6}$ | $^{94}$ Trp-$^{1}L_{a3}$ | 9.16 | -1.5 |
| $^{78}$ Tyr-$^{1}L_{b6}$ | $^{94}$ Trp-$^{1}L_{a2}$ | 9.16 | -22.5 |
| $^{78}$ Tyr-$^{1}L_{b6}$ | $^{94}$ Trp-$^{1}L_{a}$ | 9.16 | -25.7 |
| $^{78}$ Tyr-$^{1}L_{b5}$ | $^{94}$ Trp-$^{1}L_{a4}$ | 9.16 | -1.0 |
| $^{78}$ Tyr-$^{1}L_{b5}$ | $^{94}$ Trp-$^{1}L_{a3}$ | 9.16 | -10.0 |
| $^{78}$ Tyr-$^{1}L_{b5}$ | $^{94}$ Trp-$^{1}L_{a2}$ | 9.16 | -1.6 |
| $^{78}$ Tyr-$^{1}L_{b5}$ | $^{94}$ Trp-$^{1}L_{a}$ | 9.16 | -33.7 |
| $^{78}$ Tyr-$^{1}L_{b4}$ | $^{94}$ Trp-$^{1}L_{a7}$ | 9.16 | -1.4 |
| $^{78}$ Tyr-$^{1}L_{b4}$ | $^{94}$ Trp-$^{1}L_{a6}$ | 9.16 | -1.4 |
| $^{78}$ Tyr-$^{1}L_{b4}$ | $^{94}$ Trp-$^{1}L_{a5}$ | 9.16 | -20.3 |
| $^{78}$ Tyr-$^{1}L_{b4}$ | $^{94}$ Trp-$^{1}L_{a4}$ | 9.16 | -2.1 |
| $^{78}$ Tyr-$^{1}L_{b4}$ | $^{94}$ Trp-$^{1}L_{a3}$ | 9.16 | -2.5 |
| $^{78}$ Tyr-$^{1}L_{b4}$ | $^{94}$ Trp-$^{1}L_{a2}$ | 9.16 | -49.1 |
| $^{78}$ Tyr-$^{1}L_{b4}$ | $^{94}$ Trp-$^{1}L_{a}$ | 9.16 | -40.4 |
| $^{78}$ Tyr-$^{1}L_{b3}$ | $^{94}$ Trp-$^{1}L_{a4}$ | 9.16 | -1.0 |
| $^{78}$ Tyr-$^{1}L_{b3}$ | $^{94}$ Trp-$^{1}L_{a2}$ | 9.16 | -10.5 |
| $^{78}$ Tyr-$^{1}L_{b3}$ | $^{94}$ Trp-$^{1}L_{a}$ | 9.16 | -41.6 |
| $^{78}$ Tyr-$^{1}L_{b2}$ | $^{94}$ Trp-$^{1}L_{a9}$ | 9.16 | -12.0 |
| $^{78}$ Tyr-$^{1}L_{b2}$ | $^{94}$ Trp-$^{1}L_{a7}$ | 9.16 | -1.9 |
| $^{78}$ Tyr-$^{1}L_{b2}$ | $^{94}$ Trp-$^{1}L_{a6}$ | 9.16 | -13.8 |
| $^{78}$ Tyr-$^{1}L_{b2}$ | $^{94}$ Trp-$^{1}L_{a5}$ | 9.16 | -3.0 |
| $^{78}$ Tyr-$^{1}L_{b2}$ | $^{94}$ Trp-$^{1}L_{a4}$ | 9.16 | -3.4 |
| $^{78}$ Tyr-$^{1}L_{b2}$ | $^{94}$ Trp-$^{1}L_{a3}$ | 9.16 | -31.2 |
| $^{78}$ Tyr-$^{1}L_{b2}$ | $^{94}$ Trp-$^{1}L_{a2}$ | 9.16 | -5.3 |
| $^{78}$ Tyr-$^{1}L_{b2}$ | $^{94}$ Trp-$^{1}L_{a}$ | 9.16 | -5.7 |
| $^{78}$ Tyr-$^{1}L_{b}$ | $^{94}$ Trp-$^{1}L_{a7}$ | 9.16 | -1.7 |
| $^{78}$ Tyr-$^{1}L_{b}$ | $^{94}$ Trp-$^{1}L_{a6}$ | 9.16 | -10.7 |
| $^{78}$ Tyr-$^{1}L_{b}$ | $^{94}$ Trp-$^{1}L_{a5}$ | 9.16 | -2.7 |

| | | | |
|---|---|---|---|
| $^{78}$ Tyr-$^{1}L_{b}$ | $^{94}$ Trp-$^{1}L_{a4}$ | 9.16 | -30.0 |
| $^{78}$ Tyr-$^{1}L_{b}$ | $^{94}$ Trp-$^{1}L_{a3}$ | 9.16 | -2.9 |
| $^{78}$ Tyr-$^{1}L_{b}$ | $^{94}$ Trp-$^{1}L_{a2}$ | 9.16 | -14.8 |
| $^{78}$ Tyr-$^{1}L_{b}$ | $^{94}$ Trp-$^{1}L_{a}$ | 9.16 | -5.1 |
| $^{78}$ Tyr-$^{1}L_{a}$ | $^{94}$ Trp-$^{1}L_{b7}$ | 9.16 | -23.7 |
| $^{78}$ Tyr-$^{1}L_{a}$ | $^{94}$ Trp-$^{1}L_{b6}$ | 9.16 | -2.0 |
| $^{78}$ Tyr-$^{1}L_{a}$ | $^{94}$ Trp-$^{1}L_{b5}$ | 9.16 | -37.3 |
| $^{78}$ Tyr-$^{1}L_{a}$ | $^{94}$ Trp-$^{1}L_{b4}$ | 9.16 | -15.1 |
| $^{78}$ Tyr-$^{1}L_{a}$ | $^{94}$ Trp-$^{1}L_{b3}$ | 9.16 | -9.8 |
| $^{71}$ Trp-$^{1}L_{b}$ | $^{71}$ Trp-$^{1}L_{a8}$ | 0 | -9.4 |
| $^{71}$ Trp-$^{1}L_{b7}$ | $^{71}$ Trp-$^{1}L_{a9}$ | 0 | -43.2 |
| $^{71}$ Trp-$^{1}L_{b7}$ | $^{71}$ Trp-$^{1}L_{a8}$ | 0 | -2.5 |
| $^{71}$ Trp-$^{1}L_{b7}$ | $^{71}$ Trp-$^{1}L_{a7}$ | 0 | -36.1 |
| $^{71}$ Trp-$^{1}L_{b7}$ | $^{71}$ Trp-$^{1}L_{a6}$ | 0 | -5.8 |
| $^{71}$ Trp-$^{1}L_{b7}$ | $^{71}$ Trp-$^{1}L_{a5}$ | 0 | -19.7 |
| $^{71}$Trp-$^{1}L_{b6}$ | $^{71}$ Trp-$^{1}L_{a9}$ | 0 | -11.7 |
| $^{71}$ Trp-$^{1}L_{b6}$ | $^{71}$ Trp-$^{1}L_{a}$ | 0 | -9.7 |
| $^{71}$ Trp-$^{1}L_{b6}$ | $^{71}$ Trp-$^{1}L_{a8}$ | 0 | -10.3 |
| $^{71}$ Trp-$^{1}L_{b6}$ | $^{71}$ Trp-$^{1}L_{a7}$ | 0 | -3.2 |
| $^{71}$ Trp-$^{1}L_{b6}$ | $^{71}$ Trp-$^{1}L_{a6}$ | 0 | -13.0 |
| $^{71}$ Trp-$^{1}L_{b6}$ | $^{71}$ Trp-$^{1}L_{a5}$ | 0 | -5.1 |
| $^{71}$ Trp-$^{1}L_{b6}$ | $^{71}$ Trp-$^{1}L_{a4}$ | 0 | -5.8 |
| $^{71}$ Trp-$^{1}L_{b6}$ | $^{71}$ Trp-$^{1}L_{a3}$ | 0 | -5.4 |
| $^{71}$ Trp-$^{1}L_{b6}$ | $^{71}$ Trp-$^{1}L_{a2}$ | 0 | -39.1 |
| $^{71}$ Trp-$^{1}L_{b5}$ | $^{71}$ Trp-$^{1}L_{a9}$ | 0 | -26.2 |
| $^{71}$ Trp-$^{1}L_{b5}$ | $^{71}$ Trp-$^{1}L_{a8}$ | 0 | -5.0 |
| $^{71}$ Trp-$^{1}L_{b4}$ | $^{71}$ Trp-$^{1}L_{a9}$ | 0 | -4.3 |
| $^{71}$ Trp-$^{1}L_{b4}$ | $^{71}$ Trp-$^{1}L_{a8}$ | 0 | -13.5 |
| $^{71}$ Trp-$^{1}L_{b4}$ | $^{71}$ Trp-$^{1}L_{a7}$ | 0 | -18.3 |
| $^{71}$ Trp-$^{1}L_{b4}$ | $^{71}$ Trp-$^{1}L_{a6}$ | 0 | -7.9 |
| $^{71}$ Trp-$^{1}L_{b3}$ | $^{71}$ Trp-$^{1}L_{a9}$ | 0 | -8.3 |
| $^{71}$ Trp-$^{1}L_{b3}$ | $^{71}$ Trp-$^{1}L_{a8}$ | 0 | -6.7 |

| | | | |
|---|---|---|---|
| $^{71}$ Trp-$^{1}L_{b2}$ | $^{71}$ Trp-$^{1}L_{a9}$ | 0 | -8.9 |
| $^{71}$ Trp-$^{1}L_{b2}$ | $^{71}$ Trp-$^{1}L_{a8}$ | 0 | -7.1 |
| $^{94}$ Trp-$^{1}L_{b7}$ | $^{94}$ Trp-$^{1}L_{a9}$ | 0 | -10.5 |
| $^{94}$ Trp-$^{1}L_{b7}$ | $^{94}$ Trp-$^{1}L_{a8}$ | 0 | -1.2 |
| $^{94}$ Trp-$^{1}L_{b7}$ | $^{94}$ Trp-$^{1}L_{a7}$ | 0 | -2.9 |
| $^{94}$ Trp-$^{1}L_{b7}$ | $^{94}$ Trp-$^{1}L_{a6}$ | 0 | -2.8 |
| $^{94}$ Trp-$^{1}L_{b7}$ | $^{94}$ Trp-$^{1}L_{a5}$ | 0 | -4.7 |
| $^{94}$ Trp-$^{1}L_{b7}$ | $^{94}$ Trp-$^{1}L_{a4}$ | 0 | -45.3 |
| $^{94}$ Trp-$^{1}L_{b7}$ | $^{94}$ Trp-$^{1}L_{a3}$ | 0 | -5.0 |
| $^{94}$ Trp-$^{1}L_{b7}$ | $^{94}$ Trp-$^{1}L_{a2}$ | 0 | -18.3 |
| $^{94}$ Trp-$^{1}L_{b7}$ | $^{94}$ Trp-$^{1}L_{a}$ | 0 | -8.9 |
| $^{94}$ Trp-$^{1}L_{b6}$ | $^{94}$ Trp-$^{1}L_{a7}$ | 0 | -11.5 |
| $^{94}$ Trp-$^{1}L_{b6}$ | $^{94}$ Trp-$^{1}L_{a6}$ | 0 | -14.5 |
| $^{94}$ Trp-$^{1}L_{b6}$ | $^{94}$ Trp-$^{1}L_{a5}$ | 0 | -2.5 |
| $^{94}$ Trp-$^{1}L_{b6}$ | $^{94}$ Trp-$^{1}L_{a4}$ | 0 | -2.8 |
| $^{94}$ Trp-$^{1}L_{b6}$ | $^{94}$ Trp-$^{1}L_{a3}$ | 0 | -2.9 |
| $^{94}$ Trp-$^{1}L_{b6}$ | $^{94}$ Trp-$^{1}L_{a2}$ | 0 | -44.4 |
| $^{94}$ Trp-$^{1}L_{b6}$ | $^{94}$ Trp-$^{1}L_{a}$ | 0 | -4.7 |
| $^{94}$ Trp-$^{1}L_{b5}$ | $^{94}$ Trp-$^{1}L_{a9}$ | 0 | -3.0 |
| $^{94}$ Trp-$^{1}L_{b5}$ | $^{94}$ Trp-$^{1}L_{a8}$ | 0 | -12.4 |
| $^{94}$ Trp-$^{1}L_{b5}$ | $^{94}$ Trp-$^{1}L_{a7}$ | 0 | -5.8 |
| $^{94}$ Trp-$^{1}L_{b5}$ | $^{94}$ Trp-$^{1}L_{a6}$ | 0 | -35.5 |
| $^{94}$ Trp-$^{1}L_{b5}$ | $^{94}$ Trp-$^{1}L_{a5}$ | 0 | -9.1 |
| $^{94}$ Trp-$^{1}L_{b5}$ | $^{94}$ Trp-$^{1}L_{a3}$ | 0 | -9.7 |
| $^{94}$ Trp-$^{1}L_{b4}$ | $^{94}$ Trp-$^{1}L_{a9}$ | 0 | -2.1 |
| $^{94}$ Trp-$^{1}L_{b4}$ | $^{94}$ Trp-$^{1}L_{a8}$ | 0 | -1.8 |
| $^{94}$ Trp-$^{1}L_{b4}$ | $^{94}$ Trp-$^{1}L_{a7}$ | 0 | -14.0 |
| $^{94}$ Trp-$^{1}L_{b4}$ | $^{94}$ Trp-$^{1}L_{a6}$ | 0 | -3.8 |
| $^{94}$ Trp-$^{1}L_{b4}$ | $^{94}$ Trp-$^{1}L_{a5}$ | 0 | -36.3 |
| $^{94}$ Trp-$^{1}L_{b4}$ | $^{94}$ Trp-$^{1}L_{a4}$ | 0 | -27.2 |
| $^{94}$ Trp-$^{1}L_{b4}$ | $^{94}$ Trp-$^{1}L_{a3}$ | 0 | -6.8 |
| $^{94}$ Trp-$^{1}L_{b3}$ | $^{94}$ Trp-$^{1}L_{a9}$ | 0 | -4.0 |

| | | | |
|---|---|---|---|
| $^{94}$ Trp-$^{1}L_{b3}$ | $^{94}$ Trp-$^{1}L_{a8}$ | 0 | -23.2 |
| $^{94}$ Trp-$^{1}L_{b3}$ | $^{94}$ Trp-$^{1}L_{a7}$ | 0 | -7.7 |
| $^{94}$ Trp-$^{1}L_{b3}$ | $^{94}$ Trp-$^{1}L_{a6}$ | 0 | -17.3 |
| $^{94}$ Trp-$^{1}L_{b2}$ | $^{94}$ Trp-$^{1}L_{a9}$ | 0 | -4.3 |
| $^{94}$ Trp-$^{1}L_{b2}$ | $^{94}$ Trp-$^{1}L_{a8}$ | 0 | -3.4 |
| $^{94}$ Trp-$^{1}L_{b2}$ | $^{94}$ Trp-$^{1}L_{a7}$ | 0 | -18.2 |
| $^{94}$ Trp-$^{1}L_{b2}$ | $^{94}$ Trp-$^{1}L_{a6}$ | 0 | -7.8 |
| $^{94}$ Trp-$^{1}L_{b}$ | $^{94}$ Trp-$^{1}L_{a9}$ | 0 | -15.7 |
| $^{94}$ Trp-$^{1}L_{b}$ | $^{94}$ Trp-$^{1}L_{a8}$ | 0 | -41.5 |

