

CORE EXCITATION SPECTROSCOPY
OF AMINO ACIDS AND PEPTIDES

INNER SHELL EXCITATION SPECTROSCOPY
OF AMINO ACIDS AND SMALL PEPTIDES

By

Hua Jiang, B.Eng.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfillment of the Requirements

for the Degree of

Master of Science

McMaster University

© Hua Jiang, August 2005

MASTER OF SCIENCE (2005)
(Chemistry)

McMaster University
Hamilton, Ontario

TITLE: Inner Shell Excitation Spectroscopy of Amino Acids and Small Peptides

AUTHOR: Hua Jiang, B.Eng. (Tianjin University)

SUPERVISOR: Professor A. P. Hitchcock

NUMBER OF PAGES: 108, xvi

ABSTRACT

In this thesis, two inner shell excitation spectroscopy techniques, inner shell electron energy loss spectroscopy (ISEELS) and near edge X-ray absorption fine structure (NEXAFS) spectroscopy, were used to measure the C 1s, N 1s and O 1s spectra of the amino acids, glycine, alanine, cysteine, phenylalanine, proline, threonine, tryptophan and the peptides, Gly-Ala, Lys-Trp-Lys (KWK), and Arg-Gly-Asp (RGD). The spectra are analysed with the aide of *ab initio* computations using the GSCF3 method. The characteristic spectral features of the specific side chains of amino acids are identified. Differences in the spectra of the gas and solid are related to differences between the neutral gas phase molecule and the zwitterionic solid form. A rationalization of observations of high degree of variability in the N 1s spectra of amino acids is proposed. The characteristic spectral signatures of peptide bonds have been identified further by comparing the spectra of small peptides to the spectra of their subunit amino acids. A modified “building block” approach is showed to be very useful in modeling the inner shell excitation spectra of peptides through linear combinations of the spectra of the amino acids residues and peptide bonds.

ACKNOWLEDGEMENTS

I would like to express my deep and sincere gratitude to my supervisor, Prof. Adam P. Hitchcock, for his support and scientific guidance. His understanding, encouragement and numerous suggestions have provided a good basis for the present thesis.

I would also like to thank my supervisory committee member, Dr. John Valliant for his suggestions throughout my Master's program, and also to Dr. Kalai Saravanamuttu for reading and commenting on a draft of this thesis.

Thanks to all the former and current lab members: Dr. Glyn Cooper, Dr. Cynthia Morin, Jian Wang, Li Li and Jacob Stewart-Ornstein for their helpful assistance.

I wish to thank my parents, brothers and sister for their encouragement from my homecountry.

Especially, I would like to give my special thanks to my husband Zhenhong whose patient love enabled me to complete this work. Also, thanks my son Kyle since he had to endure my drastically curtailed involvement in his school work which resulted from the long hours I used to spend in my study.

TABLE OF CONTENTS

Preliminaries	Page
Descriptive Note	ii
Abstract	iii
Acknowledgements	iv
Table of contents	v
List of Figures	viii
List of Tables	xiii
List of Abbreviations and Symbols	xv
Chapter 1. INTRODUCTION	1
1.1 Motivation for This Work	1
1.2 Introduction to Amino Acids and Peptides	8
1.3 Inner Shell Excitation Spectroscopy	12
1.4 Near Edge X-ray Absorption Fine Structure (NEXAFS) Spectroscopy	13
1.5 Inner Shell Electron Energy Loss Spectroscopy	15
1.6 Principles of NEXAFS and ISEELS	16
1.7 Comparing ISEELS and NEXAFS	17
1.8 Interpreting Inner Shell Excitation Spectra	22
1.9 Building Block Principle	23
1.10 Scanning Transmission X-ray Microscopy	27
1.11 Outline of Thesis Chapters	28
Chapter 2. EXPERIMENTAL	30

2.1	ISEELS	30
2.2	NEXAFS Spectra	42
Chapter 3. CORE EXCITATION CALCULATIONS		45
3.1	Introduction	45
3.2	GSCF3 Calculation.	46
3.2	The Application of GSCF3 Calculation.	47
Chapter 4. INNER SHELL SPECTRA OF SELECTED AMINO ACIDS		52
4.1	C 1s Spectra of Gly, Ala, Thr, Phe and Trp	52
4.2	N 1s spectra of gly, Ala, Cys, Pro, thr and Phe	61
4.3	O 1s spectra of Gly, Ala, Cys, Thr, Pro, and Phe	65
4.4	S 2p Spectrum of Cys.	70
4.5	Molecular Conformational Dependence of Core Excitation Spectra of Cysteine	71
4.6	Comparison of C 1s and O 1s Spectra of Gas and Solid Phase Pro and Cys	74
4.7	Comparison of the N 1s Spectra of Gas and Solid Phase Pro and Cys	76
Chapter 5. INNER SHELL SPECTRA OF SMALL PEPTIDES.		83
5.1	Introduction	83
5.2	ISEELS and NEXAFS Spectra of Gly-Ala	84
5.3	NEXAFS Spectra of Solid KWK.	88
5.3	NEXAFS Spectra of Solid RGD	93
Chapter 6. SUMMARY AND FUTURE WORK		96

6.1	Summary	97
6.2	Future Work	98
	References	102

LIST OF FIGURES

Chapter 1

Fig. 1.1.1	C 1s, N 1s and O 1s NEXAFS spectra of albumin, collagen, fibroin, serine and fibrinogen recorded using scanning transmission X-ray microscopy at the Advanced Light Source (Berkeley, CA)	7
Fig. 1.2.1	Chemical structure of α -amino acids	8
Fig. 1.2.2	The formation of zwitterions of amino acids	11
Fig. 1.2.3	The formation of a dipeptide	12
Fig. 1.9.1	Demonstration of the building block principle of inner shell excitation spectra. Gas-phase experimental ISEELS C 1s spectra of benzene, alanine and phenylalanine. The sum of the alanine and benzene spectra is also shown [CG&04]	26

Chapter 2

Fig. 2.1.1	Schematic of the McMaster ISEELS spectrometer	31
Fig. 2.1.2	The evolution of the C 1s spectrum of cysteine with temperature of gas cell	36
Fig. 2.1.3	ISEELS_11g Main Display Panel	38
Fig. 2.1.4	The evolution of the C1s spectrum of glycine with the heating time	38
Fig. 2.1.5	The main stack of ISEELS.	41
Fig. 2.2.1	Schematic of the ALS beamline 5.3.2 and STXM.	43

Chapter 4

Fig. 4.1.1	C 1s optical oscillator strength spectra of gaseous glycine, alanine, cysteine, threonine, proline, phenylalanine and tryptophan derived from inner shell electron energy loss spectra (ISEELS) recorded with 2.5 keV final electron energy and 2° scattering angle	53
------------	---	----

Fig. 4.1.2	Calculated C 1s oscillator strength spectra of alanine, threonine, phenylalanine and tryptophan compared to experimental gaseous spectra. The computed spectra are plotted on absolute oscillator strength scales (total for the full molecule) with offsets. The computed energy scales are plotted with a shift relative to experiment of 3.0 eV (Ala), 3.1 eV (Thr), 2.5 eV (Phe) and 2.0 eV(Trp) in order to align the first peak.	54
Fig. 4.2.1	N 1s optical oscillator strength spectra of gaseous glycine, alanine, cysteine, threonine, praline and phenylalanine derived from inner shell electron energy loss spectra (ISEELS) recorded with 2.5 keV final electron energy and 2° scattering angle.	62
Fig. 4.2.2	Calculated N 1s oscillator strength spectra of alanine, threonine, and phenylalanine compared to experimental gaseous spectra. The computed spectra are plotted on absolute oscillator strength scales (total for the full molecule) with offsets. The computed energy scales are plotted with a shift relative to experiment of 3.0 eV (Ala), 2.9 eV (Thr), and 3.2 eV(Phe) in order to align the broaden peak of the computed spectra	63
Fig. 4.3.1	O 1s optical oscillator strength spectra of gaseous glycine, alanine, cysteine, threonine, praline and phenylalanine derived from inner shell electron energy loss spectra (ISEELS)) recorded with 2.5 keV final electron energy and 2° scattering angle.	66
Fig. 4.3.2	Calculated O 1s oscillator strength spectra of alanine, threonine, and phenylalanine compared to experimental gaseous spectra. The computed spectra are plotted on absolute oscillator strength scales (total for the full molecule) with offsets. The computed energy scales are plotted with a shift relative to experiment of 1.0 eV (Ala), 1.1 eV (Thr), and 1.2 eV(Phe) in order to align the 2.0 first peak	67
Fig. 4.4.1	Experimental S 2p oscillator strength spectra of gaseous cysteine recorded by inner shell electron energy loss spectra (ISEELS) recorded with 2.5 keV final electron energy and 2° scattering angle. IP, 169.28 eV is taken from experimental XPS data for CH3SH [BCJ80].	70
Fig. 4.5.1	Computed C 1s spectra of two conformers of cysteine using	

	Kosugi's Gaussian Self-Consistent Field Version 3 (GSCF3). The two conformers of cysteine are shown in this Figure.	72
Fig. 4.5.2	Computed N 1s spectra of two conformers of cysteine using GSCF3.	73
Fig. 4.5.3	Computed O 1s spectra of two conformers of cysteine using GSCF3.	73
Fig. 4.6.1	Comparison of the C 1s ISEELS spectra of gaseous proline recorded by ISEELS and NEXAFS spectra of solid praline recorded by scanning transmission X-ray microscopy(STXM). STXM images at 288.2 eV for the crystalline in the regions measured are displayed on the right	75
Fig. 4.6.2	Comparison of the O 1s ISEELS spectra of gaseous proline recorded by ISEELS and NEXAFS spectra of solid proline recorded by scanning transmission X-ray microscopy(STXM).	75
Fig. 4.7.1	Comparison of the experimental N 1s ISEELS of gaseous proline and NEXAFS spectrum of solid proline, i.e., zwitterionic form of proline. The computed spectra are plotted on absolute oscillator strength scales (total for the full molecule) with offsets. The computed energy scales are plotted with a shift relative to experiment of 4.0 eV in order to align the broad peak	77
Fig. 4.7.2	Comparison of the experimental N 1s ISEELS of gaseous cysteine and NEXAFS spectrum of solid cysteine, i.e., zwitterionic form of cysteine. The computed spectra are plotted on absolute oscillator strength scales (total for the full molecule) with offsets. The computed energy scales are plotted with a shift relative to experiment of 4.0 eV in order to align the broad peak.	80
Chapter 5		
Fig. 5.2.1	Comparison of the ISEELS C 1s spectra of gaseous Gly-Ala with the ISEELS C1s spectra of glycine, alanine, the weighted sum of glycine and alanine, and the NEXASF spectra of condensed Gly-ala. An enlarged section emphasizing the energy region in the 286-292 eV range is shown in the right panel. A STXM image at 288.6 eV for the sample in the region measured is displayed on the right.	85

Fig. 5.2.2	Comparison of the ISEELS N 1s spectra of gaseous Gly-Ala with the ISEELS N1s spectra of glycine, alanine and the NEXAFS spectra of condensed Gly-Ala.	87
Fig. 5.2.3	Comparison of the ISEELS O 1s spectra of gaseous Gly-Ala with the ISEELS N1s spectra of glycine, alanine and the NEXAFS spectra of condensed Gly-Ala.	87
Fig. 5.3.1	C 1s NEXAFS spectra of solid Lys-Trp-Lys (KWK) in comparison to the spectra of its subunits, tryptophan and lysine, and the simulated spectrum of KWK using modified building block principle. Trp* is the trp residue without –COOH and –NH ₂ group. The spectrum of peptide bond is cited from the previous paper [1]. The simulated spectra of KWK is assigned to the spectrum of KWK*, which is the sum of the spectrum of trp*, peptide bond and lys. An enlarged section emphasizing the difference of the spectra of KWK and KWK* is shown in the right panel. STXM images at 288.6 eV for the sample in the regions measured are displayed on the right	89
Fig. 5.3.2	N 1s NEXAFS spectra of solid Lys-Trp-Lys (KWK) in comparison to the spectra of its subunits, tryptophan and lysine.	91
Fig. 5.3.3	O 1s NEXAFS spectra of solid Lys-Trp-Lys (KWK) in comparison to the spectra of its subunits, tryptophan and lysine, and the simulated spectrum of KWK using modified building block principle. The spectrum of peptide bond is cited from the previous paper [GC&04]. The simulated spectra of KWK is assigned to the spectrum of KWK*, which is the sum of the spectrum of peptide bond and lys. The trp residue does not contribute to the simulated spectrum since the trp residue does not contain O atom. An enlarged section emphasizing the difference of the spectra of KWK and KWK* is shown in the right panel	92
Fig. 5.4.1	C 1s NEXAFS spectra of solid Arg-Gly-Asp(RGD), in comparison with the NEXAFS spectra of solid Arg, Gly, Asp, and the simulated spectrum of RGD using modified building block principle. The spectrum of peptide bond is cited from the previous paper [GC&04]. The simulated spectra of RGD is assigned to the spectrum of RGD*, which is the sum of the spectrum of arg, peptide bond and asp. The gly residue does not contribute to the simulated spectrum, while the asp residue contributes since it contains another –COOH group. An	

enlarged section emphasizing the difference of the spectra of RGD and RGD* is shown in the right panel. A STXM image at 288.6 eV for the sample in the region measured is displayed on the right 94

Fig. 5.4.2 The N1s and O1s NEXAFS spectra of solid Arg-Gly-Asp (RGD) recorded by scanning transmission X-ray microscopy at ALS 95

LIST OF TABLES

Chapter 1

Table 1.1.1	Summary of selected literature on inner shell excitation of amino acids	3
Table 1.1.2	Summary of literature on inner shell excitation of peptides and proteins	6
Table 1.2.1	The structures and some physical properties of the 20 natural amino acids	9

Chapter 3

Table 3.2.1	Molecular geometries used for the GSCF3 calculations.	49
-------------	---	----

Chapter 4

Table 4.1.1	Energies, Term Values and proposed assignments for C 1s spectral features of gaseous Ala, Thr, Pro, Phe, and Trp	58
Table 4.2.1	Energies, Term Values and proposed assignments for N 1s spectral features of gaseous Ala, Thr, Phe, Cys and Pro	64
Table 4.3.1	Energies, Term Values and proposed assignments for O 1s spectral features of gaseous Ala, Cys, Thr, Phe and Pro.	68
Table 4.7.1	Selected eigenvalues, oscillator strengths and orbital characters for computed C 1s, N 1s and O 1s core excited states of conformer I and II of gaseous cystiene, solid cysteine, and computed N 1s core excited state of gaseous proline and solid proline.	81

Chapter 5

Table 5.2.1	Energies and Term Values and proposed assignments for C1s, N1s and O 1s spectral features of gaseous Gly-Ala	87
Table 5.3.1	Energies, and proposed assignments for C 1s, N 1s and O 1s	

	spectral features of solid KWK.	93
Table 5.4.1	Energies, and proposed assignments for C 1s, N 1s and O 1s spectral features of solid RGD.	96

LIST OF ABBREVIATIONS AND SYMBOLS

Å	Angstrom
ALS	Advanced Light Source
°C	Degree Celsius
DD	Double Deflector
E	Energy
EELS	Electron Energy Loss Spectroscopy
EXAFS	Extended X-ray Absorption Fine Structure
F	Absorbed Flux
f	Focal Length
FBA	First Born Aproximation
FWHM	Full Width at Half Maximum
GSCF3	Gaussian Self-Consistent Field Version 3
ISEELS	Inner Shell Electron Energy Loss Spectroscopy
I	Transmitted Photon Intensity
I_0	Incident Photon Intensity
IE	Ionization Energy
IP	Ionization Potential
IVO	Improved Virtual Orbital
K	1000
KWK	Lys-Trp-Lys
λ	Wavelength of light
$\mu(E)$	Energy Dependent Mass Absorption Coefficient
LUMO	Lowest Uncupied Molecular Orbital
M	Initial State
M^*	Excited State
MO	Molecular Orbital
NEXAFS	Near Edge X-Ray Absorption Fine Structure

NLS	National Synchrotron Light Source
OD	Optical Density
RGD	Arg-Gly-Asp
ρ	Sample Density
σ	Linear Absorption Coefficient
s	Second
STXM	Scanning Transmission X-Ray Microscopy
t	Sample Thickness
X-PEEM	X-Ray Photoemission Electron Microscopy

CHAPTER 1

INTRODUCTION

This chapter presents the motivation for this work and a general introduction to the properties of amino acids and peptides. The fundamental principles of inner shell excitation spectroscopy including near edge X-ray absorption fine structure (NEXAFS) spectroscopy and inner shell electron energy loss spectroscopy (ISEELS) are then described. The interpretation of inner shell excitation spectra is discussed in detail. This chapter closes with an outline of the whole thesis.

§ 1.1 Motivation for this work

Proteins are the most abundant macromolecules in living cells. They occur in all cells and in all parts of cells. There are thousands of different proteins which differ in their size, shape and functionality. They are constructed from the common 20 amino acids which act as subunits joined by peptide bonds. Each type of protein differs in its sequence and number of amino acids; therefore, it is the sequence of the chemically different side chains that makes each protein distinct. The analysis of proteins including amino acid composition, sequence and partial or complete atomic-resolution structure, have been developed using chromatography, mass spectrometry, NMR, X-ray diffraction and others.

Inner shell excitation by inelastic electron scattering [H00] or X-ray absorption [S92] is a powerful probe of molecular electronic and geometric structures. This technique has been used increasingly in the analysis of biological systems[HM&03,

MH&04, M05]. The first motivation of the research reported in this thesis is to improve our understanding of the inner shell excitation spectra of amino acids in order to get the characteristic features for amino acids through detailed comparative studies of the gas and solid species. Studies to date [KO&02, ZS&05] suggest that the inner shell excitation spectrum is not sufficient to differentiate each amino acid, but that there are characteristic features for different ‘classes’ of amino acid. **Table 1.1.1** summarizes all inner shell excitation studies of amino acids in the gaseous and solid states. Most studies have been done on solids using the Near-Edge X-ray Absorption Fine Structure (NEXAFS) technique. However, the NEXAFS studies, which to date deal only with solid amino acids, have some limitations. First, amino acids exist in different charge states depending on their environment. In the solid state, they exist in zwitterionic form, with ammonium (protonated amino) and carboxylate (deprotonated carboxylic acid) groups [FF86]. NEXAFS spectra are sensitive to the electronic structure changes associated with protonation and deprotonation. In contrast, previous photoemission experiments have shown that the carboxyl and amine functional groups are well preserved for gaseous amino acids [LFC80]. So the measurements of the inner shell spectra of gas phase amino acids – of which only glycine [GC&03, CG&04], alanine [CG&04] and phenylalanine [CG&04] have been published - yield the spectra of the neutral species in which the N-end is amine, and the C-end is a carboxylic acid. The results of this study further confirm the amine/acid structure in the gas phase structures. In addition, the thin film samples which have been examined previously [KO&02] were prepared by spin coating or

Table 1.1.1 Summary of selected literature on inner shell excitation of amino acids

Amino Acids	Reference	Edge	Phase	Technique	Comment
Glycine	[BO&97]	C	solid	NEXAFS	
	[HK&98]	C, N, O	solid/ Cu(110)	NEXAFS	
	[NH&00]	C, N, O	Solid/ Cu(110)	NEXAFS	glycine and analogues on copper
	[TN&01]	O	solid	NEXAFS	
	[GC&03]	C, N, O	gas solid	NEXAFS, ISEELS	comparison of glycine, diglycine and triglycine
	[ZZ&04]	C, N, O	solid	NEXAFS	
	[TK&04]		solid	NEXAFS	Glycine on Al ₂ O ₃
	[CG&04]	C	gas	ISEELS	
	[KT&05]	N, O	solid	NEXAFS	radiation damage
	[JS&05]	C, N, O	gas	ISEELS	
alanine	[CG&04]	C	gas	ISEELS	
	[JS&05]	C, N, O	gas	ISEELS	
arginine	[BO&97]	C	solid	NEXAFS	
arsparagine	[TN&01]	N, O	solid	NEXAFS	
cystiene	[JS&05]	C, N, O	gas, solid	ISEELS NEXAFS	
histidine	[BO&97]	C	solid	NEXAFS	
	[ZZ&05]	C, N, O	solid	XPS	Histidine on Gold
phenylalanine	[BO&97]	C	solid	NEXAFS	
	[JS&05]	C, N, O	gas	ISEELS	
proline	[JS&05]	C, N, O	gas,solid	ISEELS NEXAFS	
	[CG&04]	C	gas	ISEELS	
serine	[TN&01]	O	solid	NEXAFS	
threonine	[JS&05]	C, N, O	gas	ISEELS	
tryptophan	[BO&97]	C	solid	NEXAFS	
	[JS&05]	C	gas	ISEELS	
tyrosine	[PU05]	C, O	solid	NEXAFS	Amino Acid analogues on Au
	[TN&01]	O	solid	NEXAFS	
All 20 amino acids	[KO&02]	C	solid	STXM	TFA solvent contamination
All 20 amino acids	[ZS&05]	C, N, O	solid	NEXAFS	comprehensive

solution casting, both of which can lead to contamination by solvents. Another problem is X-ray beam damage which can modify the spectrum. In contrast, when one studies a gas sample, although individual molecules are also damaged by electron beams or X-rays, they are constantly replaced with undamaged molecules. However, gas phase spectroscopy of amino acids is difficult due to their tendency to decompose when heated [LFC80]. Several amino acids have been found not to decompose upon heating to temperatures which provide sufficient gas density for electron spectroscopy [K76, CH79]. Therefore, investigating inner shell excitation spectroscopy of gaseous amino acids is expected to provide useful complementary information to X-ray absorption studies of condensed phase amino acids. Through comparison of the spectra of gaseous and condensed amino acids, we explore how the zwitterionic structure and intermolecular solid state interactions (e.g. hydrogen bonding) are reflected in inner shell excitation spectra.

Another motivation is to study the effect of peptide bond formation on the inner shell spectra by comparing the spectra of single amino acids with those of small peptides. This essential work extended earlier investigations of peptide bond effects by others who compared the ISEELS and NEXAFS spectra of glycine and diglycine [GC&03, CG&04]. The third goal is to test the building block concept [S92] as applied to peptides, in order to evaluate the validity and possible complications to such an approach. Finally this study contributes to understanding the extent to which the intrinsic inner shell spectral contrast of peptides and proteins might be used for molecule-specific analysis. Where this is

possible, spatially resolved detection of specific peptides in a complex synthetic, or even natural, biological milieu might be carried out using modern high resolution X-ray microscopy techniques.

Previously reported studies of the inner shell spectra of peptides are listed in **Table 1.1.2**. Boese et al [BO&97] measured the NEXAFS C 1s spectra of Gly-Try, His-Phe, Gly-Trp, Arg-Gly, Gly-Gly-Gly and compared the spectra of these peptides with weighted sums of the monomer spectra. They suggested that the building block concept [S92] might apply, namely that the spectrum of a peptide can be reproduced by simply adding the spectral components of the constituent amino acids. Gordon et al. [GC&03] reported the spectra of Gly-gly, Gly-gly-gly and derived the inner shell signature of the peptide bond. Cooper et al [CG&04] studied the C 1s, N 1s and O 1s spectra of Gly-gly by inner shell electron energy loss spectroscopy (ISEELS) and NEXAFS spectroscopy and thereby estimated the spectral signatures of the peptide bond in the gas and solid phases.

Table 1.1.2 Summary of literature on inner shell excitation of peptides and proteins

Ref	Sample	Edge	Phase	Technique	Comment
[BO&97]	Gly-Try, His-Phe, Gly-Trp, Arg-Gly, Gly-Gly- Gly	C	solid	NEXAFS	Building block tested
[GC&03]	diglycine, triglycine	C, N, O	solid gas	NEXAFS, ISEELS	
[CG&04]	diglycine	C, N, O	solid	ISEELS NEXASF	peptide bond spectroscopy
[ZZ&04]	diglycine, triglycine	C, N, O	solid	NEXAFS	
[JS&05]	Gly-ala, KWK, RGD	C, N, O	solid	NEXAFS	
[LS&01]	Albumin	C, N, O	solid	STXM	quantitative densitometry of the protein
[GC&03]	Fibrinogen	C, N, O	gas	NEXAFS	
[HM&04]	Albumin, Fibrinogen	C, N, O	solid	STXM	proteins on polymer
[MH&04]	Fibrinogen	C,N, O	solid	PEEM,	
[HH&05]	Fibroin, Sericine	C, N, O	solid	STXM	Quantitative mapping of fibroin β -sheet orientation in B. Mori cocoon fibers;different spectra of proteins

In addition, this study continues a systematic program of investigation of the inner shell spectroscopy of amino acids and peptides. There are currently considerable discrepancies, especially in the N 1s region, among the spectra of amino acids, peptides and proteins reported in the literature by various groups [CG&04, ZZ&04]. **Figure 1.1.1** shows the NEXAFS C 1s, N 1s and O 1s spectra of some proteins, including fibrinogen [MH&04], albumin [HM&04], collagen, fibroin and sericine, which were recorded with

scanning transmission X-ray microscopy by Hitchcock et al. The C 1s and O 1s spectra of these proteins are quite similar, while the N1s spectra are variable. Therefore, reliable, quantitative gas phase reference spectra of amino acids and peptides could be useful to help understand possible reasons for these discrepancies.

