

**Spectroscopic Characterization of HIV-1 Transactivator of Transcription's  
Interactions with Divalent Metals**

by

THACH NGOC VO

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfilment of the requirements of the degree of

MASTER OF SCIENCE

Department of Chemistry

University of Manitoba

Winnipeg

Copyright © 2008 by Thach Vo

**THE UNIVERSITY OF MANITOBA  
FACULTY OF GRADUATE STUDIES  
\*\*\*\*\*  
COPYRIGHT PERMISSION**

**Spectroscopic Characterization of HIV-1 Transactivator of  
Transcription's Interactions with Divalent Metals**

**BY**

**Thach Ngoc Vo**

**A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of**

**Manitoba in partial fulfillment of the requirement of the degree**

**Of**

**Master of Science**

**Thach Ngoc Vo © 2008**

**Permission has been granted to the University of Manitoba Libraries to lend a copy of this thesis/practicum, to Library and Archives Canada (LAC) to lend a copy of this thesis/practicum, and to LAC's agent (UMI/ProQuest) to microfilm, sell copies and to publish an abstract of this thesis/practicum.**

**This reproduction or copy of this thesis has been made available by authority of the copyright owner solely for the purpose of private study and research, and may only be reproduced and copied as permitted by copyright laws or with express written authorization from the copyright owner.**

## Abstract

The HIV-1 transactivator of transcription (Tat) protein is an important regulator of viral replication via enhancement of transcription elongation. Tat binds to a specific region of the viral RNA transcript during which it recruits other viral factors and phosphorylates the carboxyl terminal domain of the RNA polymerase II. This promotes transcription elongation leading to an increase in viral RNA transcripts. The Tat protein is comprised of 101 amino acids with the first 72 amino acids are known to be essential for transactivation. Structurally, the Tat protein is an intrinsically disordered protein that is highly susceptible to aggregation via disulfide-linkage due the presence of seven cysteine residues. The main focus of the thesis is to study the structural changes of a His-tag Tat protein during the interaction with Zn(II) and Cd(II) using circular dichroism and nuclear magnetic resonance spectroscopy. The oligomerization states of the Tat protein were probed by SDS-PAGE in the presence of a reducing agent, TCEP. Analysis of SDS-PAGE data shows the Tat protein to be mainly monomeric in the presence of TCEP. Conformational studies by circular dichroism suggest that the Tat protein is in a random coil conformation and significant structural changes were observed upon addition of Zn(II) and Cd(II) at pH 5.  $^1\text{H}$ - $^{15}\text{N}$  HSQC NMR results of apo-Tat and Tat-Zn(II) complex reveal dramatic line broadening in the cysteine-rich regions accompanied by significant chemical shift changes in the His-tag and Arg-rich regions.  $^{113}\text{Cd}$  NMR chemical shift analysis of Tat in the presence of increasing concentrations of Cd(II) indicates a weak protein-metal interaction with a

dissociation constant of 1.8 mM. The analysis of the structural changes of Tat in the presence of Zn(II) and Cd(II) has revealed important information regarding which region within the Tat protein is responsible for metal-binding and that the metal interactions can induce a global conformation change. These findings contribute to an understanding of the structural studies of the Tat protein and its metal-binding environment, which can provide a pathway to a therapeutic solution to HIV.

## Acknowledgements

This dissertation would not have been written without Dr. Joe O'Neil who not only served as my supervisor but also provided invaluable advice and expertise. It is through his wisdom, guidance, enthusiasm, and encouragements that have shaped and inspired me to believe that higher learning will always be rewarding.

I would also like to express my gratitude to Drs. Shaheen Shajonia and Jamie Galka for all their advice, support, and helpful discussions throughout my research. I would like to thank Simon Baturin for moral support, advice, and humor that have got me through many difficult moments. I also thank past and present lab colleagues for their friendship.

I am deeply indebted to Dr. Kirk Marat for contributing to my knowledge of nuclear magnetic resonance and in maintaining the spectrometers. I would like to extend a special thank you to my committee members Drs. Sean McKenna and Brian Mark for reading my thesis and serving on my committee.

I gratefully acknowledge the financial support from the Department of Chemistry, Faculty of Graduate Studies, University of Manitoba, and The Natural Science and Engineering Research Council of Canada.

I owe my loving thanks to my family who has always provided unconditional love and support throughout my academic journey. Finally, I would like to give special thanks to my wife, Mai, for always being there when I needed her the most.

# Table of Contents

|  |            |
|--|------------|
| <b>Abstract</b>  | <b>ii</b>  |
| <b>Acknowledgements</b>                                      | <b>iv</b>  |
| <b>List of Figures</b>                                       | <b>x</b>   |
| <b>List of Tables</b>  | <b>xiv</b> |
| <b>Abbreviations</b>   | <b>xv</b>  |
| <br>   |            |
| <b>1. Introduction .....</b>                                 | <b>1</b>   |
| 1.1 Human Immunodeficiency Virus-1 .....                     | 1          |
| 1.1.1 HIV-1 Replication Cycle.....                           | 1          |
| 1.1.2 Role of Tat in HIV-1 Transcription .....               | 5          |
| 1.1.3 Primary Sequence of the Tat Protein and Function ..... | 9          |
| 1.1.4 Structure of the Tat Protein .....                     | 14         |
| 1.1.5 Natively Disordered Proteins .....                     | 15         |
| 1.1.6 Other Activities of Tat .....                          | 16         |
| 1.2 History of Zinc-binding Proteins.....                    | 20         |
| 1.2.1 Significance of Zinc in Biological Activities .....    | 20         |
| 1.2.2 Function of Zinc in Cells .....                        | 21         |
| 1.2.3 Metabolism of Zinc in Cells.....                       | 23         |
| 1.2.4 Comparison between Zinc and Cadmium .....              | 24         |

|   |           |
|---|-----------|
| 1.2.5 Zinc fingers and Comparison with Cys-region of Tat to Zinc Finger |           |
| Sequences .....   | 25        |
| 1.2.6 Zn(II) induced dimerization of Tat protein .....                  | 26        |
| 1.2.7 Zinc Coordination Geometry.....                                   | 27        |
| 1.3 SDS-Polyacrylamide Gel Electrophoresis.....                         | 30        |
| 1.4 Circular Dichroism Spectropolarimetry .....                         | 31        |
| 1.5 Nuclear Magnetic Resonance Spectroscopy .....                       | 33        |
| 1.5.1 Heteronuclear Single Quantum Correlation Spectroscopy .....       | 34        |
| 1.5.2 Heteronuclear Multiple Bond Correlation Spectroscopy.....         | 36        |
| 1.5.3 Nuclear Overhauser Effect Spectroscopy.....                       | 39        |
| 1.5.4 <sup>113</sup> Cadmium NMR Spectroscopy .....                     | 44        |
| 1.6 Ultra-Violet Absorption Spectroscopy using Cd(II) .....             | 48        |
| 1.7 Goals of the Research .....   | 49        |
| <b>2. Materials and Methods .....</b>                                   | <b>50</b> |
| 2.1 Plasmid Construction.....   | 50        |
| 2.2 Expression of Unlabeled Tat in <i>E. coli</i> .....                 | 50        |
| 2.3 Expression of Labeled <sup>15</sup> N-Tat in <i>E. coli</i> .....   | 51        |
| 2.4 Purification of His-tagged Tat.....                                 | 53        |
| 2.5 Alkylation of the Cysteine-rich Region of Tat .....                 | 54        |
| 2.6 Circular Dichroism Spectropolarimetry .....                         | 55        |
| 2.6.1 pH Titration of Tat by CD.....                                    | 55        |