Table A.5: Calculated interaction energies of electronic transitions of aromatic side chains in frame 16 of (PDB:1JNJ) of the NMR structure

| Chromophore/Transitions | Chromophore/Transitions | Distance (Å) | Interaction Energy ($cm^{-1}$) |
|---|---|---|---|
| $^{22}$Phe-$^{1}L_{a}$ | $^{70}$Phe-$^{1}L_{b}$ | 7.47 | -1.0 |
| $^{22}$ Phe-$^{1}L_{b6}$ | $^{67}$ Tyr-$^{1}L_{a}$ | 8.63 | -2.0 |
| $^{22}$ Phe-$^{1}L_{b5}$ | $^{67}$ Tyr-$^{1}L_{a}$ | 8.63 | 3.0 |
| $^{22}$ Phe-$^{1}L_{b4}$ | $^{67}$ Tyr-$^{1}L_{a}$ | 8.63 | -4.0 |
| $^{22}$ Phe-$^{1}L_{b}$ | $^{67}$ Tyr-$^{1}L_{a}$ | 8.63 | 6.8 |
| $^{22}$ Phe-$^{1}L_{b3}$ | $^{67}$ Tyr-$^{1}L_{a}$ | 8.63 | 5.6 |
| $^{22}$ Phe-$^{1}L_{b2}$ | $^{67}$ Tyr-$^{1}L_{a}$ | 8.63 | -4.9 |
| $^{22}$ Phe-$^{1}L_{a}$ | $^{67}$ Tyr-$^{1}L_{b8}$ | 8.63 | -1.6 |
| $^{22}$ Phe-$^{1}L_{a}$ | $^{67}$ Tyr-$^{1}L_{b7}$ | 8.63 | -1.6 |
| $^{22}$ Phe-$^{1}L_{a}$ | $^{67}$ Tyr-$^{1}L_{b6}$ | 8.63 | -3.2 |
| $^{22}$ Phe-$^{1}L_{a}$ | $^{67}$ Tyr-$^{1}L_{b5}$ | 8.63 | -2.0 |
| $^{22}$Phe-$^{1}L_{a}$ | $^{67}$ Tyr-$^{1}L_{b4}$ | 8.63 | -5.2 |
| $^{22}$ Phe-$^{1}L_{a}$ | $^{67}$ Tyr-$^{1}L_{b3}$ | 8.63 | -1.9 |
| $^{22}$Phe-$^{1}L_{a}$ | $^{67}$ Tyr-$^{1}L_{b2}$ | 8.63 | -6.7 |
| $^{22}$Phe-$^{1}L_{a}$ | $^{67}$Tyr-$^{1}L_{b}$ | 8.63 | -6.0 |
| $^{30}$ Phe-$^{1}L_{a}$ | $^{62}$ Phe-$^{1}L_{b}$ | 6.09 | 3.6 |
| $^{30}$ Phe-$^{1}L_{a}$ | $^{62}$ Phe-$^{1}L_{b3}$ | 6.09 | 3.0 |

| | | | |
|---|---|---|---|
| $^{30}$ Phe-$^{1}L_{b2}$ | $^{62}$ Phe-$^{1}L_{a}$ | 6.09 | 2.3 |
| $^{30}$ Phe-$^{1}L_{b4}$ | $^{62}$ Phe-$^{1}L_{a}$ | 6.09 | 1.9 |
| $^{30}$ Phe-$^{1}L_{a}$ | $^{62}$ Phe-$^{1}L_{b5}$ | 6.09 | 1.6 |
| $^{30}$ Phe-$^{1}L_{b6}$ | $^{62}$ Phe-$^{1}L_{a}$ | 6.09 | 1.0 |
| $^{30}$ Phe-$^{1}L_{a}$ | $^{62}$ Phe-$^{1}L_{b6}$ | 6.09 | -1.1 |
| $^{30}$ Phe-$^{1}L_{b5}$ | $^{62}$ Phe-$^{1}L_{a}$ | 6.09 | -1.4 |
| $^{30}$ Phe-$^{1}L_{a}$ | $^{62}$ Phe-$^{1}L_{b4}$ | 6.09 | -2.1 |
| $^{30}$Phe-$^{1}L_{a}$ | $^{62}$ Phe-$^{1}L_{b2}$ | 6.09 | -2.6 |
| $^{30}$ Phe-$^{1}L_{b3}$ | $^{62}$ Phe-$^{1}L_{a}$ | 6.09 | -2.7 |
| $^{30}$ Phe-$^{1}L_{b}$ | $^{62}$ Phe-$^{1}L_{a}$ | 6.09 | -3.2 |
| $^{30}$ Phe-$^{1}L_{b}$ | $^{66}$ Tyr-$^{1}L_{b2}$ | 8.69 | -1.0 |
| $^{30}$ Phe-$^{1}L_{b}$ | $^{66}$ Tyr-$^{1}L_{a}$ | 8.69 | 1.0 |
| $^{56}$Phe-$^{1}L_{a}$ | $^{62}$ Phe-$^{1}L_{b2}$ | 5.19 | 2.6 |
| $^{56}$ Phe-$^{1}L_{a}$ | $^{62}$ Phe-$^{1}L_{b4}$ | 5.19 | 2.1 |
| $^{56}$ Phe-$^{1}L_{b}$ | $^{62}$ Phe-$^{1}L_{b2}$ | 5.19 | 1.4 |
| $^{56}$Phe-$^{1}L_{b2}$ | $^{62}$ Phe-$^{1}L_{b}$ | 5.19 | 1.4 |
| $^{56}$ Phe-$^{1}L_{b}$ | $^{62}$ Phe-$^{1}L_{a}$ | 5.19 | 1.4 |
| $^{56}$Phe-$^{1}L_{b3}$ | $^{62}$ Phe-$^{1}L_{b2}$ | 5.19 | 1.1 |
| $^{56}$ Phe-$^{1}L_{b2}$ | $^{62}$ Phe-$^{1}L_{b3}$ | 5.19 | 1.1 |
| $^{56}$ Phe-$^{1}L_{b}$ | $^{62}$ Phe-$^{1}L_{b4}$ | 5.19 | 1.1 |
| $^{56}$ Phe-$^{1}L_{b4}$ | $^{62}$ Phe-$^{1}L_{b}$ | 5.19 | 1.1 |
| $^{56}$ Phe-$^{1}L_{b3}$ | $^{62}$ Phe-$^{1}L_{a}$ | 5.19 | 1.1 |
| $^{56}$ Phe-$^{1}L_{a}$ | $^{62}$ Phe-$^{1}L_{b6}$ | 5.19 | 1.0 |
| $^{56}$ Phe-$^{1}L_{b2}$ | $^{62}$Phe-$^{1}L_{a}$ | 5.19 | -1.0 |
| $^{56}$ Phe-$^{1}L_{b2}$ | $^{62}$ Phe-$^{1}L_{b2}$ | 5.19 | -1.0 |
| $^{56}$ Phe-$^{1}L_{b3}$ | $^{62}$ Phe-$^{1}L_{b3}$ | 5.19 | -1.3 |
| $^{56}$ Phe-$^{1}L_{a}$ | $^{62}$ Phe-$^{1}L_{b5}$ | 5.19 | -1.5 |
| $^{56}$ Phe-$^{1}L_{b}$ | $^{62}$ Phe-$^{1}L_{b3}$ | 5.19 | -1.6 |
| $^{56}$ Phe-$^{1}L_{b3}$ | $^{62}$ Phe-$^{1}L_{b}$ | 5.19 | -1.6 |
| $^{56}$ Phe-$^{1}L_{b}$ | $^{62}$ Phe-$^{1}L_{b}$ | 5.19 | -1.9 |
| $^{56}$ Phe-$^{1}L_{a}$ | $^{62}$ Phe-$^{1}L_{b3}$ | 5.19 | -2.9 |
| $^{56}$ Phe-$^{1}L_{a}$ | $^{62}$ Phe-$^{1}L_{b}$ | 5.19 | -3.5 |

| | | | |
|---|---|---|---|
| $^{56}$ Phe-$^{1}L_{a}$ | $^{63}$ Tyr-$^{1}L_{b2}$ | 7.22 | 1.1 |
| $^{56}$ Phe-$^{1}L_{a}$ | $^{63}$ Tyr-$^{1}L_{b}$ | 7.22 | 1.0 |
| $^{56}$ Phe-$^{1}L_{b}$ | $^{60}$ Trp-$^{1}L_{a}$ | 7.54 | -1.0 |
| $^{62}$ Phe-$^{1}L_{b}$ | $^{63}$ Tyr-$^{1}L_{a}$ | 8.38 | -1.1 |
| $^{70}$ Phe-$^{1}L_{a}$ | $^{78}$Tyr-$^{1}L_{b8}$ | 4.59 | 4.7 |
| $^{70}$Phe-$^{1}L_{a}$ | $^{78}$Tyr-$^{1}L_{b7}$ | 4.59 | 4.7 |
| $^{70}$Phe-$^{1}L_{a}$ | $^{78}$ Tyr-$^{1}L_{b6}$ | 4.59 | 9.4 |
| $^{70}$Phe-$^{1}L_{a}$ | $^{78}$ Tyr-$^{1}L_{b5}$ | 4.59 | 6.0 |
| $^{70}$ Phe-$^{1}L_{a}$ | $^{78}$ Tyr-$^{1}L_{b4}$ | 4.59 | 15.1 |
| $^{70}$Phe-$^{1}L_{a}$ | $^{78}$Tyr-$^{1}L_{b3}$ | 4.59 | 5.7 |
| $^{70}$ Phe-$^{1}L_{a}$ | $^{78}$ Tyr-$^{1}L_{b2}$ | 4.59 | 19.5 |
| $^{70}$ Phe-$^{1}L_{a}$ | $^{78}$ Tyr-$^{1}L_{b}$ | 4.59 | 17.5 |
| $^{70}$ Phe-$^{1}L_{b6}$ | $^{95}$ Trp-$^{1}L_{a5}$ | 6.48 | 1.2 |
| $^{70}$ Phe-$^{1}L_{b6}$ | $^{95}$ Trp-$^{1}L_{a4}$ | 6.48 | 1.4 |
| $^{70}$ Phe-$^{1}L_{b6}$ | $^{95}$ Trp-$^{1}L_{a3}$ | 6.48 | 1.3 |
| $^{70}$ Phe-$^{1}L_{b6}$ | $^{95}$ Trp-$^{1}L_{a2}$ | 6.48 | 2.2 |
| $^{70}$ Phe-$^{1}L_{b6}$ | $^{95}$Trp-$^{1}L_{a}$ | 6.48 | 2.4 |
| $^{70}$ Phe-$^{1}L_{b5}$ | $^{95}$ Trp-$^{1}L_{a7}$ | 6.48 | -1.1 |
| $^{70}$ Phe-$^{1}L_{b5}$ | $^{95}$ Trp-$^{1}L_{a6}$ | 6.48 | -1.1 |
| $^{70}$ Phe-$^{1}L_{b5}$ | $^{95}$ Trp-$^{1}L_{a5}$ | 6.48 | -1.8 |
| $^{70}$ Phe-$^{1}L_{b5}$ | $^{95}$ Trp-$^{1}L_{a4}$ | 6.48 | -2.1 |
| $^{70}$ Phe-$^{1}L_{b5}$ | $^{95}$ Trp-$^{1}L_{a3}$ | 6.48 | -1.9 |
| $^{70}$ Phe-$^{1}L_{b5}$ | $^{95}$Trp-$^{1}L_{a2}$ | 6.48 | -3.2 |
| $^{70}$ Phe-$^{1}L_{b5}$ | $^{95}$ Trp-$^{1}L_{a}$ | 6.48 | -3.5 |
| $^{70}$ Phe-$^{1}L_{b4}$ | $^{95}$ Trp-$^{1}L_{b}$ | 6.48 | -1.1 |
| $^{70}$ Phe-$^{1}L_{b4}$ | $^{95}$ Trp-$^{1}L_{a7}$ | 6.48 | 1.5 |
| $^{70}$ Phe-$^{1}L_{b4}$ | $^{95}$ Trp-$^{1}L_{a6}$ | 6.48 | 1.5 |
| $^{70}$ Phe-$^{1}L_{b4}$ | $^{95}$ Trp-$^{1}L_{a5}$ | 6.48 | 2.5 |
| $^{70}$ Phe-$^{1}L_{b4}$ | $^{95}$ Trp-$^{1}L_{a4}$ | 6.48 | 2.8 |
| $^{70}$ Phe-$^{1}L_{b4}$ | $^{95}$ Trp-$^{1}L_{a3}$ | 6.48 | 2.6 |
| $^{70}$ Phe-$^{1}L_{b4}$ | $^{95}$ Trp-$^{1}L_{a2}$ | 6.48 | 4.4 |
| $^{70}$ Phe-$^{1}L_{b4}$ | $^{95}$ Trp-$^{1}L_{a}$ | 6.48 | 4.7 |

| | | | |
|---|---|---|---|
| $^{70}$ Phe-$^{1}L_{b3}$ | $^{95}$ Trp-$^{1}L_{b3}$ | 6.48 | 1.1 |
| $^{70}$ Phe-$^{1}L_{b3}$ | $^{95}$ Trp-$^{1}L_{b2}$ | 6.48 | 1.2 |
| $^{70}$ Phe-$^{1}L_{b3}$ | $^{95}$ Trp-$^{1}L_{b}$ | 6.48 | 1.6 |
| $^{70}$ Phe-$^{1}L_{b3}$ | $^{95}$ Trp-$^{1}L_{a9}$ | 6.48 | -1.1 |
| $^{70}$ Phe-$^{1}L_{b3}$ | $^{95}$ Trp-$^{1}L_{a7}$ | 6.48 | -2.2 |
| $^{70}$ Phe-$^{1}L_{b3}$ | $^{95}$ Trp-$^{1}L_{a6}$ | 6.48 | -2.1 |
| $^{70}$ Phe-$^{1}L_{b3}$ | $^{95}$ Trp-$^{1}L_{a5}$ | 6.48 | -3.5 |
| $^{70}$ Phe-$^{1}L_{b3}$ | $^{95}$ Trp-$^{1}L_{a4}$ | 6.48 | -3.9 |
| $^{70}$ Phe-$^{1}L_{b3}$ | $^{95}$ Trp-$^{1}L_{a3}$ | 6.48 | -3.7 |
| $^{70}$ Phe-$^{1}L_{b3}$ | $^{95}$ Trp-$^{1}L_{a2}$ | 6.48 | -6.1 |
| $^{70}$ Phe-$^{1}L_{b3}$ | $^{95}$ Trp-$^{1}L_{a}$ | 6.48 | -6.6 |
| $^{70}$ Phe-$^{1}L_{b2}$ | $^{95}$ Trp-$^{1}L_{b2}$ | 6.48 | -1.0 |
| $^{70}$ Phe-$^{1}L_{b2}$ | $^{95}$ Trp-$^{1}L_{b}$ | 6.48 | -1.4 |
| $^{70}$ Phe-$^{1}L_{b2}$ | $^{95}$ Trp-$^{1}L_{a9}$ | 6.48 | 1.0 |
| $^{70}$ Phe-$^{1}L_{b2}$ | $^{95}$ Trp-$^{1}L_{a7}$ | 6.48 | 1.9 |
| $^{70}$ Phe-$^{1}L_{b2}$ | $^{95}$ Trp-$^{1}L_{a6}$ | 6.48 | 1.8 |
| $^{70}$ Phe-$^{1}L_{b2}$ | $^{95}$ Trp-$^{1}L_{a5}$ | 6.48 | 3.0 |
| $^{70}$ Phe-$^{1}L_{b2}$ | $^{95}$ Trp-$^{1}L_{a4}$ | 6.48 | 3.4 |
| $^{70}$ Phe-$^{1}L_{b2}$ | $^{95}$ Trp-$^{1}L_{a3}$ | 6.48 | 3.2 |
| $^{70}$ Phe-$^{1}L_{b2}$ | $^{95}$ Trp-$^{1}L_{a2}$ | 6.48 | 5.4 |
| $^{70}$ Phe-$^{1}L_{b2}$ | $^{95}$ Trp-$^{1}L_{a}$ | 6.48 | 5.8 |
| $^{70}$ Phe-$^{1}L_{b}$ | $^{95}$ Trp-$^{1}L_{b5}$ | 6.48 | 1.0 |
| $^{70}$ Phe-$^{1}L_{b}$ | $^{95}$ Trp-$^{1}L_{b3}$ | 6.48 | 1.3 |
| $^{70}$ Phe-$^{1}L_{b}$ | $^{95}$ Trp-$^{1}L_{b2}$ | 6.48 | 1.4 |
| $^{70}$ Phe-$^{1}L_{b}$ | $^{95}$ Trp-$^{1}L_{b}$ | 6.48 | 1.9 |
| $^{70}$ Phe-$^{1}L_{b}$ | $^{95}$ Trp-$^{1}L_{a9}$ | 6.48 | -1.4 |
| $^{70}$ Phe-$^{1}L_{b}$ | $^{95}$ Trp-$^{1}L_{a8}$ | 6.48 | -1.1 |
| $^{70}$ Phe-$^{1}L_{b}$ | $^{95}$ Trp-$^{1}L_{a7}$ | 6.48 | -2.6 |
| $^{70}$ Phe-$^{1}L_{b}$ | $^{95}$ Trp-$^{1}L_{a6}$ | 6.48 | -2.5 |
| $^{70}$ Phe-$^{1}L_{b}$ | $^{95}$ Trp-$^{1}L_{a5}$ | 6.48 | -4.2 |
| $^{70}$ Phe-$^{1}L_{b}$ | $^{95}$ Trp-$^{1}L_{a4}$ | 6.48 | -4.7 |
| $^{70}$ Phe-$^{1}L_{b}$ | $^{95}$ Trp-$^{1}L_{a3}$ | 6.48 | -4.5 |