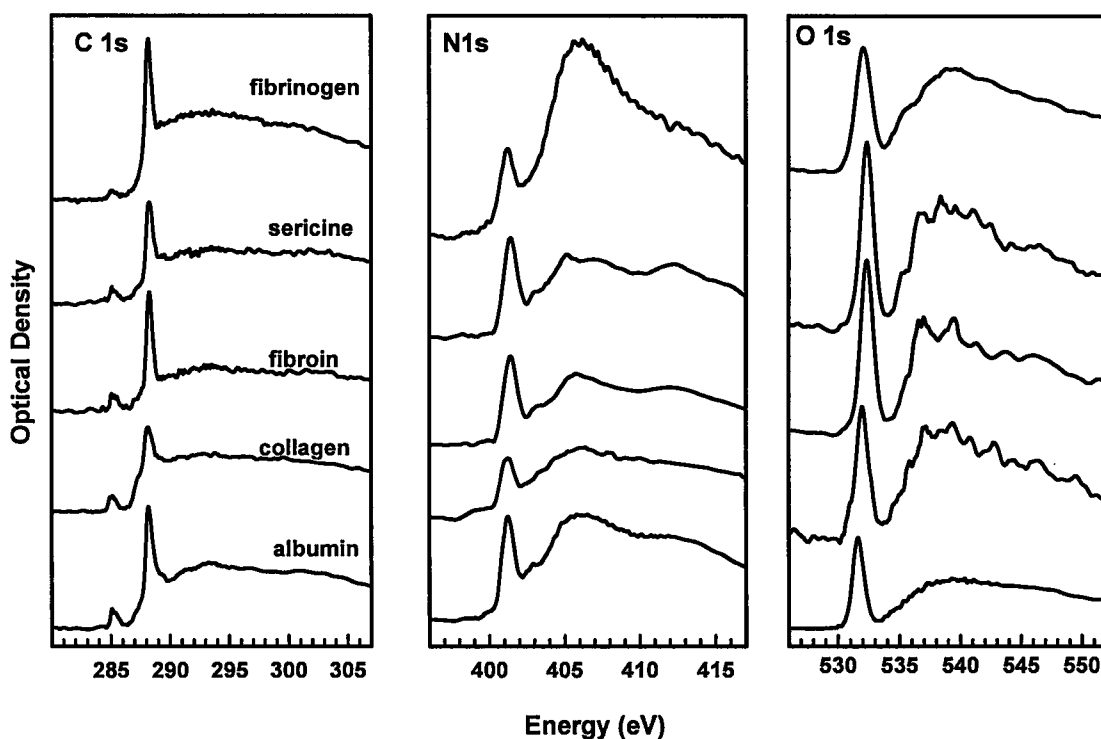


Figure 1.1.1 C 1s, N 1s and O 1s NEXAFS spectra of albumin, collagen, fibroin, sericine and fibrinogen recorded using scanning transmission X-ray microscopy at the Advanced Light Source (Berkeley, CA).

§ 1.2 Introduction to Amino Acids and Peptides

Amino acids are the building blocks of proteins. All organisms contain proteins which are constructed primarily from a set of 20 naturally occurring α -amino acids. Peptides are small proteins consisting of a few (2 to ~20) amino acids.

§ 1.2.1 Chemical Nature of the Amino Acids

The general chemical structure of amino acids is shown in **Figure 1.2.1**. The α -amino acids in peptides and proteins (excluding proline) consist of a carboxylic acid (-COOH) and an amino (-NH₂) functional group attached to the same tetrahedral carbon atom, which is defined as the α -carbon atom. Distinct R-groups, which differ from one amino acid from another, are attached to the α -carbon. All amino acids except glycine (R=H) are chiral. For the sake of convenience and clarity, the names of amino acids are often abbreviated. **Table 1.2.1** lists the structures and abbreviation of the 20 fundamental amino acids [FF86]

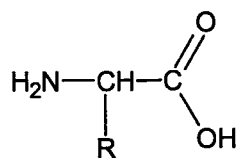
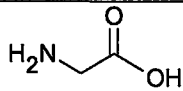
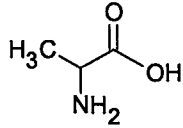
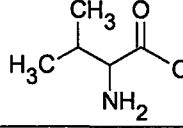
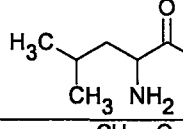
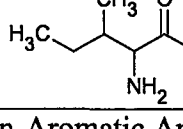
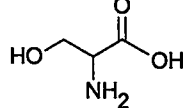
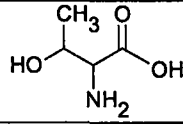
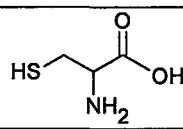
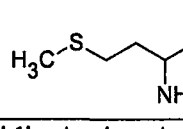
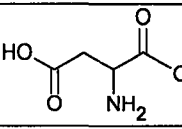
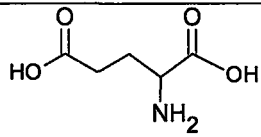
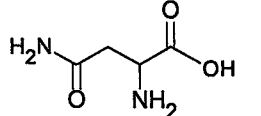
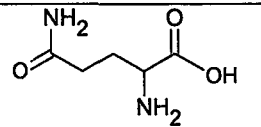
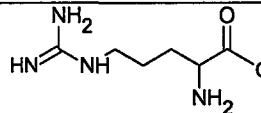
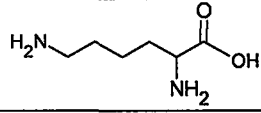
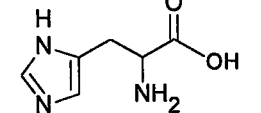
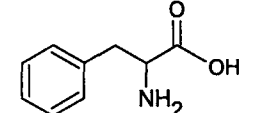
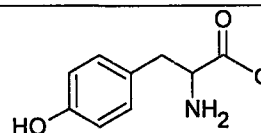
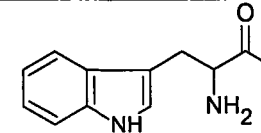
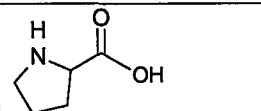


Figure 1.2.1 Structures of α -amino acids

Table 1.2.1 The structures and some physical properties of the 20 natural amino acids

Amino Acids	Symbol	Structure	Isoelectric Point	p <i>K</i> _a (α acid)	p <i>K</i> _a (α amine)
Amino Acids with Aliphatic R-Groups					
Glycine	Gly-G		6.0	2.3	9.8
Alanine	Ala-A		6.0	2.3	9.8
Valine	Val-V		6.0	2.2	9.7
Leucine	Leu-L		6.0	2.3	9.7
Isolucine	Ile-I		6.0	2.3	9.7
Non-Aromatic Amino Acids with -OH					
Serine	Ser-S		5.7	2.2	9.2
Threonine	Thr-T		5.6	2.1	9.1
Amino Acids with Sulfur-Containing R-Groups					
Cysteine	Cys-C		5.0	1.8	10.2
Methionine	Met-M		5.7	2.3	9.2
Acidic Amino Acids and their Amides					
Aspartic Acid	Asp-D		3.0	2.1	9.8

Glutamic Acid	Glu-E		3.2	2.1	9.5
Asparagine	Asn-N		5.4	2.0	8.8
Gluamine	Gln-Q		5.7	2.2	9.1
Basic Amino Acids					
Arginine	Arg-R		10.8	2.0	9.0
Lysine	Lys-K		9.7	2.2	8.9
Histidine	His-H		7.6	1.8	9.1
Amino Acids with Aromatic Rings					
Phenylalanine	Phe-F		5.9	2.6	9.2
Tyrosine	Tyr-Y		5.7	2.2	9.1
Tryptophan	Try-W		5.9	2.4	9.4
Imino Acid					
Proline	Pro-P		6.3	2.0	10.6

In the condensed state, the acidic carboxylic acid and the basic amino group undergo an intramolecular acid-base reaction to form a neutral dipolar ion called a zwitterion [FF86], with the structure shown in **Figure 1.2.2**.

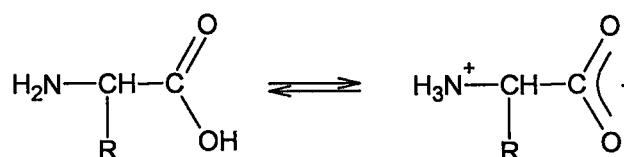


Figure 1.2.2 The formation of zwitterions of amino acids

In a reaction mixture with a low pH, both the amine and the carboxylic acid groups protonate giving the amine group a positive charge. In a reaction mixture with a high pH acidic protons are removed and the amino acid has a net negative charge. **Table 1.2.1** includes the isoelectric point and the pK_a values of the α -acid and α -amine[FF86].

§ 1.2.2 Peptides

Peptide bond formation involves a condensation reaction leading to the polymerization of amino acids into peptides and proteins. The simplest peptide, a dipeptide, contains a single peptide bond formed by the condensation of the carboxyl group of one amino acid with the amino group of the second with the concomitant elimination of water (**Figure 1.2.3**) [FF86]. Each amino acid in a peptide is called a unit, or a residue. The presence of the carbonyl group in a peptide bond allows electron resonance stabilization to occur such that the C-N bond has some double-bond character

due to partial overlap of the *p* orbitals of the carbonyl group with the 2*p* level of the nitrogen. The C-N bond length in the peptide bond is shorter than that of a usual amine C-N single bond [FF86].

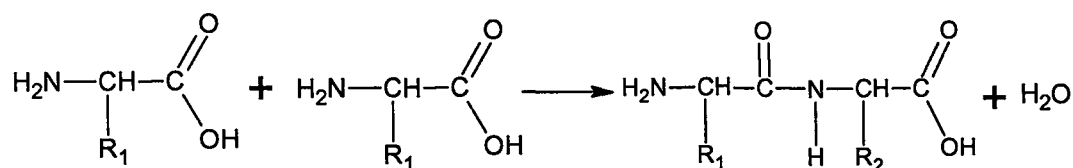


Figure 1.2.3 The formation of a dipeptide

§ 1.3 Inner Shell Excitation Spectroscopy

Inner shell excitation has been widely used in the analysis of complex molecules [HM94, HT&89, IH88, UH&97, UT&97, U97, ZS&05, ZZ&05]. In recent years, it is getting increasingly popular for characterization of biologically relevant systems. Its basis is the excitation of an inner shell electron to unoccupied molecular orbitals. Inner shell excitation spectra can be measured in many different ways. The two used in this work are near edge X-ray absorption fine structure (NEXAFS) [S92] and inner shell electron energy loss spectroscopy (ISEELS) [H00]. Inner shell excitation is easily differentiated from valence excitation by the difference in excitation energies. Valence shell excitation and ionization, which involve the electrons in orbitals primarily constructed from the outer principle quantum number orbitals of the constituent atoms, occur at excitation energies less than approximately 50 eV. Inner shell excitations, on the

other hand, involve excitation of electrons from deeper energy levels. The spectral structures associated with different types of inner shell levels are well separated. In addition to the difference in energies, inner shell excitation spectra are often much easier to interpret than valence shell spectra of the same molecule due in part to the lower congestion, but also because the inner shell electrons are spatially localized on a single atomic core and are very atomic-like in character. This gives rise to a ‘point probe’ character to inner shell excitation. The intensity of a given transition is related to the proportion of dipole-allowed character on the core-excited atom in the unoccupied orbital to which the inner shell electron is excited [IH88, H00].

§ 1.4 Near Edge X-ray Absorption Fine Structure (NEXAFS) Spectroscopy

NEXAFS is a type of photoabsorption spectroscopy. This technique was developed in the 1980’s in conjunction with the development of synchrotron radiation which is used as a high quality tunable X-ray source. Initially NEXAFS referred to application of X-ray absorption to studies of the electronic and geometric structure of molecules bonded to surfaces, in particular organic molecules containing low Z atoms such as hydrogen, carbon, nitrogen, oxygen and fluorine [S92]. However, it is currently the preferred acronym to refer to soft X-ray absorption spectroscopy. NEXAFS excites a specific core level by using photons at the energy of the core excitation edge and probes the electronic structure associated with the local bonding environment of the core excited atom.

§ 1.4.1 X-Ray Absorption and X-ray Absorption Edge

When a beam of X-rays passes through a sample, they are absorbed to an extent which depends on the nature of the substance, the thickness of the sample, and the density of the sample. The intensity of the X-rays passing through the sample is determined by the nature of the sample and the X-ray absorption coefficients of the components. The absorbed photons cause excitation of the inner shell electrons which are either promoted to unoccupied energy levels to form a short lived excited state or they are removed completely to form an ionized state. Traditionally, X-ray absorption spectroscopy was described in terms of absorption edges which occur whenever the energy of the incident photons is just sufficient to ionize the atom, *i.e.* cause excitation of an inner shell electron of the absorbing atom to the continuum ionized state[S92]. Thus, the energies of the absorbed radiation at these edges correspond to the bonding energies of electrons in the K, L, M, etc, shells of the absorbing elements. The threshold energy removing an electron is called its ionization energy (IE) or ionization potential (IP). The absorption edges are labeled in the order of increasing energy, K, L₁, L₂, L₃, M₁, M₂..., corresponding to core ionized states arising from the excitation of an electron to form 1s(²S_{1/2}), 2s(²S_{1/2}), 2p(²P_{1/2}), 2p(²P_{3/2}), 3s(²S_{1/2}) ion cores..., respectively. There is an absorption edge associated with each inner shell energy level of an atom. All elements have an X-ray absorption edge in the soft X-ray energy range (50-2000 eV).

Based on the excitation energy, and thus the kinetic energy of any photoelectron that is produced, X-ray absorption spectroscopy is divided into two parts: NEXAFS and EXAFS (extended X-ray absorption fine structure). NEXAFS spectroscopy refers to the

absorption fine structure close to the absorption edge, in the region between 10 eV below the ionization threshold up to 50eV above the edge. This region usually shows the largest variations and is often dominated by intense, narrow resonances (as much as 10 times stronger than the absorption edge jump). EXAFS [G95] is not described further since it was not studied in this work.

§ 1.5 Inner Shell Electron Energy Loss Spectroscopy

ISEELS is a kind of electron-impact spectroscopy which is based on the analysis of velocities of inelastically scattered electrons following a collision between an electron as the impacting particle and an atom, molecule or solid as the target. When electrons enter the molecular field, they interact with the molecule through electrostatic forces. As a result, some electrons are scattered. There are two categories of scattering: elastic scattering and inelastic scattering. Elastic scattering is defined as the process where no change is observed in the kinetic energy of the scattered electrons after collision. By contrast, in inelastic scattering part of the kinetic energy of the incident electron is lost inside the target giving rise to excitation or ionization of the target. Although the energy loss that is associated with inelastic scattering of electrons was first observed as early as 1914 by *Frank* and *Hertz* [FH14], progress towards electron energy loss spectroscopy was initially very slow. In the 1960's dramatic improvement in vacuum technique and electronic optics technology allowed electron energy loss spectroscopy (EELS) to develop extensively [BB81].

In EELS, an incident beam of mono-energetic electrons of energy E_0 is passed through the sample. The electrons scatter elastically and inelastically after the collision with an atom or molecule in a field-free region. The yield of inelastically scattered electrons is measured with an electron spectrometer as a function of the energy loss at a small and fixed scattering angle θ . The energy distribution of the inelastically scattered electrons gives detailed spectroscopic information about the excited states of the target.

§ 1.6 Principles of NEXAFS and ISEELS

§ 1.6.1 NEXAFS

An X-ray photon of specific energy excites a core-level electron. The electron absorbs the energy of the photo and is excited to an unoccupied orbital. The process is represented by the equation:



The excited state (M^*) is more energetic than the initial state (M) by an amount of energy equal to E . X-ray absorption is a resonant process, so that a transition only occurs when

$$h\nu = E_{M^*} - E_M \quad (1.6.2)$$

In NEXAFS spectroscopy, the X-ray absorption cross section for the excitation or ionization of core electrons is measured.

§ 1.6.2 ISEELS

The basic electron energy loss process can be represented as:



where E_0 is the energy of the incident electrons, E_1 is the residual energy of the inelastically scattered electron beam after exciting a target molecule from the ground state M to an excited state M^* with energy E_n , where $E_n = E_0 - E_1$, is the energy loss, i.e. the excitation energy.

In ISEELS spectroscopy, the flux of inelastically scattered electrons with a certain energy loss is measured. The ISEELS technique is a non-resonant one, in that the incident electron can lose any fraction of its energy to the target molecule. This contrasts strongly with photoabsorption, which is a resonant process in which photons are either absorbed at very specific energies, or are transmitted unabsorbed when the photon energy does not match the transition energy.

§ 1.7 Comparing ISEELS and NEXAFS

These two inner shell excitation techniques have different physical mechanisms. NEXAFS is an optical spectroscopy, in which there are selection rules which forbid certain transitions. ISEELS is an electron spectroscopy, which has different physical interactions and thus different selection rules. By varying the incident energy and/or the scattering angle in ISEELS, it is possible to observe transitions not observed in optical spectroscopy.

In order to compare ISEELS with NEXAFS spectra, the relationship between electron scattering and photoabsorption should be understood. Clearly, the energy loss E_n of the inelastically scattered electron which excites a given transition is equivalent to the resonant photon energy for that transition [S92]. In 1930 *Bethe* [B30] showed that, under certain experimental conditions, the electron energy-loss cross section can be directly related to the X-ray or optical absorption cross section. This relationship can be derived as follows.

The basis of *Bethe's* scattering theory is the First Born Approximation (FBA) and as a result the following is often termed the Bethe-Born theory. The first Born approximation assumes that the interaction between the electron and the target is weak and therefore the incident wave is negligibly distorted by the interaction. The inelastic scattering of a fast electron can be described by evaluating the probability for transition from an initial state $|i\rangle$, to a final state $\langle f|$. Within the First Born Approximation, the differential cross-section $d\sigma/dK$ (DCS) represents the collision probability of an incident electron being scattered by a target molecule, where σ is the cross section and K is the momentum transfer. The number of electrons that have lost an energy E_n in a collision with an atom and are scattered in a direction characterized by the momentum $k_f = k_0 - K$, where k_0 , k_f are wave vectors of the incident electron before and after scattering, is proportional to the double-differential cross section [S92]

$$\frac{d^2\sigma_e}{dE_n dK} = f(K, E_n) \left| \left\langle f \left| \sum_{j=1}^Z e^{iK \cdot r_j} \right| i \right\rangle \right|^2 \quad (1.7.1)$$

where $f(K, E_n)$ is a momentum and energy loss dependent function and the summation is over all electrons with coordinates r_j in the target atom. In the limit $K \cdot r \ll 1$, where r is the radius of the initial core state, the exponential factor in (1.7.1) can be simplified by expansion in a power series:

$$e^{iK \cdot r_j} = 1 + iK \cdot r_j + \frac{1}{2}(iK \cdot r_j)^2 + \dots \quad (1.7.2)$$

Because the states $|i\rangle$ and $\langle f|$ are orthogonal, the matrix element of the first term in the series is zero. The second term is the dipole matrix element; the third term contains the quadruple matrix element and so on. In ISEELS, when the momentum transfer K is small, the dipole term will dominate over the higher order multipole terms. If we define $r = \sum r_j$, and then integrate over the finite angular acceptance of the spectrometer up to a maximum angle θ_m from the incident direction, we get

$$\frac{d\sigma_e}{dE_n} = f(E_0, E_n, \theta_{\max}) \left| \langle f | e_K \cdot r | i \rangle \right|^2 \quad (1.7.3)$$

where E_0 is the incident beam energy and e_K is the unit vector in the direction of the momentum transfer K . By comparing this equation with the expression for X-ray absorption [S92], we get

$$\sigma_x = f(h\nu, k) \left| \langle f | e_E \cdot r | i \rangle \right|^2 \quad (1.7.4)$$

where k is wave vector magnitude of the photoelectron and e_E is the unit vector in the direction of the electric field vector E of the X-rays. The momentum transfer K depends on the scattering angle θ and the energy loss E_n . Its magnitude can be derived by applying the conservation of momentum to the electron-molecule collision [BH81],

$$K^2 = |K|^2 = |k_0 - k_1|^2 = k_0^2 + k_1^2 - 2k_0k_1 \cos \theta \quad (1.7.5)$$

So, the dipole limit, $K \rightarrow 0$, is approached when $\theta \approx 0^\circ$ and the impact velocity is high. In this case, in the limit $K=0$, one obtains,

$$\frac{d\sigma_e}{dE_n} = BE_0(E_n)^{-3} \sigma_x \quad (1.7.6)$$

where the constant B is given by $8mc^2(e^2/\hbar c)$. The equation (1.7.6) provides a quantitative relationship between X-ray photoabsorption (σ) and electron scattering ($d\sigma_e/dE_n$) in small momentum transfer conditions. Under these conditions, the resulting interaction between the electron and the target is weak and the electric dipole process dominates. ISEELS spectra in this thesis were recorded under a fixed final electron energy of 2.5 keV, and thus the impact energy was between 2.8 keV (for C 1s) and 3.0 keV (for O 1s). The scattering angle was small at 2° . Under these conditions, the first (electric-dipole) term in the Bethe-Born expansion dominates for most inner shell transitions, and thus the

resulting spectra exhibit the same features as electric-dipole NEXAFS spectra, with very comparable relative intensities.

With the increasing improvement of synchrotron radiation, NEXAFS has become the technique of choice for the study of inner shell excitation. However, gas phase ISEELS is still a powerful technique with many useful complementary aspects to NEXAFS. Even though it is now the ‘poor cousin’, ISEELS does have some advantages. First, there is the experimental convenience of a home-based instrument. Second, transition energies can often be measured with higher accuracy in ISEELS than NEXAFS. This is because it is easy to measure voltages accurately and thus the transition energy in ISEELS is readily referenced via a voltage scale to the elastic peak (or to a secondary standard, which is a transition in another molecule). In fact the most accurate NEXAFS energy scales are based on calibration to values provided by ISEELS studies [SB84, MI&87] carried out to provide secondary reference standards for use in the soft X-ray spectral region. ISEELS also can provide more accurate intensity scales since it is a non-resonant technique and thus it does not suffer from absorption saturation effects which are a major challenge in optical spectroscopy [CC&91, OC&97]. Finally, ISEELS intensity scales can be converted to absolute photoabsorption oscillator strength scales. To balance the score card though, there are significant advantages to NEXAFS, including much higher spectral resolution, much faster data acquisition for equivalent quality, implementation at high spatial resolution, and ready adaptation to all phases of matter.

§ 1.8 Interpreting Inner Shell Excitation Spectra

The features in inner shell spectra near the ionization potential (IP) correspond to electronic transitions of inner shell electrons to either unfilled molecular or Rydberg orbitals. When the excitation is to unfilled molecular orbitals, the transitions are called valence transitions and typically include core $\rightarrow\pi^*$ and core $\rightarrow\sigma^*$ transitions. The MOs are labeled in terms of π or σ symmetry (perpendicular or along a given bond, respectively) and denote unfilled MOs by an asterisk. For unsaturated molecules the lowest unoccupied MO (LUMO) is usually a π^* orbital, with σ^* orbitals occurring at higher energy. When the excitation is from a core level to a Rydberg orbital, it is called a Rydberg transition. The dominant features in energy loss spectra are attributed to valence excitations. Unfilled π^* orbitals typically lie above the vacuum level in the ground state of neutral molecules, but in the core excited molecule the π^* orbitals are pulled below the vacuum level by Coulomb attraction between the excited electron and the core hole. This Coulombic shift results in a $1s \rightarrow \pi^*$ transition energy which is less than the core level IP. Features that occur below the IP are referred to as “bound” or “discrete” states or resonances, while features that occur above the IP are referred to as “unbound” or “shape” resonances, which appear as broad features superimposed on the signal from direct core level ionization. The $1s \rightarrow \sigma^*$ excitations predicted from a consideration of a minimal basis description of molecular electronic structure often can be matched to observed inner shell shape resonances [LH&02].

In the inner shell excitation spectra of unsaturated molecules, the most pronounced feature by far is the core $\rightarrow \pi^*$ transition, whose energy depends strongly on

the type of atoms which are π -bonded. Thus $\pi^*_{\text{C=C}}$ peaks occur around 285 eV, $\pi^*_{\text{C=N}}$ peaks occur around 287 eV and $\pi^*_{\text{C=O}}$ peaks occur between 286 and 291 eV, depending on the number of heteroatoms that are bonded to the carbonyl [S92, U97]. These shifts are largely related to changes in the position of the core level binding energy or ionization potential (IP), which depends strongly on the oxidation state and electronegativity of the neighbour atoms. For example, fluorination of ethylene shifts the IP of C=C to higher energy by about 5.4 eV compared to ethylene [S92]. This is why π^* resonance positions are a valuable tool for chemical analysis: on an absolute energy scale, their positions mainly depends on the bonded atoms and neighbour atoms. σ^* resonance positions show a stronger and more complicated energy dependence than π^* peaks. The energy positions of σ^* resonances are very sensitive to internuclear distance [S92]. In particular, the energy shifts to higher energy as the bond length decreases. The bond length dependence of σ^* features is larger because the σ orbital lie along the internuclear axis, in contrast to the π^* orbital.

§ 1.9 Building Block Principle

Inner shell excitation spectroscopy is predominantly a local probe of electronic structure at the site of the core excited atom. In a very much oversimplified simplified picture, sometimes called the building block model for inner-shell excitation [S92], the electronic structures of molecules can be viewed as an assembly of the electronic structure of smaller sub-components, such as those of the individual bonds. One may therefore envision that the inner shell spectra of complex molecules are a superposition of

those of smaller building blocks, such as diatomics for each bond. An analogy to a spectroscopy more familiar to most chemists may help. This is the “structural sub-unit” approach to the interpretation of nuclear magnetic resonance (nmr) spectra. In nmr a nucleus of non-zero nuclear spin (I) forms $2I+1$ sublevels when placed in a magnetic field. Thus for spin $1/2$ nuclei such as ^1H , there are two allowed orientations of the nuclear spin vector, a lower energy state with the spin vector parallel to the field and a higher energy state, with the spin vector anti-parallel to the field. Nmr spectra consist of transitions between these nuclear spin energy levels. For a given magnetic field strength (eg. that in a 100 MHz proton spectrometer) the resonant energy, which is in the microwave region, is a strong function of the type of nucleus. Thus in the magnetic field of a 100 MHz nmr machine, ^2D resonates at 15.25 MHz, ^{13}C resonates at 25.15 MHz, ^{31}P resonates at 40.48 MHz etc. [<http://www.webelements.com/webelements/properties/text/definitions/nmr-frequency.html>]. This large scale chemical shift is very analogous to that of the shifts among the various inner shell energy levels (C 1s ~290 eV, N 1s ~410 eV, O 1s ~540 eV, etc). Of course the chemically valuable information in nmr arises from the very small shifts in the resonant frequency associated with electronic structure (specifically the s-electron density at the nucleus). These shifts are of the order of 10-2000 ppm of the nominal resonant frequency, depending on the nucleus. The chemical shifts are very strongly dependent on the nature of the chemical group – e.g. methyl, methylene, aromatic, hydride etc protons all occur at characteristic chemical shifts in proton nmr. A consequence of this “local probe” description of nmr is that the ‘zeroth order’ description of an nmr spectrum starts by linking bands at particular chemical shifts with likely

chemical substructures. Of course there are refinements, such as the J-J coupling pattern, but the basic chemical identification power of nmr is associated with chemical shift in a ‘building block’ sense. The building block principle in inner shell excitation spectroscopy functions in a very similar manner, although the way in which one defines the sub-unit is more closely related to the nature of bonds between atoms or groups (C-C, versus C=C, versus C≡C, etc). Thus, while recognizing that the building block model is not a panacea, it does provide a starting point for interpreting inner shell excitation spectra, to which further refinements can be added as needed.