|  |           |
|--|-----------|
| 2.7 SDS Polyacrylamide Gel Electrophoresis.....                          | 56        |
| 2.8 Nuclear Magnetic Resonance (NMR) Spectroscopy .....                  | 57        |
| 2.8.1 Heteronuclear Single Quantum Correlation Spectroscopy .....        | 57        |
| 2.8.2 Heteronuclear Multiple-Bond Correlation Spectroscopy.....          | 58        |
| 2.8.3 Nuclear Overhauser Effect Spectroscopy .....                       | 59        |
| 2.8.4 <sup>113</sup> Cadmium NMR Spectroscopy .....                      | 59        |
| 2.9 Cadmium Titration by Ultraviolet Absorption Spectroscopy .....       | 63        |
| <b>3. Results .....</b>  | <b>64</b> |
| 3.1 Protein Expression, Purification, and Protein Concentration.....     | 64        |
| 3.1.1 Protein Purification via Cobalt Metal-affinity Chromatography..... | 65        |
| 3.1.2 Ultraviolet Absorption Spectroscopy of HIV-1 Tat.....              | 66        |
| 3.2 SDS-PAGE Analysis.....   | 68        |
| 3.2.1 Analysis of Tat Fractions from Cobalt Metal-affinity Column. ....  | 68        |
| 3.2.2 Analysis of the Effect of TCEP on Tat.....                         | 70        |
| 3.2.3 Determining the Optimal Concentration of TCEP .....                | 72        |
| 3.2.4 Analysis of Alkylated Tat.....                                     | 75        |
| 3.3 Circular Dichroism .....   | 77        |
| 3.3.1 Deconvolution of the CD Spectrum of HIV-1 Tat .....                | 77        |
| 3.3.2 Effects of pH on Tat CD Spectra .....                              | 80        |
| 3.3.3 Effects of pH on Tat-Zn(II) Complex .....                          | 86        |
| 3.3.4 Zn(II) Titration of Tat at pH 4.0.....                             | 88        |

|   |            |
|---|------------|
| 3.3.5 Zn(II) Titration of Tat at pH 5.0.....  | 90         |
| 3.3.6 Zn(II) Titration of Tat at pH 7.2.....  | 92         |
| 3.3.7 Cd(II) Titration of Tat at pH 4.0 .....   | 94         |
| 3.3.8 Cd(II) Titration of Tat at pH 5.0 .....   | 97         |
| 3.3.9 Cd(II) Titration of Tat at pH 7.1 .....   | 99         |
| <b>3.4 Nuclear Magnetic Resonance Spectroscopy (NMR).....</b>   | <b>102</b> |
| 3.4.1 Natural Abundance 2D <sup>1</sup> H- <sup>15</sup> N HSQC and 1D <sup>1</sup> H NMR of Unlabeled HIV-1<br>Tat.....                              | 102        |
| 3.4.2 2D <sup>1</sup> H- <sup>15</sup> N HSQC NMR on <sup>15</sup> N-labeled HIV-1 Tat and Sequence<br>Assignments .....                              | 106        |
| 3.4.3 2D <sup>1</sup> H- <sup>15</sup> N HSQC NMR on <sup>15</sup> N-Labeled Tat in the Absence and Presence of<br>Zn(II) .....                       | 109        |
| 3.4.4 <sup>1</sup> H- <sup>15</sup> N HMBC Spectra of Apo-Tat and the Tat-Zn Complex .....  | 126        |
| 3.4.5 2D <sup>1</sup> H- <sup>15</sup> N HSQC NMR on <sup>15</sup> N-Labeled Tat in the presence of 0.5 MEQ of<br>Zn(II) at different pH values ..... | 131        |
| 3.4.6 2D <sup>1</sup> H- <sup>13</sup> C HMQC NMR on <sup>13</sup> C-Labeled Tat and Tat-Zn(II) Complex .....   | 134        |
| 3.4.7 NOESY/ROESY NMR .....   | 138        |
| 3.4.8 <sup>113</sup> Cd NMR Titration Experiment .....  | 152        |
| <b>3.5 UV Absorption Analysis of Tat in the presence of Cadmium.....</b>  | <b>156</b> |
| <b>4. Discussion .....</b>  | <b>159</b> |
| 4.1 Protein Purification and Expression .....   | 159        |
| 4.2 SDS-PAGE .....  | 161        |

|   |            |
|---|------------|
| 4.2.1 Alkylation of Tat.....  | 162        |
| 4.3 Circular Dichroism .....  | 164        |
| 4.3.1. Secondary Structural Analysis of Apo-Tat and Content predicted by<br>Dichroweb .....       | 164        |
| 4.3.2 Effects of pH Titration .....   | 166        |
| 4.3.3 Effects of Zn(II) at pH 4.0, 5.0, and 7.2.....  | 169        |
| 4.3.4 Effects of Cd(II) at pH 4.0, 5.0, and 7.1.....  | 170        |
| 4.4 Nuclear Magnetic Resonance Spectroscopy .....   | 172        |
| 4.4.1 $^1\text{H}$ - $^{15}\text{N}$ Heteronuclear Single Quantum Correlation Spectroscopy.....   | 172        |
| 4.4.2 Effect of Zn(II) .....  | 173        |
| 4.4.3 $^1\text{H}$ - $^1\text{H}$ Nuclear Overhauser Effect Spectroscopy.....                     | 178        |
| 4.4.4 $^1\text{H}$ - $^{15}\text{N}$ Heteronuclear Multiple Quantum Correlation Spectroscopy..... | 182        |
| 4.4.5 $^1\text{H}$ - $^{13}\text{C}$ HMQC Analysis of Apo-Tat and Tat-Zn(II) Complex.....         | 183        |
| 4.4.6 $^{113}\text{Cd}$ NMR Spectroscopy.....   | 184        |
| 4.5 UV Spectroscopy of Metal-binding using Cd(II) .....   | 186        |
| Future Perspectives.....  | 187        |
| <b>5. Conclusions.....</b>  | <b>190</b> |
| <b>References.....</b>  | <b>193</b> |
| <b>Appendix A.....</b>  | <b>210</b> |
| <b>Appendix B.....</b>  | <b>211</b> |

## List of Figures

|  |    |
|--|----|
| Figure 1.1: The life cycle of HIV-1 Virus .....  | 3  |
| Figure 1.2: Model for the activation of RNAPII by Tat.....   | 6  |
| Figure 1.3: Model of interaction between Tat, TAR, and p-TEFb .....  | 8  |
| Figure 1.4: The HIV-1 Tat sequence corresponding to the BH10 Tat isolate .....   | 10 |
| Figure 1.5: Viral life cycle during HIV-mediated apoptosis.....  | 20 |
| Figure 1.6: Zinc coordination geometry for the catalytic, structural, and co-catalytic<br>sties of Zinc-binding.....                           | 28 |
| Figure 1.7: CD spectra illustrating the key characteristics of secondary structures:<br>$\alpha$ -helix, $\beta$ -sheet, and random coil ..... | 33 |
| Figure 1.8: Pulse sequence for a standard HSQC experiment .....  | 36 |
| Figure 1.9: Pulse sequence for a standard HMBC experiment .....  | 38 |
| Figure 1.10: Pulse sequence of a standard NOESY experiment .....   | 41 |
| Figure 1.11: Effect of rotational correlation time on NOE.....   | 44 |
| Figure 1.12: $^{113}\text{Cd}$ Chemical shift ranges for cadmium metallothioneins.....   | 47 |
| Figure 2.1: Model for the binding curve .....  | 62 |
| Figure 3.1: Growth curve for <i>E. coli</i> cells.....   | 64 |
| Figure 3.2: Concept of cobalt metal-affinity column purification.....  | 66 |
| Figure 3.3: UV Absorption spectrum of HIV-1 Tat .....  | 67 |
| Figure 3.4: SDS-PAGE of the Tat protein fractions after purification.....  | 69 |
| Figure 3.5: SDS-PAGE results of Tat protein in the presence of TCEP .....  | 71 |

|  |     |
|--|-----|
| Figure 3.6: SDS-PAGE analysis of Tat in the presence of increasing concentration of TCEP .....   | 74  |
| Figure 3.7: SDS-PAGE analysis of Tat and alkylated Tat protein .....   | 76  |
| Figure 3.8: Far-UV CD spectrum of HIV-1 apo-Tat.....   | 78  |
| Figure 3.9: CD spectra of Tat at different pH .....  | 83  |
| Figure 3.10: CD spectra of apo-Tat in the absence of 0.5 MEQ Zn (II) at various pH conditions .....  | 87  |
| Figure 3.11: CD spectra of apo-Tat in the presence of increasing concentrations of Zn(II) at pH 4.0 .....  | 89  |
| Figure 3.12: CD spectra of apo-Tat in the presence of increasing concentrations of Zn(II) at pH 5.0 .....  | 91  |
| Figure 3.13: CD spectra of apo-Tat in the presence of increasing concentrations of Zn(II) at pH 7.2 .....  | 93  |
| Figure 3.14: CD spectra of apo-Tat in the presence of increasing concentrations of Cd(II) at pH 4.0.....   | 96  |
| Figure 3.15: CD spectra of apo-Tat in the presence of increasing concentrations of Cd(II) at pH 5.0.....   | 98  |
| Figure 3.16: CD spectra of apo-Tat in the presence of increasing concentrations of Cd(II) at pH 7.1.....   | 101 |
| Figure 3.17: One-dimensional $^1\text{H}$ NMR and Two-dimensional natural abundance $^1\text{H}$ - $^{15}\text{N}$ HSQC spectra of apo-Tat at pH 3.5 ..... | 103 |