| | | | |
|---|---|---|---|
| $^{70}$ Phe-$^{1}L_{b}$ | $^{95}$ Trp-$^{1}L_{a2}$ | 6.48 | -7.4 |
| $^{70}$ Phe-$^{1}L_{b}$ | $^{95}$ Trp-$^{1}L_{a}$ | 6.48 | -8.0 |
| $^{70}$ Phe-$^{1}L_{a}$ | $^{95}$Trp-$^{1}L_{b7}$ | 6.48 | -5.1 |
| $^{70}$ Phe-$^{1}L_{a}$ | $^{95}$Trp-$^{1}L_{b6}$ | 6.48 | -2.7 |
| $^{70}$Phe-$^{1}L_{a}$ | $^{95}$ Trp-$^{1}L_{b4}$ | 6.48 | -6.9 |
| $^{10}$ Tyr-$^{1}L_{b8}$ | $^{10}$ Tyr-$^{1}L_{a}$ | 0 | -3.3 |
| $^{10}$ Tyr-$^{1}L_{b7}$ | $^{10}$ Tyr-$^{1}L_{a}$ | 0 | -3.3 |
| $^{10}$ Tyr-$^{1}L_{b6}$ | $^{10}$ Tyr-$^{1}L_{a}$ | 0 | -6.7 |
| $^{10}$ Tyr-$^{1}L_{b5}$ | $^{10}$ Tyr-$^{1}L_{a}$ | 0 | -4.3 |
| $^{10}$ Tyr-$^{1}L_{b3}$ | $^{10}$ Tyr-$^{1}L_{a}$ | 0 | -4.1 |
| $^{10}$ Tyr-$^{1}L_{b8}$ | $^{26}$ Tyr-$^{1}L_{b}$ | 6.56 | -1.0 |
| $^{10}$ Tyr-$^{1}L_{b8}$ | $^{26}$ Tyr-$^{1}L_{b2}$ | 6.56 | -1.2 |
| $^{10}$ Tyr-$^{1}L_{b7}$ | $^{26}$ Tyr-$^{1}L_{b}$ | 6.56 | -1.0 |
| $^{10}$ Tyr-$^{1}L_{b7}$ | $^{26}$ Tyr-$^{1}L_{b2}$ | 6.56 | -1.2 |
| $^{10}$ Tyr-$^{1}L_{b6}$ | $^{26}$ Tyr-$^{1}L_{b6}$ | 6.56 | -1.1 |
| $^{10}$ Tyr-$^{1}L_{b6}$ | $^{26}$ Tyr-$^{1}L_{b4}$ | 6.56 | -1.9 |
| $^{10}$ Tyr-$^{1}L_{b6}$ | $^{26}$ Tyr-$^{1}L_{b}$ | 6.56 | -2.1 |
| $^{10}$ Tyr-$^{1}L_{b6}$ | $^{26}$ Tyr-$^{1}L_{b2}$ | 6.56 | -2.4 |
| $^{10}$ Tyr-$^{1}L_{b6}$ | $^{26}$ Tyr-$^{1}L_{a}$ | 6.56 | 1.4 |
| $^{10}$ Tyr-$^{1}L_{b5}$ | $^{26}$ Tyr-$^{1}L_{b4}$ | 6.56 | -1.2 |
| $^{10}$ Tyr-$^{1}L_{b5}$ | $^{26}$ Tyr-$^{1}L_{b}$ | 6.56 | -1.4 |
| $^{10}$ Tyr-$^{1}L_{b5}$ | $^{26}$ Tyr-$^{1}L_{b2}$ | 6.56 | -1.5 |
| $^{10}$ Tyr-$^{1}L_{b4}$ | $^{26}$ Tyr-$^{1}L_{b6}$ | 6.56 | -1.9 |
| $^{10}$ Tyr-$^{1}L_{b4}$ | $^{26}$ Tyr-$^{1}L_{b5}$ | 6.56 | -1.2 |
| $^{10}$ Tyr-$^{1}L_{b4}$ | $^{26}$ Tyr-$^{1}L_{b4}$ | 6.56 | -3.0 |
| $^{10}$ Tyr-$^{1}L_{b4}$ | $^{26}$ Tyr-$^{1}L_{b}$ | 6.56 | -3.5 |
| $^{10}$ Tyr-$^{1}L_{b4}$ | $^{26}$ Tyr-$^{1}L_{b3}$ | 6.56 | -1.1 |
| $^{10}$ Tyr-$^{1}L_{b4}$ | $^{26}$ Tyr-$^{1}L_{b2}$ | 6.56 | -3.9 |
| $^{10}$ Tyr-$^{1}L_{b4}$ | $^{26}$ Tyr-$^{1}L_{a}$ | 6.56 | 2.2 |
| $^{10}$ Tyr-$^{1}L_{b3}$ | $^{26}$ Tyr-$^{1}L_{b4}$ | 6.56 | -1.1 |
| $^{10}$ Tyr-$^{1}L_{b3}$ | $^{26}$ Tyr-$^{1}L_{b}$ | 6.56 | -1.3 |
| $^{10}$ Tyr-$^{1}L_{b3}$ | $^{26}$ Tyr-$^{1}L_{b2}$ | 6.56 | -1.4 |

| | | | |
|---|---|---|---|
| $^{10}$ Tyr-$^{1}L_{b2}$ | $^{26}$ Tyr-$^{1}L_{b8}$ | 6.56 | -1.2 |
| $^{10}$ Tyr-$^{1}L_{b2}$ | $^{26}$ Tyr-$^{1}L_{b7}$ | 6.56 | -1.2 |
| $^{10}$ Tyr-$^{1}L_{b2}$ | $^{26}$ Tyr-$^{1}L_{b6}$ | 6.56 | -2.4 |
| $^{10}$ Tyr-$^{1}L_{b2}$ | $^{26}$ Tyr-$^{1}L_{b5}$ | 6.56 | -1.5 |
| $^{10}$ Tyr-$^{1}L_{b2}$ | $^{26}$ Tyr-$^{1}L_{b4}$ | 6.56 | -3.9 |
| $^{10}$ Tyr-$^{1}L_{b2}$ | $^{26}$ Tyr-$^{1}L_{b}$ | 6.56 | -4.5 |
| $^{10}$ Tyr-$^{1}L_{b2}$ | $^{26}$ Tyr-$^{1}L_{b3}$ | 6.56 | -1.4 |
| $^{10}$ Tyr-$^{1}L_{b2}$ | $^{26}$ Tyr-$^{1}L_{b2}$ | 6.56 | -5.0 |
| $^{10}$ Tyr-$^{1}L_{b2}$ | $^{26}$ Tyr-$^{1}L_{a}$ | 6.56 | 2.9 |
| $^{10}$ Tyr-$^{1}L_{b}$ | $^{26}$ Tyr-$^{1}L_{b8}$ | 6.56 | -1.0 |
| $^{10}$ Tyr-$^{1}L_{b}$ | $^{26}$ Tyr-$^{1}L_{b7}$ | 6.56 | -1.0 |
| $^{10}$ Tyr-$^{1}L_{b}$ | $^{26}$ Tyr-$^{1}L_{b6}$ | 6.56 | -2.1 |
| $^{10}$ Tyr-$^{1}L_{b}$ | $^{26}$ Tyr-$^{1}L_{b5}$ | 6.56 | -1.4 |
| $^{10}$ Tyr-$^{1}L_{b}$ | $^{26}$ Tyr-$^{1}L_{b4}$ | 6.56 | -3.5 |
| $^{10}$ Tyr-$^{1}L_{b}$ | $^{26}$ Tyr-$^{1}L_{b}$ | 6.56 | -4.0 |
| $^{10}$ Tyr-$^{1}L_{b}$ | $^{26}$ Tyr-$^{1}L_{b3}$ | 6.56 | -1.3 |
| $^{10}$ Tyr-$^{1}L_{b}$ | $^{26}$ Tyr-$^{1}L_{b2}$ | 6.56 | -4.5 |
| $^{10}$ Tyr-$^{1}L_{b}$ | $^{26}$ Tyr-$^{1}L_{a}$ | 6.56 | 2.6 |
| $^{10}$ Tyr-$^{1}L_{a}$ | $^{26}$ Tyr-$^{1}L_{b6}$ | 6.56 | 1.0 |
| $^{10}$ Tyr-$^{1}L_{a}$ | $^{26}$ Tyr-$^{1}L_{b4}$ | 6.56 | 1.6 |
| $^{10}$ Tyr-$^{1}L_{a}$ | $^{26}$ Tyr-$^{1}L_{b}$ | 6.56 | 1.8 |
| $^{10}$ Tyr-$^{1}L_{a}$ | $^{26}$ Tyr-$^{1}L_{b2}$ | 6.56 | 2.1 |
| $^{10}$ Tyr-$^{1}L_{b2}$ | $^{63}$ Tyr-$^{1}L_{a}$ | 12.88 | 1.0 |
| $^{10}$ Tyr-$^{1}L_{a}$ | $^{63}$ Tyr-$^{1}L_{b4}$ | 12.88 | -1.2 |
| $^{10}$ Tyr-$^{1}L_{a}$ | $^{63}$ Tyr-$^{1}L_{b2}$ | 12.88 | -1.6 |
| $^{10}$ Tyr-$^{1}L_{a}$ | $^{63}$ Tyr-$^{1}L_{b}$ | 12.88 | -1.4 |
| $^{10}$ Tyr-$^{1}L_{a}$ | $^{67}$ Tyr-$^{1}L_{b8}$ | 8.94 | 1.2 |
| $^{10}$ Tyr-$^{1}L_{a}$ | $^{67}$ Tyr-$^{1}L_{b7}$ | 8.94 | 1.2 |
| $^{10}$ Tyr-$^{1}L_{a}$ | $^{67}$ Tyr-$^{1}L_{b6}$ | 8.94 | 2.5 |
| $^{10}$ Tyr-$^{1}L_{a}$ | $^{67}$ Tyr-$^{1}L_{b5}$ | 8.94 | 1.6 |
| $^{10}$ Tyr-$^{1}L_{a}$ | $^{67}$ Tyr-$^{1}L_{b4}$ | 8.94 | 4.1 |
| $^{10}$ Tyr-$^{1}L_{a}$ | $^{67}$ Tyr-$^{1}L_{b3}$ | 8.94 | 1.5 |

| | | | |
|---|---|---|---|
| $^{10}$ Tyr-$^{1}L_{a}$ | $^{67}$ Tyr-$^{1}L_{b2}$ | 8.94 | 5.3 |
| $^{10}$ Tyr-$^{1}L_{a}$ | $^{67}$ Tyr-$^{1}L_{b}$ | 8.94 | 4.8 |
| $^{10}$ Tyr-$^{1}L_{b6}$ | $^{95}$ Trp-$^{1}L_{b}$ | 10.02 | -1.2 |
| $^{10}$Tyr-$^{1}L_{b4}$ | $^{95}$ Trp-$^{1}L_{b5}$ | 10.02 | -1.0 |
| $^{10}$ Tyr-$^{1}L_{b4}$ | $^{95}$ Trp-$^{1}L_{b3}$ | 10.02 | -1.3 |
| $^{10}$ Tyr-$^{1}L_{b4}$ | $^{95}$ Trp-$^{1}L_{b2}$ | 10.02 | -1.4 |
| $^{10}$ Tyr-$^{1}L_{b4}$ | $^{95}$ Trp-$^{1}L_{b}$ | 10.02 | -1.9 |
| $^{10}$ Tyr-$^{1}L_{b2}$ | $^{95}$ Trp-$^{1}L_{b5}$ | 10.02 | -1.3 |
| $^{10}$ Tyr-$^{1}L_{b2}$ | $^{95}$ Trp-$^{1}L_{b3}$ | 10.02 | -1.7 |
| $^{10}$ Tyr-$^{1}L_{b2}$ | $^{95}$ Trp-$^{1}L_{b2}$ | 10.02 | -1.8 |
| $^{10}$ Tyr-$^{1}L_{b2}$ | $^{95}$ Trp-$^{1}L_{b}$ | 10.02 | -2.5 |
| $^{10}$ Tyr-$^{1}L_{b2}$ | $^{95}$ Trp-$^{1}L_{a2}$ | 10.02 | -1.0 |
| $^{10}$ Tyr-$^{1}L_{b2}$ | $^{95}$ Trp-$^{1}L_{a}$ | 10.02 | -1.1 |
| $^{10}$ Tyr-$^{1}L_{b}$ | $^{95}$ Trp-$^{1}L_{b5}$ | 10.02 | -1.1 |
| $^{10}$ Tyr-$^{1}L_{b}$ | $^{95}$ Trp-$^{1}L_{b3}$ | 10.02 | -1.5 |
| $^{10}$ Tyr-$^{1}L_{b}$ | $^{95}$ Trp-$^{1}L_{b2}$ | 10.02 | -1.6 |
| $^{10}$ Tyr-$^{1}L_{b}$ | $^{95}$ Trp-$^{1}L_{b}$ | 10.02 | -2.2 |
| $^{10}$ Tyr-$^{1}L_{b}$ | $^{95}$ Trp-$^{1}L_{a}$ | 10.02 | -1.0 |
| $^{10}$ Tyr-$^{1}L_{a}$ | $^{95}$ Trp-$^{1}L_{b7}$ | 10.02 | -1.0 |
| $^{10}$ Tyr-$^{1}L_{a}$ | $^{95}$ Trp-$^{1}L_{b5}$ | 10.02 | -2.0 |
| $^{10}$ Tyr-$^{1}L_{a}$ | $^{95}$Trp-$^{1}L_{b4}$ | 10.02 | -1.4 |
| $^{10}$ Tyr-$^{1}L_{a}$ | $^{95}$ Trp-$^{1}L_{b3}$ | 10.02 | -2.7 |
| $^{10}$ Tyr-$^{1}L_{a}$ | $^{95}$ Trp-$^{1}L_{b2}$ | 10.02 | -2.9 |
| $^{10}$ Tyr-$^{1}L_{a}$ | $^{95}$ Trp-$^{1}L_{b}$ | 10.02 | -3.8 |
| $^{26}$ Tyr-$^{1}L_{b6}$ | $^{26}$ Tyr-$^{1}L_{a}$ | 0 | 1.9 |
| $^{26}$ Tyr-$^{1}L_{b5}$ | $^{26}$ Tyr-$^{1}L_{a}$ | 0 | 1.2 |
| $^{26}$ Tyr-$^{1}L_{b4}$ | $^{26}$ Tyr-$^{1}L_{a}$ | 0 | 3.0 |
| $^{26}$ Tyr-$^{1}L_{b}$ | $^{26}$ Tyr-$^{1}L_{a}$ | 0 | 3.5 |
| $^{26}$ Tyr-$^{1}L_{b3}$ | $^{26}$ Tyr-$^{1}L_{a}$ | 0 | 1.1 |
| $^{26}$ Tyr-$^{1}L_{b2}$ | $^{26}$ Tyr-$^{1}L_{a}$ | 0 | 3.9 |
| $^{26}$ Tyr-$^{1}L_{b8}$ | $^{63}$ Tyr-$^{1}L_{a}$ | 6.81 | 1.0 |
| $^{26}$ Tyr-$^{1}L_{b7}$ | $^{63}$ Tyr-$^{1}L_{a}$ | 6.81 | 1.0 |