Although there are well known failures typically related to multiple bond conjugation [S92], the building block approach is frequently used to interpret inner shell excitation spectra. Large molecules can be simply treated as an assembly of diatomic building blocks. In this way, we may therefore imagine that the inner shell excitation spectra of complex molecules are a superposition of the spectra the constituent building blocks. Many experimental results [BO&97, KO&02] have shown a building block approach can work well. However, this approach has some limitations, particularly when there is conjugation among multiple bonded sites as in butadiene ($\text{H}_2\text{C}=\text{CH}-\text{CH}=\text{CH}_2$) with two interacting double bonds, or in aromatic molecules such as benzene [S92].

Furthermore, for much larger molecules composed of various functional subgroups, it is more useful to apply the building block picture to interpret the spectra by addition of the spectral “fingerprint” of the subgroups. An example of this approach, relevant to the interpretation of the inner shell spectra of peptides, is presented in Figure 1.9.1 [CG&04], in which the C 1s spectrum of phenylalanine is compared with the C 1s

spectrum of the alanine, benzene, and the sum of C 1s of alanine and benzene. Phenylalanine can be treated as the molecular sum of alanine and benzene. If the building block principle is applied, the spectrum of phenylalanine should be similar to the sum of the spectra of alanine and benzene. This is indeed the case. The C 1s spectrum of phenylalanine is dominated by the C 1s \rightarrow π^* ring transition at 285.1 eV, which is characteristic for the phenyl ring functional group, and consistent with the C 1s spectrum of benzene. The second peak, at 288.5 eV in phenylalanine, is characteristic of the carboxyl group, and is the same as a corresponding feature in the spectrum of alanine. This example shows that the building block principle for inner shell excitation spectra of complex molecules can work very well.

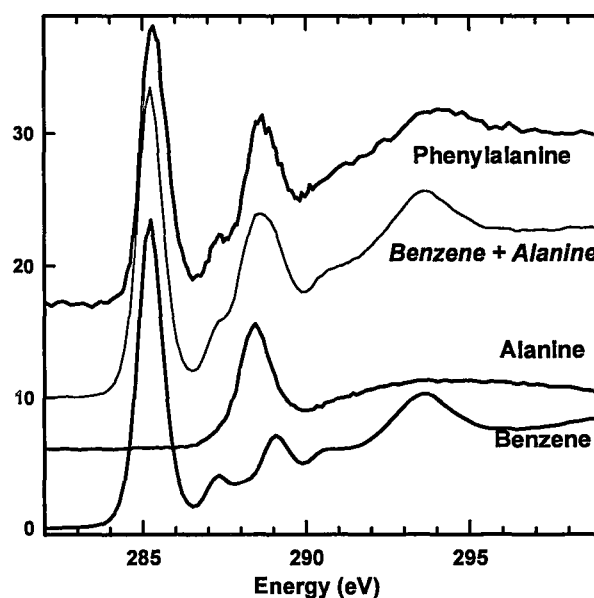


Figure 1.9.1 Demonstration of the building block principle of inner shell excitation spectra. Gas-phase experimental ISEELS C 1s spectra of benzene, alanine and phenylalanine. The sum of the alanine and benzene spectra is also shown [CG&04].

§ 1.10 Scanning Transmission X-ray Microscopy

Scanning Transmission X-ray Microscopy (STXM), developed by Kirz and co-workers at the National Synchrotron Light Source (NSLS) [AZC&92, KJ&95, CJW96], is a synchrotron based, soft X-ray technique which provides chemical speciation at 50 nm spatial resolution based on contrast changes associated with differences in X-ray absorption spectra of the species. STXM performs near-edge X-ray absorption fine structure spectroscopy (NEXAFS) of microscopic areas in thin samples and makes x-ray images based on absorption contrast. In scanning transmission X-ray microscopy, a monochromated soft X-ray beam is focused by a Fresnel zone plate [A99] onto the sample and the transmitted X-rays are detected. Transmission images are obtained by a x-y raster scan of the sample placed at the focal point of the zone plate (z). NEXAFS spectra are obtained at a single point, along a line (linescan spectra) or over all areas of an image (image sequence or ‘stack’ mode [JW&00]) by acquiring the appropriate signal (point, line, or image) at multiple photon energies. The primary signal measured in STXM is transmitted intensity (I) as a function of energy (spectra), or position (images). For quantitative analysis, the transmitted signal is converted to an optical density (OD) according to

$$OD = \ln (I_0/I) \quad (1.10.1)$$

where for a given X-ray energy, I_0 is the incident x-ray flux, I is the transmitted flux through the sample. Based on the Beer-Lambert law [H48],

$$I = I_0 e^{-\mu \rho t} \quad (1.10.2)$$

Therefore, the OD is related to the sample properties by:

$$\text{OD} = \mu(E) \cdot \rho \cdot t \quad (1.10.3)$$

where $\mu(E)$ is the mass absorption coefficient at X-ray energy E , ρ is the density and t is the sample thickness. The mass absorption coefficient can be derived from measurements of the NEXAFS spectra of the pure material. Practically, spectra are obtained in single beam mode, *i.e.*, by first recording an energy scan (I) from the spot of interest and subsequently the incident flux (I_0) measured with the same detector and optical path but with the sample out of the beam. In ideal cases (samples with a hole within 50 μm of the region of interest), the time difference between I and I_0 measurement is less than a second.

§1.11 Outline of Thesis Chapters

Chapter 2 describes the experimental apparatus and outlines NEXAFS and ISEELS procedures used in this thesis. The instrumentation and spectral acquisition of ISEELS is discussed in detail, while NEXAFS is presented only briefly, as these spectra were acquired by my research supervisor, while I participated remotely via the internet. Chapter 3 presents the GSCF3 *ab initio* computational method used to predict the inner shell spectra of selected species to assist interpretation. In Chapter 4, the experimental and computed spectra of selected amino acids including glycine, cysteine, alanine, phenylalanine, proline, threonine and tryptophan are presented and discussed. This analysis includes consideration of possible molecular conformational dependence of core excitation spectra of amino acids, comparison of spectra for gaseous and solid amino

acids, and of the possible origins of the variation of N1s spectra of amino acids. In Chapter 5, the experimental spectra of three small peptides, Gly-Ala, Lys-Trp-Lys (KWK) and Arg-Gly-Asp (RGD) are presented and their analysis in term of building block model is discussed. A summary and proposals for future work are presented in Chapter 6.

CHAPTER 2

EXPERIMENTAL

This chapter outlines the ISEELS instrumentation, with highlights to key features, and spectral acquisition. It also describes the instrument and methods of STXM which was used to measure NEXAFS spectra of samples.

§ 2.1 ISEELS

Inner shell excitation spectra of gaseous amino acids and peptides in the regions of C 1s, N 1s and O 1s were measured using the gas-phase McMaster ISEELS spectrometer that has been described previously [H90, H00]. The instrument was originally build by Steel [S82], and has been modified by Newbury [N86], Wen [W92], and Urquhart [U97]. A schematic of this spectrometer is shown in **Figure 2.1.1**. It consists of an electron source, electron scattering column, collision cell, electrostatic analyzer, and a channeltron for single electron counting. The gun chamber and main chamber are contained in a high vacuum environment evacuated by two diffusion pumps. Helmholtz coils are wound on an external aluminum cube, fixed to the exterior of the spectrometer in order to compensate for the magnetic field of the earth and the structural steel of the building, thereby providing a low magnetic field in the region where the electrons have low kinetic energy in the ISEELS spectrometer (deceleration lens and analyzer region).

McMaster ISEELS Spectrometer

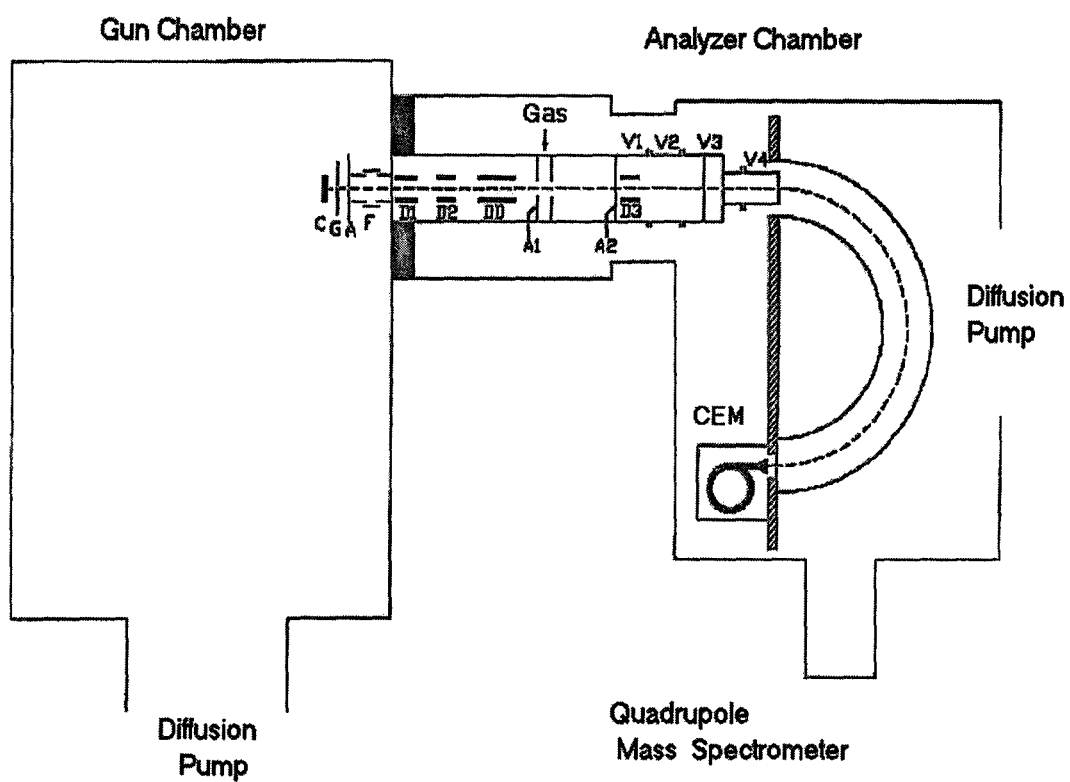


Figure 2.1.1 Schematic of the McMaster ISEELS spectrometer

§ 2.1.1 Electron Source

The electron beam in this spectrometer is generated by an electron gun fabricated for use in black and white televisions. Although designed to operate at much higher potential (25 kV), the gun is operated stably in this spectrometer at 2.5 kV accelerating voltage and gives a parallel, sub-millimeter diameter beam when it is new. These guns are often destroyed by chemically reactive molecules due to gas entering the gun region through the electron beam path. In this work, amino acids, especially cysteine, were found to damage the electron guns.

§ 2.1.2 Electron Scattering Column

The electron beam is accelerated to impact energy by a potential difference, which is 2.5 kV plus the energy loss and pass energy of the analyzer ($2.5 \text{ kV} + E_{\text{loss}} + E_{\text{pass}}$). The electron scattering column consists of a number of quadrupole deflectors, apertures and a four-element deceleration lens. The quadrupole deflectors guide the electron beam to the center of the collision cell. Small angle rather than zero-degree scattering is used to prevent the main electron beam from entering the analyzer and creating a strong background. The incident beam is deflected by the double deflector system DD, which allows the main beam to go through the center of collision cell, but be intercepted at aperture plate, A_2 . DD is twice as long as D2 and located at the halfway point between D2 and collision cell. The beam is first deflected through an angle by D2 and then deflected back in the opposite direction by twice that angle by applying the D2 voltage in the opposite polarity to the DD plate. Typically, the scattered angle is $2^\circ (\pm 1^\circ)$. That

means only electrons inelastically scattered at 2° can pass through the angular selection aperture plate, A_2 . The inelastically scattered electrons are decelerated to the pass energy of the analyzer (E_{pass}) by the lens system consisting of four tube lens elements, V_1 , V_2 , V_3 and V_4 .

§ 2.1.3 Electrostatic Analyzer

The analyzer is a 180° hemispherical electrostatic energy analyzer, which consists of an inner and outer bowl. This analyzer is operated in constant residual energy mode, where the pass energy of the analyzer (E_{pass}) is kept constant and the impact energy ($2.5 \text{ kV} + E_{\text{loss}} + E_{\text{pass}}$) of the electron beam is varied in order to vary the energy loss. Only electrons that enter the analyzer with a kinetic energy equal to the pass energy will travel through the analyzer in a circular path with radius R_{pass} . The pass energy is given by

$$V = V_{\text{inner}} - V_{\text{outer}} = V_{\text{pass}} \left(\frac{R_{\text{outer}}}{R_{\text{inner}}} - \frac{R_{\text{inner}}}{R_{\text{outer}}} \right) \quad (2.1.1)$$

where V_{inner} and V_{outer} are the voltages applied to the analyzer bowls; and R_{outer} and R_{inner} are the radii of the inner and outer hemispherical bowls. For the McMaster ISEELS spectrometer, $R_{\text{outer}} = 11.5 \text{ cm}$ and $R_{\text{inner}} = 7.0 \text{ cm}$. Typically, the pass energies are in the 40 – 50 V range.

The energy resolution (ΔE_{anal}) of the analyser depends on several parameters given in the following formula:

$$\Delta E_{\text{anal}} = E_{\text{pass}} \left(\frac{r_a}{R_{\text{pass}}} + \alpha^2 \right) \quad (2.1.2)$$

where r_a is the radius of the entrance and exit apertures, R_{pass} is the radius of the electron path through the analyser and α is the maximum acceptance half-angle. For the ISEELS analyzer, $r_a = 0.5$ mm and $R_{\text{pass}} = 9.2$ cm. The working resolution of the spectrometer is a function of the electron beam energy bandwidth (ΔE_{gun}) and the resolution of the electrostatic analyzer (ΔE_{anal}) [W92], given by the formula:

$$\Delta E \approx \sqrt{\Delta E_{\text{gun}}^2 + \Delta E_{\text{anal}}^2} \quad (2.1.3)$$

The electron beam bandwidth (ΔE_{gun}) of the BaSrO cathode of the gun currently used in ISEELS is about 0.4 eV (FWHM). Energy resolution is also affected by the total beam current. If the current is too large, space charging effects degrade resolution. Typically, the total energy resolution is 0.7 eV full width half maximum (FWHM) at a beam current of 20 μA and 0.5 eV fwhm at a beam current of ~ 1 μA .

§ 2.1.3 Sample Introduction System

The McMaster ISEELS spectrometer has several ways to introduce samples, including a conventional leak valve line for permanent gases and high vapor pressure liquids, and a solid introduction system. For the latter system, the collision cell can be heated to about 200°C with an internal quartz bulb heater embedded in the wall of the gas cell, which allows measurements of solids with very low vapor pressure, including species with melting points above 300°C.

In this work, 30-50 mg of an amino acid or peptide sample is put in an aluminum or Teflon™ tube attached directly to the gas cell of the spectrometer. The temperature of the gas cell was carefully adjusted to be high enough to provide adequate density in the

excitation region (estimated to be 10^{-6} torr over a path length of 1 cm) but not so high as to cause rapid thermal decomposition or to evaporate the sample over a short period into the apparatus, thereby leading to creation of insulating deposits which required cleaning to recover spectrometer performance. Multiple spectra are recorded at different temperatures to check the evolution of amino acids. At a certain temperature, the peaks become strong enough to record inner shell spectra. **Figure 2.1.2** is an example which shows the evolution of the signal from cysteine with temperature and time. At 127°C, a weak peak at 288.6 eV occurred which indicated the sample started evaporating. This peak became stronger as the temperature increased to 140°C. The optimal condition to obtain a stable, reproducible spectrum for cysteine was found to be 150°C. Above this temperature, however, the sample decomposed. As soon as that began, a strong peak for CO₂ was apparent immediately, as shown in the topmost spectrum. The feature at 285 eV is also related to sample decomposition.

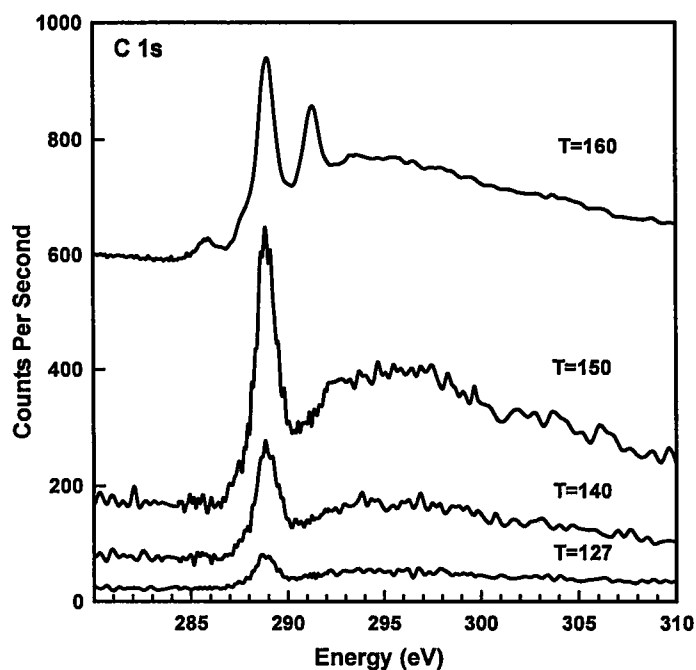


Figure 2.1.2 The evolution of the C 1s spectrum of cysteine with temperature of gas cell

§ 2.1.4 Spectral Acquisition

ISEELS spectral acquisition is controlled by a custom Labview program operated on a personal computer with Windows 98 operating system. It was originally written by Hitchcock, updated in 2001 by Bob Lessard and further updated in 2004 by Jacob Stewart-Ornstein. The front panel of current version of the acquisition code is ISEELS_11g, shown in **Figure 2.1.3**. In this Labview 7 program, the spectra can be saved at intervals in order to track any shifting of the machine performance or changes of the sample. The save interval was usually every 5 scans, so as to check whether or when the sample decomposed. **Figure 2.1.4** is an example of the time history for a measurement of the C 1s spectrum of glycine. Only some of the spectra, which were saved at every 5

scans, are shown in this figure. The weak peak in the first spectrum indicated the temperature of the collision cell had just reached the point at which the sample started to evaporate. In the middle 3 hour period, 7 – 10 pm, the spectra were very stable. After ~ 3 hours of stable acquisition, the main 288.6 eV peak decreased in intensity which was caused by depletion of the sample. Also in the last few spectra, a very large change occurred in the spectrum, due to the onset of thermal decomposition. The last spectrum only shows a very weak peak which indicates the sample had run out. All of the stable scans prior to the onset of decomposition were added to get the actual spectrum of the glycine sample from this particular run. In most cases the same spectrum would be recorded 3 or more times, either in the same charge or with multiple sample loadings. The final spectra for each species reported in this thesis are the sum of all similar spectra of the same edge.

The energy scales of all spectra were calibrated by acquiring the spectra of a mixture of analyte and a suitable reference compound. The C 1s and O 1s spectra of all molecules were calibrated with the C $1s \rightarrow \pi^*$ transition of CO₂ at 290.74(4) eV [BD&82, SB84] and the O $1s \rightarrow \pi^*$ transition of CO₂ at 535.4 (2) eV [HI87]. The N 1s spectra were calibrated with the N $1s \rightarrow \pi^*$ transition of N₂ at 401.10 (2) eV [BD&82, SB84].

The excitation spectra associated with a particular inner shell edge are isolated from the underlying valence-shell and core ionization continua by subtracting a smooth curve determined from a curve fit of the function $a(E-b)^c$ to the pre-edge experimental signal. The background subtracted spectra are converted to absolute oscillator strength scales using previously described methods which involve a curved background

subtraction, correction for kinematic factors relating electron scattering to photoabsorption, and normalization of the far continuum to atomic oscillator strength [HM94].

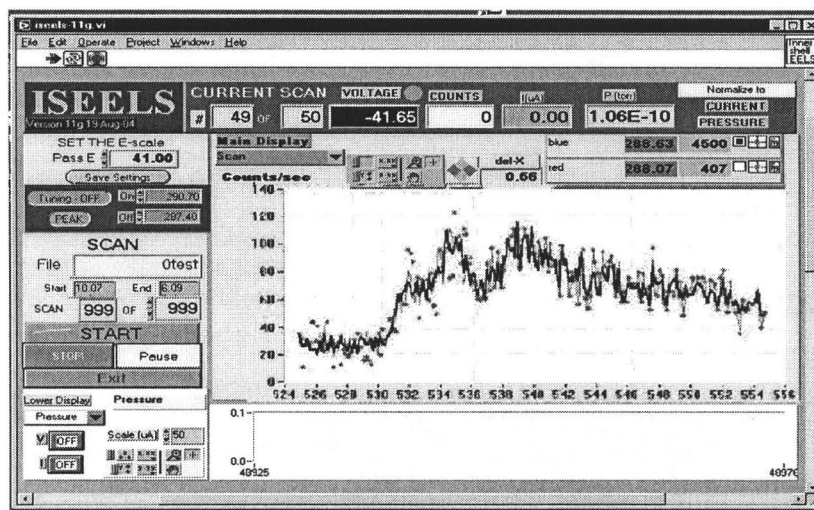


Figure 2.1.3 ISEELS_11g Main Display Panel

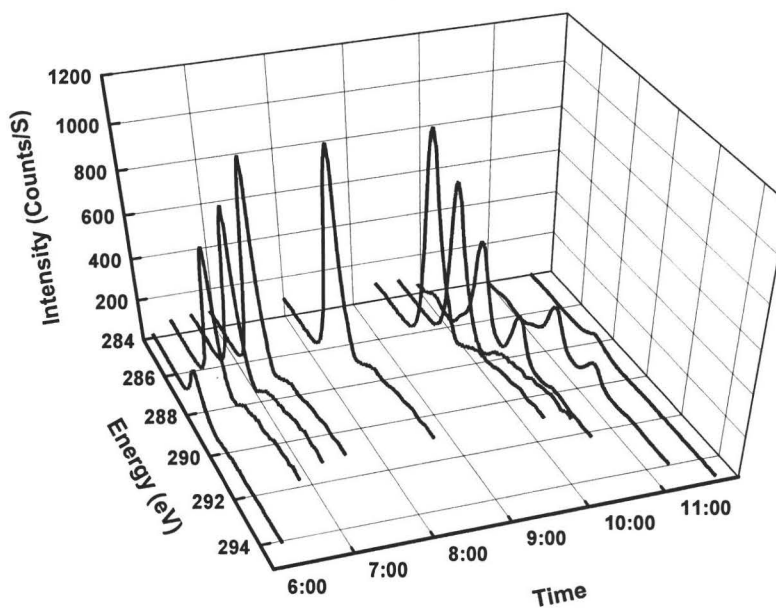


Figure 2.1.4 The evolution of the C 1s spectrum of glycine with the heating time

§ 2.1.5 ISEELS Instrumental Upgrades and Maintenance

During the last two years, several upgrades to the ISEELS spectrometer were performed to improve its performance and reliability.

Due to the tendency of amino acids to decompose when heated, ISEELS experiments in this work are complicated. In order to investigate inner shell excitation of more amino acids and obtain higher quality spectra for undamaged amino acids, the methods were explored to increase volatility and reduce decomposition. First, the aluminum sample container was replaced with a TeflonTM tube. For testing the effectiveness of this replacement, the same experimental conditions were kept including the same condition of the instrument and the same position of the quartz bulb. For tryptophan, the first signal was obtained at 150 °C with aluminum tube, while the first signal was obtained at 140°C with the TeflonTM tube. This suggests that it is easier for amino acids to evaporate using a TeflonTM tube rather than an aluminum tube. Furthermore, the TeflonTM tube kept the amino acids without decomposition when heated for a longer time than aluminum. This may be because aluminum can act as a catalyst to speed up the decomposition of the amino acids.

Another method to improve the volatility of amino acids was to explore the potential of high surface area materials, when coated by amino acids, to increase their volatility. Molecular sieves and carbon black, materials possessing high surface area, and the amino acid glycine, chosen for its simplicity and well established spectra, were used to study this method. By comparing the experimental results, molecular sieves and carbon black are unsuitable as high surface area materials for studying amino acids in the gas

phase. The problem with molecular sieves is that they are very apt to absorb water. Even when some preprocessing of molecular sieves was done, like heating, it is hard to avoid exposure to moist air when loading the sample, as the collision cell is under atmosphere pressure. No usable spectra were obtained using molecular sieves. Another possible barrier to their use might be adhesion of amino acids to the sieves, preventing their vaporization. Furthermore, the molecular sieves seemed to break down after long periods of heating. Carbon black did somewhat better. It at least provided a reasonable glycine signal, but also has a fatal flaw, its strong tendency to release CO₂ on heating. This makes obtaining reliable carbon or oxygen spectra extremely difficult. Both materials failed to increase volatility of glycine and so provided no benefit. However, it does illustrate some of the problems faced by this research. It is difficult to find a material that possesses all the necessary properties, high surface area, resistance to long periods of heating, low water absorption, low adhesion to amino acids and capable of producing a clean background. Unless such a material can be found it is not worthwhile pursuing the project to increase volatility by increasing surface area.