|  |     |
|--|-----|
| Figure 3.18: $^1\text{H}$ - $^{15}\text{N}$ HSQC NMR spectrum of isotopically-labeled HIV-1 Tat protein<br>.....   | 107 |
| Figure 3.19: $^1\text{H}$ - $^{15}\text{N}$ HSQC spectra of apo-Tat and Tat in the presence of 0.5 mole<br>equivalents of Zn(II) at pH 4.02.....                   | 112 |
| Figure 3.20: $^1\text{H}$ - $^{15}\text{N}$ HSQC cross peak intensities for residues of apo-Tat and Tat in<br>the presence of 0.5 mole equivalents of Zn(II) ..... | 114 |
| Figure 3.21: Chemical shift difference plots for (a) $^1\text{HN}$ and (b) $^{15}\text{N}$ chemical shifts<br>.....  | 115 |
| Figure 3.22: $^1\text{H}$ - $^{15}\text{N}$ HSQC spectra of Tat in the presence of 0.5 and 1.0 mole<br>equivalents of Zn(II) at pH 4.02.....                       | 117 |
| Figure 3.23: $^1\text{H}$ - $^{15}\text{N}$ HSQC cross peak intensities for residues of Tat in the presence<br>of 0.5 and 1.0 mole equivalents of Zn(II).....      | 119 |
| Figure 3.24: Chemical shift difference plots for (a) $^1\text{HN}$ and (b) $^{15}\text{N}$ chemical shifts<br>.....  | 120 |
| Figure 3.25: $^1\text{H}$ - $^{15}\text{N}$ HSQC spectra of Tat in the presence of 1.0 and 2.0 mole<br>equivalents of Zn(II) at pH 4.02.....                       | 122 |
| Figure 3.26: $^1\text{H}$ - $^{15}\text{N}$ HSQC cross peak intensities for residues of Tat in the presence<br>of 1.0 and 2.0 mole equivalents of Zn(II) .....     | 124 |
| Figure 3.27: Chemical shift difference plots for (a) $^1\text{HN}$ and (b) $^{15}\text{N}$ chemical shifts<br>.....  | 125 |
| Figure 3.28: $^1\text{H}$ - $^{15}\text{N}$ HMBC spectra of apo-Tat and Tat in the presence of increasing<br>concentration of Zn(II).....                          | 128 |

|  |     |
|--|-----|
| Figure 3.29: $^1\text{H}$ - $^{15}\text{N}$ HSQC overlay of apo-Tat (red) and Tat in the presence of 0.5 mole equivalents of Zn at different pH.....                                     | 132 |
| Figure 3.30: Aliphatic (A) and aromatic (B) regions of a $^1\text{H}$ - $^{13}\text{C}$ HMQC spectrum of apo-Tat and Tat in the presence of 0.5 mole equivalents of Zn(II) .....         | 136 |
| Figure 3.31: Expanded regions of the $^1\text{H}$ - $^{13}\text{C}$ HMQC spectra of the (A) aliphatic and (B and C) aromatic regions of apo-Tat.....                                     | 137 |
| Figure 3.32: $^1\text{H}$ - $^1\text{H}$ NOESY (A) and ROESY (B) spectra of apo-Tat .....  | 140 |
| Figure 3.33: $^1\text{H}$ - $^1\text{H}$ NOESY spectra of Tat at pH 4.0 with different mixing times  | 143 |
| Figure 3.34: $^1\text{H}$ - $^1\text{H}$ NOESY spectra of apo-Tat (A) and Tat in the presence of 0.5 mole equivalents of Zn(II) (B) at pH 4.0.....                                       | 146 |
| Figure 3.35: $^1\text{H}$ - $^1\text{H}$ NOESY spectrum of apo-Tat that was dialyzed against Ni-free NTA resin .....   | 149 |
| Figure 3.36: $^1\text{H}$ - $^{15}\text{N}$ HSQC cross peak intensities for residues of Tat dialyzed against Ni-free NTA resin and control Tat.....                                      | 150 |
| Figure 3.37: $^1\text{H}$ - $^{15}\text{N}$ HSQC cross peak intensities for residues of apo-Tat and Tat in the presence of 0.5 mole equivalents of Zn(II) dialyzed against Ni-free NTA.. | 151 |
| Figure 3.38: $^{113}\text{Cd}$ NMR spectra of $\text{CdCl}_2$ titrations with HIV-1 Tat protein at pD 5.0 .....  | 154 |
| Figure 3.39: Ultraviolet absorption spectra of apo-Tat with 0.5-2.0 mole equivalents of Cd(II) at pH 4.5.....  | 157 |
| Figure 4.1: Illustration of the equilibrium between multiple conformations. ....   | 166 |

|   |     |
|---|-----|
| Figure 4.2: Schematic representation bond between glutamic or aspartic acid with<br>the imidazole ring..... | 175 |
| Figure 4.3: Model representation of a Tat metal-linked dimer. ....  | 177 |

## List of Tables

|   |    |
|---|----|
| Table 2.1: M9 Minimal Medium ingredients .....  | 52 |
| Table 3.1: Predicted secondary structure content of HIV-1 Tat at pH 4.17.....                       | 80 |
| Table 3.2: Predicted secondary structure content of HIV-1 Tat at pH 8.06.....                       | 82 |
| Table 3.3: Comparison between secondary structural content of HIV-1 Tat at pH<br>4.17 and 8.06..... | 82 |

## Abbreviations

|                     |  |
|---------------------|--|
| AIDS                | Acquired Immune Deficiency Syndrome        |
| ALS                 | amyotrophic lateral sclerosis              |
| Casp-8              | Caspase-8                                  |
| Cd(II)              | Cadmium                                    |
| CBP                 | CREB-binding protein                       |
| CD                  | circular dichroism                         |
| CDK9                | cyclin-dependent kinase 9                  |
| CDKs                | cyclin-dependent kinases                   |
| CREB                | cAMP response element binding protein      |
| CTD                 | carboxyl terminal domain                   |
| CV                  | column volumes                             |
| Da                  | Dalton                                     |
| deg                 | degrees                                    |
| dmol                | decimole                                   |
| D <sub>2</sub> O    | deuterated water                           |
| dd H <sub>2</sub> O | double-deionized water                     |
| DNA                 | deoxyribonucleic acid                      |
| DNase 1             | deoxyribonuclease I                        |
| DPIV                | dipeptidyl peptidase IV                    |
| DSIF                | DRB sensitivity inducing factors           |
| DSS                 | 2,2-Dimethyl-2-silapentane-5-sulfonic acid |

|                |   |
|----------------|---|
| <i>E. coli</i> | <i>Escherichia coli</i>                               |
| EIAV           | Equine Infectious Anemia Variant                      |
| FADD           | Fas-Associated protein with Death Domain              |
| FVIIa          | coagulation factor VIIa                               |
| g              | gram  |
| <i>g</i>       | acceleration due to gravity                           |
| Gdn-HCl        | guanidine hydrogen chloride                           |
| gp120          | glycoprotein 120                                      |
| HAART          | highly active antiretroviral therapy                  |
| HIV            | Human immunodeficiency virus                          |
| His-tag        | hexahistidine affinity tag                            |
| HMBC           | Heteronuclear Multiple-Bond Correlation               |
| HSPGs          | heparin sulfate proteoglycans                         |
| HSQC           | Heteronuclear Single Quantum Coherence                |
| Hz             | Hertz   |
| INEPT          | Insensitive Nucleus Enhanced by Polarization Transfer |
| IPTG           | isopropyl- <sup>β</sup> -D-thiogalactopyranoside      |
| IRE            | iron responsive element                               |
| K              | Kelvin  |
| kDa            | kilodalton  |
| KIX            | KID-interacting domain of CBP                         |
| KS             | Kaposi's sarcoma                                      |