| | | | |
|---|---|---|---|
| $^{26}$ Tyr-$^{1}L_{b6}$ | $^{63}$ Tyr-$^{1}L_{a}$ | 6.81 | 2.0 |
| $^{26}$ Tyr-$^{1}L_{b5}$ | $^{63}$ Tyr-$^{1}L_{a}$ | 6.81 | 1.3 |
| $^{26}$ Tyr-$^{1}L_{b4}$ | $^{63}$ Tyr-$^{1}L_{a}$ | 6.81 | 3.2 |
| $^{26}$ Tyr-$^{1}L_{b}$ | $^{63}$ Tyr-$^{1}L_{a}$ | 6.81 | 3.7 |
| $^{26}$ Tyr-$^{1}L_{b3}$ | $^{63}$ Tyr-$^{1}L_{a}$ | 6.81 | 1.2 |
| $^{26}$ Tyr-$^{1}L_{b2}$ | $^{63}$ Tyr-$^{1}L_{a}$ | 6.81 | 4.2 |
| $^{26}$ Tyr-$^{1}L_{a}$ | $^{63}$ Tyr-$^{1}L_{b6}$ | 6.81 | -1.3 |
| $^{26}$ Tyr-$^{1}L_{a}$ | $^{63}$ Tyr-$^{1}L_{b4}$ | 6.81 | -2.1 |
| $^{26}$ Tyr-$^{1}L_{a}$ | $^{63}$ Tyr-$^{1}L_{b2}$ | 6.81 | -2.7 |
| $^{26}$ Tyr-$^{1}L_{a}$ | $^{63}$ Tyr-$^{1}L_{b}$ | 6.81 | -2.4 |
| $^{26}$ Tyr-$^{1}L_{b6}$ | $^{66}$ Tyr-$^{1}L_{a}$ | 11.14 | -1.1 |
| $^{26}$ Tyr-$^{1}L_{b4}$ | $^{66}$ Tyr-$^{1}L_{a}$ | 11.14 | -1.8 |
| $^{26}$ Tyr-$^{1}L_{b}$ | $^{66}$ Tyr-$^{1}L_{a}$ | 11.14 | -2.1 |
| $^{26}$ Tyr-$^{1}L_{b2}$ | $^{66}$ Tyr-$^{1}L_{a}$ | 11.14 | -2.4 |
| $^{26}$ Tyr-$^{1}L_{a}$ | $^{66}$ Tyr-$^{1}L_{b6}$ | 11.14 | -1.8 |
| $^{26}$ Tyr-$^{1}L_{a}$ | $^{66}$ Tyr-$^{1}L_{b5}$ | 11.14 | -1.1 |
| $^{26}$ Tyr-$^{1}L_{a}$ | $^{66}$ Tyr-$^{1}L_{b4}$ | 11.14 | -2.9 |
| $^{26}$ Tyr-$^{1}L_{a}$ | $^{66}$ Tyr-$^{1}L_{b3}$ | 11.14 | -1.1 |
| $^{26}$ Tyr-$^{1}L_{a}$ | $^{66}$ Tyr-$^{1}L_{b2}$ | 11.14 | -3.8 |
| $^{26}$ Tyr-$^{1}L_{a}$ | $^{66}$ Tyr-$^{1}L_{b}$ | 11.14 | -3.4 |
| $^{26}$ Tyr-$^{1}L_{b}$ | $^{67}$ Tyr-$^{1}L_{b2}$ | 10.23 | 1.1 |
| $^{26}$ Tyr-$^{1}L_{b}$ | $^{67}$ Tyr-$^{1}L_{b}$ | 10.23 | 1.0 |
| $^{26}$ Tyr-$^{1}L_{b2}$ | $^{67}$ Tyr-$^{1}L_{b2}$ | 10.23 | 1.2 |
| $^{26}$ Tyr-$^{1}L_{b2}$ | $^{67}$ Tyr-$^{1}L_{b}$ | 10.23 | 1.1 |
| $^{26}$ Tyr-$^{1}L_{a}$ | $^{95}$ Trp-$^{1}L_{b5}$ | 15 | -1.0 |
| $^{26}$ Tyr-$^{1}L_{a}$ | $^{95}$ Trp-$^{1}L_{b3}$ | 15 | -1.4 |
| $^{26}$ Tyr-$^{1}L_{a}$ | $^{95}$ Trp-$^{1}L_{b2}$ | 15 | -1.5 |
| $^{26}$ Tyr-$^{1}L_{a}$ | $^{95}$ Trp-$^{1}L_{b}$ | 15 | -2.0 |
| $^{63}$ Tyr-$^{1}L_{b8}$ | $^{63}$ Tyr-$^{1}L_{a}$ | 0 | -5.9 |
| $^{63}$ Tyr-$^{1}L_{b7}$ | $^{63}$ Tyr-$^{1}L_{a}$ | 0 | -5.9 |
| $^{63}$ Tyr-$^{1}L_{b5}$ | $^{63}$ Tyr-$^{1}L_{a}$ | 0 | -7.6 |
| $^{63}$ Tyr-$^{1}L_{b3}$ | $^{63}$ Tyr-$^{1}L_{a}$ | 0 | -7.1 |

| | | | |
|---|---|---|---|
| $^{63}$ Tyr-$^{1}L_{b4}$ | $^{66}$ Tyr-$^{1}L_{a}$ | 11.93 | 1.1 |
| $^{63}$ Tyr-$^{1}L_{b2}$ | $^{66}$ Tyr-$^{1}L_{a}$ | 11.93 | 1.5 |
| $^{63}$ Tyr-$^{1}L_{b}$ | $^{66}$ Tyr-$^{1}L_{a}$ | 11.93 | 1.3 |
| $^{63}$ Tyr-$^{1}L_{a}$ | $^{66}$Tyr-$^{1}L_{b4}$ | 11.93 | -1.2 |
| $^{63}$ TYR-1La | $^{66}$ Tyr-$^{1}L_{b2}$ | 11.93 | -1.5 |
| $^{63}$ TYR-1La | $^{66}$ Tyr-$^{1}L_{b}$ | 11.93 | -1.4 |
| $^{63}$ Tyr-$^{1}L_{b4}$ | $^{67}$ Tyr-$^{1}L_{a}$ | 12.31 | -1.0 |
| $^{63}$ Tyr-$^{1}L_{b2}$ | $^{67}$ Tyr-$^{1}L_{a}$ | 12.31 | -1.3 |
| $^{63}$ Tyr-$^{1}L_{b}$ | $^{67}$ Tyr-$^{1}L_{a}$ | 12.31 | -1.1 |
| $^{63}$ Tyr-$^{1}L_{b4}$ | $^{60}$ Trp-$^{1}L_{a}$ | 10.52 | 1.0 |
| $^{63}$ Tyr-$^{1}L_{b2}$ | $^{60}$ Trp-$^{1}L_{a}$ | 10.52 | 1.3 |
| $^{63}$ Tyr-$^{1}L_{b2}$ | $^{60}$ Trp-$^{1}L_{a2}$ | 10.52 | 1.2 |
| $^{63}$ Tyr-$^{1}L_{b}$ | $^{60}$ Trp-$^{1}L_{a}$ | 10.52 | 1.1 |
| $^{63}$ Tyr-$^{1}L_{b}$ | $^{60}$ Trp-$^{1}L_{a2}$ | 10.52 | 1.0 |
| $^{66}$ Tyr-$^{1}L_{b8}$ | $^{66}$ Tyr-$^{1}L_{a}$ | 0 | 8.3 |
| $^{66}$ Tyr-$^{1}L_{b7}$ | $^{66}$ Tyr-$^{1}L_{a}$ | 0 | 8.3 |
| $^{66}$ Tyr-$^{1}L_{b6}$ | $^{66}$ Tyr-$^{1}L_{a}$ | 0 | 16.6 |
| $^{66}$ Tyr-$^{1}L_{b5}$ | $^{66}$ Tyr-$^{1}L_{a}$ | 0 | 10.6 |
| $^{66}$ Tyr-$^{1}L_{b4}$ | $^{66}$ Tyr-$^{1}L_{a}$ | 0 | 26.7 |
| $^{66}$ Tyr-$^{1}L_{b3}$ | $^{66}$ Tyr-$^{1}L_{a}$ | 0 | 10.0 |
| $^{66}$ Tyr-$^{1}L_{b2}$ | $^{66}$ Tyr-$^{1}L_{a}$ | 0 | 34.4 |
| $^{66}$ Tyr-$^{1}L_{b}$ | $^{66}$ Tyr-$^{1}L_{a}$ | 0 | 30.8 |
| $^{66}$ Tyr-$^{1}L_{b4}$ | $^{67}$ Tyr-$^{1}L_{b2}$ | 9.96 | 1.2 |
| $^{66}$ Tyr-$^{1}L_{b4}$ | $^{67}$ Tyr-$^{1}L_{b}$ | 9.96 | 1.1 |
| $^{66}$ Tyr-$^{1}L_{b4}$ | $^{67}$ Tyr-$^{1}L_{a}$ | 9.96 | -1.0 |
| $^{66}$ Tyr-$^{1}L_{b2}$ | $^{67}$ Tyr-$^{1}L_{b4}$ | 9.96 | 1.2 |
| $^{66}$ Tyr-$^{1}L_{b2}$ | $^{67}$ Tyr-$^{1}L_{b2}$ | 9.96 | 1.6 |
| $^{66}$ Tyr-$^{1}L_{b2}$ | $^{67}$ Tyr-$^{1}L_{b}$ | 9.96 | 1.4 |
| $^{66}$ Tyr-$^{1}L_{b2}$ | $^{67}$ Tyr-$^{1}L_{a}$ | 9.96 | -1.3 |
| $^{66}$ Tyr-$^{1}L_{b}$ | $^{67}$ Tyr-$^{1}L_{b4}$ | 9.96 | 1.1 |
| $^{66}$ Tyr-$^{1}L_{b}$ | $^{67}$ Tyr-$^{1}L_{b2}$ | 9.96 | 1.4 |
| $^{66}$ Tyr-$^{1}L_{b}$ | $^{67}$ Tyr-$^{1}L_{b}$ | 9.96 | 1.2 |

| | | | |
|---|---|---|---|
| $^{66}$ Tyr-$^{1}L_{b}$ | $^{67}$ Tyr-$^{1}L_{a}$ | 9.96 | -1.1 |
| $^{67}$ Tyr-$^{1}L_{b4}$ | $^{67}$ Tyr-$^{1}L_{a}$ | 0 | 1.1 |
| $^{67}$ Tyr-$^{1}L_{b2}$ | $^{67}$ Tyr-$^{1}L_{a}$ | 0 | 1.5 |
| $^{67}$ Tyr-$^{1}L_{b}$ | $^{67}$ Tyr-$^{1}L_{a}$ | 0 | 1.3 |
| $^{67}$ Tyr-$^{1}L_{a}$ | $^{78}$Tyr-$^{1}L_{b2}$ | 16.09 | -1.2 |
| $^{67}$ Tyr-$^{1}L_{a}$ | $^{78}$ Tyr-$^{1}L_{b}$ | 16.09 | -1.0 |
| $^{78}$ Tyr-$^{1}L_{b8}$ | $^{78}$ Tyr-$^{1}L_{a}$ | 0 | -3.2 |
| $^{78}$ Tyr-$^{1}L_{b7}$ | $^{78}$ Tyr-$^{1}L_{a}$ | 0 | -3.2 |
| $^{78}$ Tyr-$^{1}L_{b6}$ | $^{78}$ Tyr-$^{1}L_{a}$ | 0 | -6.4 |
| $^{78}$ Tyr-$^{1}L_{b5}$ | $^{78}$ Tyr-$^{1}L_{a}$ | 0 | -4.1 |
| $^{78}$ Tyr-$^{1}L_{b3}$ | $^{78}$ Tyr-$^{1}L_{a}$ | 0 | -3.8 |
| $^{78}$ Tyr-$^{1}L_{b8}$ | $^{95}$ Trp-$^{1}L_{a2}$ | 8.68 | -1.2 |
| $^{78}$ Tyr-$^{1}L_{b8}$ | $^{95}$ Trp-$^{1}L_{a}$ | 8.68 | -1.3 |
| $^{78}$ Tyr-$^{1}L_{b7}$ | $^{95}$ Trp-$^{1}L_{a2}$ | 8.68 | -1.2 |
| $^{78}$ Tyr-$^{1}L_{b7}$ | $^{95}$ Trp-$^{1}L_{a}$ | 8.68 | -1.3 |
| $^{78}$ Tyr-$^{1}L_{b6}$ | $^{95}$ Trp-$^{1}L_{a5}$ | 8.68 | -1.4 |
| $^{78}$ Tyr-$^{1}L_{b6}$ | $^{95}$ Trp-$^{1}L_{a4}$ | 8.68 | -1.6 |
| $^{78}$ Tyr-$^{1}L_{b6}$ | $^{95}$ Trp-$^{1}L_{a3}$ | 8.68 | -1.5 |
| $^{78}$ Tyr-$^{1}L_{b6}$ | $^{95}$ Trp-$^{1}L_{a2}$ | 8.68 | -2.5 |
| $^{78}$ Tyr-$^{1}L_{b6}$ | $^{95}$ Trp-$^{1}L_{a}$ | 8.68 | -2.7 |
| $^{78}$ Tyr-$^{1}L_{b5}$ | $^{95}$ Trp-$^{1}L_{a4}$ | 8.68 | -1.0 |
| $^{78}$ Tyr-$^{1}L_{b5}$ | $^{95}$ Trp-$^{1}L_{a3}$ | 8.68 | -1.0 |
| $^{78}$ Tyr-$^{1}L_{b5}$ | $^{95}$ Trp-$^{1}L_{a2}$ | 8.68 | -1.6 |
| $^{78}$ Tyr-$^{1}L_{b5}$ | $^{95}$ Trp-$^{1}L_{a}$ | 8.68 | -1.7 |
| $^{78}$ Tyr-$^{1}L_{b4}$ | $^{95}$ Trp-$^{1}L_{a7}$ | 8.68 | -1.4 |
| $^{78}$ Tyr-$^{1}L_{b4}$ | $^{95}$ Trp-$^{1}L_{a6}$ | 8.68 | -1.4 |
| $^{78}$ Tyr-$^{1}L_{b4}$ | $^{95}$ Trp-$^{1}L_{a5}$ | 8.68 | -2.3 |
| $^{78}$ Tyr-$^{1}L_{b4}$ | $^{95}$ Trp-$^{1}L_{a4}$ | 8.68 | -2.6 |
| $^{78}$ Tyr-$^{1}L_{b4}$ | $^{95}$ Trp-$^{1}L_{a3}$ | 8.68 | -2.5 |
| $^{78}$ Tyr-$^{1}L_{b4}$ | $^{95}$ Trp-$^{1}L_{a2}$ | 8.68 | -4.1 |
| $^{78}$ Tyr-$^{1}L_{b4}$ | $^{95}$ Trp-$^{1}L_{a}$ | 8.68 | -4.4 |
| $^{78}$ Tyr-$^{1}L_{b3}$ | $^{95}$ Trp-$^{1}L_{a4}$ | 8.68 | -1.0 |