A manual for the ISEELS spectrometer was written by Tulumello in 2002. Through dealing with many experimental details and solving many problems in ISEELS, the troubleshooting section for the manual was summarized and written by me. The major problems in ISEELS are: no signal on A1 or A2, which are the aperture plates; no signal detected at entrance, or center, or cone of analyzer and others. Troubleshooting solutions were developed. As well, a checklist after cleaning was added in the manual. Due to low volatility of amino acids, the main stack, shown in **Figure 2.1.5**, can become

contaminated after each experiment. Such contamination prevents the smooth operations of ISEELS due to charging, as well as leaving their own spectra behind, interfering with the correct recording of a spectrum of a different compound. To restore stable operation and to have minimal background, I had to clean the stack which consists of hundreds of parts. This involved complete disassembly, removal of all contamination from all of the spectrometer surfaces by scrubbing parts with an abrasive cleaner, rinsing with water and removing abrasive from any holes, polishing surfaces with fine sand paper, rinsing with methanol, then acetone, reassembly of the stack, and testing. It is really a time-consuming work and success strongly depends on experience. At first, I needed at least five days to clean it, and now, after dozens of cleanings, I can finish it in one day.

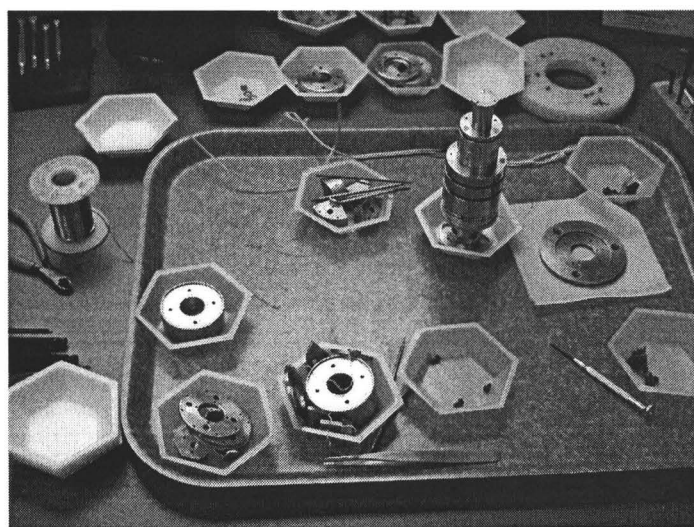
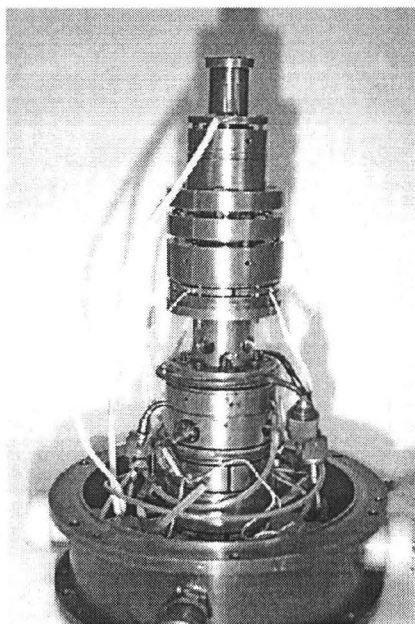


Figure 2.1.5 The Main Stack of ISEELS

Other improvements made to the ISEELS include:

- redesigned the circuit of the beam control box,
- redesigned and rebuilt the gas inlet system,
- rebuilt some cables,
- replacement of the diaphragm valve of the gas inlet system,
- replacement of some parts of the electrometer,
- replacement of the heater of the diffusion pump.

The routine maintenance of ISEELS includes change of the electron gun, change of the oil of the rotary and diffusion pumps, replacement of water line filters and ion gauge filaments and fixing electronics problems of the ISEELS power supplies.

§ 2.2 NEXAFS Spectra

The NEXAFS spectra reported in this work were recorded using the bending magnet scanning transmission X-ray microscope located on beamline 5.3.2 at the Advanced Light Source [WA&02, KT&03].

§ 2.2.1 STXM Microscope and Beamline

Figure 2.2.1 is a sketch of the beamline and STXM at beamline 5.3.2 at the Advanced Light Source.

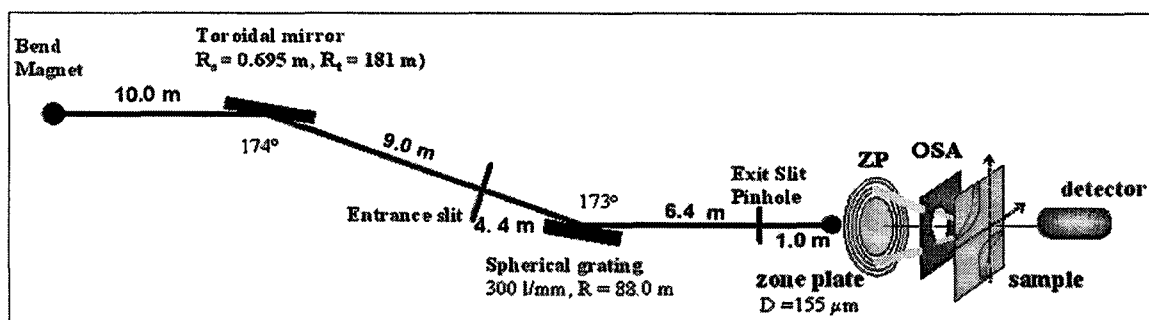


Fig. 2.1.1 Schematic of the ALS beamline 5.3.2 and STXM [WA&02].

Beamline 5.3.2 was optimized for the energy range from 250 to 600 eV, to cover the C 1s, N 1s, and O 1s absorption edges. The actual usable photon energy range is considerably larger – at least 160 – 1160 eV.

§ 2.2.2 Sample Preparation and Spectral Acquisition

The samples were made by dissolving a few crystals (~1 mg) in 1 ml of distilled water. A small drop (~1 μL of that dilute solution) is placed on a Si_3N_4 window and dried with a hot air gun to form a film sufficiently thin (~100-400 nm) to allow X-ray transmission with an optical density of about 1 at the strongest C 1s peak. The regions used to acquire C 1s spectra were typically ~100 nm thick while regions with 300-400 nm thickness were used to record the N 1s and O 1s spectra. Spectra were acquired in image sequence mode [JW&00] since this exposes a single pixel on the sample for a total time of much less than 1 second (typically $200 \times 1 \text{ ms} = 0.2 \text{ s}$, for a 200-energy measurement) and thus has less radiation damage than the linescan mode (typically $200 \times 10 \text{ ms} = 2 \text{ s}$) or point mode (typically $200 \times 200 \text{ ms} = 40 \text{ s}$) of spectral acquisition. The regions of these

samples examined did not exhibit detectable polarization effect (only azimuthal dependence can be probed). Aside from changing optical densities due to thickness variations, the spectra of all regions of the samples were the same. After each stack spectral measurement, the extent of radiation damage was checked by imaging at the $\pi^*_{C=O}$ resonance (the signal which changes most rapidly with damage). Only those measurements with negligible radiation damage were used in generating the final spectra. The NEXAFS spectra were calibrated by measuring the spectrum of gaseous CO₂, N₂ and O₂, introduced into the STXM chamber, and taking the literature values of the sharp Rydberg lines [BD&82, SB84, IH87].

The transmitted X-ray intensity is converted to absorbance (optical density) by equation (1.10.1). In order to compare the NEXAFS and ISEELS spectra quantitatively, one needs to convert the absorbance spectra to absolute oscillator strength spectra. The relationship [HL&82] between atomic cross section, σ (cm²/atom) and mass absorption coefficient μ (E) (cm²/g) is:

$$\sigma = (M / N_0) \mu(E) \quad (2.2.1)$$

where M is atomic weight, N₀ is Avogadro's number, which is 6.02×10²³. Here we define the parameter:

$$\varphi_i = M / N_0 \quad (2.2.2)$$

which converts the NEXAFS absorbance scale to the optical oscillator strength scale for a specific absorption edge. For example, for C 1s, the scale factor is 19.94; for N1s, it is 23.26 and for O1s, it is 26.56.

CHAPTER 3

CORE EXCITATION CALCULATIONS

This chapter presents the ab initio calculation method, Gaussian Self-Consistent Field Version 3 (GSCF3), which is valuable for aiding the interpretation of the inner shell excitation spectra of molecules. The Improved Virtual Orbital (IVO) method, as implemented in GSCF3, is used to perform quantum calculations on core-excited molecules. The geometries and basis sets of the calculated molecules reported in thesis are presented.

§ 3.1 Introduction

In order to aid the spectral assignments and gain insight into how the electronic structure changes are reflected in inner shell excitation spectra, calculations can be performed using Kosugi's Gaussian Self-Consistent Field Version 3 (GSCF3) [KK80, K87], which is an *ab initio* code designed to be highly optimized for computation of inner shell excitation spectra. The program preserves a user specified core hole through variational optimization and can use a wide range of different basis sets. After comparing ground and ionized states in fixed geometry, the inner shell excited states and transition energies are computed with the improved virtual orbital approximation (IVO) [WA69]. This method has been shown to be quite accurate in predicting term values and intensities of core excitations [KK80, K87]. GSCF3 has been applied to the inner shell excitation spectroscopy of many species, including organic molecules [UH&97, GC&04], [UG&05]), organometallics [UT&97] and polymers [UH&97].

§ 3.2 GSCF3 Calculation

Calculations are performed in three steps. In step one, the eigenvectors (MOs) and eigenvalues of the ground state are calculated, and the core MO that will lose the electron is determined. In the second step, the core ion state is computed by removing the user-specified core electron and allowing the system to relax and reorganize in the presence of the core hole. The difference between the total energy of the ground and core-ionized states corresponds to the core level ionization potential (IP). This is predicted with a typical accuracy of 1 eV. The last step provides the core excitation energies and transition probabilities in terms of the IVO approximation [KK80, K87]. In this calculation, the basis sets used are those of Huzinaga et al. [HA&84]. The size of the basis set affects the absolute accuracy of the computed core excitation energies. However, the core state term values are more accurate and relatively independent of basis set choices. The core excitation term values and optical oscillator strengths are generated by the third step of the GSCF3 calculation. The output of GSCF3 is then used to generate simulated core excitation spectra by summing Gaussian lines at an energy given by the term value, an area given by the oscillator strength for excitation to each improved virtual orbital, and a width chosen as a function of the term value. A typical feature width setting is 0.8 eV for term values above 2 eV, 2.0 eV for term values between 2 and -2 eV, 4.0 eV for term values between -2 and -10 eV and 6.0 eV for term values below -10 eV. A separate calculation is performed for each chemically distinct site in each molecule.

Simulated spectra are then generated by summing the components in stoichiometric proportion.

§ 3.3 The Application of GSCF3 Calculation

In this thesis, the inner shell excitation spectra of alanine (C 1s, N 1s and O 1s), threonine (C 1s, N 1s and O 1s), cysteine (C 1s, N 1s and O 1s), phenylalanine (C 1s, N 1s and O 1s), tryptophan (C 1s) and proline (N 1s) were calculated. In our implementation of the GSCF3 method the ground state and the core ionized state (in the same geometry) are computed with mid sized basis sets, followed by generation of the electronic excitation information within the improved virtual orbital (IVO) approximation. The basis set consist of primitive basis functions that were taken from (73/6), (63/5), (6), and (53353) contracted Gaussian-type functions of Huzinaga et al. [HA&84] for core excited carbon, nitrogen, and oxygen atom, the other carbon, nitrogen, oxygen atom, hydrogen atom and sulfur atom, respectively. The contraction scheme was (411121/21111) on the core excited carbon, nitrogen, and oxygen atom, where the polarization functions were $\xi_d=1.34$ and 0.29 for carbon, $\xi_d=0.412$ and 1.986 for nitrogen and $\xi_d=0.535$ and 2.704 for oxygen.

The ground state geometries of alanine, threonine, phenylalanine, tryptophan and proline were obtained from an *ab initio* geometry optimization performed by the Hartree-Fock SCF component of the quantum chemistry software package Spartan [S99] at the 6-31G* basis set level. These geometries are listed in **Table 3.3.1**. Amino acids exist in the gas phase as an equilibrium among many different molecular conformations. We selected two conformers of cysteine to study possible conformation dependence of the inner shell

spectra. In order to study the difference between the N 1s spectra of gas phase and solid phase amino acids, the N 1s spectra of the zwitterion forms of cysteine and proline were calculated. These geometries are listed in **Table 3.3.1**.

Table 3.3.1 Molecular Geometries used for the GSCF3 Calculations**A. Alanine**

Atom	Position(Å)		
C1	-1.0157	0.0053	-0.6406
C2	0.0405	-0.1084	0.4529
C3	0.7416	1.2230	0.6813
H	0.5643	-2.0433	0.0662
H	1.5472	-0.9606	-0.6256
H	-0.4989	-0.3874	1.3536
H	1.4442	1.1188	1.4987
H	0.0311	2.0045	0.9179
H	1.2857	1.5302	-0.2058
H	-2.236	-1.0462	-1.5834
N	1.0126	-1.1581	0.1992
O1	-1.3341	0.9964	-1.2119
O2	-1.5826	-1.1742	-0.9026

B. Cysteine

Atom	Conformer I Position(Å)			Conformer II Position(Å)			Zwitterion Position(Å)			
	C1	0.4318	-1.1192	-0.2854	1.0742	0.6595	-0.3998	0.3659	0.1346	0.4819
C2	-0.3451	0.0352	0.3226	-0.1396	0.3191	0.4505	1.6391	0.4143	-0.4076	
C3	0.2513	1.3836	-0.0723	-0.9629	-0.8722	-0.0611	-0.2796	-1.2258	0.269	
O1	1.4009	-1.0341	-0.9651	1.6169	-0.3849	-1.0131	2.3706	-0.5346	-0.5564	
O2	-0.1106	-2.2892	0.0391	1.5329	1.7516	-0.4855	1.685	1.5949	-0.7831	
N	-1.7084	-0.0507	-0.1852	-0.9886	1.4946	0.473	-0.5453	1.292	0.1997	
S	1.8321	1.7943	0.7287	-0.3784	-2.5305	0.4066	-1.0502	-1.3993	-1.0858	
H	0.3508	1.4466	-1.1454	2.3865	-0.0828	-1.4875	-1.0229	1.6658	0.9965	
H	-0.4288	2.1663	0.2406	0.2677	0.0609	1.4294	-1.2095	1.0569	-0.5238	
H	-0.2814	-0.082	1.404	-1.6148	1.4638	1.2532	0.6718	0.2217	1.517	
H	0.3945	-2.9813	-0.3775	-0.438	2.3273	0.544	-202479	-1.7938	-0.9764	
H	-2.1224	-0.93	0.0558	-1.104	-0.8093	-1.1307	0.5119	-1.9527	0.356	
H	-2.2752	0.6692	0.2198	-1.945	-0.8045	0.3879	-1.0311	-1.4348	1.0192	
H	2.6108	0.9914	0.0205	0.693	-2.5925	-0.367	0.1493	1.9605	-0.2061	
		Bond Length(Å)			Bond Length(Å)			Bond Length(Å)		
C1-O2	1.188			1.187			1.208			
C1-O2	1.327			1.33			1.21			
C-C2	1.521			1.519			1.578			
C2-N	1.458			1.45			1.5			
N-H	1.001			1.002			1.046			
C2-C3	1.536			1.526			1.521			
C3-S	1.819			1.82			1.834			
S-H	1.323			1.32			1.326			

C. Threonine

Atom	Position(Å)		
C1	-0.2122	-1.7842	0.4318
C2	-0.2627	0.714	0.3662
C3	0.369	-0.5549	-0.2394
C4	0.4702	1.9841	-0.0498
H	0.4590	2.1246	-1.1252
H	1.5075	1.9546	0.2757
H	-0.0023	2.8453	0.4075
H	-1.2935	0.769	0.0389
H	1.4269	-0.5729	-0.0041
H	-1.7898	-2.769	0.5358
H	-0.7328	-0.656	-1.9558
H	0.7381	0.0014	-2.1702
H	0.5367	0.5224	2.1327
N	0.2274	-0.7018	-1.676
O3	-0.334	0.606	1.7651
O1	0.371	-2.4912	1.1844
O2	-1.4786	-1.9915	0.0824

D. Proline

Atom	Proline Position(Å)			Zwitterion of proline Position(Å)			
	C1	-0.3272	0.5218	-0.7225	-0.1486	0.6803	-0.7292
C2	-1.0865	0.6068	0.6175	-0.7026	0.8359	0.6982	
C3	-0.87704	-0.7788	1.2739	-0.8695	-0.5968	1.2264	
C4	0.1328	-1.4813	0.3616	0.2202	-1.3717	0.4955	
C5	0.9708	1.3268	-0.6939	1.1931	1.4476	-0.9985	
O1	1.1314	2.353	-0.1171	1.1412	2.6361	-0.7816	
O2	1.9251	0.791	-1.4545	2.1027	0.697	-1.3869	
N	-0.1605	-0.8993	-0.9384	0.1932	-0.7845	-0.8721	
H	-0.7106	1.419	1.2223	0.0034	1.3963	1.2962	
H	-2.1368	0.7939	0.4367	-1.6348	1.3842	0.7052	
H	-1.8077	-1.3331	1.2869	-1.845	-0.9972	0.9645	
H	-0.5227	-0.7041	2.2951	-0.7614	-0.6698	2.3012	
H	1.1542	-1.2767	0.6922	1.1999	-1.1618	0.9013	
H	0.528	-1.1089	-1.6322	1.1381	-0.7114	-1.2939	
H	0.0046	-2.5566	0.3352	0.0752	-2.4412	0.4449	
H	-0.9005	0.9584	-1.5357	-0.8706	0.9363	-1.4909	
H	2.6824	1.3683	-1.4271	-0.4346	-1.2792	-1.4799	
H	2.6108	0.9914	0.0205	0.693	-2.5925	-0.367	
		Bond Length(Å)			Bond Length(Å)		
C1-O1	1.188			1.209			
C1-O2	1.333			1.21			
C1-C2	1.528			1.569			
C2-C3	1.543			1.539			
C3-C4	1.548			1.536			
C4-C5	1.532			1.524			
C5-N	1.454			1.489			
N-H	1.093			1.037			
N-C2	1.447			1.511			

E. Phenylalanine

Atom	Position(Å)		
C1	0.9817	-1.2418	2.6923
C2	0.8029	-0.8073	1.2487
C3	-0.4599	0.0473	1.0979
C4	-0.6204	0.6157	-0.2974
C5	-0.8869	1.6598	-2.8788
C6	0.0849	1.7516	-0.6852
C7	-1.465	0.0167	-1.2248
C8	-1.5977	0.5326	-2.5054
C9	-0.0441	2.2694	-1.9621
H	1.5612	-2.5423	0.4967
H	0.6234	-1.7358	-0.5439
H	1.6755	-0.1893	1.0216
H	-0.4186	0.853	1.8203
H	-1.3159	-0.565	1.3537
H	0.734	2.2405	0.0212
H	-2.0256	-0.8573	-0.9427
H	-2.2589	0.0552	-3.2067
H	0.5072	3.1502	-2.24
H	-0.9909	2.063	-3.8702
H	2.0482	-2.3536	3.742
H	0.726	-1.9971	0.4168
H	0.3606	-0.8489	3.6241
N	1.9783	-2.1166	2.8222
O1	0.9817	-1.2418	2.6923
O2	0.9817	-1.2418	2.6923

F. Trptophan

Atom	Position(Å)		
C1	2.8637	1.143	-1.9952
C2	1.4916	0.8122	-1.4046
C3	1.6407	-0.0875	-0.1651
C4	0.5239	0.029	0.8365
C5	0.6328	0.6585	2.0265
C6	-0.824	-0.5047	0.7912
C7	-1.4446	-0.1446	1.994
C8	-1.5481	-1.2595	-0.1393
C9	-2.7561	-0.5091	2.2977
H	-3.4396	-1.2536	1.3674
H	-2.839	-1.6283	0.1553
H	0.3487	0.8981	-3.0877
H	0.9194	-0.5731	-2.7754
H	1.1005	1.7713	-1.0786
H	2.5691	0.1669	0.3284
H	1.7409	-1.1179	-0.5014
H	1.4778	1.1685	2.4438
H	-0.6948	0.9406	3.6393
H	-3.2144	-0.22	3.2272
H	-4.4527	-1.5537	1.5691
H	-1.1051	-1.5283	-1.0783
H	-3.4038	-2.2076	-0.5532
H	3.7087	1.5686	-3.5956
N1	0.5479	0.2538	-2.3482
N2	-0.5413	0.574	2.7302
O1	3.8647	1.2835	-1.3692
O2	2.8332	1.3197	-3.3151

CHAPTER 4

INNER SHELL SPECTRA OF SELECTED AMINO ACIDS

This chapter reports the C 1s, N 1s and O 1s spectra of the gas phase amino acids, glycine(g), alanine(g), cysteine (g), proline (g), threonine (g), phenylalanine(g) and tryptophan (g) recorded with ISEELS and the C 1s, N 1s and O 1s spectra of cysteine (s), proline (s) recoded with NEXAFS. All spectra have been converted to quantitative optical oscillator strength scales, suitable for use as references for quantitative analysis. The spectra are analysed with the aide of ab initio computations using the GSCF3 method. Differences in the spectra of the gas and solid are related to differences between the neutral gas phase molecule and the zwitterionic solid form. A rationalization of observations of high degree of variability in the N 1s spectra of amino acids, peptides and proteins is proposed.

§ 4.1 C 1s Spectra of Gly, Ala, Thr, Phe and Trp

Figure 4.1.1 presents the C 1s oscillator strength spectra of gaseous glycine, alanine, cysteine, threonine, proline, phenylalanine and tryptophan derived from ISEELS spectra recorded under dipole scattering conditions. The C 1s spectra of alanine, threonine, proline, phenylalanine and tryptophan computed from GSCF3 *ab initio* calculations are presented in Figure 4.1.2, in comparison to the experimental spectra of the related gaseous amino acids. The site-specific components of the calculations are shown, along with a set of lines indicating the position and oscillator strength (height) of the major

transitions. The calculated C 1s spectrum of glycine using GSCF3 was reported elsewhere [GC&03] and thus is not included in Fig. 4.1.2. Energies, term values and proposed assignments for the experimental and calculated spectral features are listed in Table 4.1.1.

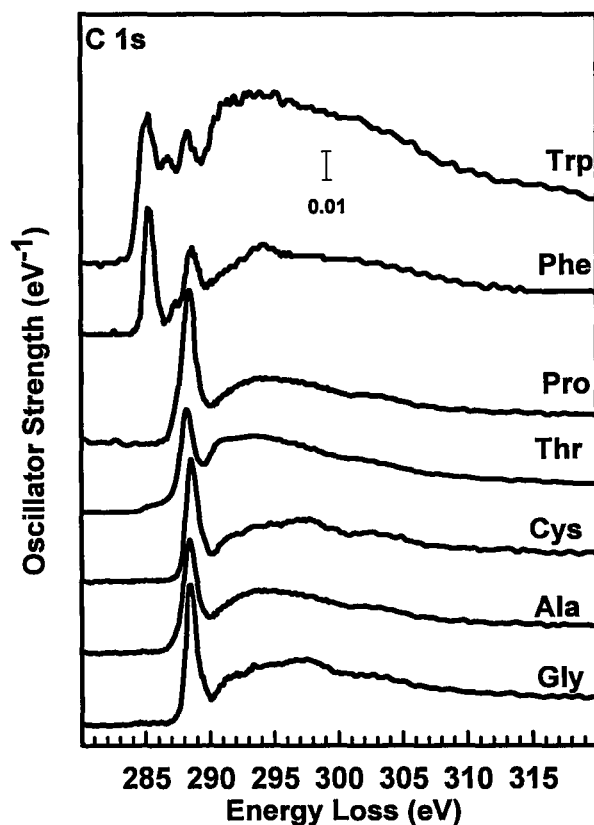


Figure 4.1.1 C 1s optical oscillator strength spectra of gaseous glycine, alanine, cysteine, threonine, proline, phenylalanine and tryptophan derived from inner shell electron energy loss spectra (ISEELS) recorded with 2.5 keV final electron energy and 2° scattering angle.

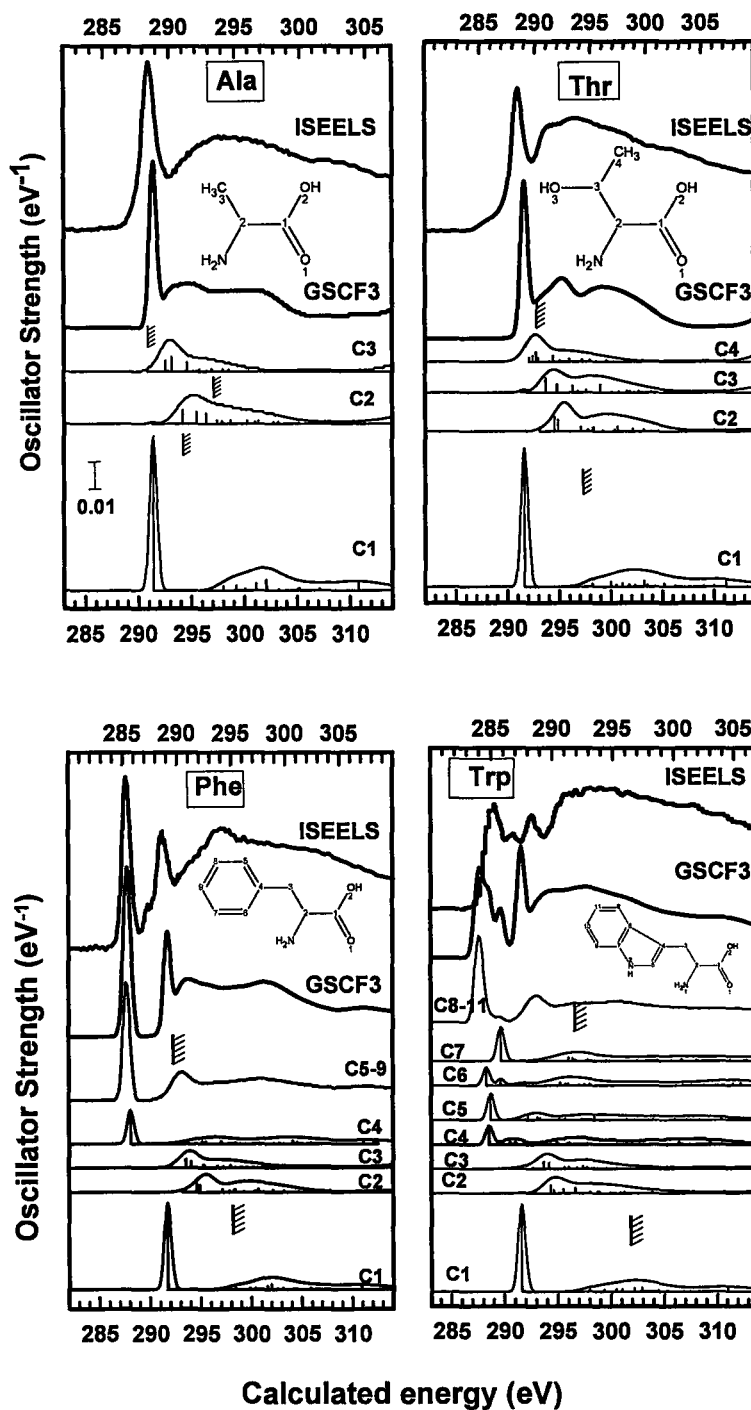


Figure 4.1.2 Calculated C 1s oscillator strength spectra of alanine, threonine, phenylalanine and tryptophan compared to experimental gaseous spectra. The computed spectra are plotted on absolute oscillator strength scales (total for the full molecule) with offsets. The computed energy scales are plotted with a shift relative to experiment of 3.0 eV (Ala), 3.1 eV (Thr), 2.5 eV (Phe) and 2.0 eV (Trp) in order to align the first peak.