|                |   |
|----------------|---|
| L              | litre                                       |
| LTR            | long terminal repeat                        |
| M              | molar                                       |
| mL             | milliliter                                  |
| mM             | millimolar                                  |
| mg             | milligram                                   |
| ms             | millisecond                                 |
| $\mu$ L        | microlitre                                  |
| $\mu$ s        | microsecond                                 |
| MEQ            | mole equivalents                            |
| mRNA           | messenger RNA                               |
| MT             | Metallothionein                             |
| NELF           | negative elongation factor                  |
| NF- $\kappa$ B | nuclear factor-kappa B                      |
| NLS            | nuclear localization segment                |
| NMR            | nuclear magnetic resonance                  |
| NOE            | nuclear Overhauser effect                   |
| NOESY          | nuclear Overhauser enhancement spectroscopy |
| ns             | nanosecond                                  |
| OD             | optical density                             |
| PCAF           | p300/CBP associated factor                  |
| PCR            | polymerase chain reaction                   |

|          |   |
|----------|---|
| ppm      | parts per million   |
| p-TEFb   | positive transcription elongation factor-b                |
| RGD      | Arginine-Glycine-Aspartate                                |
| RNA      | ribonucleic acid  |
| RNAPII   | RNA polymerase II   |
| RNase 1  | ribonuclease I  |
| ROESY    | Rotating frame NOESY                                      |
| RT       | reverse transcriptase                                     |
| s        | second  |
| SDS-PAGE | sodium dodecyl sulfate polyacrylamide gel electrophoresis |
| TAK      | Tat-TAR-kinase complex                                    |
| TAR      | transactivation responsive element                        |
| Tat      | transactivator of transcription                           |
| TB       | Terrific Broth, modified                                  |
| TCEP     | tris(2-carboxyethyl)phosphine                             |
| UV       | ultra-violet  |
| WT1      | Wilms tumor suppressor protein                            |
| Zn(II)   | Zinc  |

# 1. Introduction

## 1.1 Human Immunodeficiency Virus-1

The human immunodeficiency virus (HIV) is the viral agent responsible for the end-stage disease, acquired immune deficiency syndrome (AIDS) [1]. AIDS, first recognized in 1981, is classified as an autoimmune disease due to the modification and destruction of the immune system in response to the HIV infection. In HIV infected patients, T-lymphocytes have been reported to undergo programmed cell death (apoptosis) [2], which may account for T-helper cell depletion and also explains the initial characterization of AIDS as the depletion of CD4<sup>+</sup> or mature T-helper lymphocytes in the blood. As a result of low levels of CD4<sup>+</sup> T-cells, the infected patient becomes increasingly susceptible to opportunistic infections [3] that would otherwise have no effect on a healthy patient's immune system.

The classification of HIV can be divided into two species: HIV type 1 (HIV-1) and type 2 (HIV-2). The significant difference between HIV-1 and HIV-2 is that the former has a higher infectivity than the latter [4]. Thus, HIV-1 represents a significant majority of all HIV infections.

### *1.1.1 HIV-1 Replication Cycle*

HIV-1 infection (Figure 1.1) begins with the viral gp120 surface envelope protein binding to the host cell-surface receptor, CD4<sup>+</sup> molecule, located on the surfaces of

lymphocytes and macrophages, and is followed by fusion of the viral membrane with the cell membrane [5]. Two coreceptors, the  $\alpha$ -chemokine receptor CXCR4 and the  $\beta$ -chemokine receptor CCR5, were discovered to be required in assisting the HIV-1 mediated membrane fusion, and determine cellular tropism [6]. Immediately after membrane fusion, the viral core enters the host cell. A viral enzyme, reverse transcriptase (RT), transcribes a DNA copy from the viral RNA genome. Integration of viral DNA (provirus) into the host genome inside the nucleus is accomplished by the viral enzyme, integrase. The provirus is transcribed into RNA molecules by the host's RNA polymerase II (RNAPII). The RNA molecules can either function as mRNA in producing viral protein or as new viral genomes corresponding to infected particles prior to release from the host [7].

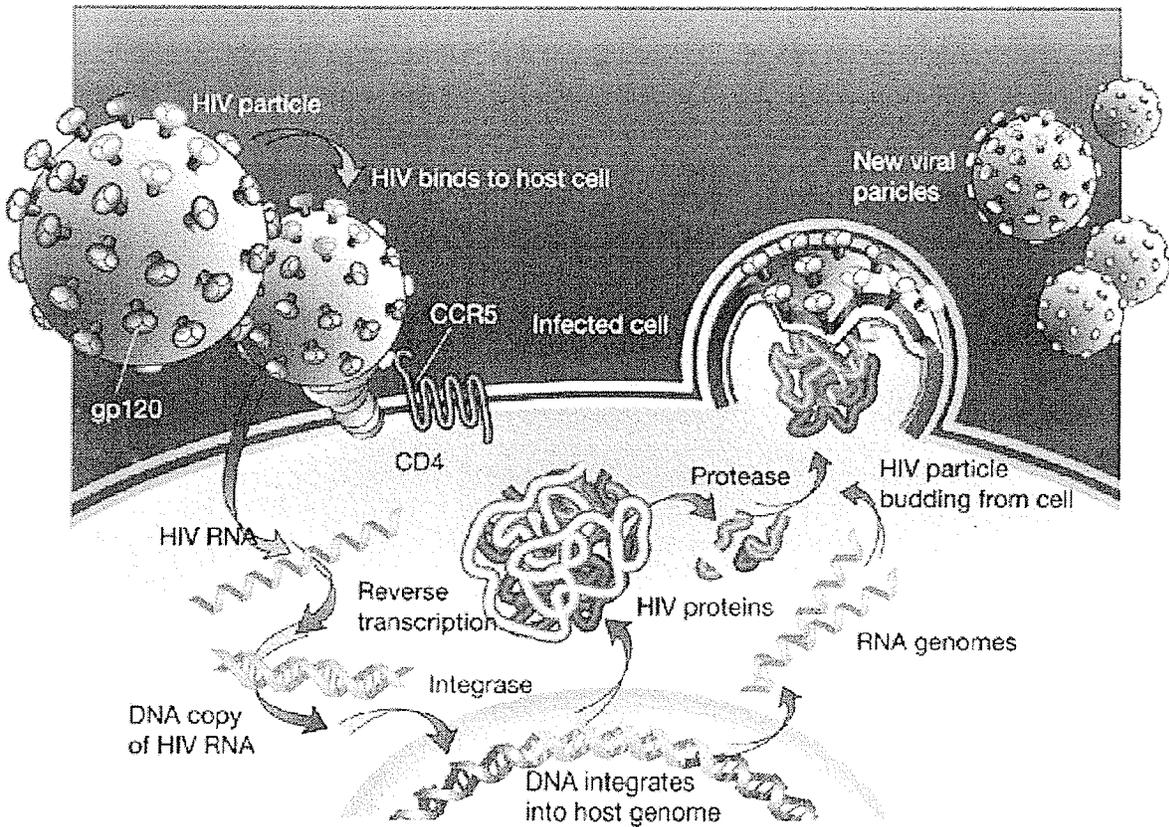


Figure 1.1: The life cycle of HIV-1 Virus. Reprinted from [8] with the permission of Weiss R.A.

Patients infected with HIV-1 contain a small number of memory CD4<sup>+</sup> T-cells with fully integrated provirus that is transcriptionally silent due to transcriptional interference or mutations. This is known as HIV-1 latency. However, the transcriptional silencing of provirus is not permanent as viral expression can be induced by the presence of a viral promoter or the rearrangement of chromatin [9].

The importance of HIV-1 latency in memory CD4<sup>+</sup> T-cells is seen in the survival of HIV-1 through avoidance of host immune response and antiretroviral drugs.

Currently, the treatment of HIV-infected patients involves reduction of viral plasma loads via highly active antiretroviral therapy (HAART), which is a combination of HIV-1 reverse-transcriptase, protease, and gp41 inhibitors. Although HAART was shown to be successful in massively reducing HIV-1 levels in the plasma of infected patients, the levels of latent virus reservoirs remain unchanged [10]. Currently, research is aimed at discovering compounds that induce HIV-1 expression in latently-infected cells during HAART treatment. In particular, a non-tumor promoting molecule, Prostratin, was shown to induce expression of HIV-1 in latently-infected cells in conjunction with HAART treatment thus increasing the efficacy of depleting active viral load [11]. However, one of the limiting factors in the use of Prostratin has been the inefficient isolation of the compound from its natural source, the bark of the mamala tree of Samoa. Other research is aimed at characterizing key factors that regulate HIV-1 expression in latently infected cells. Of those key factors, the transactivator of transcription (Tat), a regulatory protein, plays an important role in inducing HIV-1 expression in HIV-1 latent cells. In the presence of the Tat protein, activation of transcriptional elongation occurs leading to the efficient production of full-length viral RNA transcripts that would otherwise be truncated in the absence of the protein [12].