| | | | |
|---|---|---|---|
| $^{78}$ Tyr-$^{1}L_{b3}$ | $^{95}$ Trp-$^{1}L_{a2}$ | 8.68 | -1.5 |
| $^{78}$ Tyr-$^{1}L_{b3}$ | $^{95}$ Trp-$^{1}L_{a}$ | 8.68 | -1.6 |
| $^{78}$ Tyr-$^{1}L_{b2}$ | $^{95}$ Trp-$^{1}L_{a9}$ | 8.68 | -1.0 |
| $^{78}$ Tyr-$^{1}L_{b2}$ | $^{95}$ Trp-$^{1}L_{a7}$ | 8.68 | -1.9 |
| $^{78}$ Tyr-$^{1}L_{b2}$ | $^{95}$ Trp-$^{1}L_{a6}$ | 8.68 | -1.8 |
| $^{78}$ Tyr-$^{1}L_{b2}$ | $^{95}$ Trp-$^{1}L_{a5}$ | 8.68 | -3.0 |
| $^{78}$ Tyr-$^{1}L_{b2}$ | $^{95}$ Trp-$^{1}L_{a4}$ | 8.68 | -3.4 |
| $^{78}$ Tyr-$^{1}L_{b2}$ | $^{95}$ Trp-$^{1}L_{a3}$ | 8.68 | -3.2 |
| $^{78}$ Tyr-$^{1}L_{b2}$ | $^{95}$ Trp-$^{1}L_{a2}$ | 8.68 | -5.3 |
| $^{78}$ Tyr-$^{1}L_{b2}$ | $^{95}$ Trp-$^{1}L_{a}$ | 8.68 | -5.7 |
| $^{78}$ Tyr-$^{1}L_{b}$ | $^{95}$ Trp-$^{1}L_{a7}$ | 8.68 | -1.7 |
| $^{78}$ Tyr-$^{1}L_{b}$ | $^{95}$ Trp-$^{1}L_{a6}$ | 8.68 | -1.6 |
| $^{78}$ Tyr-$^{1}L_{b}$ | $^{95}$ Trp-$^{1}L_{a5}$ | 8.68 | -2.7 |
| $^{78}$ Tyr-$^{1}L_{b}$ | $^{95}$ Trp-$^{1}L_{a4}$ | 8.68 | -3.0 |
| $^{78}$ Tyr-$^{1}L_{b}$ | $^{95}$ Trp-$^{1}L_{a3}$ | 8.68 | -2.9 |
| $^{78}$ Tyr-$^{1}L_{b}$ | $^{95}$ Trp-$^{1}L_{a2}$ | 8.68 | -4.8 |
| $^{78}$ Tyr-$^{1}L_{b}$ | $^{95}$ Trp-$^{1}L_{a}$ | 8.68 | -5.1 |
| $^{78}$ Tyr-$^{1}L_{a}$ | $^{95}$ Trp-$^{1}L_{b7}$ | 8.68 | -3.7 |
| $^{78}$ Tyr-$^{1}L_{a}$ | $^{95}$ Trp-$^{1}L_{b6}$ | 8.68 | -2.0 |
| $^{78}$ Tyr-$^{1}L_{a}$ | $^{95}$ Trp-$^{1}L_{b5}$ | 8.68 | -7.3 |
| $^{78}$ Tyr-$^{1}L_{a}$ | $^{95}$ Trp-$^{1}L_{b4}$ | 8.68 | -5.1 |
| $^{78}$ Tyr-$^{1}L_{a}$ | $^{95}$ Trp-$^{1}L_{b3}$ | 8.68 | -9.8 |
| $^{60}$ Trp-$^{1}L_{b}$ | $^{60}$ Trp-$^{1}L_{a8}$ | 0 | -9.4 |
| $^{60}$ Trp-$^{1}L_{b7}$ | $^{60}$ Trp-$^{1}L_{a9}$ | 0 | -3.2 |
| $^{60}$ Trp-$^{1}L_{b7}$ | $^{60}$ Trp-$^{1}L_{a8}$ | 0 | -2.5 |
| $^{60}$ Trp-$^{1}L_{b7}$ | $^{60}$ Trp-$^{1}L_{a7}$ | 0 | -6.1 |
| $^{60}$ Trp-$^{1}L_{b7}$ | $^{60}$ Trp-$^{1}L_{a6}$ | 0 | -5.8 |
| $^{60}$ Trp-$^{1}L_{b7}$ | $^{60}$ Trp-$^{1}L_{a5}$ | 0 | -9.7 |
| $^{60}$Trp-$^{1}L_{b6}$ | $^{60}$ Trp-$^{1}L_{a9}$ | 0 | -1.7 |
| $^{60}$ Trp-$^{1}L_{b6}$ | $^{60}$ Trp-$^{1}L_{a}$ | 0 | -9.7 |
| $^{60}$ Trp-$^{1}L_{b6}$ | $^{60}$ Trp-$^{1}L_{a8}$ | 0 | -1.3 |
| $^{60}$ Trp-$^{1}L_{b6}$ | $^{60}$ Trp-$^{1}L_{a7}$ | 0 | -3.2 |

| | | | |
|---|---|---|---|
| $^{60}$ Trp-$^{1}L_{b6}$ | $^{60}$ Trp-$^{1}L_{a6}$ | 0 | -3.0 |
| $^{60}$ Trp-$^{1}L_{b6}$ | $^{60}$ Trp-$^{1}L_{a5}$ | 0 | -5.1 |
| $^{60}$ Trp-$^{1}L_{b6}$ | $^{60}$ Trp-$^{1}L_{a4}$ | 0 | -5.8 |
| $^{60}$ Trp-$^{1}L_{b6}$ | $^{60}$ Trp-$^{1}L_{a3}$ | 0 | -5.4 |
| $^{60}$ Trp-$^{1}L_{b6}$ | $^{60}$ Trp-$^{1}L_{a2}$ | 0 | -9.1 |
| $^{60}$ Trp-$^{1}L_{b5}$ | $^{60}$ Trp-$^{1}L_{a9}$ | 0 | -6.2 |
| $^{60}$ Trp-$^{1}L_{b5}$ | $^{60}$ Trp-$^{1}L_{a8}$ | 0 | -5.0 |
| $^{60}$ Trp-$^{1}L_{b4}$ | $^{60}$ Trp-$^{1}L_{a9}$ | 0 | -4.3 |
| $^{60}$ Trp-$^{1}L_{b4}$ | $^{60}$ Trp-$^{1}L_{a8}$ | 0 | -3.5 |
| $^{60}$ Trp-$^{1}L_{b4}$ | $^{60}$ Trp-$^{1}L_{a7}$ | 0 | -8.3 |
| $^{60}$ Trp-$^{1}L_{b4}$ | $^{60}$ Trp-$^{1}L_{a6}$ | 0 | -7.9 |
| $^{60}$ Trp-$^{1}L_{b3}$ | $^{60}$ Trp-$^{1}L_{a9}$ | 0 | -8.3 |
| $^{60}$ Trp-$^{1}L_{b3}$ | $^{60}$ Trp-$^{1}L_{a8}$ | 0 | -6.7 |
| $^{60}$ Trp-$^{1}L_{b2}$ | $^{60}$ Trp-$^{1}L_{a9}$ | 0 | -8.9 |
| $^{60}$ Trp-$^{1}L_{b2}$ | $^{60}$ Trp-$^{1}L_{a8}$ | 0 | -7.1 |
| $^{95}$ Trp-$^{1}L_{b7}$ | $^{95}$ Trp-$^{1}L_{a9}$ | 0 | -1.5 |
| $^{95}$ Trp-$^{1}L_{b7}$ | $^{95}$ Trp-$^{1}L_{a8}$ | 0 | -1.2 |
| $^{95}$ Trp-$^{1}L_{b7}$ | $^{95}$ Trp-$^{1}L_{a7}$ | 0 | -2.9 |
| $^{95}$ Trp-$^{1}L_{b7}$ | $^{95}$ Trp-$^{1}L_{a6}$ | 0 | -2.8 |
| $^{95}$ Trp-$^{1}L_{b7}$ | $^{95}$ Trp-$^{1}L_{a5}$ | 0 | -4.7 |
| $^{95}$ Trp-$^{1}L_{b7}$ | $^{95}$ Trp-$^{1}L_{a4}$ | 0 | -5.3 |
| $^{95}$ Trp-$^{1}L_{b7}$ | $^{95}$ Trp-$^{1}L_{a3}$ | 0 | -5.0 |
| $^{95}$ Trp-$^{1}L_{b7}$ | $^{95}$ Trp-$^{1}L_{a2}$ | 0 | -8.3 |
| $^{95}$ Trp-$^{1}L_{b7}$ | $^{95}$ Trp-$^{1}L_{a}$ | 0 | -8.9 |
| $^{95}$ Trp-$^{1}L_{b6}$ | $^{95}$ Trp-$^{1}L_{a7}$ | 0 | -1.5 |
| $^{95}$ Trp-$^{1}L_{b6}$ | $^{95}$ Trp-$^{1}L_{a6}$ | 0 | -1.5 |
| $^{95}$ Trp-$^{1}L_{b6}$ | $^{95}$ Trp-$^{1}L_{a5}$ | 0 | -2.5 |
| $^{95}$ Trp-$^{1}L_{b6}$ | $^{95}$ Trp-$^{1}L_{a4}$ | 0 | -2.8 |
| $^{95}$ Trp-$^{1}L_{b6}$ | $^{95}$ Trp-$^{1}L_{a3}$ | 0 | -2.6 |
| $^{95}$ Trp-$^{1}L_{b6}$ | $^{95}$ Trp-$^{1}L_{a2}$ | 0 | -4.4 |
| $^{95}$ Trp-$^{1}L_{b6}$ | $^{95}$ Trp-$^{1}L_{a}$ | 0 | -4.7 |
| $^{95}$ Trp-$^{1}L_{b5}$ | $^{95}$ Trp-$^{1}L_{a9}$ | 0 | -3.0 |

| | | | |
|---|---|---|---|
| $^{95}$ Trp-$^{1}L_{b5}$ | $^{95}$ Trp-$^{1}L_{a8}$ | 0 | -2.4 |
| $^{95}$ Trp-$^{1}L_{b5}$ | $^{95}$ Trp-$^{1}L_{a7}$ | 0 | -5.8 |
| $^{95}$ Trp-$^{1}L_{b5}$ | $^{95}$ Trp-$^{1}L_{a6}$ | 0 | -5.5 |
| $^{95}$ Trp-$^{1}L_{b5}$ | $^{95}$ Trp-$^{1}L_{a5}$ | 0 | -9.1 |
| $^{95}$ Trp-$^{1}L_{b5}$ | $^{95}$ Trp-$^{1}L_{a3}$ | 0 | -9.7 |
| $^{95}$ Trp-$^{1}L_{b4}$ | $^{95}$ Trp-$^{1}L_{a9}$ | 0 | -2.1 |
| $^{95}$ Trp-$^{1}L_{b4}$ | $^{95}$ Trp-$^{1}L_{a8}$ | 0 | -1.7 |
| $^{95}$ Trp-$^{1}L_{b4}$ | $^{95}$ Trp-$^{1}L_{a7}$ | 0 | -4.0 |
| $^{95}$ Trp-$^{1}L_{b4}$ | $^{95}$ Trp-$^{1}L_{a6}$ | 0 | -3.8 |
| $^{95}$ Trp-$^{1}L_{b4}$ | $^{95}$ Trp-$^{1}L_{a5}$ | 0 | -6.3 |
| $^{95}$ Trp-$^{1}L_{b4}$ | $^{95}$ Trp-$^{1}L_{a4}$ | 0 | -7.2 |
| $^{95}$ Trp-$^{1}L_{b4}$ | $^{95}$ Trp-$^{1}L_{a3}$ | 0 | -6.8 |
| $^{95}$ Trp-$^{1}L_{b3}$ | $^{95}$ Trp-$^{1}L_{a9}$ | 0 | -4.0 |
| $^{95}$ Trp-$^{1}L_{b3}$ | $^{95}$ Trp-$^{1}L_{a8}$ | 0 | -3.2 |
| $^{95}$ Trp-$^{1}L_{b3}$ | $^{95}$ Trp-$^{1}L_{a7}$ | 0 | -7.7 |
| $^{95}$ Trp-$^{1}L_{b3}$ | $^{95}$ Trp-$^{1}L_{a6}$ | 0 | -7.3 |
| $^{95}$ Trp-$^{1}L_{b2}$ | $^{95}$ Trp-$^{1}L_{a9}$ | 0 | -4.3 |
| $^{95}$ Trp-$^{1}L_{b2}$ | $^{95}$ Trp-$^{1}L_{a8}$ | 0 | -3.4 |
| $^{95}$ Trp-$^{1}L_{b2}$ | $^{95}$ Trp-$^{1}L_{a7}$ | 0 | -8.2 |
| $^{95}$ Trp-$^{1}L_{b2}$ | $^{95}$ Trp-$^{1}L_{a6}$ | 0 | -7.8 |
| $^{95}$ Trp-$^{1}L_{b}$ | $^{95}$ Trp-$^{1}L_{a9}$ | 0 | -5.7 |
| $^{95}$ Trp-$^{1}L_{b}$ | $^{95}$ Trp-$^{1}L_{a8}$ | 0 | -4.5 |