The C 1s spectra of all the non-aromatic amino acids are dominated by a strong feature at 288.6 ± 0.1 eV which is the C 1s (COOH) $\rightarrow \pi^*_{C=O}$ transition associated with the carboxyl group. The carboxyl peak of proline is shifted to lower energy about 0.1 eV, probably due to conjugation of the π^*_{COOH} orbital with the cyclic side group. The C 1s (COOH) $\rightarrow \pi^*_{C=O}$ features in the computed C 1s spectra of all these amino acids are consistent with the experimental spectra, except that the calculated energies are systematically about 2 eV higher than the experimental energies. The prominence of the $\pi^*_{C=O}$ feature is most marked in the 2 simplest amino acids, glycine and alanine, since their simple R groups (H, CH₃) make little or no spectral contribution, but it is also true for the spectra of cysteine and proline. The GSCF3 calculation for alanine shown in Fig 4.1.2 predicts the lowest energy C 1s excitation to virtual valence levels of σ symmetry to be to an orbital of σ^*_{C-H} character. This is assigned to the shoulder on the high-energy side of the main peak, at around 289.7 eV. The broad peak at 294 eV is attributed to C 1s (CH₂, CH₃) excitations to $\sigma^*_{CH_2NH_2}$, σ^*_{CNH} and σ^*_{CH} orbitals.

The C 1s spectrum of cysteine contains an additional feature at slightly lower energy than that of the $\pi^*_{C=O}$ peak. This is attributed to the C 1s $\rightarrow \sigma^*_{CS}$ transition. Studies of other organo-sulfur species [HTM89, H90, DMH90] consistently contain a feature around 286 eV ascribed to the presence of the C-S bond. The attribution of this feature in cysteine is also supported by the calculated spectrum of the carbon of the mercapto group shown in Fig. 4.5.1. Due to the modest resolution of the ISEELS spectrometer, this peak is not fully resolved from the carboxyl peak (but see below for the better resolved X-ray

spectrum). The shoulder around 289 eV is associated with the C 1s (CH₂) → σ^*_{CNH} transition.

Threonine contains a hydroxyl group in the side chain. This is a new peak at 286 eV which should correspond to the C 1s → $\sigma^*_{\text{C-OH}}$ transition associated with the hydroxyl group.

Tryptophan and phenylalanine are amino acids with aromatic side chains. The dominant feature of the C 1s spectra of tryptophan and phenylalanine is the C 1s (C-H_{ring}) → $\pi^*_{\text{C=C}}$ transition around 285.1 eV which unambiguously identifies the phenyl functional group, since all C=C bonds in amino acids are aromatic. Thus, this peak is an excellent fingerprint for amino acids with phenyl rings. The carboxyl peaks of both tryptophan and phenylalanine are also at 288.6 eV.

Based on the molecular structure, the spectrum of phenylalanine should be the sum of the spectra of benzene and alanine according to the “building block” principle [S92]. This principle was illustrated earlier [Figure 1.9.1, GC&03] and is further explored here. Comparing the computed spectrum of alanine and phenylalanine, the features for carbon sites 1, 2, 3 are the same. In phenylalanine, the sum of $\pi^*_{\text{C=C}}$ transition of these six carbon sites in the phenyl ring should be similar with the spectrum of benzene. However, since the C₄ site has a different chemical environment from the other carbons in the ring, the C 1s (C₄) → $\pi^*_{\text{C=C}}$ transition is shifted 0.5 eV to higher energy relative to the $\pi^*_{\text{C=C}}$ transition of the other five carbon sites in the ring, for which the computed $\pi^*_{\text{C=C}}$ transitions have almost the same value about 287.6 eV, which is 2.4 eV higher than the experimental value. The sum of these six $\pi^*_{\text{C=C}}$ transitions, *i.e.*, the $\pi^*_{\text{C=C}}$ transition of

the phenylalanine is a double peak structure separated by 0.5 eV. At the ISEELS spectrometer resolution, the experimental spectra of phenylalanine only shows one peak in the $\pi^*_{C=C}$ transition region. The $C\ 1s(C-H_{ring}) \rightarrow \pi^*_{C=C}$ transition is much stronger than the carboxyl peak since it arises from six carbons. The small peak at 287.3 eV is attributed to σ^*_{CH} supported by the lowest energy C 1s excitations to virtual valence levels of σ symmetry to be of σ^*_{C-H} excited from phenyl carbon site.

Tryptophan contains one phenyl and pyrrole group. The calculated spectrum (**Figure 4.1.2**) exhibits four main peaks. In the pyrrole part, the C_4 site spectrum has a $\pi^*_{C=C}$ transition at 288.56* eV (* means calculated energy) and the C_5 site spectrum also has a $\pi^*_{C=C}$ transition which is at 288.74* eV. In the total spectrum of tryptophan, these two π^* transitions overlap and merge to form a composite $C_{4,5} \rightarrow \pi^*_{C=C}$ transition at $\sim 288.65^*$ eV,. Each of the six carbons in the phenyl part exists in a different chemical environment. The $C_7 \rightarrow \pi^*_{ring}$ transition is predicted to occur at 289.7* eV, which is higher than other π^*_{ring} transitions since C_7 is bonded to a N atom, and $C_6 \rightarrow \pi^*_{ring}$ transition is at 288.59* eV. Another π^*_{ring} transition located at $\sim 287.80^*$ eV is excited from the other four carbons C_{8-11} . The calculated energies are higher than the experimental energies. Relatively, we can assign the peaks in the experimental C 1s spectrum of tryptophan at 285.1, 285.5, 286.7 and 288.6 eV to $C_{8-11} \rightarrow \pi^*_{ring}$, $C_{4,5} \rightarrow \pi^*_{C=C}$, $C_7 \rightarrow \pi^*_{ring}$ and $C_1 \rightarrow \pi^*_{COOH}$, respectively.

Table 4.1.1 Energies, Term Values and proposed assignments for C 1s spectral features of gaseous Ala, Thr, Pro, Phe and Trp

(A) <i>Ala</i>	Experiment			Theory (GSCF3)		
	assignment	Energy(eV)	^b IP(eV)	Term(eV)	IP(eV)	Term(eV)
CH3 σ^*_{CH} σ^*_{CH}	289.7	291.0(2)	1.3	291.99	-1.14 -2.58	0.017 0.022
CH-NH2 σ^*_{CNH} σ^*_{CH}	294	292.2(1)		294.06	-0.14 -1.44	0.016 0.013
COOH π^* σ^*_{OH} σ^*_{C-OH}	288.6(1)	295.0(1)	-6.4	297.58	6.16 -3.55 -4.47	0.075 0.015 0.024

(B) <i>Cys</i>	Experiment				Theory(GSCF3)		
	Gas			Solid			
	assignment	Energy(eV)	^b IP(eV)	Term(eV)	Energy(eV)	IP(eV)	Term(eV)
HS-CH2 σ^*_{CSH} σ^*_{CH}	287.4	292.2(1)	4.8	287.5	292.85	1.65 -2.57	0.017 0.014
CH-NH3 σ^*_{CNH} σ^*_{CH}			1.4		294.11	0.10 -0.62	0.011 0.014
COOH π^* σ^*_{OH} σ^*_{C-OH}	288.6	295.0(1)	6.4	288.6	297.88	6.26 -3.80 -5.22	0.075 0.023 0.010

(C) Thr	Experiment			Theory (GSCF3)		
	assignment	Energy(eV)	^b IP(eV)	Term(eV)	IP(eV)	Term(eV)
CH3 σ^*_{CH} σ^*_{CH}	289.3	^c 290.7	1.4	292.28	-0.49 -2.15	0.013 0.017
COH σ^*_{C-OH}	286	^d 291.5(3)		294.45	-0.11	0.019
CH-NH2 σ^*_{CNH} σ^*_{CH}		292.2(3)		293.97	0.03 -1.07	0.020 0.011
COOH π^* σ^*_{OH} σ^*_{C-OH}	288.6	295.0(1)	6.4	297.53	5.88 -2.48 -5.07	0.076 0.013 0.015

(D) Pro	Experiment				
	Gas			Solid	
	assignment	Energy(eV)	^b IP(eV)	Term(eV)	Energy(eV)
CH3 σ^*_{CH}	289.4	291.0(2)	4.8		289.4
COOH π^*	288.5	295.0(1)	6.5		288.5

(E) Phe	Experiment			Theory (GSCF3)		
	assignment	Energy(eV)	^a IP(eV)	Term(eV)	IP(eV)	Term(eV)
C9 π^*	285.2	290.3	5.1	292.46	4.72	0.027
C8 π^*	285.2	290.3	5.1	292.44	4.74	0.027
C7 π^*	285.2	290.3	5.1	292.40	4.68	0.026
C6 π^*	285.2	290.3	5.1	292.40	4.73	0.026
C5 π^*	285.2	290.3	5.1	292.34	4.73	0.026
C4 σ^*_{CH}	285.7	290.3	4.6	293.01	4.93	0.029
C3 σ^*_{CH}		291.0(3)		292.58	-0.96	0.018
CH-NH2 σ^*_{CNH} σ^*_{CH}		292.2(1)		293.89	-0.58 -1.78	0.016 0.013
COOH π^* σ^*_{OH} σ^*_{C-OH}	288.6	295.0(1)	6.4	297.63	5.92 -3.91 -4.33	0.076 0.013 0.021

(F) <i>Trp</i> assignment	Experiment			Theory (GSCF3)		
	Energy(eV)	^a IP(eV)	Term(eV)	IP(eV)	Term(eV)	OS
C11 π^*	285.1	^e 290.3	5.2	291.83	4.13	0.023
C10 π^*	285.1	290.3	5.2	291.83	4.13	0.023
C9 π^*	285.1	290.3	5.2	291.88	4.23	0.021
C8 π^*	285.1	290.3	5.2	292.44	4.31	0.024
C7 π^*	286.7	290.3	3.7	293.27	3.56	0.029
C6 π^*	285.5	290.3		292.40	3.81	0.016
C5 $\pi^*_{C=C}$ σ^*_{CNH}	285.5	^f 291.8	6.3	292.62	3.88 -5.85	0.023 0.013
C4 $\pi^*_{C=C}$	285.5	290.3	4.8	291.99	3.43	0.016
C3 σ^*_{CH}		291.0(3)		292.39	-1.28	0.015
CH-NH2 σ^*_{CNH}		292.2(1)		293.74	-0.58	0.017
COOH π^* σ^*_{OH} σ^*_{C-OH}	288.6	295.0(1)	6.4	297.46	5.86 -3.00 -4.70	0.076 0.011 0.014

^aCarbon atoms are assigned as shown in Figure 4.1.2.

^b Experimental XPS data for *Gly* gas [SB88] is modified for other amino acids to reflect the tendency, reported for saturated hydrocarbons [R88]

^c Estimated, based on similar data for saturated hydrocarbon [PJ74]

^d Estimated, based on similar data for alcohols [IH87]

^e Estimated, based on similar data for gaseous benzene [BCJ87]

^f Estimated, based on similar data for gaseous phenol [OFK75]

§ 4.2 N 1s spectra of Gly, Ala, Cys, Pro, Thr and Phe

Figure 4.2.1 compares the N 1s oscillator strength spectra of gaseous glycine, alanine, cysteine, threonine, proline, and phenylalanine derived from dipole regime ISEELS. This is the first report of the N 1s and O 1s spectra of all of these amino acids in the gas phase, except for glycine [GC&03, CG&04]. The N 1s spectra of alanine, threonine and phenylalanine computed from GSCF3 *an inito* calculations are presented in **Figure 4.2.2**, in comparison to the experimental spectra of the related gaseous amino acids. The computed N 1s spectrum of cysteine is presented in section §4.7, in the context of the computational exploration of possible conformation dependence. Energies, term values and proposed assignments for the experimental and calculated spectral features are listed in **Table 4.2.1**. All of these N 1s spectra exhibit a shoulder at 401.2 eV, which is attributed to N 1s \rightarrow 3s Rydberg transitions. The feature at about 402.4 eV is attributed to a mixture of N 1s \rightarrow 3p Rydberg and N 1s $\rightarrow\sigma^*_{\text{NH}}$ transitions. Rydberg transitions are not predicted by the GSCF3 calculation since the basis sets used are not sufficiently diffuse. For computed N 1s spectra of these molecules, the weak low energy peaks predicted by the GSCF3 computation are N 1s $\rightarrow\sigma^*_{\text{NH}}$ transitions. These suggest that the experimental peak at 402.4 eV should be assigned to the mixture of N 1s \rightarrow 3p Rydberg and N 1s $\rightarrow\sigma^*_{\text{NH}}$ transitions, rather than solely to a N 1s \rightarrow 3p Rydberg transition. The broad peak at 405.1 eV is assigned to N 1s $\rightarrow\sigma^*_{\text{CN}}$ resonances related to the C-N bond on the basis of the calculation. The intensity of N 1s $\rightarrow\sigma^*_{\text{CN}}$ transition in proline is much higher than others since proline has two C-N bonds.

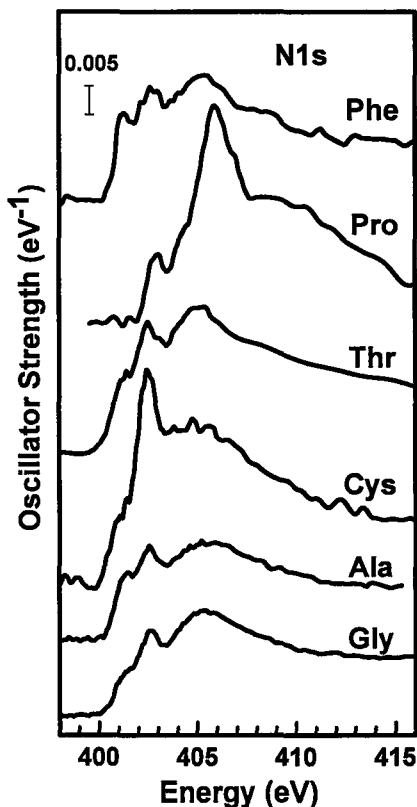


Figure 4.2.1 N 1s optical oscillator strength spectra of gaseous glycine, alanine, cysteine, threonine, praline and phenylalanine derived from inner shell electron energy loss spectra (ISEELS) recorded with 2.5 keV final electron energy and 2° scattering angle.

Comparing the computed N 1s spectra further, we can see that the positions and intensities of the computed σ_{NH}^* transitions have some differences, which in fact correlate with differences in the experimental N 1s spectra. This suggests that N 1s $\rightarrow \sigma_{\text{CN}}^*$ transitions are sensitive to changes in the side-chains, which may be at least part of the explanation for observed variability in the N 1s spectra of peptides and proteins. In addition, there are good reasons to suspect that the N 1s spectra are quite sensitive to the immediate chemical environment, such as changes on pH and thus extent of protonation

of the amine [MC&05]. The computed N 1s spectra of cysteine and proline are reported and discussed below in § 4.7 to examine this phenomenon.

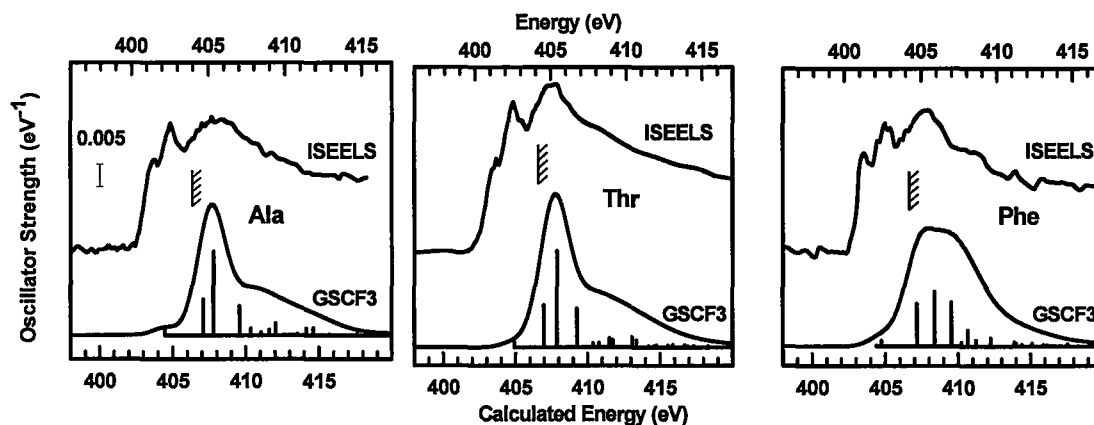


Fig. 4.2.2 Calculated N 1s oscillator strength spectra of alanine, threonine, and phenylalanine compared to experimental gaseous spectra. The computed spectra are plotted on absolute oscillator strength scales (total for the full molecule) with offsets. The computed energy scales are plotted with a shift relative to experiment of 3.0 eV (Ala), 2.9 eV (Thr), and 3.2 eV (Phe) in order to align the broadened peak of the computed spectra.

Table 4.2.1 Energies, Term Values and proposed assignments for N 1s spectral features of gaseous Ala, Thr, Phe, Cys and Pro

Amino Acids	Assignment	Experiment			Theory (GSCF3)		
		Energy(eV)	^a IP(eV)	Term(eV)	IP(eV)	Term(eV)	OS
Ala	3s	401.3	405.6	4.3	406.21	1.72	0.002
	3p/ σ^*_{NH}	402.4		3.2		-0.90	0.01
	σ^*_{CN}	405.1		0.4		-1.58	0.024
Thr	3s	401.3	405.6	4.3	406.30	1.37	0.001
	3p/ σ^*_{NH}	402.4		3.2		-0.68	0.01
	σ^*_{CN}	405.1		0.4		-1.59	0.023
Phe	3s	401.3	405.6	4.3	406.02	1.62	0.0021
	3p/ σ^*_{NH}	402.4		3.2		-1.28	0.01
	σ^*_{CN}	405.1		0.4		-2.37	0.024
Cys	3s	401.3	405.6	4.3			
	3p/ σ^*_{NH}	402.4		3.2			
	σ^*_{CN}	405.1		0.4			
Pro	3s	401.3	405.6	4.3			
	3p/ σ^*_{NH}	402.4		3.2			
	σ^*_{CN}	405.1		0.4			

^aExperimental XPS data for *Gly* gas [SB88].

§ 4.3 O 1s spectra of Gly, Ala, Cys, Thr, Pro and Phe

Figure 4.3.1 compares the O 1s oscillator strength spectra of gaseous glycine, alanine, cysteine, threonine, proline, and phenylalanine derived from ISEELS. The O 1s spectra of alanine, threonine and phenylalanine computed from GSCF3 *ab initio* calculations are presented in **Figure 4.3.2** compared to the experimental spectra of these gaseous amino acids. The computed O 1s spectra of cysteine are presented in §4.5, in the context of the conformation dependence discussion. Energies, term values and proposed assignments for the experimental and calculated spectral features are listed in **Table 4.3.1**. The experimental O 1s spectra of gaseous alanine, cysteine, threonine, proline and phenylalanine and cysteine are all similar and are dominated by two strong low lying features. These two features arise from the different chemical environments of the C=O and OH atoms in carboxyl group. By comparison to the spectra of other carboxylic acids [IH87] and supported by the GSCF3 calculations, the lower energy features at 532.2 eV is assigned to the O 1s(C=O) $\rightarrow \pi^*_{C=O}$ transitions while the higher energy features at 535.3 eV are the O 1s (OH) $\rightarrow \pi^*_{C=O}$ transitions. The feature at 540 eV in these O 1s spectra are due to O 1s $\rightarrow \sigma^*_{C-O}$ transitions. In the spectrum of threonine, the intensity of this O 1s $\rightarrow \sigma^*_{C-O}$ transition is much higher since threonine contains another –OH group, and this transition arises from the O atoms both in –COOH and –OH group. The broad peaks at 545 eV are O 1s $\rightarrow \sigma^*_{C=O}$ transitions.

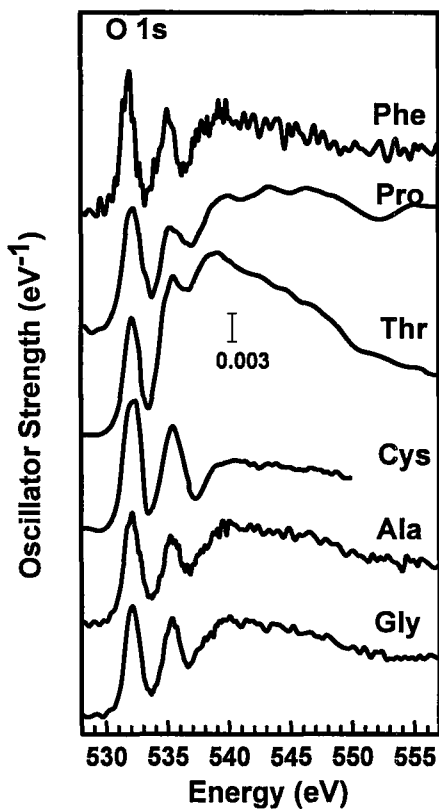


Figure 4.3.1 O 1s optical oscillator strength spectra of gaseous glycine, alanine, cysteine, threonine, proline and phenylalanine derived from inner shell electron energy loss spectra (ISEELS) recorded with 2.5 keV final electron energy and 2° scattering angle.

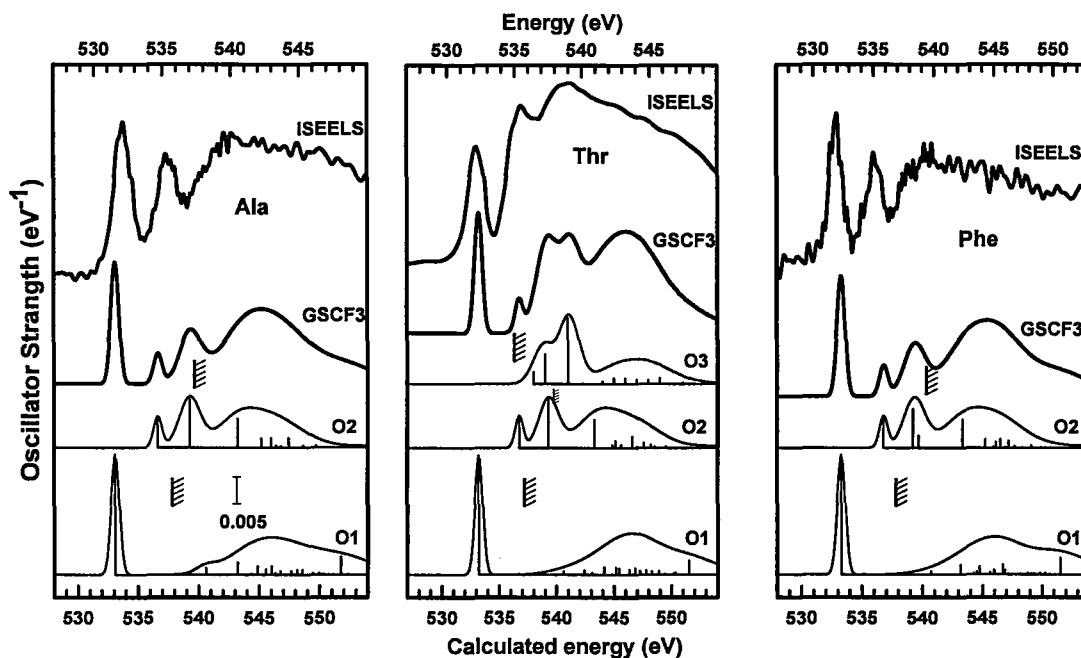


Figure 4.3.2 Calculated O 1s oscillator strength spectra of alanine, threonine, and phenylalanine compared to experimental gaseous spectra. The computed spectra are plotted on absolute oscillator strength scales (total for the full molecule) with offsets. The computed energy scales are plotted with a shift relative to experiment of 1.0 eV (Ala), 1.1 eV (Thr), and 1.2 eV (Phe) in order to align the first peak.

Table 4.3.1 Energies, Term Values and proposed assignments for O 1s spectral features of gaseous Ala, Cys, Thr, Phe and Pro.