### *1.1.2 Role of Tat in HIV-1 Transcription*

HIV-1 transcription is controlled by host cell and viral factors but the primary regulator is Tat [13]. Interestingly, Tat is the first example of the regulation of gene expression through the control of RNAPII elongation [12]. Early studies showed that Tat increases the level of RNAs transcribed from the HIV-1 long-terminal (LTR) repeat by approximately 20-50 fold [21]. This led to a substantial difference in the population of RNA transcripts in the absence and presence of Tat. Figure 1.2 shows a model for the transcription elongation process in the absence and presence of the Tat protein. The LTR found at the 5' end of the proviral DNA within the nucleus of infected cells contains the HIV-1 promoter that governs the transcriptional activity of the provirus [15]. The primary phases of the transcription cycle are pre-initiation, initiation, promoter clearance, elongation, and termination [205]. In the absence of Tat, a pre-initiation complex is assembled on the LTR by RNAPII and several general transcription factors (TFIIB, E, F). An important event in transcription initiation is the phosphorylation of Ser-5 on the carboxyl terminus of the polymerase by TFIIH. Promoter clearance and elongation are positively and negatively regulated by several viral and cellular factors including the positive transcription elongation factor b (p-TEFb). The p-TEFb is comprised of the cyclin-dependent kinase-9 (CDK9) and cyclin T1. The carboxyl-terminal domain (CTD) of the RNAPII is further phosphorylated at Ser-2 by the p-TEFb, permitting clearance of the promoter by the RNAPII. Following the phosphorylation, the RNAPII transcribes the transactivation-responsive (TAR) element, which then folds into a

stem-loop structure. After promoter clearance, the negative elongation factor (NELF) and 5,6-dichloro-1-b-D-ribofuranosylbenzimidazole (DRB) sensitivity inducing factors (DSIF) associate with the transcription complex and inhibit the kinase activity of CDK9. This results in abortive elongation whereby the transcribed nascent RNA is truncated [16] (see the left side of Figure 1.2). This abortive elongation stage halts the transcription of full-length viral RNA (premature termination).

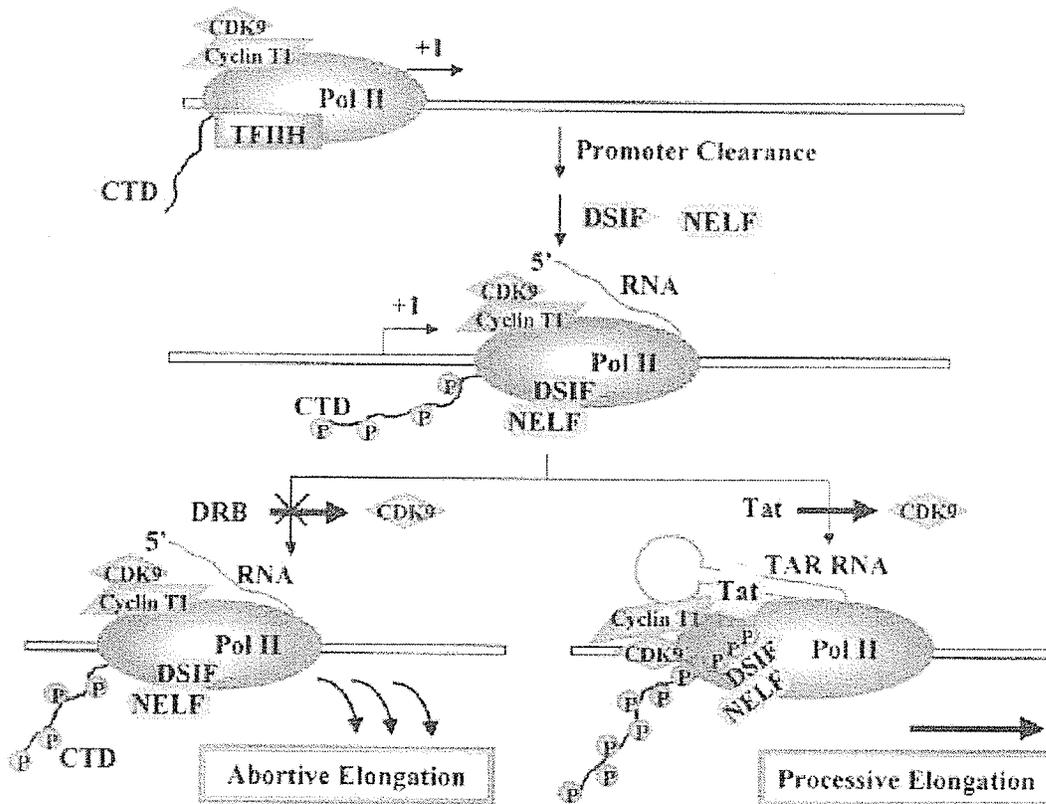


Figure 1.2: Model for the activation of RNAPII by Tat and cellular co-factors. Reprinted from [16] with permission of Ping Y-H.

The Tat protein is recruited to the transcription complex through interaction with the TAR RNA on the nascent chain and p-TEFb. The arginine-rich region of Tat binds specifically to the U-rich tri-nucleotide bulge within the stem-loop region of the TAR RNA (Figure 1.3) [17]. Subsequently, the Tat-TAR complex associates with the p-TEFb relocating the CDK9 closer to the RNAPII. The CDK9 hyperphosphorylates the CTD of RNAPII leading to the enhancement of elongation (processive elongation) (see the right side of Figure 1.2). Furthermore, p-TEFb inactivates by phosphorylation DSIF and NELF that are inhibitors of transcription elongation and facilitates the transition from abortive to processive elongation [13]. The dissociation of the Tat-TAR-kinase complex (TAK) from the RNAPII involves several cellular factors [18]. The CREB binding protein, p300/CBP, acetylates the Lys-50 residue of the Tat protein inducing the dissociation of the protein from the TAR RNA and subsequent release of TAR from the transcription complex. Several histone acetyltransferases acetylate Lys-50 of the Tat protein promoting release of p-TEFb, transferring Tat from TAR to the elongating RNAPII complex, and recruiting the histone acetyl transferase PCAF (p300/CBP associated factor) [206]. With the assistance of PCAF, the transcription complex becomes activated and continues transcribing the remainder of the HIV genome efficiently leading to the synthesis of full-length RNA transcripts.

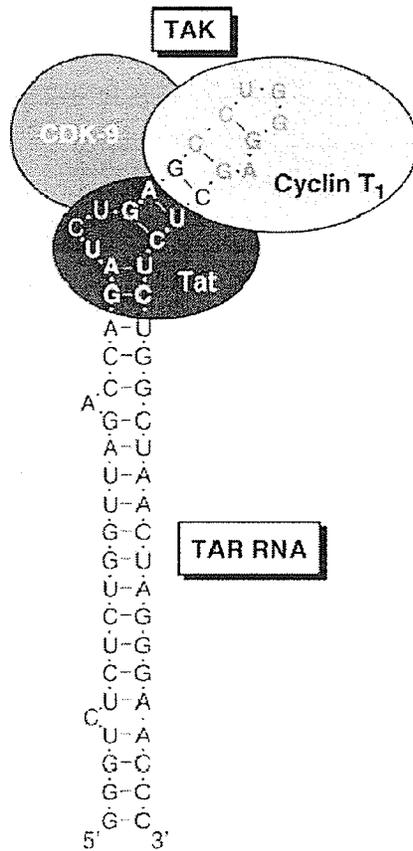


Figure 1.3: Model of interaction between Tat, TAR, and p-TEFb. Reprinted from [12] with permission of Karn J.

It is clear that HIV-1 transcription is dependent on the key activator Tat. Yet this raises the question of the origin of the Tat protein in the absence of processive elongation. It is possible that in the absence of Tat, the truncated RNA transcripts may encode the Tat gene [14]. This would allow the Tat protein to be translated in the cytoplasm and enter the nucleus to bind with the TAR RNA and would provide a plausible explanation for the activation of the transcription complex without extracellular Tat. Following its dissociation from the TAR RNA,

the Tat protein may exit from the host cell and enter and activate a latent cell. Extracellular Tat protein, upon release from infected HIV-1 cells, binds to the polyanionic heparin sulfate proteoglycans (HSPGs) of uninfected cells via its highly basic arginine-rich motif [17]. This binding induces adhesion of Tat to the cell surface and subsequent penetration into the host.