Table A.6: Calculated interaction energies of electronic transitions of aromatic side chains in frame 5 of (PDB:1XOA) of the NMR structure

| Chromophore/Transitions | Chromophore/Transitions | Distance (Å) | Interaction Energy ($cm^{-1}$) |
|---|---|---|---|
| $^{12}$Phe-$^{1}L_{a}$ | $^{81}$Phe-$^{1}L_{b}$ | 7.78 | -1.1 |
| $^{12}$ Phe-$^{1}L_{b6}$ | $^{49}$ Tyr-$^{1}L_{a}$ | 8.87 | -2.1 |
| $^{12}$ Phe-$^{1}L_{b5}$ | $^{49}$ Tyr-$^{1}L_{a}$ | 8.87 | 3.1 |
| $^{12}$ Phe-$^{1}L_{b4}$ | $^{49}$ Tyr-$^{1}L_{a}$ | 8.87 | -4.2 |

| | | | |
|---|---|---|---|
| $^{12}$ Phe-$^{1}L_{b}$ | $^{49}$ Tyr-$^{1}L_{a}$ | 8.87 | 6.0 |
| $^{12}$ Phe-$^{1}L_{b3}$ | $^{49}$ Tyr-$^{1}L_{a}$ | 8.87 | 5.5 |
| $^{12}$ Phe-$^{1}L_{b2}$ | $^{49}$ Tyr-$^{1}L_{a}$ | 8.87 | -4.0 |
| $^{12}$ Phe-$^{1}L_{a}$ | $^{49}$ Tyr-$^{1}L_{b8}$ | 8.87 | -1.6 |
| $^{12}$ Phe-$^{1}L_{a}$ | $^{49}$ Tyr-$^{1}L_{b7}$ | 8.87 | -1.6 |
| $^{12}$ Phe-$^{1}L_{a}$ | $^{49}$ Tyr-$^{1}L_{b6}$ | 8.87 | -3.2 |
| $^{12}$ Phe-$^{1}L_{a}$ | $^{49}$ Tyr-$^{1}L_{b5}$ | 8.87 | -2.3 |
| $^{12}$Phe-$^{1}L_{a}$ | $^{49}$ Tyr-$^{1}L_{b4}$ | 8.87 | -5.9 |
| $^{12}$ Phe-$^{1}L_{a}$ | $^{49}$ Tyr-$^{1}L_{b3}$ | 8.87 | -1.2 |
| $^{12}$Phe-$^{1}L_{a}$ | $^{49}$ Tyr-$^{1}L_{b2}$ | 8.87 | -6.8 |
| $^{12}$Phe-$^{1}L_{a}$ | $^{49}$Tyr-$^{1}L_{b}$ | 8.87 | -6.1 |
| $^{27}$ Phe-$^{1}L_{a}$ | $^{27}$ Phe-$^{1}L_{b}$ | 0 | 3.8 |
| $^{27}$ Phe-$^{1}L_{a}$ | $^{27}$ Phe-$^{1}L_{b3}$ | 0 | 3.1 |
| $^{27}$ Phe-$^{1}L_{b2}$ | $^{27}$ Phe-$^{1}L_{a}$ | 0 | 2.4 |
| $^{27}$ Phe-$^{1}L_{b4}$ | $^{27}$ Phe-$^{1}L_{a}$ | 0 | 1.3 |
| $^{27}$ Phe-$^{1}L_{a}$ | $^{27}$ Phe-$^{1}L_{b5}$ | 0 | 1.6 |
| $^{27}$ Phe-$^{1}L_{b6}$ | $^{27}$ Phe-$^{1}L_{a}$ | 0 | 1.0 |
| $^{27}$ Phe-$^{1}L_{a}$ | $^{27}$ Phe-$^{1}L_{b6}$ | 0 | -1.1 |
| $^{27}$ Phe-$^{1}L_{b5}$ | $^{81}$ Phe-$^{1}L_{a}$ | 0 | -1.5 |
| $^{27}$ Phe-$^{1}L_{a}$ | $^{81}$ Phe-$^{1}L_{b4}$ | 0 | -2.2 |
| $^{27}$Phe-$^{1}L_{a}$ | $^{81}$ Phe-$^{1}L_{b2}$ | 0 | -2.7 |
| $^{27}$ Phe-$^{1}L_{b3}$ | $^{81}$ Phe-$^{1}L_{a}$ | 0 | -2.4 |
| $^{27}$ Phe-$^{1}L_{b}$ | $^{81}$ Phe-$^{1}L_{a}$ | 0 | -3.3 |
| $^{27}$ Phe-$^{1}L_{b}$ | $^{70}$ Tyr-$^{1}L_{b2}$ | 8.90 | -10.0 |
| $^{27}$ Phe-$^{1}L_{b}$ | $^{70}$ Tyr-$^{1}L_{a}$ | 8.90 | 6.2 |
| $^{81}$Phe-$^{1}L_{a}$ | $^{81}$ Phe-$^{1}L_{b2}$ | 0 | 2.6 |
| $^{81}$ Phe-$^{1}L_{a}$ | $^{81}$ Phe-$^{1}L_{b4}$ | 0 | 2.0 |
| $^{81}$ Phe-$^{1}L_{b}$ | $^{81}$ Phe-$^{1}L_{b2}$ | 0 | 1.5 |
| $^{81}$Phe-$^{1}L_{b2}$ | $^{81}$ Phe-$^{1}L_{b}$ | 0 | 1.2 |
| $^{81}$ Phe-$^{1}L_{b}$ | $^{81}$ Phe-$^{1}L_{a}$ | 0 | 2.4 |
| $^{81}$Phe-$^{1}L_{b3}$ | $^{81}$ Phe-$^{1}L_{b2}$ | 0 | 1.9 |
| $^{81}$ Phe-$^{1}L_{b2}$ | $^{81}$ Phe-$^{1}L_{b3}$ | 0 | 3.2 |

| | | | |
|---|---|---|---|
| $^{81}$ Phe-$^{1}L_{b}$ | $^{81}$ Phe-$^{1}L_{b4}$ | 0 | 2.8 |
| $^{81}$ Phe-$^{1}L_{b4}$ | $^{81}$ Phe-$^{1}L_{b}$ | 0 | 1.2 |
| $^{81}$ Phe-$^{1}L_{b3}$ | $^{81}$ Phe-$^{1}L_{a}$ | 0 | 2.6 |
| $^{81}$ Phe-$^{1}L_{a}$ | $^{70}$ Tyr-$^{1}L_{b2}$ | 7.52 | 5.1 |
| $^{81}$ Phe-$^{1}L_{a}$ | $^{70}$ Tyr-$^{1}L_{b}$ | 7.52 | 6.0 |
| $^{81}$ Phe-$^{1}L_{b}$ | $^{31}$ Trp-$^{1}L_{a}$ | 8.86 | -14.0 |
| $^{49}$ Tyr-$^{1}L_{b8}$ | $^{49}$ Tyr-$^{1}L_{a}$ | 0 | -3.3 |
| $^{49}$ Tyr-$^{1}L_{b7}$ | $^{49}$ Tyr-$^{1}L_{a}$ | 0 | -3.3 |
| $^{49}$ Tyr-$^{1}L_{b6}$ | $^{49}$ Tyr-$^{1}L_{a}$ | 0 | -6.7 |
| $^{49}$ Tyr-$^{1}L_{b5}$ | $^{49}$ Tyr-$^{1}L_{a}$ | 0 | -4.3 |
| $^{49}$ Tyr-$^{1}L_{b3}$ | $^{49}$ Tyr-$^{1}L_{a}$ | 0 | -4.1 |
| $^{49}$ Tyr-$^{1}L_{b8}$ | $^{70}$ Tyr-$^{1}L_{b}$ | 7.44 | -8.0 |
| $^{49}$ Tyr-$^{1}L_{b8}$ | $^{70}$ Tyr-$^{1}L_{b2}$ | 7.44 | -11.1 |
| $^{49}$Tyr-$^{1}L_{b7}$ | $^{70}$ Tyr-$^{1}L_{b}$ | 7.44 | -1.0 |
| $^{49}$ Tyr-$^{1}L_{b7}$ | $^{70}$ Tyr-$^{1}L_{b2}$ | 7.44 | -1.1 |
| $^{49}$ Tyr-$^{1}L_{b6}$ | $^{70}$ Tyr-$^{1}L_{b6}$ | 7.44 | -1.2 |
| $^{49}$ Tyr-$^{1}L_{b6}$ | $^{70}$ Tyr-$^{1}L_{b4}$ | 7.44 | -11.9 |
| $^{49}$ Tyr-$^{1}L_{b6}$ | $^{70}$ Tyr-$^{1}L_{b}$ | 7.44 | -6.3 |
| $^{49}$ Tyr-$^{1}L_{b6}$ | $^{70}$Tyr-$^{1}L_{b2}$ | 7.44 | -2.4 |
| $^{49}$Tyr-$^{1}L_{b6}$ | $^{70}$ Tyr-$^{1}L_{a}$ | 7.44 | 10.4 |
| $^{49}$ Tyr-$^{1}L_{b5}$ | $^{70}$ Tyr-$^{1}L_{b4}$ | 7.44 | -1.1 |
| $^{49}$ Tyr-$^{1}L_{b5}$ | $^{70}$ Tyr-$^{1}L_{b}$ | 7.44 | -8.4 |
| $^{49}$ Tyr-$^{1}L_{b5}$ | $^{70}$ Tyr-$^{1}L_{b2}$ | 7.44 | -1.5 |
| $^{49}$ Tyr-$^{1}L_{b4}$ | $^{70}$ Tyr-$^{1}L_{b6}$ | 7.44 | -1.9 |
| $^{49}$ Tyr-$^{1}L_{b4}$ | $^{70}$Tyr-$^{1}L_{b5}$ | 7.44 | -1.0 |
| $^{49}$Tyr-$^{1}L_{b4}$ | $^{70}$ Tyr-$^{1}L_{b4}$ | 7.44 | -13.1 |
| $^{49}$ Tyr-$^{1}L_{b4}$ | $^{70}$ Tyr-$^{1}L_{b}$ | 7.44 | -3.6 |
| $^{49}$ Tyr-$^{1}L_{b4}$ | $^{70}$ Tyr-$^{1}L_{b3}$ | 7.44 | -10.1 |
| $^{49}$ Tyr-$^{1}L_{b4}$ | $^{70}$ Tyr-$^{1}L_{b2}$ | 7.44 | -3.9 |
| $^{49}$ Tyr-$^{1}L_{b4}$ | $^{70}$ Tyr-$^{1}L_{a}$ | 7.44 | 2.3 |
| $^{49}$ Tyr-$^{1}L_{b3}$ | $^{70}$ Tyr-$^{1}L_{b4}$ | 7.44 | -14.8 |
| $^{49}$ Tyr-$^{1}L_{b3}$ | $^{70}$Tyr-$^{1}L_{b}$ | 7.44 | -11.5 |

| | | | |
|---|---|---|---|
| $^{49}$Tyr-$^{1}L_{b3}$ | $^{70}$Tyr-$^{1}L_{b2}$ | 7.44 | -1.4 |
| $^{70}$ Tyr-$^{1}L_{b8}$ | $^{70}$ Tyr-$^{1}L_{a}$ | 0 | -5.3 |
| $^{70}$ Tyr-$^{1}L_{b7}$ | $^{70}$ Tyr-$^{1}L_{a}$ | 0 | -7.4 |
| $^{70}$ Tyr-$^{1}L_{b6}$ | $^{70}$ Tyr-$^{1}L_{a}$ | 0 | -6.9 |
| $^{70}$ Tyr-$^{1}L_{b5}$ | $^{70}$ Tyr-$^{1}L_{a}$ | 0 | -14.7 |
| $^{70}$ Tyr-$^{1}L_{b3}$ | $^{70}$ Tyr-$^{1}L_{a}$ | 0 | -3.9 |
| $^{70}$ Tyr-$^{1}L_{b8}$ | $^{31}$ Trp-$^{1}L_{a2}$ | 9.39 | -1.2 |
| $^{70}$ Tyr-$^{1}L_{b8}$ | $^{31}$ Trp-$^{1}L_{a}$ | 9.39 | -1.3 |
| $^{70}$ Tyr-$^{1}L_{b7}$ | $^{31}$ Trp-$^{1}L_{a2}$ | 9.39 | -10.2 |
| $^{70}$ Tyr-$^{1}L_{b7}$ | $^{31}$ Trp-$^{1}L_{a}$ | 9.39 | -1.3 |
| $^{70}$ Tyr-$^{1}L_{b6}$ | $^{31}$ Trp-$^{1}L_{a5}$ | 9.39 | -4.4 |
| $^{70}$ Tyr-$^{1}L_{b6}$ | $^{31}$ Trp-$^{1}L_{a4}$ | 9.39 | -1.6 |
| $^{70}$ Tyr-$^{1}L_{b6}$ | $^{31}$ Trp-$^{1}L_{a3}$ | 9.39 | -1.8 |
| $^{70}$ Tyr-$^{1}L_{b6}$ | $^{31}$ Trp-$^{1}L_{a2}$ | 9.39 | -2.5 |
| $^{70}$ Tyr-$^{1}L_{b6}$ | $^{31}$ Trp-$^{1}L_{a}$ | 9.39 | -2.7 |
| $^{70}$ Tyr-$^{1}L_{b5}$ | $^{31}$ Trp-$^{1}L_{a4}$ | 9.39 | -8.0 |
| $^{70}$ Tyr-$^{1}L_{b5}$ | $^{31}$ Trp-$^{1}L_{a3}$ | 9.39 | -1.0 |
| $^{70}$ Tyr-$^{1}L_{b5}$ | $^{31}$ Trp-$^{1}L_{a2}$ | 9.39 | -10.6 |
| $^{70}$ Tyr-$^{1}L_{b5}$ | $^{31}$ Trp-$^{1}L_{a}$ | 9.39 | -1.7 |
| $^{70}$ Tyr-$^{1}L_{b4}$ | $^{31}$ Trp-$^{1}L_{a7}$ | 9.39 | -12.4 |
| $^{70}$ Tyr-$^{1}L_{b4}$ | $^{31}$ Trp-$^{1}L_{a6}$ | 9.39 | -1.4 |
| $^{70}$ Tyr-$^{1}L_{b4}$ | $^{31}$ Trp-$^{1}L_{a5}$ | 9.39 | -2.3 |
| $^{70}$ Tyr-$^{1}L_{b4}$ | $^{31}$ Trp-$^{1}L_{a4}$ | 9.39 | -2.4 |
| $^{70}$ Tyr-$^{1}L_{b4}$ | $^{31}$ Trp-$^{1}L_{a3}$ | 9.39 | -2.5 |
| $^{70}$Tyr-$^{1}L_{b4}$ | $^{31}$ Trp-$^{1}L_{a2}$ | 9.39 | -4.1 |
| $^{70}$ Tyr-$^{1}L_{b4}$ | $^{31}$ Trp-$^{1}L_{a}$ | 9.39 | -4.4 |
| $^{70}$ Tyr-$^{1}L_{b3}$ | $^{31}$ Trp-$^{1}L_{a4}$ | 9.39 | -11.0 |
| $^{70}$ Tyr-$^{1}L_{b3}$ | $^{31}$ Trp-$^{1}L_{a2}$ | 9.39 | -1.5 |
| $^{70}$ Tyr-$^{1}L_{b3}$ | $^{31}$ Trp-$^{1}L_{a}$ | 9.39 | -9.6 |
| $^{70}$ Tyr-$^{1}L_{b2}$ | $^{31}$ Trp-$^{1}L_{a9}$ | 9.39 | -1.0 |
| $^{70}$ Tyr-$^{1}L_{b2}$ | $^{31}$ Trp-$^{1}L_{a7}$ | 9.39 | -13.9 |
| $^{70}$ Tyr-$^{1}L_{b2}$ | $^{31}$ Trp-$^{1}L_{a6}$ | 9.39 | -1.8 |