(A) <i>Ala</i>	Experiment			Theory (GSCF3)		
	assignment	Energy(eV)	^b IP(eV)	Term(eV)	IP(eV)	Term(eV)
C=O				538.7		
π^*	532.2	538.2	6.0		5.62	0.016
σ^*_{OH}	540.0		-1.8		-1.89	0.022
σ^*_{COH}					-4.49	0.008
O-H		540.0		540.5		
π^*	535.3		4.7		3.89	0.004
σ^*_{OH}					1.25	0.016
σ^*_{COH}	545.0		-5		-2.09	0.019

(B) <i>Cys</i>	Experiment			
	Gas			Solid
assignment	Energy(eV)	^b IP(eV)	Term(eV)	Energy(eV)
C=O		538.2		
π^*	532.2		6.0	532.2
σ^*_{OH}	540.0		-1.8	540.0
σ^*_{COH}				
O-H		540.0		
π^*	535.3		4.7	
σ^*_{OH}				
σ^*_{COH}	545.0		-5	

(C) <i>Thr</i>	Experiment			Theory (GSCF3)		
	assignment	Energy(eV)	^b IP(eV)	Term(eV)	IP(eV)	Term(eV)
C=O				538.5		
π^*	532.2	538.2	6.0		5.27	0.016
σ^*_{OH}	540.0		-1.8		-2.11	0.022
σ^*_{COH}					-3.93	0.030
COO-H		540.0		540.3		
π^*	535.3		4.7		3.59	0.004
σ^*_{OH}					1.02	0.016
σ^*_{COH}	545.0		-5		-2.98	0.019
O-H		540.0		539.4		
σ^*_{OH}	540.0		4.7		0.95	0.004
σ^*_{COH}	545.0		-1.8		-1.77	0.022

(D) <i>Pro</i>	Experiment			
	Gas			Solid
C=O		538.2		
π^*	532.2		6.0	532.2
σ^*_{OH}	540.0		-1.8	540.0
σ^*_{COH}				
O-H		540.0		
π^*	535.3		4.7	
σ^*_{OH}				
σ^*_{COH}	545.0		-5	

(E) <i>Phe</i>	Experiment			Theory (GSCF3)			
	assignment	Energy(eV)	^b IP(eV)	Term(eV)	IP(eV)	Term(eV)	OS
C=O			538.2		538.6		
π^*		532.2		6.0		5.32	0.016
σ^*_{OH}		540.0		-1.8		-2.11	0.022
σ^*_{COH}						-4.56	0.065
O-H			540.0		540.4		
π^*		535.3		4.7		3.60	0.004
σ^*_{OH}						1.15	0.015
σ^*_{COH}		545.0		-5		-2.96	0.019

^aExperimental XPS data for *Gly* gas [SB88].

§ 4.4 S 2p Spectrum of Cys

Figure 4.4.1 shows the experimental S 2p oscillator strength spectra of gaseous cysteine. The first peak at 164.5 eV dominates this spectrum. Its term value is about 4.7 eV based on the S 2p IP of 169.28 eV estimated from experimental XPS data for CH₃SH [BCJ80]. The 164.5 eV feature is assigned to S 2p→σ*_{CS} transitions, an assignment supported by a term value similar to that for S 1s→σ*_{CS} transition in all molecules containing an S-C bond. Typically, the features corresponding to excitations from different core level to the same virtual orbital occur at very similar term values [DM&90].

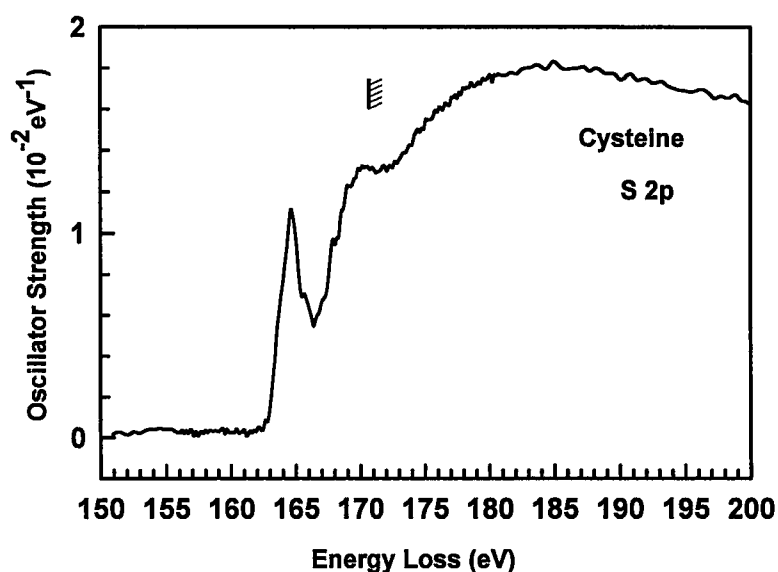


Figure 4.4.1 Experimental S 2p oscillator strength spectra of gaseous cysteine recorded by inner shell electron energy loss spectra (ISEELS) with 2.5 keV final electron energy and 2° scattering angle. The IP, 169.28 eV, is taken from experimental XPS data for CH₃SH [BCJ80].

§ 4.5 Molecular Conformational Dependence of Core Excitation Spectra of Cysteine

Cysteine is one of the most important amino acids. It often plays a role in the active site of enzymes, and, when it forms disulfide bridges, it is used to tether two strands of a multi-strand protein. **Figures 4.5.1, 4.5.2 and 4.5.3** compare the GSCF3 calculation results of the C 1s, N 1s and O 1s spectra of conformer I and conformer II of cysteine. Energies, term values and proposed assignments for the computed spectral features of conformer I and conformer II of gaseous cysteine are listed in **Table 4.5.1**. The site-specific components of the calculation are also shown. The top spectra are the experimental spectra of cysteine for comparison. The experimental and computed spectra are generally in very good agreement, aside from the energy shift, which is typical of GSCF3 results. The C 1s and O 1s spectra of each carbon site and the sum are rather similar, in particular the C 1s (COOH) $\rightarrow \pi^*$ transitions and O 1s (COOH) $\rightarrow \pi^*$ transitions of these two conformers have the same energy. The computed IP for each carbon and oxygen are also the same. There are subtle differences between the computed N 1s spectra for these two conformers. However, the shape of the spectra and the energy scales are almost the same. The difference is mainly in the excitation intensity. The reason for this difference will be discussed in section § 4.7. Based on this study and previous work [GC&03] related to the influence of molecular conformation on glycine, we can conclude that the C 1s and O 1s spectra of monomer of amino acids do not have any conformational dependence detectible at the ISEELS energy resolution, but the N 1s spectrum should have a small conformational dependence.

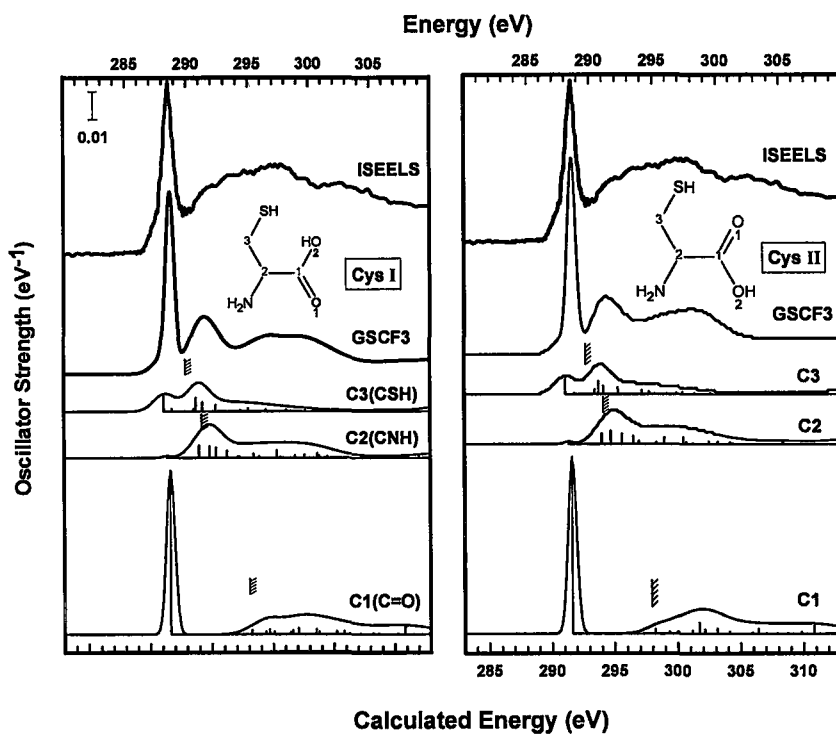


Figure 4.5.1 Computed C 1s spectra of two conformers of cysteine using GSCF3. The structures of the two conformers of cysteine are shown.

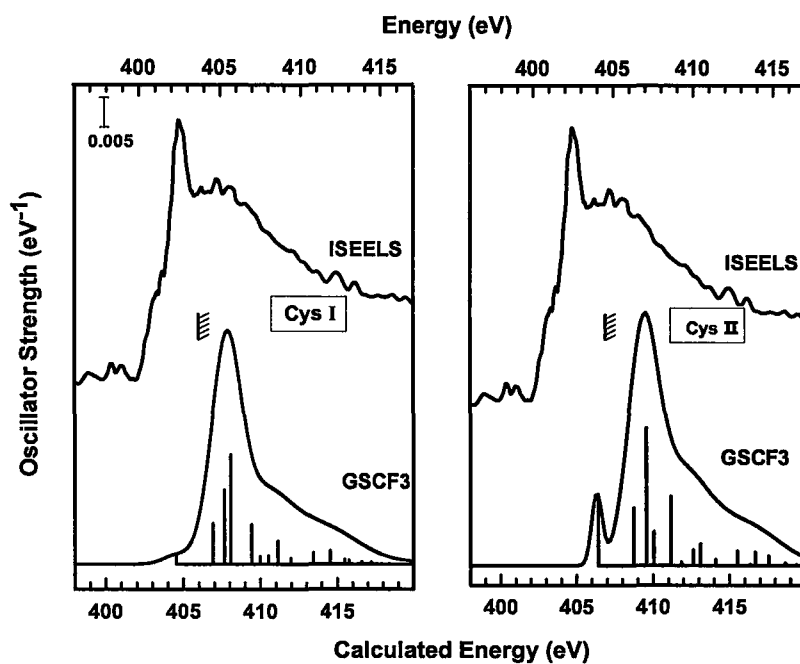


Figure 4.5.2 Computed N 1s spectra of two conformers of cysteine using GSCF3.

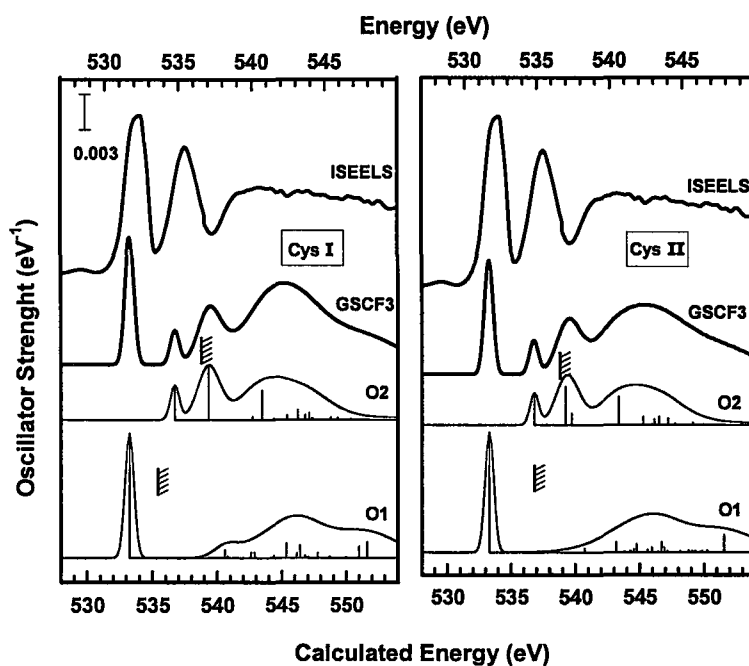


Figure 4.5.3 Computed O 1s spectra of two conformers of cysteine using GSCF3.

§ 4.6 Comparison of C 1s and O 1s Spectra of Gas and Solid Phase Pro and Cys

Figure 4.6.1 compares the gaseous and solid phase C 1s spectra of proline and cysteine. The solid phase spectra were recorded using NEXAFS in the STXM, with an energy resolution of ~ 0.2 eV, while the ISEELS data was recorded with an energy resolution of ~ 0.7 eV. The higher energy resolution in NEXAFS is probably why the low lying feature at 287.4 eV assigned to C 1s $\rightarrow \sigma^*_{CS}$ transitions in the C 1s spectrum of cysteine is much more pronounced. The C 1s (COOH) $\rightarrow \pi^*$ transition occurs at the same energy for both phases of these two species. It appears that the C 1s spectra of amino acids are relatively unperturbed on going from gas-phase to condensed phase, therefore we conclude that the zwitterion character of the condensed amino acid has little effect on the C 1s spectrum.

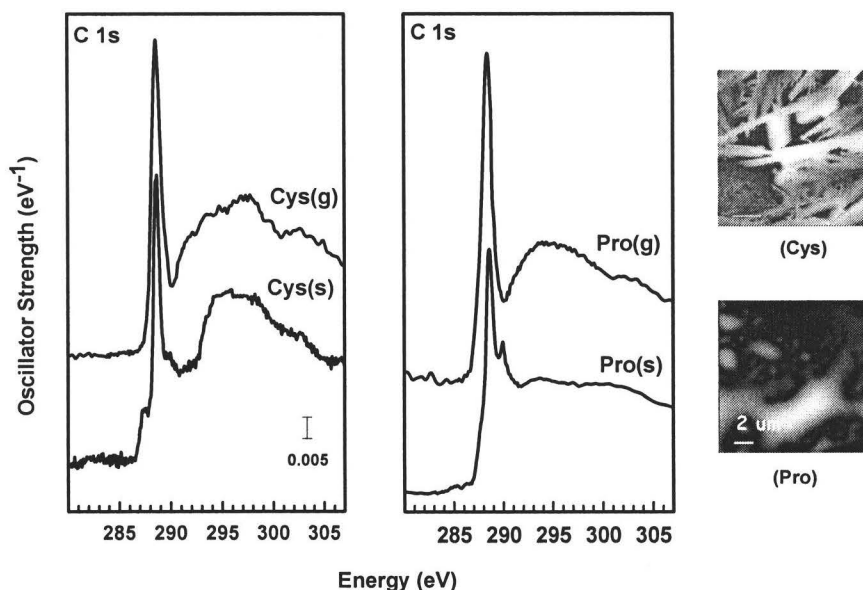


Figure 4.6.1 Comparison of the C 1s ISEELS spectra of gaseous proline recorded by ISEELS and NEXAFS spectra of solid proline recorded by scanning transmission X-ray microscopy (STXM). STXM images at 288.2 eV for the sample in the regions measured are displayed on the right. (bar=1 μ m)

Figure 4.6.2 compares the O 1s spectra of the gas and solid of these two species. In contrast to the C 1s, the O 1s spectra of gaseous and condensed phase amino acids are quite different. The O 1s spectrum of gaseous proline contains two distinct features arising from the two different oxygen atoms in the carboxyl group. These two peaks merge into a single peak at 532.2 eV in the O 1s spectrum of condensed proline, similar but to the high energy side of the lower energy feature in the gas phase spectrum. It is well known that condensed amino acids exist in zwitterionic form [FF86].

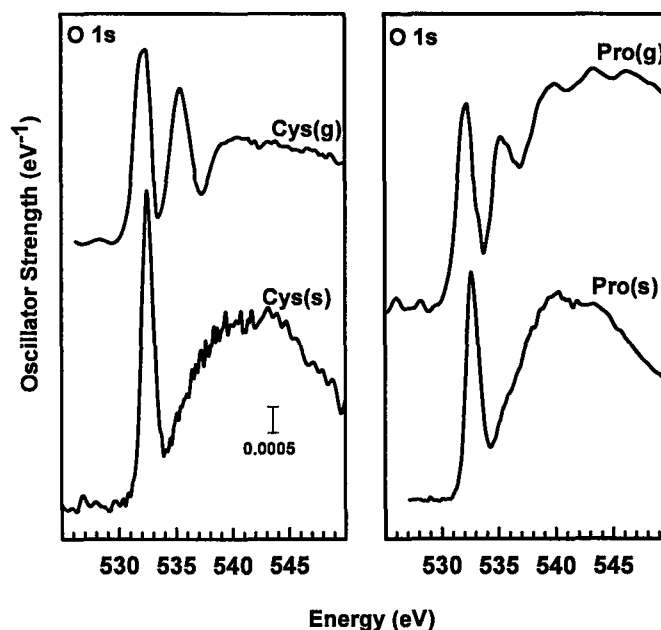


Figure 4.6.2 Comparison of the O1s ISEELS spectra of gaseous proline recorded by dipole regime ISEELS and NEXAFS spectra of solid proline recorded by scanning transmission X-ray microscopy (STXM).

The two C-O bonds of the carboxylate anion in the zwitterionic form are the same length, intermediate between the lengths of a C-O single bond and a C=O double bond [FF86].

We must use two valence-bond resonance structures to describe the real structure. That

means the two oxygen atoms in the zwitterionic form of amino acids have the same chemical environment which explains why there is only a single peak in the O 1s spectrum of condensed proline. These results are similar to those for O 1s spectra of gaseous and solid cysteine. Clearly there is a very large effect of zwitterion formation on the O 1s spectra of amino acids, with the O 1s spectra of zwitterion states exhibiting only a single O 1s \rightarrow π^* transition whereas the protonated condensed species or the neutral carboxylic acid form exhibit two O 1s \rightarrow π^* transitions. This difference can be readily used to determine the local charge state at the C-terminus of an amino acid. A similar effect would be expected in the case of small peptides.

§ 4.7 Comparison of the N 1s Spectra of Gas and Solid Phase Pro and Cys

Figure 4.7.1 compares the experimental and computed N 1s spectra of gaseous and solid proline. Energies, term values and proposed assignments for the computed spectral features of gaseous, solid cysteine and gaseous, solid proline are also listed in Table 4.7.1. There is a significant difference between the experimental N 1s spectra of gaseous and solid proline. The N 1s spectrum of gaseous proline has a prominent peak at 402.4 eV assigned to N 1s \rightarrow σ^*_{NH} /Rydberg transitions. This transition vanishes in solid phase N 1s spectrum or, at least, is considerably broadened. Here we discuss the reason why both the pure Rydberg transition and the N 1s \rightarrow σ^*_{NH} disappear in the condensed state.

Attenuation of pure Rydberg transitions in the condensed-phase has been found in many studies [R74, WB&99, UG05]. This is because Rydberg MOs have a much larger

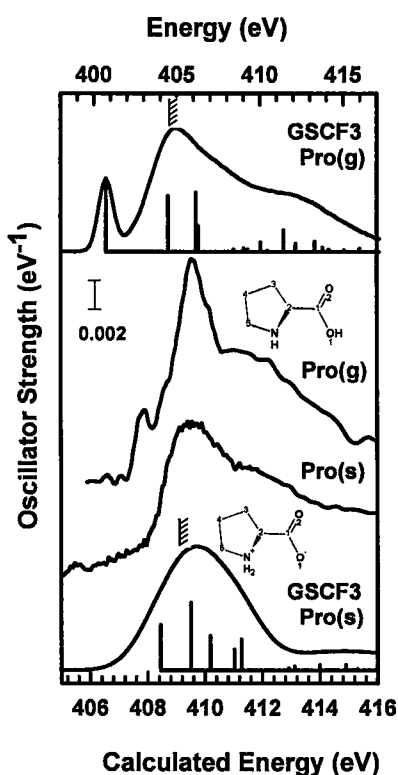


Figure 4.7.1 Comparison of the experimental N 1s ISEELS of gaseous proline and NEXAFS spectrum of solid proline, i.e., zwitterionic form of proline. The computed spectra are plotted on absolute oscillator strength scales (total for the full molecule) with offsets. The computed energy scales are plotted with a shift relative to experiment of 4.0 eV in order to align the broad peak.

average radius than that of the core or virtual valence type orbitals. Due to their size Rydberg states are much more sensitive to perturbation from the outside, and thus are considered to either be quenched [R74], to shift energies and have an excitonic character [ST&99], or to merge into a band, which, in the case of metals or semiconductors, would be considered to have conduction band character. In crystalline proline, the electronic

relaxation time is very short due to strong extra-molecular interactions, so the Rydberg transition is broadened, to the extent it cannot be observed [S92]. In contrast, core to valence excitations like $1s \rightarrow \pi^*$ and $1s \rightarrow \sigma^*$ transitions are relatively unperturbed on going from the gas phase to the solid phase, as observed in the C 1s of proline and cysteine, since the upper orbital is relatively localized. So the N $1s \rightarrow \sigma^*_{\text{NH}}$ component of the low lying spectrum of condensed proline might be expected to survive. However, we do not observe a corresponding 402.4 eV transition in solid proline. This means that the N $1s \rightarrow \sigma^*_{\text{NH}}$ transition is dramatically different for the neutral and zwitterionic forms of amino acids.

In order to investigate the origin of the variation of these spectra, we used *ab initio* calculations to explore the sensitivity of N 1s spectra to different states. The GSCF3 computed spectra of gaseous and solid proline are also presented in **Figure 4.7.1**. The GSCF3 calculation can not predict Rydberg transitions due to the small basis set. Based on the orbital characteristics, the first peak of the computed spectrum of gaseous proline is assigned to N $1s \rightarrow \sigma^*_{\text{NH}}$ transitions. Here we confirm further that in the experimental spectrum of gaseous proline the peak at 402.4 eV is a mixture of a Rydberg transition and a σ^*_{NH} transition, not just a pure Rydberg transition. For the computed N 1s spectrum of solid proline, the single peak is associated with the σ^*_{NC} transition and σ^*_{NH} transition. The σ^*_{NH} transition has a higher energy and smaller term value relative to that in the gas phase, so it overlaps with the σ^*_{NC} transition. It is well known that all amino acids are capable of forming an N \cdots H-O hydrogen bond between the two terminal groups -COOH and -NH₂. The zwitterion form has a significant stronger intramolecular hydrogen bond

than the neutral form. The strength of the hydrogen bonds may lead to the subtle difference of the N-H bond lengths. Therefore, the variation of these spectra may be due to different strengths of intramolecular hydrogen bonds.

Figure 4.7.2 shows there is also a dramatic variation between the N 1s spectra of gaseous and solid cysteine. To help explain this we calculated the N 1s spectrum of solid cysteine and the N 1s spectra of two conformers of neutral cysteine showed in **Figure 4.5.2**. In conformer I, the -NH₂ group and -OH group are *trans*, (on opposite sides of the the plane of the C-C bond), whereas they are *cis* in conformer II. The N···H-O hydrogen bond should be a little bit stronger in conformer I than in conformer II. This tiny difference appears to cause a difference of term values and intensity of σ^*_{NH} transition for these two conformers. Based on the calculated geometry listed in **Table 3.2.1**, the N-H bond lengths in solid cysteine and in gaseous phase are different, so the σ^*_{NH} transition in N 1s spectrum of solid cysteine is shifted to high energy, and overlaps with the σ^*_{NC} transition. We suggest that the N 1s spectrum is very sensitive to the environment of the amide bond. The origin of the variation of N 1s spectra of gaseous and condensed amino acids is arising not only from the absence of the Rydberg transition, but also from the different hydrogen bond effects.

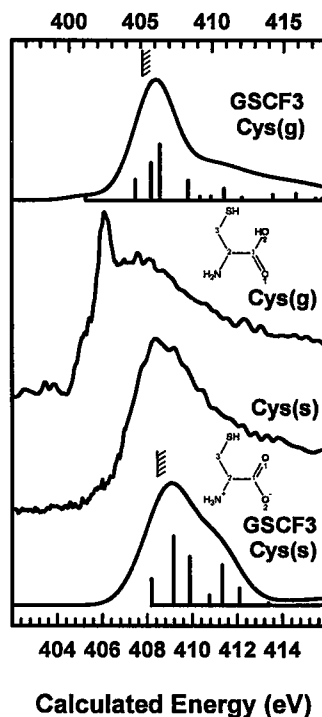


Figure 4.7.2 Comparison of the experimental N 1s ISEELS of gaseous cysteine and NEXAFS spectrum of solid cysteine, i.e., zwitterionic form of cysteine. The computed spectra are plotted on absolute oscillator strength scales (total for the full molecule) with offsets. The computed energy scales are plotted with a shift relative to experiment of 4.0 eV in order to align the broad peak.

Table 4.7.1 Selected eigenvalues, oscillator strengths and orbital characters for computed C 1s, N 1s and O 1s core excited states of conformer I and II of gaseous cystiene, solid cysteine, and computed N 1s core excited state of gaseous proline and solid proline

A. Cysteine

	Site	IP(eV)	Character	Orbital No.	Term Value (eV)	OS(eV)
I	C ₁ (C=O)	297.89	$\pi^*(C=O)$	35	6.26	0.075
			$\sigma^*(C-H)$	36	-0.33	0.004
			$\pi^*(C=O)$	33	2.75	0.001
	C ₂ (CN)	294.1	$\sigma^*(C-H)$	34	0.10	0.011
			$\sigma^*(C-N)$	37	-0.62	0.014
			$\sigma^*(C-S)$	35	1.74	0.018
II	C ₁ (C=O)	297.97	$\sigma^*(C-H)$	34	-0.63	0.005
			$\pi^*(C=O)$	35	6.32	0.075
			$\sigma^*(C-H)$	36	-0.30	0.005
	C ₂ (CN)	294.1	$\pi^*(C=O)$	33	2.75	0.001
			$\sigma^*(C-H)$	34	0.15	0.014
			$\sigma^*(C-N)$	37	-0.69	0.013
I	C ₃ (CS)	292.8	$\sigma^*(C-S)$	35	1.65	0.017
			$\sigma^*(C-H)$	34	-0.99	0.003
			$\pi^*(C=O)$	33	5.66	0.016
	O ₁ (C=O)	538.9	$\sigma^*(O-H)$	36	-1.48	0.0004
			$\sigma^*(C-OH)$	34	-1.84	0.002
			$\pi^*(C=O)$	33	3.93	0.004
II	O ₂ (OH)	540.6	$\sigma^*(O-H)$	37	1.30	0.016
			$\sigma^*(C-OH)$	34	-2.01	0.002
			$\pi^*(C=O)$	33	5.68	0.015
	O ₁ (C=O)	538.9	$\sigma^*(O-H)$	36	-1.67	0.0024
			$\sigma^*(C-OH)$	34	-1.85	0.005
			$\pi^*(C=O)$	33	4.00	0.004
I	N	406.30	$\sigma^*(O-H)$	37	1.39	0.016
			$\sigma^*(O-H)$	34	-1.65	0.001
			$\sigma^*(N-H)$	38	-0.60	0.007
	O ₂ (OH)	540.8	$\sigma^*(N-H)$	34	-1.36	0.012
			$\sigma^*(C-N)$	48	-1.78	0.018
			$\sigma^*(O-H)$	37	1.39	0.016
II	N	406.21	$\sigma^*(O-H)$	34	-1.65	0.001
			$\sigma^*(N-H)$	38	-0.65	0.006
			$\sigma^*(N-H)$	34	-1.36	0.012
Solid	N	410.17	$\sigma^*(C-N)$	48	-1.99	0.018
			$\sigma^*(N-H)$	35	-1.99	0.0081
			$\sigma^*(N-H)$	38	1.01	0.021
			$\sigma^*(N-C)$	66	0.29	0.015

B. Proline

	Site	IP(eV)	Character	Orbital No.	Term Value (eV)	OS(eV)
Gas	N	408.36	$\sigma^*(\text{N-H})$	33	-1.89	0.0022
			$\sigma^*(\text{N-H})$	38	0.24	0.0089
			$\sigma^*(\text{C-N})$	48	1.07	0.0175
Solid	N	410.17	$\sigma^*(\text{N-H})$	35	-1.99	0.0081
			$\sigma^*(\text{N-C})$	38	1.01	0.021
			$\sigma^*(\text{N-H})$	48	-0.29	0.0015

- (a) The predicted spectra were generated using Gaussian lines of area given by the computed oscillator strength (f) and widths of 0.8 eV ($\epsilon < -2\text{eV}$), 2.0 eV ($-2.0\text{ eV} < \epsilon < 2.0\text{ eV}$), 4.0 eV ($2.0\text{eV} < \epsilon < 8.0\text{ eV}$) and 6.0 eV ($\epsilon > 8.0\text{ eV}$)

CHAPTER 5

INNER SHELL SPECTRA OF SMALL PEPTIDES

The ISEELS spectra of the gas phase di-peptide, glycyl-alanine (Gly-Ala) and the NEXAFS spectra of three small peptides, Gly-Ala, Lys-Trp-Lys (KWK), and Arg-Gly-Asp (RGD) in the solid state are reported in this chapter. A detailed interpretation of these spectra is presented based on a modified building block treatment involving the spectra of the residues with compensation for spectral changes associated with the peptide bond.