### *1.1.3 Primary Sequence of the Tat Protein and Function*

The transactivator of transcription (Tat) protein is a small nuclear protein essential to the regulation of viral replication. Depending on the strain of HIV-1, the Tat protein is comprised of 83-101 amino acids encoded by two exons: the first exon encodes residues 1-72 while residues 73-101 are encoded by the second exon [21] (Figure 1.4). However, only the first exon of Tat encodes for residues that were discovered to be essential to virus production [22].

| Domains/Regions  |  | Exon 1              | Exon 2 |
|------------------|--|---------------------|--------|
| Pro-rich (1-20)  |  | MEPVDPRLPWKHPGSQPKT | 73-101 |
| Cys-rich (21-31) |  | ACTNCYCKKCC         |        |
| Core (32-48)     |  | FHCQVCFITKALGISYG   |        |
| Arg-rich (49-58) |  | RKKRRQRRRP          |        |
| Gln-rich (59-72) |  | PQGSQTHQVLSKQ       |        |
| Histidine-Tag    |  | MGSSHHHHHSSGLVPRGSH |        |

Figure 1.4: The HIV-1 Tat sequence corresponding to the BH10 Tat isolate illustrating the amino acid residues 1-72 encoded by the first exon. The second exon encodes residues 73-101. The Tat<sub>1-72</sub> protein is comprised of five distinct domains: N-terminal (1-20) that contains 3 acidic amino acids and 5 prolines, Cys-rich (21-31), Core (32-48), Arg-rich (49-58), and Gln-rich (59-72) [14]. The histidine-tag (20 amino acids) is fused to the N-terminus of the Tat<sub>1-72</sub> protein.

#### 1.1.3.1 N-terminal region

The N-terminal domain, comprised of residues 1-20 of the Tat<sub>1-72</sub> protein, contains 3 acidic and 5 proline amino acids (Pro-rich). Its functional significance can be seen in a study that showed the N-terminus of the Tat protein recognizing the KIX-domain of CBP (CREB-binding protein) *in vitro* [33]. This is important because the CBP and p300 co-factors are known to associate with the Tat protein in order to activate transcription of the HIV-1 DNA [34]. It is noteworthy that the N-terminus is also found to assist in the binding of Tat to TAR through stabilizing interactions with

Tat's basic domain [35]. Furthermore, in a study of *Tat* genes containing mutations in the N-terminal domain, it was found that these mutants tend to stall viral gene expression or HIV-1 reverse transcription [36]. Finally, it was determined that the N-terminus is important in the inhibition of dipeptidyl peptidase IV (DP IV), which is an enzyme responsible for activation and proliferation of lymphocytes [37].

#### 1.1.3.2 Cys-rich region

The cysteine-rich (Cys-rich) domain is located between amino acids 21-31 and contains seven cysteine residues. This particular region is highly conserved between different strains of HIV-1. Through direct mutagenesis analysis, it was determined that a mutation on any 6 of the 7 cysteine residues will eliminate Tat's transactivation function [38]. The Cys-rich region has been implicated in the formation of dimeric Tat in the presence of Zn(II) through direct metal-bridging of the cysteine residues [39, 41]. Unfortunately, little structural detail about the Tat protein's metal binding properties was obtained from these studies. Another functional role of the cysteine-rich region pertains to it being a part of the transactivating or co-factor binding domain (along with the core and N-terminal regions) that is responsible for the interaction with cyclin T1 (part of p-TEFb) [40].

#### 1.1.3.3 Core region

The third domain is comprised of amino acids 32-48 (Core region) containing 6 aliphatic and 4 hydrophobic amino acids. It has been previously mentioned that the

core region, along with the cysteine-rich and N-terminus regions, are responsible for the transactivation function of Tat [40]. The core region has been reported to play a crucial role in the binding of the Tat protein to tubulin/microtubules both *in vivo* and *in vitro* [42]. The outcome is the inability of the microtubules to depolymerize through Tat's modification of the microtubule's dynamics leading to apoptosis. Interestingly, a study of soluble Tat peptides that correspond to the core region of the Tat protein has shown that they inhibit HIV-1 virus replication [43]. The peptides have a high probability of binding to the interface of cyclin T1 and CDK9, which may interfere with or halt the transcription of the HIV-1 LTR.

#### 1.1.3.4 Arg-rich region

The arginine-rich (Arg-rich) region consists of amino acids 49-58, which contain the characteristic basic domain RKKRRQRRR motif responsible for binding to TAR [44, 46] during the activation of the transcription complex. The basic domain is also known for its unique property of delivering proteins or molecules to the nucleus via its nuclear localization segment (NLS) [45]. The sequence of the Tat NLS is RKKRRQRRR which corresponds to residues 49-57 of the Tat<sub>1-72</sub> protein [47]. There is also evidence that the arginine-rich region is another region (besides the core region) that assists in the binding of Tat to tubulin/microtubules by providing a countercharge (positive charge) to the negative C-terminal of tubulin [42].

#### 1.1.3.5 Gln-rich region

The last domain (the glutamine-rich region), encoded by the first exon of the *Tat* gene, is comprised of amino acids 59-72 that contains 4 glutamine residues. The glutamine-rich region has been shown to be involved in mitochondrial T-cell apoptosis [48]. Analyses of two synthesized Tat<sub>1-82</sub> proteins from two Uganda patients revealed that mutations in the glutamine-rich region were found to correlate with induction of apoptosis and binding to tubulin. It is possible that the mutations may result in a change in Tat's conformation which in turn may disrupt the orientation of the protein's binding site.

#### 1.1.3.6 Exon 2 residues

Despite the fact that the second exon of the *Tat* gene encodes residues that are not involved in the activation of HIV-1 transcription [49], several studies of this segment have been carried out in hopes of discovering any hidden functions. The second exon of the *Tat* protein has been reported to bind to the human translation elongation factor EF-1 $\delta$  that causes a significant reduction in the translation of cellular mRNAs [50]. During expression of low amounts of *Tat*, the second exon is required for optimal activation of the nuclear factor-kappa B (NF-k $\beta$ ), LTR transactivation, and HIV-1 replication in primary T cells [51]. NF-k $\beta$  plays a vital role in regulating the immune response during infection. Within the second exon of the *Tat* protein is a highly conserved tripeptide, Arg-Gly-Asp (RGD), a motif known

to function in cell adhesion mediated through specific interaction with integrin receptors  $\alpha_5\beta_1$  and  $\alpha_5\beta_3$  [52, 53].

The Tat protein used in the present study has a histidine-tag (His-tag) fused to the N-terminus and is required for purification via cobalt-metal affinity chromatography. The inclusion of the affinity tag has been shown to be beneficial to the target protein (137): it improves protein yield, prevents proteolysis, and facilitates protein refolding [31]. Despite the positive effects of the His-tag as previously mentioned, removal of the affinity tag is desirable to prevent unpredicted structural changes in order to study the protein in its physiological conformation. However, attempts at removal of the His-tag using thrombin carried out by Dr. Shaheen Shojania have been unsuccessful possibly due to the presence of an internal thrombin cleavage site between Lys-61 and Ala-62 [32].

#### *1.1.4 Structure of the Tat Protein*

The first structural study on HIV-1 Tat was conducted in 1995 using 2D homonuclear  $^1\text{H}$ - $^1\text{H}$  nuclear magnetic resonance (NMR) techniques on the Z2 strain [18]. Using 2D NOESY NMR constraints, Tat was determined to adopt a condensed molecular centre composed of the core region, the glutamine-rich region, and the N-terminus. In contrast to previous studies of the analogous equine infectious anemia virus (EIAV) Tat peptide, the conformation of the basic arginine-rich region did not resemble a rigid  $\alpha$ -helix [218]. The cysteine-rich region was determined to be highly flexible. The secondary structure of Tat was investigated by Peloponese and co-

workers using the simulated annealing procedure on NMR constraints of HIV-1 Tat Bru, which led to the synthesis of Tat-specific binding molecules [54]. Interestingly, structural analysis of the HIV-1 Tat Bru showed that it did not contain regions adopting the  $\alpha$ -helical conformation, which is in contrast to the study on HIV-1 Tat Z2. Rather, the Tat Bru protein was mainly comprised of  $\beta$ -turns, corresponding to 32% of residues. Both of these observations were confirmed by circular dichroism (CD) data analysis. In recent years, structural studies on HIV-1 Tat using multinuclear NMR to measure relaxation parameters revealed that the protein predominately exists in a natively disordered state while under reduced conditions and low pH [26].