| | | | |
|---|---|---|---|
| $^{70}$Tyr-$^{1}L_{b2}$ | $^{31}$ Trp-$^{1}L_{a5}$ | 9.39 | -8.0 |
| $^{70}$ Tyr-$^{1}L_{b2}$ | $^{31}$ Trp-$^{1}L_{a4}$ | 9.39 | -1.4 |
| $^{70}$Tyr-$^{1}L_{b2}$ | $^{31}$ Trp-$^{1}L_{a3}$ | 9.39 | -13.2 |
| $^{70}$Tyr-$^{1}L_{b2}$ | $^{31}$ Trp-$^{1}L_{a2}$ | 9.39 | -5.3 |
| $^{70}$Tyr-$^{1}L_{b2}$ | $^{31}$ Trp-$^{1}L_{a}$ | 9.39 | -7.5 |
| $^{70}$Tyr-$^{1}L_{b}$ | $^{31}$ Trp-$^{1}L_{a7}$ | 9.39 | -2.7 |
| $^{70}$ Tyr-$^{1}L_{b}$ | $^{31}$ Trp-$^{1}L_{a6}$ | 9.39 | -1.6 |
| $^{70}$ Tyr-$^{1}L_{b}$ | $^{31}$ Trp-$^{1}L_{a5}$ | 9.39 | -12.7 |
| $^{70}$ Tyr-$^{1}L_{b}$ | $^{31}$ Trp-$^{1}L_{a4}$ | 9.39 | -3.0 |
| $^{70}$ Tyr-$^{1}L_{b}$ | $^{31}$ Trp-$^{1}L_{a3}$ | 9.39 | -12.9 |
| $^{70}$ Tyr-$^{1}L_{b}$ | $^{31}$ Trp-$^{1}L_{a2}$ | 9.39 | -8.8 |
| $^{70}$ Tyr-$^{1}L_{b}$ | $^{31}$ Trp-$^{1}L_{a}$ | 9.39 | -5.1 |
| $^{70}$ Tyr-$^{1}L_{a}$ | $^{31}$ Trp-$^{1}L_{b7}$ | 9.39 | -3.7 |
| $^{70}$ Tyr-$^{1}L_{a}$ | $^{31}$ Trp-$^{1}L_{b6}$ | 9.39 | -14.0 |
| $^{70}$ Tyr-$^{1}L_{a}$ | $^{31}$ Trp-$^{1}L_{b5}$ | 9.39 | -3.7 |
| $^{70}$ Tyr-$^{1}L_{a}$ | $^{31}$ Trp-$^{1}L_{b4}$ | 9.39 | -1.5 |
| $^{70}$ Tyr-$^{1}L_{a}$ | $^{31}$ Trp-$^{1}L_{b3}$ | 9.39 | -9.8 |
| $^{28}$ Trp-$^{1}L_{b}$ | $^{28}$ Trp-$^{1}L_{a8}$ | 0 | -9.4 |
| $^{28}$ Trp-$^{1}L_{b7}$ | $^{28}$ Trp-$^{1}L_{a9}$ | 0 | -3.2 |
| $^{28}$ Trp-$^{1}L_{b7}$ | $^{28}$ Trp-$^{1}L_{a8}$ | 0 | -12.5 |
| $^{28}$ Trp-$^{1}L_{b7}$ | $^{28}$ Trp-$^{1}L_{a7}$ | 0 | -8.1 |
| $^{28}$ Trp-$^{1}L_{b7}$ | $^{28}$ Trp-$^{1}L_{a6}$ | 0 | -5.8 |
| $^{28}$ Trp-$^{1}L_{b7}$ | $^{28}$ Trp-$^{1}L_{a5}$ | 0 | -9.7 |
| $^{28}$Trp-$^{1}L_{b6}$ | $^{28}$ Trp-$^{1}L_{a9}$ | 0 | -11.7 |
| $^{28}$ Trp-$^{1}L_{b6}$ | $^{28}$ Trp-$^{1}L_{a}$ | 0 | -9.7 |
| $^{28}$ Trp-$^{1}L_{b6}$ | $^{28}$ Trp-$^{1}L_{a8}$ | 0 | -10.3 |
| $^{28}$ Trp-$^{1}L_{b6}$ | $^{28}$ Trp-$^{1}L_{a7}$ | 0 | -13.2 |
| $^{28}$ Trp-$^{1}L_{b6}$ | $^{28}$ Trp-$^{1}L_{a6}$ | 0 | -4.0 |
| $^{28}$ Trp-$^{1}L_{b6}$ | $^{28}$ Trp-$^{1}L_{a5}$ | 0 | -6.1 |
| $^{28}$ Trp-$^{1}L_{b6}$ | $^{28}$ Trp-$^{1}L_{a4}$ | 0 | -9.8 |
| $^{28}$ Trp-$^{1}L_{b6}$ | $^{28}$ Trp-$^{1}L_{a3}$ | 0 | -5.4 |
| $^{28}$ Trp-$^{1}L_{b6}$ | $^{28}$ Trp-$^{1}L_{a2}$ | 0 | -9.1 |

| | | | |
|---|---|---|---|
| $^{28}$ Trp-$^{1}L_{b5}$ | $^{28}$ Trp-$^{1}L_{a9}$ | 0 | -6.2 |
| $^{28}$ Trp-$^{1}L_{b5}$ | $^{28}$ Trp-$^{1}L_{a8}$ | 0 | -5.0 |
| $^{28}$ Trp-$^{1}L_{b4}$ | $^{28}$ Trp-$^{1}L_{a9}$ | 0 | -14.3 |
| $^{28}$ Trp-$^{1}L_{b4}$ | $^{28}$ Trp-$^{1}L_{a8}$ | 0 | -3.5 |
| $^{28}$ Trp-$^{1}L_{b4}$ | $^{28}$ Trp-$^{1}L_{a7}$ | 0 | -8.3 |
| $^{28}$ Trp-$^{1}L_{b4}$ | $^{28}$ Trp-$^{1}L_{a6}$ | 0 | -7.9 |
| $^{28}$ Trp-$^{1}L_{b3}$ | $^{28}$ Trp-$^{1}L_{a9}$ | 0 | -8.3 |
| $^{28}$ Trp-$^{1}L_{b3}$ | $^{28}$ Trp-$^{1}L_{a8}$ | 0 | -6.7 |
| $^{28}$ Trp-$^{1}L_{b2}$ | $^{28}$ Trp-$^{1}L_{a9}$ | 0 | -8.9 |
| $^{28}$ Trp-$^{1}L_{b2}$ | $^{28}$ Trp-$^{1}L_{a8}$ | 0 | -7.1 |
| $^{31}$ Trp-$^{1}L_{b7}$ | $^{31}$ Trp-$^{1}L_{a9}$ | 0 | -11.5 |
| $^{31}$ Trp-$^{1}L_{b7}$ | $^{31}$ Trp-$^{1}L_{a8}$ | 0 | -3.2 |
| $^{31}$ Trp-$^{1}L_{b7}$ | $^{31}$ Trp-$^{1}L_{a7}$ | 0 | -3.9 |
| $^{31}$ Trp-$^{1}L_{b7}$ | $^{31}$ Trp-$^{1}L_{a6}$ | 0 | -12.8 |
| $^{31}$ Trp-$^{1}L_{b7}$ | $^{31}$ Trp-$^{1}L_{a5}$ | 0 | -4.7 |
| $^{31}$ Trp-$^{1}L_{b7}$ | $^{31}$ Trp-$^{1}L_{a4}$ | 0 | -5.3 |
| $^{31}$ Trp-$^{1}L_{b7}$ | $^{31}$ Trp-$^{1}L_{a3}$ | 0 | -5.0 |
| $^{31}$ Trp-$^{1}L_{b7}$ | $^{31}$ Trp-$^{1}L_{a2}$ | 0 | -8.3 |
| $^{31}$ Trp-$^{1}L_{b7}$ | $^{31}$ Trp-$^{1}L_{a}$ | 0 | -8.9 |
| $^{31}$ Trp-$^{1}L_{b6}$ | $^{31}$ Trp-$^{1}L_{a7}$ | 0 | -11.5 |
| $^{31}$ Trp-$^{1}L_{b6}$ | $^{31}$ Trp-$^{1}L_{a6}$ | 0 | -1.5 |
| $^{31}$ Trp-$^{1}L_{b6}$ | $^{31}$ Trp-$^{1}L_{a5}$ | 0 | -6.5 |
| $^{31}$ Trp-$^{1}L_{b6}$ | $^{31}$ Trp-$^{1}L_{a4}$ | 0 | -5.8 |
| $^{31}$ Trp-$^{1}L_{b6}$ | $^{31}$ Trp-$^{1}L_{a3}$ | 0 | -8.6 |
| $^{31}$ Trp-$^{1}L_{b6}$ | $^{31}$ Trp-$^{1}L_{a2}$ | 0 | -14.4 |
| $^{31}$ Trp-$^{1}L_{b6}$ | $^{31}$ Trp-$^{1}L_{a}$ | 0 | -14.2 |
| $^{31}$ Trp-$^{1}L_{b5}$ | $^{31}$ Trp-$^{1}L_{a9}$ | 0 | -3.0 |
| $^{31}$ Trp-$^{1}L_{b5}$ | $^{31}$ Trp-$^{1}L_{a8}$ | 0 | -5.3 |
| $^{31}$ Trp-$^{1}L_{b5}$ | $^{31}$ Trp-$^{1}L_{a7}$ | 0 | -3.7 |
| $^{31}$ Trp-$^{1}L_{b5}$ | $^{31}$ Trp-$^{1}L_{a6}$ | 0 | -9.8 |
| $^{31}$ Trp-$^{1}L_{b5}$ | $^{31}$ Trp-$^{1}L_{a5}$ | 0 | -11.2 |
| $^{31}$ Trp-$^{1}L_{b5}$ | $^{31}$ Trp-$^{1}L_{a3}$ | 0 | -14.5 |

| | | | |
|---|---|---|---|
| $^{31}$ Trp-$^{1}L_{b4}$ | $^{31}$ Trp-$^{1}L_{a9}$ | 0 | -5.5 |
| $^{31}$ Trp-$^{1}L_{b4}$ | $^{31}$ Trp-$^{1}L_{a8}$ | 0 | -11.4 |
| $^{31}$ Trp-$^{1}L_{b4}$ | $^{31}$ Trp-$^{1}L_{a7}$ | 0 | -14.0 |
| $^{31}$ Trp-$^{1}L_{b4}$ | $^{31}$ Trp-$^{1}L_{a6}$ | 0 | -3.9 |
| $^{31}$ Trp-$^{1}L_{b4}$ | $^{31}$ Trp-$^{1}L_{a5}$ | 0 | -6.2 |
| $^{31}$ Trp-$^{1}L_{b4}$ | $^{31}$ Trp-$^{1}L_{a4}$ | 0 | -7.2 |
| $^{31}$ Trp-$^{1}L_{b4}$ | $^{31}$ Trp-$^{1}L_{a3}$ | 0 | -6.7 |
| $^{31}$ Trp-$^{1}L_{b3}$ | $^{31}$ Trp-$^{1}L_{a9}$ | 0 | -4.4 |
| $^{31}$ Trp-$^{1}L_{b3}$ | $^{31}$ Trp-$^{1}L_{a8}$ | 0 | -13.2 |
| $^{31}$ Trp-$^{1}L_{b3}$ | $^{31}$ Trp-$^{1}L_{a7}$ | 0 | -7.7 |
| $^{31}$ Trp-$^{1}L_{b3}$ | $^{31}$ Trp-$^{1}L_{a6}$ | 0 | -2.8 |
| $^{31}$ Trp-$^{1}L_{b2}$ | $^{31}$ Trp-$^{1}L_{a9}$ | 0 | -14.3 |
| $^{31}$ Trp-$^{1}L_{b2}$ | $^{31}$ Trp-$^{1}L_{a8}$ | 0 | -3.9 |
| $^{31}$ Trp-$^{1}L_{b2}$ | $^{31}$ Trp-$^{1}L_{a7}$ | 0 | -8.5 |
| $^{31}$ Trp-$^{1}L_{b2}$ | $^{31}$ Trp-$^{1}L_{a6}$ | 0 | -7.8 |
| $^{31}$ Trp-$^{1}L_{b}$ | $^{31}$ Trp-$^{1}L_{a9}$ | 0 | -5.1 |
| $^{31}$ Trp-$^{1}L_{b}$ | $^{31}$ Trp-$^{1}L_{a8}$ | 0 | -4.5 |

Table A.7: Calculated interaction energies of electronic transitions of aromatic side chains in frame 33 of (PDB:1E8L) of the NMR structure

| Chromophore/Transitions | Chromophore/Transitions | Distance (Å) | Interaction Energy ($cm^{-1}$) |
|---|---|---|---|
| $^{3}$Phe-$^{1}L_{a}$ | $^{38}$Phe-$^{1}L_{b}$ | 7.87 | -1.3 |
| $^{34}$ Phe-$^{1}L_{a}$ | $^{38}$ Phe-$^{1}L_{b}$ | 4.13 | 4.6 |
| $^{34}$ Phe-$^{1}L_{a}$ | $^{38}$ Phe-$^{1}L_{b3}$ | 4.13 | 4.2 |
| $^{34}$ Phe-$^{1}L_{b2}$ | $^{38}$ Phe-$^{1}L_{a}$ | 4.13 | 2.5 |
| $^{34}$ Phe-$^{1}L_{b4}$ | $^{38}$ Phe-$^{1}L_{a}$ | 4.13 | 1.1 |
| $^{34}$ Phe-$^{1}L_{a}$ | $^{38}$ Phe-$^{1}L_{b5}$ | 4.13 | 1.5 |
| $^{34}$ Phe-$^{1}L_{b6}$ | $^{38}$ Phe-$^{1}L_{a}$ | 4.13 | 2.0 |
| $^{34}$ Phe-$^{1}L_{a}$ | $^{38}$ Phe-$^{1}L_{b6}$ | 4.13 | -1.2 |
| $^{34}$ Phe-$^{1}L_{b5}$ | $^{38}$ Phe-$^{1}L_{a}$ | 4.13 | -2.4 |

| | | | |
|---|---|---|---|
| $^{34}$ Phe-$^{1}L_{a}$ | $^{38}$ Phe-$^{1}L_{b4}$ | 4.13 | -2.4 |
| $^{34}$Phe-$^{1}L_{a}$ | $^{38}$ Phe-$^{1}L_{b2}$ | 4.13 | -2.7 |
| $^{34}$ Phe-$^{1}L_{b3}$ | $^{38}$ Phe-$^{1}L_{a}$ | 4.13 | -3.0 |
| $^{34}$ Phe-$^{1}L_{b}$ | $^{38}$ Phe-$^{1}L_{a}$ | 4.13 | -3.3 |
| $^{38}$ Phe-$^{1}L_{b6}$ | $^{111}$ Trp-$^{1}L_{a5}$ | 9.35 | 1.3 |
| $^{38}$ Phe-$^{1}L_{b6}$ | $^{111}$ Trp-$^{1}L_{a4}$ | 9.35 | 3.4 |
| $^{38}$ Phe-$^{1}L_{b6}$ | $^{111}$ Trp-$^{1}L_{a3}$ | 9.35 | 1.3 |
| $^{38}$ Phe-$^{1}L_{b6}$ | $^{111}$ Trp-$^{1}L_{a2}$ | 9.35 | 12.2 |
| $^{38}$ Phe-$^{1}L_{b6}$ | $^{111}$Trp-$^{1}L_{a}$ | 9.35 | 2.4 |
| $^{38}$ Phe-$^{1}L_{b5}$ | $^{111}$ Trp-$^{1}L_{a7}$ | 9.35 | -2.1 |
| $^{38}$ Phe-$^{1}L_{b5}$ | $^{111}$ Trp-$^{1}L_{a6}$ | 9.35 | -10.9 |
| $^{38}$ Phe-$^{1}L_{b5}$ | $^{111}$ Trp-$^{1}L_{a5}$ | 9.35 | -1.8 |
| $^{38}$ Phe-$^{1}L_{b5}$ | $^{111}$ Trp-$^{1}L_{a4}$ | 9.35 | -2.1 |
| $^{38}$ Phe-$^{1}L_{b5}$ | $^{111}$ Trp-$^{1}L_{a3}$ | 9.35 | -11.9 |
| $^{38}$ Phe-$^{1}L_{b5}$ | $^{111}$Trp-$^{1}L_{a2}$ | 9.35 | -3.2 |
| $^{38}$ Phe-$^{1}L_{b5}$ | $^{111}$ Trp-$^{1}L_{a}$ | 9.35 | -3.5 |
| $^{38}$ Phe-$^{1}L_{b4}$ | $^{111}$ Trp-$^{1}L_{b}$ | 9.35 | -1.1 |
| $^{38}$ Phe-$^{1}L_{b4}$ | $^{111}$ Trp-$^{1}L_{a7}$ | 9.35 | 1.5 |
| $^{38}$ Phe-$^{1}L_{b4}$ | $^{111}$ Trp-$^{1}L_{a6}$ | 9.35 | 11.5 |
| $^{38}$ Phe-$^{1}L_{b4}$ | $^{111}$ Trp-$^{1}L_{a5}$ | 9.35 | 2.5 |
| $^{38}$ Phe-$^{1}L_{b4}$ | $^{111}$ Trp-$^{1}L_{a4}$ | 9.35 | 2.8 |
| $^{38}$ Phe-$^{1}L_{b4}$ | $^{111}$ Trp-$^{1}L_{a3}$ | 9.35 | 2.6 |
| $^{38}$ Phe-$^{1}L_{b4}$ | $^{111}$ Trp-$^{1}L_{a2}$ | 9.35 | 24.4 |
| $^{38}$ Phe-$^{1}L_{b4}$ | $^{111}$ Trp-$^{1}L_{a}$ | 9.35 | 4.7 |
| $^{38}$ Phe-$^{1}L_{b3}$ | $^{111}$ Trp-$^{1}L_{b3}$ | 9.35 | 1.1 |
| $^{38}$ Phe-$^{1}L_{b3}$ | $^{111}$ Trp-$^{1}L_{b2}$ | 9.35 | 1.2 |
| $^{38}$ Phe-$^{1}L_{b3}$ | $^{111}$ Trp-$^{1}L_{b}$ | 9.35 | 1.2 |
| $^{38}$ Phe-$^{1}L_{b3}$ | $^{111}$ Trp-$^{1}L_{a9}$ | 9.35 | -1.1 |
| $^{38}$ Phe-$^{1}L_{b3}$ | $^{111}$ Trp-$^{1}L_{a7}$ | 9.35 | -2.3 |
| $^{38}$ Phe-$^{1}L_{b3}$ | $^{111}$ Trp-$^{1}L_{a6}$ | 9.35 | -2.1 |
| $^{38}$ Phe-$^{1}L_{b3}$ | $^{111}$ Trp-$^{1}L_{a5}$ | 9.35 | -33.5 |
| $^{38}$ Phe-$^{1}L_{b3}$ | $^{111}$ Trp-$^{1}L_{a4}$ | 9.35 | -3.9 |