§ 5.1 Introduction

In this chapter we report the C 1s, N 1s and O 1s inner shell excitation spectra of gas-phase glycyl-alanine (Gly-Ala) by inner shell electron energy loss spectroscopy (ISEELS) and the near edge X-ray absorption fine structure (NEXAFS) spectra of solid Gly-Ala, Lys-Trp-Lys (KWK), and Arg-Gly-Asp(RGD). Gly-Ala was examined since it is sufficiently volatile to allow measurement of the gas phase spectrum without significant decomposition. This has allowed a probe of the differences between the inner shell spectra of the non-charge separated gas phase structure and that of the zwitterionic solid state structure. RGD (Arg-Gly-Asp) is a tripeptide recognition motif that is important in cellular adhesion [R96]. RGD provides an important biochemical cue for promoting cell receptor-mediated adhesion. The integrin superfamily of cell adhesion receptors binds the RGD sequence as a ligand to promote the cell adhesion response. Thus there is considerable interest in this sequence due to its involvement in binding to specific receptors [AX&01].

§ 5.2 ISEELS and NEXAFS Spectra of Gly-Ala

Gly-Ala is a dipeptide which contains two C=O bonds. One is in the terminal carboxyl group CO₂H and the other one participates in the CONH peptide group. These two carbonyls have different chemical environments and thus they should have different C 1s excitation energies. If the building block principle [S92] is simply used to predict the spectrum of Gly-Ala, we would expect its spectrum to be similar to the sum of the spectra of glycine and alanine, except for a correction required due to the chemical changes associated with the loss of H₂O. These changes have been documented previously in comparisons of the inner shell spectra of Gly and Gly-Gly [GC&03]. **Figure 5.2.1** shows the ISEELS C 1s spectrum of gaseous Gly-Ala in comparison with the C 1s spectra of glycine, alanine, the weighted sum of glycine and alanine (Gly+Ala), and the NEXAFS spectrum of solid Gly-Ala. An enlarged section emphasizing the energy region in the 286-292 eV range is shown in the right panel. Energies, term values and proposed assignments for the experimental spectral features of gaseous Gly-Ala are listed in **Table 5.2.1**. The C 1s spectrum of Gly-Ala is dominated by a pronounced peak at 288.28(6) eV, which is assigned as the C 1s → $\pi^*_{\text{C=ONH}}$ amide transition. It is shifted to lower energy by about 0.3 eV relative to the $\pi^*_{\text{C=O}}$ peak of glycine and alanine. This shift occurs because one of the two oxygen atoms in the monomers is replaced by a nitrogen atom in the dipeptide. The C 1s → $\pi^*_{\text{C=OOH}}$ carboxyl transition of Gly-Ala should be at the same position as that of glycine, 288.6 eV. Because the carboxyl carbon in Gly-Ala is isolated from the peptide bond by one C-C bond, the peptide bond should not affect the C 1s (COOH) excitation.

However, due to the limited energy resolution, we only observe it as an asymmetry in the line shape. The ISEELS spectrum of gaseous and the NEXAFS spectrum of solid Gly-Ala look very similar. However, the C 1s $\rightarrow \pi^*_{\text{C=ONH}}$ transition is slightly shifted to lower energy about 0.05 eV. By comparing the C 1s spectrum of Gly-Ala with the weighted sum of glycine and alanine, we can see that the “building block” principle [S92] works reasonably well.

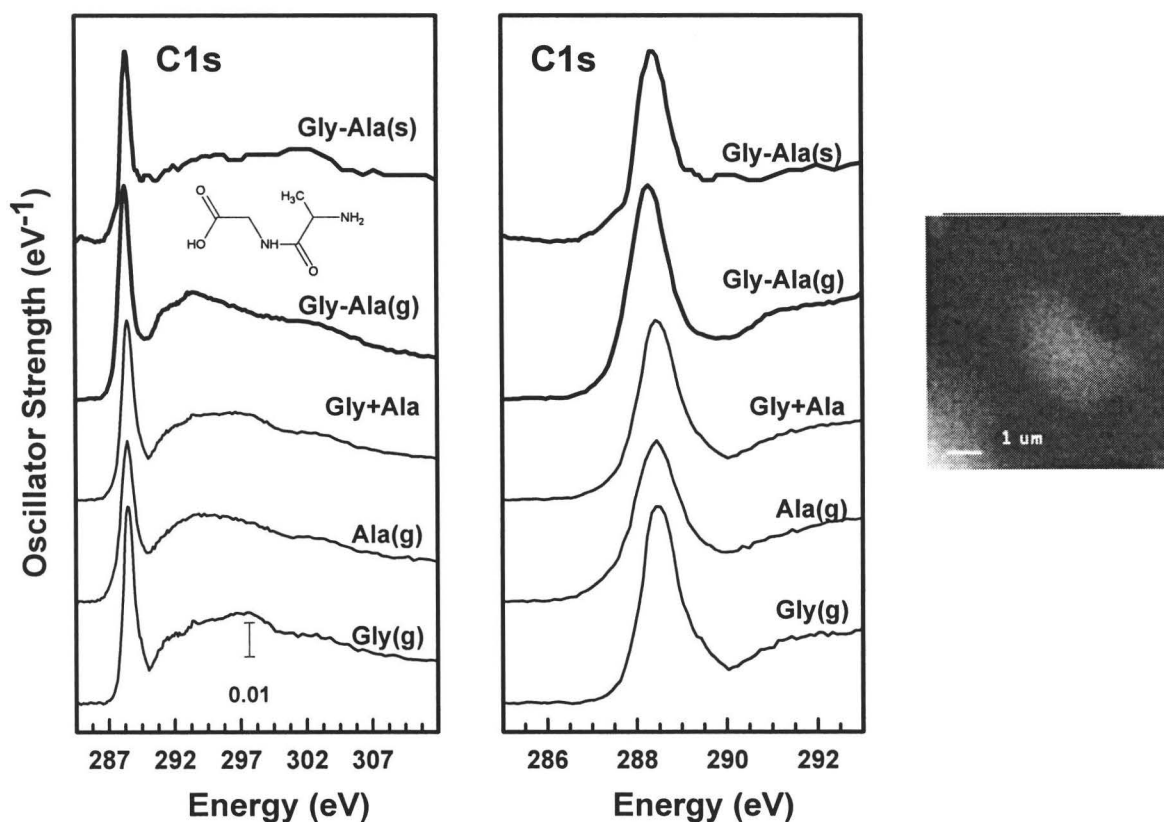


Figure 5.2.1 Comparison of the ISEELS C 1s spectra of gaseous Gly-Ala with the ISEELS C 1s spectra of glycine, alanine, the weighted sum of glycine and alanine, and the NEXAFS spectra of condensed Gly-ala. An enlarged section emphasizing the energy region in the 286-292 eV range is shown in the right panel. A STXM image at 288.6 eV for the sample in the region measured is displayed on the right.

Figure 5.2.2 shows the ISEELS N 1s spectrum of gaseous Gly-Ala in comparison with the N 1s spectra of glycine, alanine and NEXAFS spectrum of solid Gly-Ala. Both the ISEELS and NEXAFS N 1s spectra of gly-ala contain a narrow and distinct peak at 401.2 eV, as observed in the N 1s of proteins [CG&03], which is assigned to the N 1s $\rightarrow \pi^*_{\text{C=ONH}}$ amide transition associated with the peptide bond. This transition arises since the C-N bond has some double-bond character due to partial overlap of C 2p orbitals of the carbonyl π^* level with the 2p level of the nitrogen. We can not observe this peak in the N 1s spectrum of the isolated amino acids, Gly and Ala. Therefore, this N 1s $\rightarrow \pi^*_{\text{C=ONH}}$ amide transition is a characteristic feature of peptides and proteins [CG&03]. The N 1s $\rightarrow \pi^*_{\text{C=ONH}}$ amide transition in the ISEELS spectrum is broader than that in the NEXAFS spectrum. This is because N 1s \rightarrow Rydberg transitions also occur in the gas phase, in the same energy region, as is observed in the N1s spectra of Gly and Ala. The corresponding states are quenched in the solid state and thus not observed in the NEXAFS spectrum.

Figure 5.2.3 shows the ISEELS O 1s spectrum of gaseous Gly-Ala in comparison with the O1s spectra of glycine, alanine, the weighted sum of glycine and alanine (Gly+Ala), and NEXAFS spectrum of solid Gly-Ala. The main peak centered at 532.3 eV both in ISEELS and NEXAFS spectra should contain contributions from both the O1s (CONH) $\rightarrow \pi^*_{\text{C=ONH}}$ amide transition and the O1s (COO⁻) $\rightarrow \pi^*_{\text{C=O}}$ carboxyl transitions. However, this peak is slightly shifted by ~ 0.3 eV to lower energy when compared to the O1s $\rightarrow \pi^*_{\text{C=OOH}}$ carboxyl transitions in the O1s spectra of Gly and Ala. In the ISEELS O 1s spectrum of Gly-ala, the O1s (OH) $\rightarrow \pi^*_{\text{C=O}}$ transition at 536 eV should be observed as in the O 1s spectra of gaseous gly and ala. But this transition was not obtained in the

ISEELS O 1s spectrum of Gly-Ala. It may be because the Gly-Ala has decomposed.

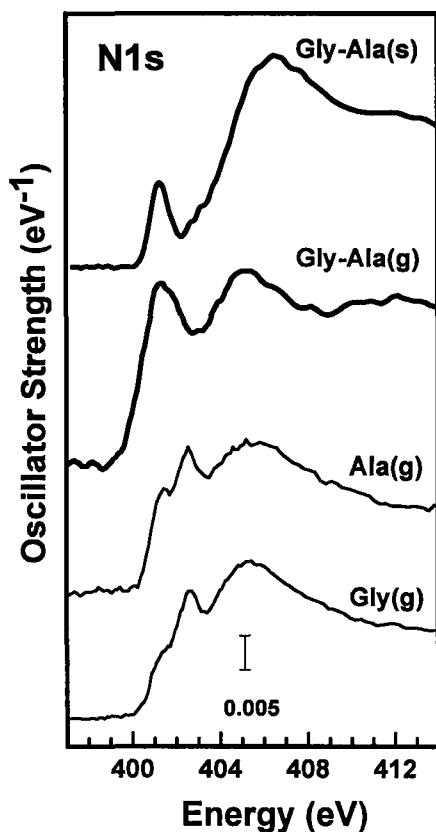


Figure 5.2.2 Comparison of the ISEELS N 1s spectra of gaseous Gly-Ala with the ISEELS N1s spectra of glycine, alanine and the NEXAFS spectra of condensed Gly-Ala.

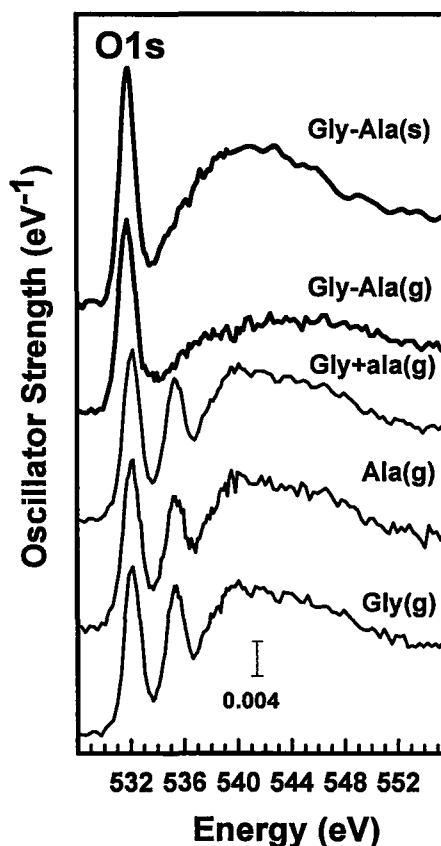


Figure 5.2.3 Comparison of the ISEELS N 1s spectra of gaseous Gly-Ala with the ISEELS N1s spectra of glycine, alanine and the NEXAFS spectra of condensed Gly-Ala.

Table 5.2.1 Energies and Term Values and proposed assignments for C 1s, N 1s and O 1s spectral features of gaseous Gly-Ala

Energy(eV)	assignment
288.3	π^*_{CONH}
288.6	π^*_{COOH}
401.5	π^*_{CONH}
405.5	$\sigma^*_{\text{NC}}, \sigma^*_{\text{NH}}$
532.3	π^*_{CONH}
532.6	π^*_{COOH}
545.0	σ^*_{CO}

§ 5.3 NEXAFS Spectra of Solid KWK

So far we have not succeeded in recording gas-phase spectra of tri-peptides, since they are heavier and thus require a higher temperature, and at the same time they appear to decompose more readily than single amino acids or dipeptides. **Figure 5.3.1** shows the C 1s NEXAFS spectra of solid Lys-Trp-Lys (KWK) in comparison to the spectra of its subunits, tryptophan and lysine, and the sum of the spectra of Trp* and Lys* using modified building block method, where the * indicates the residue spectra have been modified to take into account changes that occur with peptide bond formation. Energies, term values and proposed assignments for the experimental spectral features of KWK are listed in **Table 5.3.1**.

The procedure of this modified building block principle includes three steps. First of all, we split KWK to four parts. They are: the Trp residue which is the part of Trp without $-\text{NH}_2$ and $-\text{COOH}$ group; one Lys residue which is the C-terminus of KWK; another Lys residue which is the N-terminus of KWK, and two peptide bonds. Then, we derive the spectra of these four parts, respectively. The spectrum of Trp is derived by subtracting the spectrum of COOH and NH_2 from the spectrum of Trp. The spectrum of C-terminus Lys should be the same as the spectrum of Lys. The N-terminus Lys does not contain any unsaturated C bond, so the C 1s spectrum of this residue will not be considered. The spectrum of the peptide bond is taken from earlier work [GC&03]. The final step is to generate the simulated spectrum of KWK by summing the components in stoichiometric proportion. Looking at the comparison of the experimental C 1s spectrum of KWK with this simulated spectrum by summing the spectra of the component amino

acids and peptide bond, we can see this “modified building block” principle holds to a good approximation.

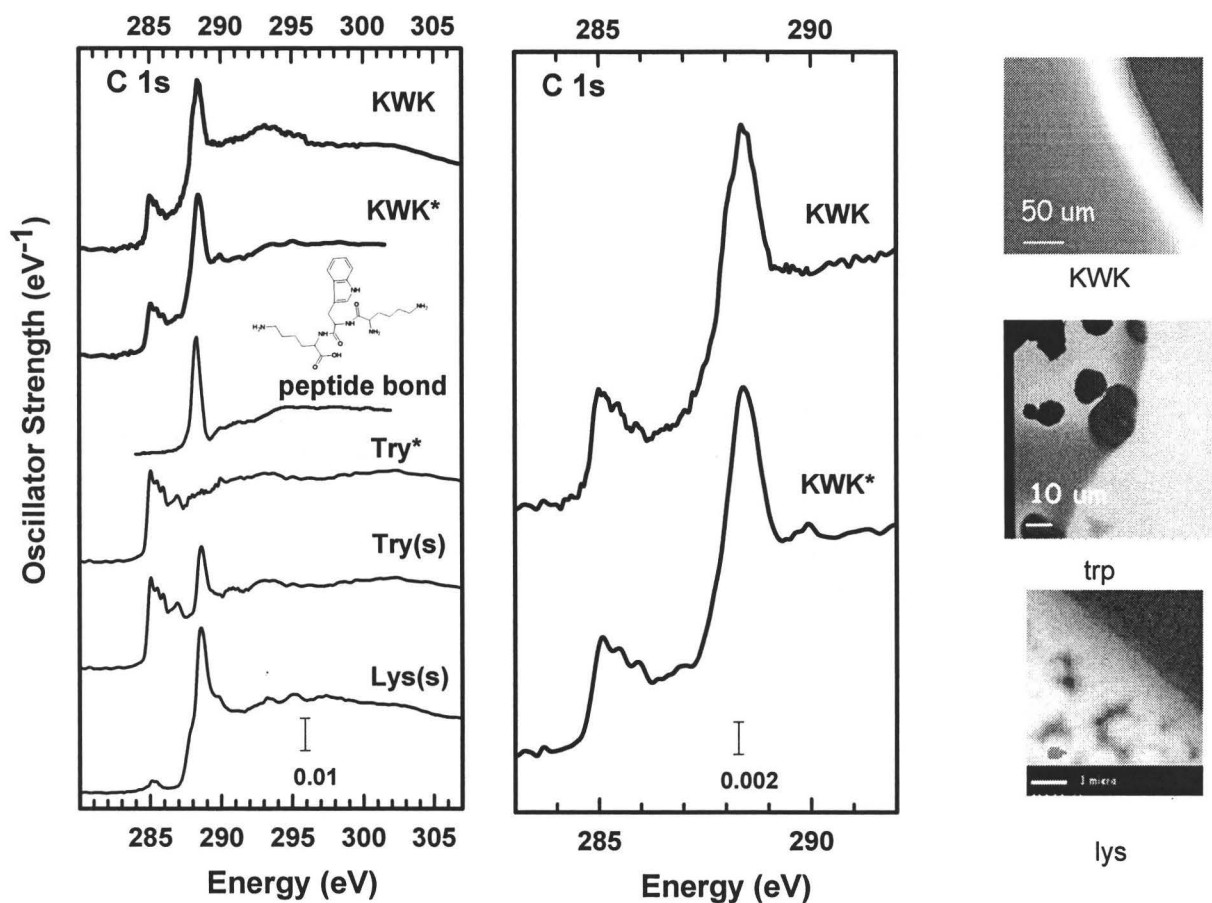


Figure 5.3.1 C 1s NEXAFS spectra of solid Lys-Trp-Lys (KWK) in comparison to the spectra of its subunits, tryptophan and lysine, and the simulated spectrum of KWK using modified building block principle. Trp* is the trp residue without –COOH and –NH₂ group. The spectrum of peptide bond is from the previous paper [GC&04]. The simulated spectra of KWK is that of KWK*, which is the sum of the spectrum of Trp*, peptide bond and Lys. An enlarged section emphasizing the difference of the spectra of KWK and KWK* is shown in the right panel. STXM images at 288.6 eV for the sample in the regions measured are displayed on the right. (bar = 1 μm)

From our previous interpretation of the C 1s spectrum of tryptophan, the first peak at 285.1 eV, observed in the C 1s spectra of both tryptophan and KWK, is assigned to C 1s (C-H_{ring}) → $\pi^*_{C=C}$ transitions. The presence of this peak reveals the Trp component of KWK. The second peak at 288.2 eV is shifted to lower energy by ~0.3 eV relative to the peak in the spectrum of tryptophan or lysine. It is attributed to C 1s → $\pi^*_{C=ONH}$ amide transitions due to the peptide bond. Its position is the same as that in Gly-Ala.

Figure 5.3.2 shows the N 1s NEXAFS spectra of solid Lys-Trp-Lys (KWK) in comparison to the spectra of its subunits, tryptophan and lysine. The N 1s spectrum of KWK shows two features. The narrow peak at 401.5 eV corresponding to the N 1s → $\pi^*_{C=ONH}$ amide transition for peptide bond is also observed in the N 1s spectrum of Gly-Ala but not in the N 1s spectra of tryptophan and lysine. A weak peak observed at 398.6 eV in both KWK and Trp is attributed to N 1s → π^*_{ring} transitions since the π^* phenyl orbital delocalizes partly on to the nitrogen atom in the tryptophan residue of KWK. The broader peak at 405.5 eV is assigned to σ^*_{NH} transitions while the dominant strong broad band peaking at 409 eV is attributed to N 1s → σ^*_{NC} transitions. Both of these σ^* features occur at the same position as corresponding features in the N 1s spectra of tryptophan and lysine. The simple building block principle does not work for the N 1s spectra of KWK since it does not predict the π^*_{CNO} feature for the peptide bond. Due to the dramatic change of the environment of the N atom during the formation of the peptide bond, it is less easy to modify the simple building block principle to more accurately predict N1s spectra of peptides and proteins.

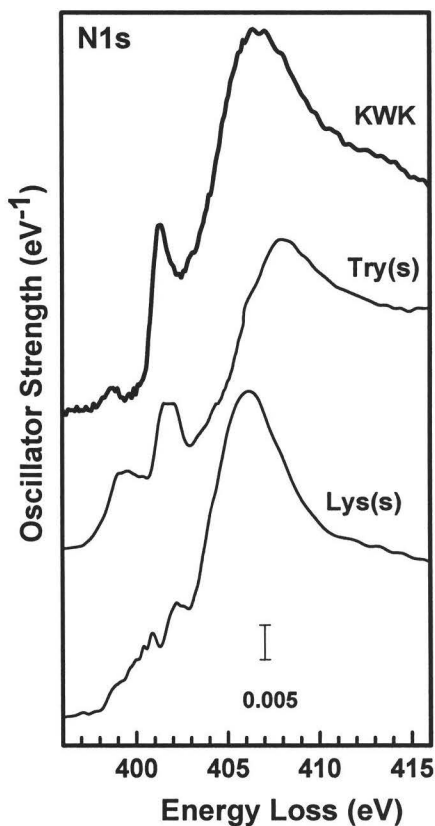


Figure 5.3.2 N 1s NEXAFS spectra of solid Lys-Trp-Lys (KWK) in comparison to the gas phase spectra of its subunits, tryptophan and lysine.

Figure 5.3.3 shows the O 1s NEXAFS spectra of solid Lys-Trp-Lys (KWK) in comparison to the spectra of its subunits, tryptophan and lysine, and the simulated spectrum of KWK. The simulated O 1s spectrum of KWK* is derived using a modified building block principle similar to that described above for the C 1s edge. Comparing the experimental O 1s spectrum of KWK and simulated spectrum of KWK, they are quite similar and this modified building block concept gives a good estimation of the O 1s

spectra of KWK as well. The O 1s spectrum of KWK is dominated by the peak at 532.3 eV, which is also slightly shifted to the lower energy ~ 0.3 eV relative to the $O1s \rightarrow \pi^*_{C=OOH}$ carboxyl transitions of lysine. Actually, this peak is an unresolved peak which contains both the $O1s (CONH) \rightarrow \pi^*_{C=ONH}$ amide and the $O1s (COO^-) \rightarrow \pi^*_{C=OO^-}$ carboxyl transitions.

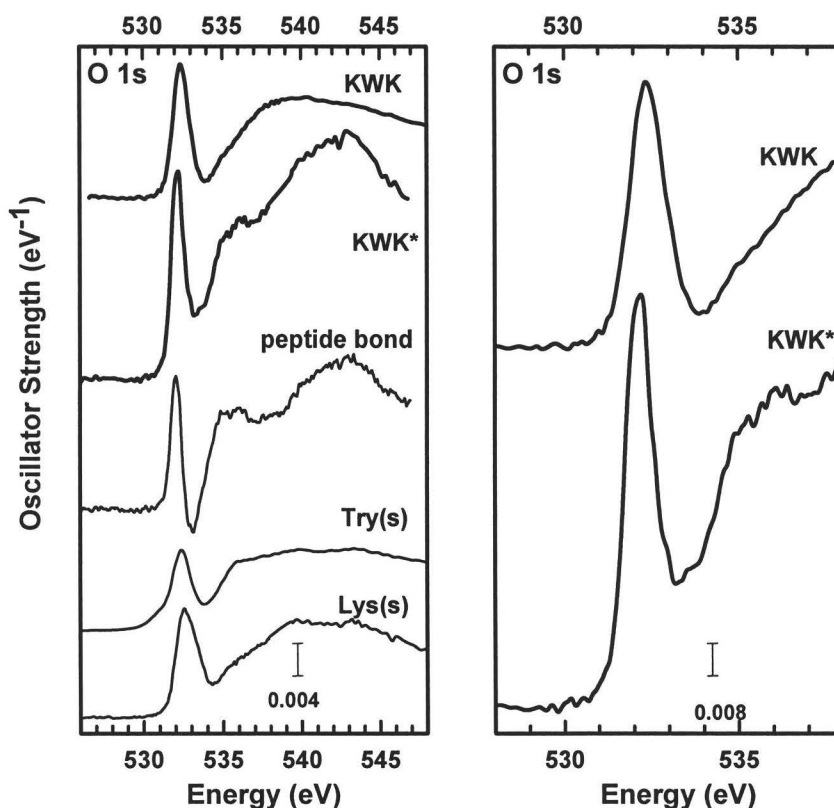


Figure 5.3.3 O 1s NEXAFS spectra of solid Lys-Trp-Lys (KWK) in comparison to the spectra of its subunits, tryptophan and lysine, and the simulated spectrum of KWK using modified building block principle. The spectrum of the peptide bond is from [GC&04]. The simulated spectra of KWK is that of KWK*, which is the sum of the spectrum of peptide bond and Lys. The Trp residue does not contribute to the simulated spectrum since the Trp residue does not contain O atom. An enlarged section emphasizing the difference of the spectra of KWK and KWK* is shown in the right panel.