### *1.1.5 Natively Disordered Proteins*

According to the *structure-function paradigm*, the function of a protein is directly linked to its three-dimensional structure. However, a re-examination of the *structure-function paradigm* showed that numerous proteins important for maintaining regulatory functions in cells have been discovered to be void of intrinsic globular structure [23]. The term “natively disordered” refers to a class of proteins that exists as a mixture of extended conformers whereby the protein rapidly samples each conformer [24]. It is also noteworthy that these proteins have the potential to form protein-protein interactions. A characteristic property of most unstructured proteins is their ability to confer flexibility by varying the protein recognition motif during binding with cellular factors. Furthermore, unfolded proteins generally provide a

larger surface area than folded proteins. This enables the possibility to recognize multiple binding sites. The dynamics and flexibility of disordered proteins can be characterized by using NMR relaxation methods [25]. In an NMR study of the HIV-1 Tat protein at pH 4, it was shown that the protein is natively unfolded through a series of NMR spin relaxation experiments [26]. By comparing the backbone chemical shifts of Tat, corrected for local sequence effects, to the random coil values it was possible to determine the conformation of the Tat protein. The chemical shift indexes of Tat show that most resonance falls within the range for classification as random coil. Furthermore, the absence of small patches of  $\alpha$ -helical and  $\beta$ -sheet regions within more than three consecutive residues provides further evidence of a random coil conformation.

#### *1.1.6 Other Activities of Tat*

Along with its vital role as an HIV-1 transcription regulator, Tat has been implicated in other cellular activities: immune suppression, neural degeneration, apoptosis, and progression of Kaposi's sarcoma (KS).

##### **1.1.6.1 Immune Suppression**

The Tat protein has been shown to suppress the immune response. In the presence of Tat, the cytotoxic effects of the protein have been shown to inhibit the antigen-induced proliferation of T-lymphocytes [86]. The exact mechanism of Tat induced immune suppression is unclear, however, recent evidence points to programmed cell

death of lymphocytes in HIV-infected patients. It is speculated that the Tat protein accumulates in the lymphoid tissues and appears to affect gene expression. In HIV-infected patients, Tat stimulates uninfected T-cells to prematurely activate cyclin-dependent kinases (CDKs) that lead to apoptosis [87]. Furthermore, the depletion of CD4<sup>+</sup> T-cells has been associated with accelerated CD95-mediated T-cell apoptosis involving both Tat and the gp120 receptor [88].

#### 1.1.6.2 Neural Degeneration

HIV-associated dementia is a debilitating disease that typically occurs in 20% of HIV-1 infected patients. The characteristic symptoms of the disease consist of cognitive, behavioral, and motor dysfunction [89]. The Tat protein is a well-known causative agent in HIV-1 neurotoxicity. In particular, Tat was found to produce dose-dependent depolarization of neurons in human fetal cultures and rat hippocampal slice preparations [90]. Furthermore, exposure of Tat to cell cultures containing macrophages and glial cells resulted in the production of proinflammatory cytokines [91]. Specifically, Tat stimulates interleukin (IL)-1 $\beta$  production in macrophages. The production of cytokines may lead to the cerebral dysfunction that is indicative of HIV dementia.

#### 1.1.6.3 Induction of Kaposi's Sarcoma

In addition to controlling HIV-1 replication, extracellular Tat regulates the development and progression of Kaposi's sarcoma (KS), which is characterized by

prominent angiogenesis [27]. Angiogenesis is defined as the sudden growth of new blood vessels that are typically involved in supplying nutrients to cancerous tissues. Multiple factors cooperate in the induction or progression of KS in HIV-infected individuals. One of the major contributors, extracellular Tat, is responsible for the initial growth of spindle cells derived from AIDS-KS cells [28]. These spindle cells have the propensity to activate endothelial cells that may stimulate angiogenesis and lead to the formation of cancerous tumors. The precise mechanism of how Tat induces angiogenesis still eludes researchers today. However, Albini et al. noted that the Tat protein Arg-rich region is similar to other potent angiogenic growth factors [29] and that it is possible the protein may mimic one of these growth factors by binding to a growth factor tyrosine kinase receptor. In particular, studies show that the stimulation and activation of endothelial cells involves the interaction between Tat and Flk-1/kinase insert domain receptor, which is a tyrosine kinase receptor [29, 30].

#### 1.1.6.4 Apoptosis

Presently, the mechanism of HIV-mediated cell death *in vivo* is not completely understood. However, there are several potential triggering events and signaling pathways that may lead to apoptosis. Figure 1.5 shows two mechanisms by which HIV can induce apoptosis. In panel A, the surface glycoprotein, gp120, is shown priming the cell for apoptosis through interaction with CD4 receptors in the absence of viral gene expression. Panel B shows the induction of apoptosis within the

infected cell carried out by the products of the viral genes *Tat*, *Nef*, and *Vpr* [92]. The Tat protein has been linked to the increased expression and up-regulation of the Fas Ligand, which in turn associates with the adaptor protein, Fas-Associated protein with Death Domain (FADD) [93]. The binding of Fas to FADD recruits the protease, Caspase-8 (casp-8). Activation of casp-8 cleaves casp-3 that leads to initiation of the caspase cascade. Interestingly, as shown in panel C, exogenous Tat protein has the unique ability to induce apoptosis of a nearby bystander cell (*in trans*) via endocytosis of Tat by the uninfected CD4<sup>+</sup> lymphocyte following which the protein is localized to the nucleus via the basic region [94].

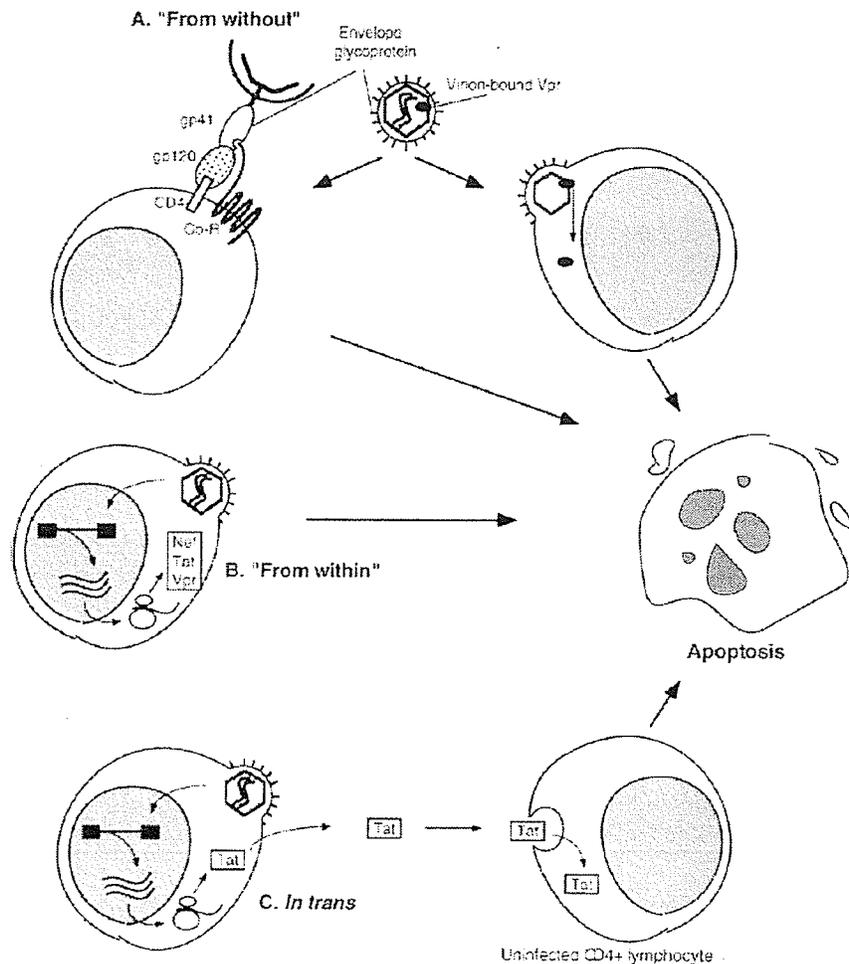


Figure 1.5: Illustration of the multiple steps in the viral life cycle during HIV-mediated apoptosis. Reprinted with permission of Roshal et al. [92]

## 1.2 History of Zinc-binding Proteins

### 1.2.1 Significance of Zinc in Biological Activities

Over half of the proteins responsible for growth of all life forms require metal ions to function [96]. An example can be seen in the function of respiration, whereby iron is an important metal in facilitating the binding of oxygen molecules to hemoglobin.