| | | | |
|---|---|---|---|
| $^{38}$ Phe-$^{1}L_{b3}$ | $^{111}$ Trp-$^{1}L_{a3}$ | 9.35 | -3.7 |
| $^{38}$ Phe-$^{1}L_{b3}$ | $^{111}$ Trp-$^{1}L_{a2}$ | 9.35 | -6.1 |
| $^{38}$ Phe-$^{1}L_{b3}$ | $^{111}$ Trp-$^{1}L_{a}$ | 9.35 | -36.6 |
| $^{38}$ Phe-$^{1}L_{b2}$ | $^{123}$ Trp-$^{1}L_{b2}$ | 10.21 | -1.0 |
| $^{38}$ Phe-$^{1}L_{b2}$ | $^{123}$ Trp-$^{1}L_{b}$ | 10.21 | -1.4 |
| $^{38}$ Phe-$^{1}L_{b2}$ | $^{123}$ Trp-$^{1}L_{a9}$ | 10.21 | 21.0 |
| $^{38}$ Phe-$^{1}L_{b2}$ | $^{123}$ Trp-$^{1}L_{a7}$ | 10.21 | 1.9 |
| $^{38}$ Phe-$^{1}L_{b2}$ | $^{123}$ Trp-$^{1}L_{a6}$ | 10.21 | 1.8 |
| $^{38}$ Phe-$^{1}L_{b2}$ | $^{123}$ Trp-$^{1}L_{a5}$ | 10.21 | 3.0 |
| $^{38}$ Phe-$^{1}L_{b2}$ | $^{123}$ Trp-$^{1}L_{a4}$ | 10.21 | 3.4 |
| $^{38}$ Phe-$^{1}L_{b2}$ | $^{123}$ Trp-$^{1}L_{a3}$ | 10.21 | 34.2 |
| $^{38}$ Phe-$^{1}L_{b2}$ | $^{123}$ Trp-$^{1}L_{a2}$ | 10.21 | 5.4 |
| $^{38}$ Phe-$^{1}L_{b2}$ | $^{123}$ Trp-$^{1}L_{a}$ | 10.21 | 55.8 |
| $^{38}$ Phe-$^{1}L_{b}$ | $^{123}$ Trp-$^{1}L_{b5}$ | 10.21 | 11.0 |
| $^{38}$ Phe-$^{1}L_{b}$ | $^{123}$ Trp-$^{1}L_{b3}$ | 10.21 | 1.3 |
| $^{38}$ Phe-$^{1}L_{b}$ | $^{123}$ Trp-$^{1}L_{b2}$ | 10.21 | 1.4 |
| $^{38}$ Phe-$^{1}L_{b}$ | $^{123}$ Trp-$^{1}L_{b}$ | 10.21 | 31.9 |
| $^{38}$ Phe-$^{1}L_{b}$ | $^{123}$ Trp-$^{1}L_{a9}$ | 10.21 | -1.4 |
| $^{38}$ Phe-$^{1}L_{b}$ | $^{123}$ Trp-$^{1}L_{a8}$ | 10.21 | -1.1 |
| $^{38}$ Phe-$^{1}L_{b}$ | $^{123}$ Trp-$^{1}L_{a7}$ | 10.21 | -2.6 |
| $^{38}$ Phe-$^{1}L_{b}$ | $^{123}$ Trp-$^{1}L_{a6}$ | 10.21 | -62.5 |
| $^{38}$ Phe-$^{1}L_{b}$ | $^{123}$ Trp-$^{1}L_{a5}$ | 10.21 | -4.3 |
| $^{38}$ Phe-$^{1}L_{b}$ | $^{123}$ Trp-$^{1}L_{a4}$ | 10.21 | -4.7 |
| $^{38}$ Phe-$^{1}L_{b}$ | $^{123}$ Trp-$^{1}L_{a3}$ | 10.21 | -4.5 |
| $^{38}$ Phe-$^{1}L_{b}$ | $^{123}$ Trp-$^{1}L_{a2}$ | 10.21 | -7.4 |
| $^{38}$ Phe-$^{1}L_{b}$ | $^{123}$ Trp-$^{1}L_{a}$ | 10.21 | -8.0 |
| $^{38}$ Phe-$^{1}L_{a}$ | $^{123}$Trp-$^{1}L_{b7}$ | 10.21 | -5.1 |
| $^{38}$ Phe-$^{1}L_{a}$ | $^{123}$Trp-$^{1}L_{b6}$ | 10.21 | -2.7 |
| $^{38}$Phe-$^{1}L_{a}$ | $^{123}$ Trp-$^{1}L_{b4}$ | 10.21 | -6.9 |
| $^{28}$ Trp-$^{1}L_{b}$ | $^{28}$ Trp-$^{1}L_{a8}$ | 0 | -29.4 |
| $^{28}$ Trp-$^{1}L_{b7}$ | $^{28}$ Trp-$^{1}L_{a9}$ | 0 | -31.2 |
| $^{28}$ Trp-$^{1}L_{b7}$ | $^{28}$ Trp-$^{1}L_{a8}$ | 0 | -25.5 |

| | | | |
|---|---|---|---|
| $^{28}$ Trp-$^{1}L_{b7}$ | $^{28}$ Trp-$^{1}L_{a7}$ | 0 | -69.1 |
| $^{28}$ Trp-$^{1}L_{b7}$ | $^{28}$ Trp-$^{1}L_{a6}$ | 0 | -55.8 |
| $^{28}$ Trp-$^{1}L_{b7}$ | $^{28}$ Trp-$^{1}L_{a5}$ | 0 | -19.7 |
| $^{28}$Trp-$^{1}L_{b6}$ | $^{28}$ Trp-$^{1}L_{a9}$ | 0 | -11.7 |
| $^{28}$ Trp-$^{1}L_{b6}$ | $^{28}$ Trp-$^{1}L_{a}$ | 0 | -19.7 |
| $^{28}$ Trp-$^{1}L_{b6}$ | $^{28}$ Trp-$^{1}L_{a8}$ | 0 | -10.3 |
| $^{28}$ Trp-$^{1}L_{b6}$ | $^{28}$ Trp-$^{1}L_{a7}$ | 0 | -13.2 |
| $^{28}$ Trp-$^{1}L_{b6}$ | $^{28}$ Trp-$^{1}L_{a6}$ | 0 | -13.0 |
| $^{28}$ Trp-$^{1}L_{b6}$ | $^{28}$ Trp-$^{1}L_{a5}$ | 0 | -15.1 |
| $^{62}$ Trp-$^{1}L_{b6}$ | $^{62}$ Trp-$^{1}L_{a4}$ | 0 | -15.8 |
| $^{62}$ Trp-$^{1}L_{b6}$ | $^{62}$ Trp-$^{1}L_{a3}$ | 0 | -15.4 |
| $^{62}$ Trp-$^{1}L_{b6}$ | $^{62}$ Trp-$^{1}L_{a2}$ | 0 | -19.1 |
| $^{62}$ Trp-$^{1}L_{b5}$ | $^{62}$ Trp-$^{1}L_{a9}$ | 0 | -6.2 |
| $^{62}$ Trp-$^{1}L_{b5}$ | $^{62}$ Trp-$^{1}L_{a8}$ | 0 | -35.0 |
| $^{62}$ Trp-$^{1}L_{b4}$ | $^{62}$ Trp-$^{1}L_{a9}$ | 0 | -34.3 |
| $^{62}$ Trp-$^{1}L_{b4}$ | $^{62}$ Trp-$^{1}L_{a8}$ | 0 | -83.5 |
| $^{62}$ Trp-$^{1}L_{b4}$ | $^{62}$ Trp-$^{1}L_{a7}$ | 0 | -8.3 |
| $^{62}$ Trp-$^{1}L_{b4}$ | $^{62}$ Trp-$^{1}L_{a6}$ | 0 | -70.9 |
| $^{62}$ Trp-$^{1}L_{b3}$ | $^{62}$ Trp-$^{1}L_{a9}$ | 0 | -82.3 |
| $^{62}$ Trp-$^{1}L_{b3}$ | $^{62}$ Trp-$^{1}L_{a8}$ | 0 | -67.7 |
| $^{62}$ Trp-$^{1}L_{b2}$ | $^{62}$ Trp-$^{1}L_{a9}$ | 0 | -88.9 |
| $^{63}$ Trp-$^{1}L_{b2}$ | $^{63}$ Trp-$^{1}L_{a8}$ | 0 | -70.1 |
| $^{63}$ Trp-$^{1}L_{b7}$ | $^{63}$ Trp-$^{1}L_{a9}$ | 0 | -9.5 |
| $^{63}$Trp-$^{1}L_{b7}$ | $^{63}$ Trp-$^{1}L_{a8}$ | 0 | -19.2 |
| $^{63}$ Trp-$^{1}L_{b7}$ | $^{63}$ Trp-$^{1}L_{a7}$ | 0 | -32.9 |
| $^{63}$ Trp-$^{1}L_{b7}$ | $^{63}$ Trp-$^{1}L_{a6}$ | 0 | -12.8 |
| $^{63}$ Trp-$^{1}L_{b7}$ | $^{63}$ Trp-$^{1}L_{a5}$ | 0 | -34.7 |
| $^{63}$ Trp-$^{1}L_{b7}$ | $^{63}$ Trp-$^{1}L_{a4}$ | 0 | -45.3 |
| $^{63}$ Trp-$^{1}L_{b7}$ | $^{63}$ Trp-$^{1}L_{a3}$ | 0 | -35.0 |
| $^{63}$ Trp-$^{1}L_{b7}$ | $^{63}$ Trp-$^{1}L_{a2}$ | 0 | -78.3 |
| $^{63}$ Trp-$^{1}L_{b7}$ | $^{63}$ Trp-$^{1}L_{a}$ | 0 | -58.9 |
| $^{111}$ Trp-$^{1}L_{b6}$ | $^{111}$ Trp-$^{1}L_{a7}$ | 0 | -41.5 |

| | | | |
|---|---|---|---|
| $^{111}$ Trp-$^{1}L_{b6}$ | $^{111}$ Trp-$^{1}L_{a6}$ | 0 | -11.5 |
| $^{111}$ Trp-$^{1}L_{b6}$ | $^{111}$ Trp-$^{1}L_{a5}$ | 0 | -13.5 |
| $^{111}$ Trp-$^{1}L_{b6}$ | $^{111}$ Trp-$^{1}L_{a4}$ | 0 | -23.8 |
| $^{111}$ Trp-$^{1}L_{b6}$ | $^{111}$ Trp-$^{1}L_{a3}$ | 0 | -55.6 |
| $^{111}$ Trp-$^{1}L_{b6}$ | $^{111}$ Trp-$^{1}L_{a2}$ | 0 | -34.4 |
| $^{111}$ Trp-$^{1}L_{b6}$ | $^{111}$ Trp-$^{1}L_{a}$ | 0 | -13.7 |
| $^{111}$ Trp-$^{1}L_{b5}$ | $^{111}$ Trp-$^{1}L_{a9}$ | 0 | -33.0 |
| $^{111}$ Trp-$^{1}L_{b5}$ | $^{111}$ Trp-$^{1}L_{a8}$ | 0 | -45.4 |
| $^{111}$ Trp-$^{1}L_{b5}$ | $^{111}$ Trp-$^{1}L_{a7}$ | 0 | -50.8 |
| $^{111}$ Trp-$^{1}L_{b5}$ | $^{111}$ Trp-$^{1}L_{a6}$ | 0 | -15.5 |
| $^{111}$ Trp-$^{1}L_{b5}$ | $^{111}$ Trp-$^{1}L_{a5}$ | 0 | -90.1 |
| $^{111}$ Trp-$^{1}L_{b5}$ | $^{111}$ Trp-$^{1}L_{a3}$ | 0 | -97.7 |
| $^{111}$ Trp-$^{1}L_{b4}$ | $^{111}$ Trp-$^{1}L_{a9}$ | 0 | -24.1 |
| $^{111}$ Trp-$^{1}L_{b4}$ | $^{111}$ Trp-$^{1}L_{a8}$ | 0 | -19.3 |
| $^{123}$ Trp-$^{1}L_{b4}$ | $^{123}$ Trp-$^{1}L_{a7}$ | 0 | -42.0 |
| $^{123}$ Trp-$^{1}L_{b4}$ | $^{123}$ Trp-$^{1}L_{a6}$ | 0 | -38.8 |
| $^{123}$ Trp-$^{1}L_{b4}$ | $^{123}$ Trp-$^{1}L_{a5}$ | 0 | -63.3 |
| $^{123}$ Trp-$^{1}L_{b4}$ | $^{123}$ Trp-$^{1}L_{a4}$ | 0 | -42.2 |
| $^{123}$ Trp-$^{1}L_{b4}$ | $^{123}$ Trp-$^{1}L_{a3}$ | 0 | -26.8 |
| $^{123}$ Trp-$^{1}L_{b3}$ | $^{123}$ Trp-$^{1}L_{a9}$ | 0 | -74.0 |
| $^{123}$ Trp-$^{1}L_{b3}$ | $^{123}$ Trp-$^{1}L_{a8}$ | 0 | -36.2 |
| $^{123}$ Trp-$^{1}L_{b3}$ | $^{123}$ Trp-$^{1}L_{a7}$ | 0 | -90.7 |
| $^{123}$ Trp-$^{1}L_{b3}$ | $^{123}$ Trp-$^{1}L_{a6}$ | 0 | -70.3 |
| $^{123}$ Trp-$^{1}L_{b2}$ | $^{123}$ Trp-$^{1}L_{a9}$ | 0 | -44.3 |
| $^{123}$ Trp-$^{1}L_{b2}$ | $^{123}$ Trp-$^{1}L_{a8}$ | 0 | -36.4 |
| $^{123}$Trp-$^{1}L_{b2}$ | $^{123}$ Trp-$^{1}L_{a7}$ | 0 | -81.2 |
| $^{123}$ Trp-$^{1}L_{b2}$ | $^{123}$ Trp-$^{1}L_{a6}$ | 0 | -78.8 |
| $^{123}$ Trp-$^{1}L_{b}$ | $^{123}$ Trp-$^{1}L_{a9}$ | 0 | -57.7 |
| $^{123}$ Trp-$^{1}L_{b}$ | $^{123}$ Trp-$^{1}L_{a8}$ | 0 | -84.5 |