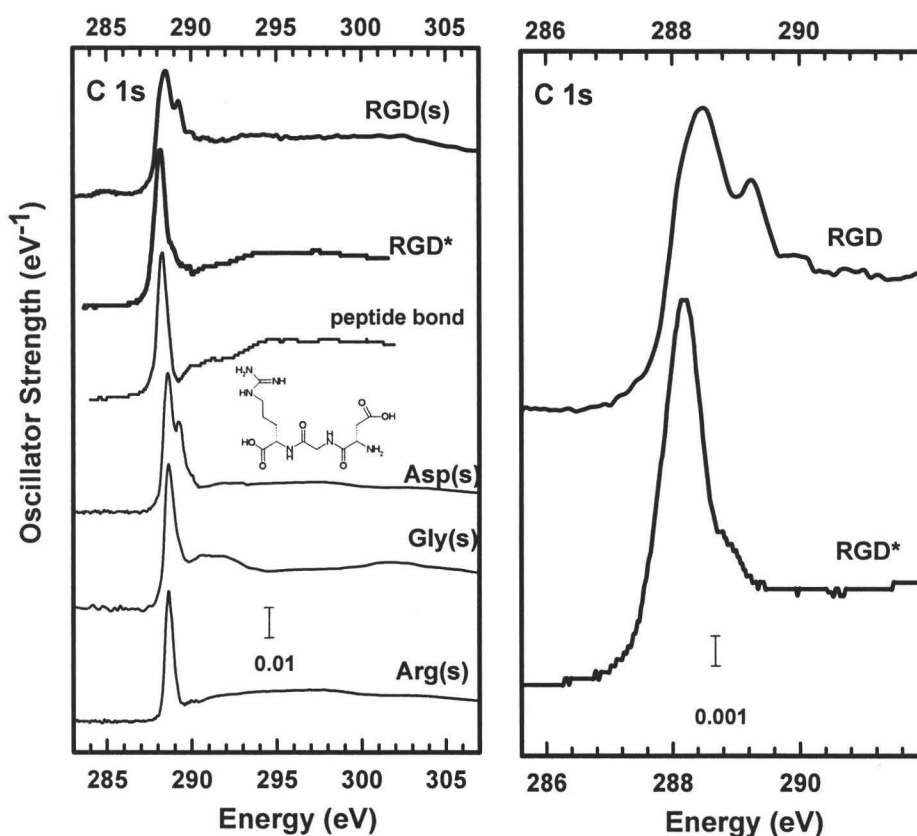
Table 5.3.1 Energies, and proposed assignments for C 1s, N 1s and O 1s spectral features of solid KWK.

Energy(eV)	assignment
285.0	π^* ring
286.7	π^* ring
288.3	π^* _{CONH}
288.6	π^* _{COOH}
398.6	π^* ring
401.5	π^* _{CONH}
405.5	σ^* _{NC} , σ^* _{NH}
532.3	π^* _{CONH}
532.6	π^* _{COOH}
545.0	σ^* _{CO}

§ 5.4 NEXAFS Spectra of Solid RGD

Figure 5.4.1 shows the C 1s spectrum of RGD in comparison with the C 1s spectra of arginine [KO&02], aspartic acid [KO&02], glycine [CG&03] and the simulated spectrum of RGD. Energies, term values and proposed assignments for the experimental spectral features of RGD are listed in **Table 5.4.1**. Based on the published interpretation of the C 1s spectrum of arginine [KO&02], the lower energy peak at 288.6 eV is C 1s→ π^* _{C=OOH} carboxyl transition, as in all the amino acids. The higher energy peak at 289.24 eV is assigned to the C 1s(C=N)→ π^* _{C=N} transition since the C=N carbon is bonded to three nitrogen atoms, which shifts the C 1s(C=N) IP to higher energy. This second π^* peak at 289.24 eV is the fingerprint of the H₂NHN=C-NH- group [KO&02]. The C 1s spectra of glycine and aspartic acid show only the C 1s→ π^* _{C=OOH} carboxyl transition. Looking at the comparison of the C 1s spectrum of RGD with that simulated by summing the spectra of the modified component amino acids and the peptide bond signal, this “modified building block” principle does a reasonable job of estimating the C 1s spectrum of RGD. The fingerprint peak for H₂NHN=C-NH- group at 289.24 eV is

quite distinct and it is at the same position as in the C 1s spectrum of arginine. We assume that we can use this specific feature to differentiate RGD from other peptides without arginine. The $C\ 1s \rightarrow \pi^*_{C=OOH}$ carboxyl and the $C\ 1s \rightarrow \pi^*_{C=ONH}$ amide transitions overlap in a single peak at 288.3 eV, as in the other peptides.



5.4.1 C 1s NEXAFS spectra of solid Arg-Gly-Asp (RGD), in comparison with the NEXAFS spectra of solid Arg [KO&02], Gly [CG&03], Asp [KO&02], and the simulated spectrum of RGD using modified building block principle. The spectrum of the peptide bond is from [GC&04]. The simulated spectrum RGD* is the sum of the spectrum of Arg, peptide bond and Asp. The Gly residue does not contribute to the simulated spectrum, while the Asp residue contributes since it contains another $-COOH$ group. An enlarged section emphasizing the difference of the spectra of RAGD and RGD* is shown in the right panel. A STXM image at 288.6 eV for the sample in the region measured is displayed on the right.

Figure 5.4.2 shows the N 1s, O 1s spectra of RGD. Energies, term values and proposed assignments for the experimental spectral features of RGD are listed in **Table 5.4.2**. In comparison with the N 1s and O 1s of KWK, they are rather similar except that the N1s spectrum of RGD does not have the weak peak at 398.6 eV, as expected since it does not contain Trp. Thus, we have similar assignments as for the C 1s, N 1s and O 1s spectra of KWK.

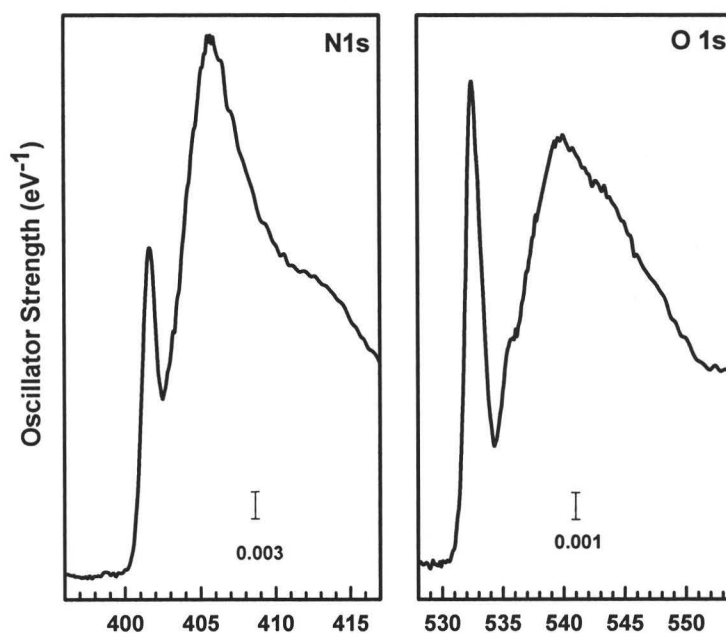


Fig. 5.4.2 The N 1s and O 1s NEXAFS spectra of solid Arg-Gly-Asp (RGD) recorded by scanning transmission X-ray microscopy at ALS.

Table 5.4.1 Energies and proposed assignments for C 1s, N 1s and O 1s spectral features of solid RGD.

Energy(eV)	assignment
288.3	π^*_{CONH}
288.6	π^*_{COOH}
289.24	C 1s(C=N) \rightarrow $\pi^*_{\text{C=N}}$
401.5	π^*_{CONH}
405.5	$\sigma^*_{\text{NC}}, \sigma^*_{\text{NH}}$
532.3	π^*_{CONH}
532.6	π^*_{COOH}
545.0	σ^*_{CO}

CHAPTER 6

SUMMARY AND FUTURE WORK

§ 6.1 Summary

In this thesis, the C 1s, N 1s and O 1s spectra of gaseous glycine, alanine, cysteine threonine, phenylalanine, proline and the C 1s spectrum of tryptophan have been reported and the spectra were analyzed, aided by GSCF3 calculations. As found elsewhere [KO&02, ZS&05] all amino acids have a C 1s \rightarrow $\pi^*_{\text{C=O}}$ carboxylic acid transition at 288.6 \pm 0.01 eV and a corresponding O 1s \rightarrow $\pi^*_{(\text{C=O})}$ transition at 535.3 \pm 0.01 eV. These two transitions are the characteristic features for the carboxylic acid group in amino acids and may be used as fingerprints. Other spectral features can be used to identify specific amino acids. The S 2p \rightarrow σ^*_{CS} peak (165 eV) and C 1s \rightarrow σ^*_{CS} transition (287.4 eV) are useful for identifying sulfur-containing amino acids. Aromatic amino acids can be distinguished using the characteristic C 1s \rightarrow π^*_{ring} at 285.1 eV.

The experimental and GSCF3 computed spectra are in good agreement given the limitations of the computational method. Thus the computational results help considerably with spectral assignments. The calculations show that the core spectra of amino acids have little conformational dependence. Any energetically sensible conformer can be chosen for the GSCF3 calculation. In contrast, there are strong differences between the NEXAFS spectra of solid amino acids and ISEELS spectra of gaseous amino acids due to zwitterionic effects. C 1s spectra have little difference, but O 1s and N 1s spectra have strong effects. For the O 1s spectra, the difference is readily explained by the

large change in structure from neutral acid to charged carboxylate. For the N 1s spectra, differences may arise from different intramolecular hydrogen bonds and by attenuation of Rydberg transitions from gas phase to the condensed phase.

Through comparisons of the spectra of small peptides to those of their subunit amino acids, the characteristic spectral signatures of peptide bonds have been identified further. The C 1s $\rightarrow \pi^*_{\text{C=O}}$ transition and O 1s $\rightarrow \pi^*_{\text{C=O}}$ transitions shift to lower energy by about 0.3 eV and there is a distinct N 1s $\rightarrow \pi^*_{\text{CNOH}}$ feature for the peptide bond. The modified “building block” approach [S92] is very useful in modeling the inner shell excitation spectroscopy of peptides through comparison to sums of the spectra of the modified constituent single amino acids. It is shown that, aside from shifts of excitations to the π^*_{carbonyl} level, the spectra of the residues of the peptides remain largely unperturbed in the peptides. With this model, we have interpreted the inner shell spectra of three peptides using the characteristic features of single amino acids. In favourable cases, it should be possible to identify small peptides using the characteristic spectral features of the constituent amino acids. The C 1s spectrum of RGD has a characteristic peak at 289.24 eV associated with arginine, while that of KWK shows a specific feature at 285.1 eV corresponding to phenyl ring.

§ 6.2 Future Work

In this study, I have collected the ISEELS spectra of all amino acids which have a relatively high volatility. Through comparing ISEELS and NEXAFS spectra of amino acids, I have explored the difference of the spectra of gas phase and solid phase amino

acids. Very recently, Zubavichus et al. [ZS&05] have reported all relevant absorption edges, i.e. C 1s, N 1s and O 1s spectra of all 20 common amino acids. But there are still issues with those spectra, for example, all spectra show a 285.1 eV peak which is not expected, except for aromatic amino acids; the O 1s spectral calibration differs by about 0.5 eV, etc. Therefore, there is a need to continue to study the inner shell spectra of amino acids both in gas and solid phase. As well, the inner shell excitation spectra of amino acids which are in different charge states by changing the pH of the solution should be studied.

Future work will improve these approaches to record, interpret and predict the spectra of other amino acids, peptides and proteins. In my follow-up PhD project I propose to study:

a) Further improved gas phase data

- alternate volatilization methods in ISEELS by using a high power laser to vaporizing the sample.
- gas phase NEXAFS using sensitive total ion yield detector at Canadian Light Source

b) Effect of local environment on spectra

- investigating the effect of different pH on the NEXAFS spectra of amino acids and peptides [as in MC&05]

c) Application of peptide spectroscopy to selective mapping in biological samples

- Add RGD peptide to a biofilm at variable concentrations and explore detection limits
- Explore NEXAFS mapping of other peptides which are differentiable from protein, and which are of interest for specific biochemical or applications reasons

In general, the results of this research have benefits in many areas where NEXAFS microscopy and spectroscopy are being applied to biological, environmental and bio-materials samples. In addition, many nanotechnology researchers are using biological macromolecules and bio-mimetic approaches to fabricate artificial structures at the nano-scale. For example, my thesis results could be useful in the study of peptide/protein orientation on surfaces. It is important to characterize protein orientation because it affects the function of proteins bound to different substrates, like polymer films [M05] or metals [BL&03]. In NEXAFS spectroscopy, molecular orientation can be determined by analyzing the polarization dependence of either the π^* or σ^* transition intensities [S92]. For a well oriented, anisotropic system such as a molecule aligned at a surface, the intensities of specific NEXAFS transitions depend on the incidence angle of the photons (α) and the tilt angle of the molecular plane (β) with respect to the substrate surface. A member of my research group (Daniel Hernández-Cruz) in collaboration with the group of Michel Pézolet (Chimie, Laval) is applying polarization dependent NEXAFS microscopy in STXM to quantitatively map the magnitude and direction of orientation of β -sheet regions in the fibroin protein in a variety of silk worm and spider silks. Improved understanding of the spectroscopy of amino acids, peptides and proteins may contribute to these areas.

In summary, this work has explored the inner shell excitation spectroscopy of amino acids and peptides. While considerable progress has been achieved, further work

is needed to improve our understanding of the inner shell spectroscopy of these important molecules, and their analytical applications.

REFERENCES

- [A99] D. Attwood, *Soft X-Rays and Extreme Ultraviolet Radiation: Principles and Application*, Cambridge University Press, 1999
- [AX&01] N. Assa-Munt, X. Jia, P. Laakkonen, E. Ruoslahti, *Biochemistry* (2001) 40, 2373
- [AZ&92] Ade, X. Zhang, S. Cameron, C. Costello, J. Kirz, and S. Williams, *Science* (1992) 258, 972
- [B30] H. Bethe, *Ann. Phys.* (1930) 5, 325
- [BB81] C. R. Brundle, A. D. Baker, *Electron Spectroscopy: Theory, Technique and Application*, V1-4 Academic Press, London (1981)
- [BCJ80] A.A. Bake, H.W. Chen, W.L. Jolly, *J. Electron. Spectrosc. Rel. Phenom.* (1980) 20, 333
- [BD&82] C. E. Brion, S. Daviel, R. N. S. Sodhi, A. P. Hitchcock, *AIP Conf. Proc.* (1982) 94, 429
- [BH81] C. E. Brion, A. Hamnett, *Adv. Chem. Phys.* (1981) 45, 1
- [BL&03] N.V. Bhagavan, E.M. Lai, P.A. Rios, J. Yang, A.M. Ortega-Lopez, H. Shinoda, S.A. Honda, C.N. Rios, C.E. Sugiyama, C.E. Ha, *Clin. Chem.*, (2003) 49, 581-585
- [BO&97] J. Boese, A. Osanna, C. Jacobsen, J. Kirz, *J. Electron. Spectrosc. Rel. Phenom.* (1997) 85, 9
- [CH79] P. H. Cannington, N. S. J. Ham, *J. Electron. Spectrosc. Rel. Phenom.* (1979) 15, 79

- [CC&91] W. F. Chan, G. Cooper, C. E. Brion, *Phys.Rev. A* (1991) 44, 186
- [CG&04] G. Cooper, M. Gordon, D.Tulumello, C.C. Turci, K. Kaznatcheev and A.P. Hitchcock, *J. Electron. Spectrosc. Rel. Phenom.* (2004) 137-140, 795
- [CJW96] H. Chapman, C. Jacobsen and S. Williams, *Ultramicroscopy*, (1996) 62, 191
- [DMH90] C. Dezarnaud, M. Tronc, A.P. Hitchcock, *Chem. Phys.* (1990) 142, 455
- [FF86] R. J. Fessenden, J. S. Fessenden, *Organic Chemistry*, Third Edition, Brooks, California, 1986.
- [FH14] J. Frank, G. Hertz, *Verh. Dtsch. Phys. Ges.* (1914) 16, 10
- [G95] S.J. Gurman, *J.Synchrotron Rad.* (1995) 2, 56-63
- [GC&03] M. L. Gordon, G.Cooper, T.Araki, C. Morin, C.C. Truci, K. Kaznatcheev, A.P. Hitchcock, *J. Phys. Chem. A* (2003) 107, 6144
- [H00] A. P. Hitchcock, *J. Electron. Spectrosc.Rel. Phenom.* (2000) 112, 9
- [H48] D.C. Harris, *Quantitative Chemical Analysis*, 3rd Edition, W.H. Freeman and Company, New York, 1948
- [H90] A. P. Hitchcock, *Phys. Scr.* (1990) T31, 159
- [H00] A. P. Hitchcock, *J. Electron. Spectrosc.Rel. Phenom.*(2000) 112, 9
- [HA&84] S. Huzinaga, J. Andzelm, M. Klobokowski, E. Radzio-Andzelm, Y. Sasaki, H. Tatewaki, *Gaussian Basis Sets for Molecular Calculations*, Amsterdam, Elsevier, 1984.
- [HB81] A. P. Hitchcock, C.E. Brion, *J. Phys.B* (1981) 14, 4399
- [HH&05] D. Hernández Cruz, A. P. Hitchcock, M. M. West, ME. Rousseau,

- M. Pézolet, unpublished
- [HI87] A. P. Hitchcock, I. J. Ishii, *J. Electron. Spectrosc. Rel. Phenom.* (1987) 42, 11
- [HL&82] B. L. Henke, P. Lee, T. J. Tanaka, R. L. Shimabukuro, B. K. Fujikawa, *Atomic Data and Nuclear Data Tables*, (1982) 27, 1-144
- [HM94] A. P. Hitchcock, D. C. Mancini, *J. Electron. Spectrosc. Rel. Phenom.* (1994) 67, 1
- [HM&02] A. P. Hitchcock, C. Morin, Y. M. Heng, R. M. Cornelius, J. L. Brash, *J. Biomater Sci. Polymer Edn.* (2002) 13, 919
- [HK&98] J. Hasselström, O. Karis, M. Weinelt, N. Wassdahl, A. Nilsson, M. Nyberg, L. G.M. Pettersson, M. G. Samant, J. Stöhr, *Surf. Sci.* (1998) 407, 221
- [HT&89] A.P. Hitchcock, M. Tronc, A. Modelli, *J. Phys. Chem.* (1989) 93, 3068
- [IH88] I. J. Ishii, A. P. Hitchcock, *J. Electron. Spectrosc. Rel. Phenom.* (1988) 46, 55
- [JW&00] C. Jacobsen, S. Wirick, G. Flynn, C. Zimba, *J. Microsc.* (2000) 197, 173
- [JS&05] H. Jiang, J. Stewart-Ornstein, G. Cooper, A. P. Hitchcock, *J. Phys. Chem. B* (in preparation)
- [K76] L. J. Klasinc, *J. Electron. Spectrosc. Rel. Phenom.* (1976) 8, 161
- [K87] N. Kosugi, *Thero. Chem. Acta.* (1987) 72, 149
- [KJH95] J. Kirz, C. Jacobsen, M. Howells, *Q. Rev. Biophys.* (1995) 28, 33-130
- [KK80] N. Kosugi, H. Kuruda, *Chem. Phys. Lett.* (1980) 74, 290

- [KO&02] K. Kaznachev, A. Osanna, C. Jacobsen, O. Plashkevych, O. Vonsahtras, H. Agren, V. Carravetta, A.P. Hitchcock, *J. Phys. Chem.* (2002) 106, 3153
- [KT&03] A.L.D. Kilcoyne, T. Tylicszak, W.F. Steele, S. Fakra, P. Hitchcock, K. Franck, E. Anderson, B. Harteneck, E.G. Rightor, G.E. Mitchell, A. P. Hitchcock, L. Yang, T. Warwick, H. Ade, *J. Synchrotron Rad.* (2003) 10, 125.
- [KT&05] F. Kaneko, M. Tanaka, S. Narita, T. Matsui, K. Nakagawa, A. Agui, K. Fujii, A. Yokoya, *J. Electron. Spectrosc. Rel. Phenom.* (2005) 144, 291
- [LFC80] D. W. Larsen, S. E. Friberg, H. Christenson, *J. Am. Chem. Soc.* (1980) 102, 6566
- [LH&02] P.W. Langhoff, R.J. Hinde, J.A. Boatz, and J.A. Sheehy, *Chem. Phys. Lett.* (2002) 358, 231
- [LS&01] Jr B.W. Loo, I. M. Sauerwald, A. P. Hitchcock, S. S. Rothman, *J. Microscopy* (2001) 204, 69-86
- [M05] C. Morin, Ph.D. Thesis, McMaster University, 2005
- [MC&05] B. M. Messer, C. D. Cappa, D. Smith, W. S. Drisdell, C. P. Schwartzl, R. C. Cohen, R. J. Saykally, *J. Phys. Chem. B* (in press)
- [MH&04] C. Morin, A. P. Hitchcock, R. M. Cornlius, J. L. Brash, S. G. Urquhart, A. Scholl, A. Doran, *J. Electron. Spectrosc. Rel. Phenom.* (2004) 137, 785
- [MI&87] R. N. S. Sodhi, I. Ishii, A.P. Hitchcock, M.B. Robin, *J. Chem. Phys.* (1987) 201, 4344

- [MJ00] M. Nyberg, J. Hasselström, O. Karis, N. Wassdahl, M. Weinelt, A. Nilsson, L. G. M. Pettersson, *J. Chem. Phys.* (2000) 112, 5420
- [MK01] M. Tanaka, K. Nakagawa, T. Koketsu, A. Agui, A. Yokoya, *J. Synchrotron Rad.* (2001) 8, 1009
- [N86] D. C. Newbery, M.Sc Thesis, McMaster University, 1986
- [OC&97] T. R. Olney, N. M. Cann, G. Cooper, C. E. Brion, *Chem. Phys.* (1997) 223, 59
- [OFK75] T. Ohta, T. Fujikawa, K. Kuroda, *Bull. Chem. Soc. Jpn.* (1975) 48, 2017
- [PJ74] W. B. Perry, W. L. Jolly, *Inorg. Chem.* (1994) 13, 1211
- [PU05] R. M. Petoral Jr., K. Uvdal, *Physica Scripta.* (2005) T115, 851–854
- [R66] S. A. Ricd, J. J. ortner, *J. Chem. Phys.* (1966) 44, 4470
- [R74] M. B. Robin, *Higher Excited States Polyatomic Molecules*, Vol.1 (Academic press, Inc., New York, 1974)
- [R88] M. B. Robin, *J. Electron. Spectrosc. Rel. Phenom.* (1988) 47, 53
- [R96] E. Ruoslahti, *Annu. Rev. Cell Dev. Biol.* (1996) 12, 697–715
- [S82] T. Steel, B.Sc Thesis, McMaster University, 1982
- [S92] J. Stöhr, *NEXAFS Spectroscopy*, Springer Tracts in Surface Science 25; Springer: Berlin, 1992.
- [SB84] R. N. S. Sodhi, C. E. Brion, *J. Electron. Spectrosc. Rel. Phenom.* (1984) 34, 363
- [SB88] A. R. Slaughter, M. S. Banana, *J. Phys. Chem.* (1988) 92, 2165

- [ST&99] I. T. Steinberger, C. M. Teodorescu, D. Gravel, R. Flesch, B. Wassermann, G. Reinhardt, C. W. Hutchings, A. P. Hitchcock, E. Rühl, *Phys. Rev. B* (1999) 60, 3995
- [T83] W. Thiel, *J. Electron. Spectrosc. Rel. Phenom.* (1983) 31, 151
- [TK&04] G. Tzvetkov, G. Koller, Y. Zubavichus, O. Fuchs, M. B. Casu, C. Heske, E. Umbach, M. Grunze, M. G. Ramsey, F P. Netzer, *Langmuir* (2004) 20 (24), 10551-10559
- [U97] S. G. Urquhart, Ph.D. Thesis, McMaster University, 1997.
- [UH&97] S. G. Urquhart, A. P. Hitchcock, A. P. Smith, H. Ade, E. G. Righter, *J. Phys. Chem. B* (1997) 101, 2267-2276
- [UT&97] S. G. Urquhart, C. C. Turci, T. Tyliczszak, M. A. Brook, A. P. Hitchcock *Organometallics* (1997) 16, 2080-2088
- [UG&05] S. G. Urquhart, R.Gillies, *J. Phys. Chem. A* (2005) 109, 2151-2159
- [W92] A. T. Wen, Ph. D. Thesis, McMaster University, 1992.
- [WA69] W. J. Goddard, W. A. Hunt, *Chem. Phys. Lett.* (1969) 3, 414
- [WA&02] T. Warwick, H. Ade, A.L.D. Kilcoyne, M. Kritscher, T. Tyliczszak, S. Fakra, A. P. Hitchcock, H.A. Padmore, *J. Synchrotron Rad.* (2002) 9, 254
- [WB99] K. Weiss, P. S. Bagus, and Ch. Wo"ll, *J. Chem. Phys.* (1999) 111, 15
- [ZS&05] Y.Zubavichus, A. Shaporenko, M. Grunze, M. Zharnikov, *J. Phys. Chem. A* (2005) 109, 6998-7000
- [ZZ&04] Y. Zubavichus, M. Zharnikov, A.Schaporenkko, M. Grunze, *J. Electron. Spectrosc. Rel. Phenom.* (2004) 134, 25

[ZZ&05] Y. Zubavichus, M. Zharnikov, Y. J. Yang, O. Fuchs, C. Heske, E. Umbach,
G. Tzvetkov, F. P. Netzer, M. Grunze, *J. Phys. Chem. B* (2005) 109, 884.