Zinc is one of the most abundant divalent ions in living organisms [57] and on average, a person will contain 2.3 g of zinc. Zinc has been widely known to play an integral part in nearly 300 enzymes. In particular, it has been shown to be vital for enzyme catalysis and gene expression. It has also been reported that zinc helps in stabilizing the mechanistic folding of several proteins and nucleic acids. In 1950, the very first zinc enzyme, carbonic anhydrase, was reported by Bert Vallee [58].

### *1.2.2 Function of Zinc in Cells*

Zinc is a multifunctional transition element that plays a vital role in homeostasis [72]. In particular, zinc is essential for the catalytic, co-catalytic, and structural function for approximately 300 enzymes such as oxidoreductases, transferases, hydrolases, isomerases, and ligases. One of the most important functions of zinc is the role it plays in the immune response. A diet consisting of very low zinc content has been shown to cause a dramatic decrease in CD4<sup>+</sup> and CD8<sup>+</sup> thymocytes (pre-T cell) within the thymus due to cell apoptosis [76]. The result of massive loss of thymocytes is the reduction of lymphocytes in the blood (lymphopenia). Recently, zinc has been implicated as a leading agent in the development of amyotrophic lateral sclerosis (ALS), which is described as a neurodegenerative disease that causes deterioration of the motor neurons [77]. It is speculated that the cause of ALS is the inability or a reduction in the binding of zinc in the enzyme, superoxide dismutase, due to mutations. Zinc can also modify calcium-dependent cellular processes by displacing or relocating the calcium through competition binding.

Inhibition studies reported by Petersen et al. showed that zinc has a higher binding affinity for the calcium-binding loop than calcium itself within the protease domain of coagulation factor VIIa (FVIIa) [78]. The zinc binding induced a conformational change of the enzyme leading to reduced activity and lower binding affinity for tissue factors.

Zinc plays an essential role in many biological and cellular processes including replication, transcription, and translation. Using laser-induced cytofluorometry for quantitative analysis of the distribution of DNA, it was shown that cell division within the eukaryotic *Euglena gracilis* decreases in the absence of zinc [79]. Zinc is also a key factor in the interference of mRNA translation by preventing the binding of human iron regulatory protein 1 with the RNA, iron responsive element (IRE), by metal-induced aggregation of the protein [97]. An example of the effects that zinc imposes on transcription can be seen in the studies of the transcription factor, GAL4 protein. Binding of zinc to GAL4 stabilizes the structure and enables the protein to recognize specific DNA sequences found upstream of genes for galactose-metabolizing enzymes [98]. Interestingly, the binding of zinc to GAL4 forms a  $Zn(II)_2Cys_6$  binuclear cluster which was confirmed by a  $^{113}Cd$  NMR study showing the presence of two cadmium-binding sites [99].

### *1.2.3 Metabolism of Zinc in Cells*

The recommended daily intake of zinc ranges from 3 mg/day for infants to 25 mg/day for women during lactation. Several factors affect the level of zinc absorption in the human body. For instance, zinc absorption increases during intake of a diet rich in proteins whereas a decrease in absorption is usually accompanied by diets low in proteins [80]. It is noteworthy that high intake of iron, calcium, and copper tends to decrease the absorption of zinc. Furthermore, during periods of low dietary zinc supply, the absorption of zinc increases.

Upon entering the intestine, orally ingested zinc absorption is regulated by an intestinal zinc-binding protein, metallothionein (MT), which transports the metal from the intestinal lumen to the plasma [81]. Once absorbed, the plasma zinc is transferred to the liver and is redistributed to all other tissues in the body. In particular, the bone and muscles are the major consumers of plasma zinc. Zinc has been shown to exist in lymphocytes and leukocytes and may play a vital role in cell immunity [76]. Finally, excessive zinc is excreted by the urine regulated by the signaling effects of glucagon [100]. Other routes of zinc excretion consist of the bile and pancreatic secretion.

Unlike iron, copper, mercury, and other metals, zinc is considered nontoxic even upon excessive accumulation due to the efficiency with which the cells and tissue excrete zinc [73]. However, cases of accidental ingestion of zinc chloride have been reported during which patients exhibit symptoms related to gastric corrosion

such as lethargy and vomiting of gastric fluids due to chemical reactions between zinc chloride and the gastric walls [102].

Many chronic illnesses are derived from the direct effect of zinc deprivation. Deficient levels of zinc in cells and tissues have been shown to be directly linked to impairment of cell growth due to the inability to divide and differentiate. In a study of the *Euglena gracilis*, a deficiency in the metabolism of zinc led to a retardation of the cell growth rate [74]. Another effect of insignificant levels of zinc is the tendency to bleeding and clotting disturbances reported in a study in which patients were deprived of zinc supplement through strict low zinc diet [75]. Results showed that induced zinc deprivation led to an impairment of platelet aggregation. It is noteworthy that platelet aggregation is an essential component of blood clotting.

#### *1.2.4 Comparison between Zinc and Cadmium*

Both zinc and cadmium are closely related metals with very similar chemical and physical properties. Physically, zinc's appearance is that of a dull grey metal but when polished, it is observed to be bluish white in color [82]. Cadmium appears as a silvery-white metal in nature. Zinc and cadmium have melting temperatures of 420 °C and 312 °C, respectively. In the periodic table, both zinc and cadmium are located within the transition metals with atomic numbers of 30 and 48, respectively. It is noteworthy that the electronic shell configuration of zinc is [Ar] 3d<sup>10</sup> 4s<sup>2</sup> while that of cadmium is [Kr] 4d<sup>10</sup> 5s<sup>2</sup>. The ionic radius of cadmium is larger than zinc with 0.0797 nm for the former and 0.074 nm for the latter.

A significant difference between zinc and cadmium is the roles they play in biology. Zinc is an essential trace element that regulates the function and structure of enzymes whereas cadmium does not have any physiological relevance other than that it is very toxic.

The mechanism of cadmium toxicity in the cell involves the activation of signaling pathways and modification of metabolic processes leading to cell death. In particular, cadmium is a well-known carcinogen [83] through induction of oxidative stress, gene regulation, inhibition of DNA repair and apoptosis interference. The main cause of apoptosis during oral intake of cadmium is the accumulation of the metal in both the liver and kidney [84]. It is the acute Cd(II) hepatotoxicity that leads to cells undergoing necrosis.

#### *1.2.5 Zinc fingers and Comparison with Cys-region of Tat to Zinc Finger Sequences*

Classically, the zinc finger domains within proteins consist of an amino acid sequence of Cys-Cys-His-His. Currently, approximately 24,000 classical zinc finger proteins have been discovered with the emergence of different subfamilies: Cys-Cys-Cys-Cys, Cys-Cys-His-Cys, Cys-Cys-His-His, and many more [104]. Zinc coordination within the zinc finger domain stabilizes unique tertiary structures that confer specific binding to different molecules like DNA, RNA, and proteins. One of the most studied zinc finger proteins is the transcription factor TFIID that is required for the initiation of the 5S RNA synthesis by RNA polymerase. The transcription factor contains 12 repeating structures with each consisting of

approximately 30 residues. It is noteworthy that each repeat has two invariable cysteine and two invariable histidine amino acids that have been shown to coordinate with at least 2 zinc ions with a high binding affinity [103]. Presently, there are numerous studies on zinc finger proteins that provide insight into the treatment of cancer. In particular, zinc fingers are found within the DNA-binding protein, Wilms tumor suppressor protein (WT1), which also contains the classical Cys-Cys-His-His finger domain [105]. Several biological functions of zinc domain proteins are found in the cell such as gene expression, signal transduction, cell growth, differentiation, and development.

The Tat protein contains a cysteine-rich region with seven cysteine and one histidine residues within 11 amino acids. It is possible that the cysteine-rich domain can function as a zinc finger that binds zinc with a high affinity. Furthermore, studies have suggested that the Tat protein forms a metal-linked dimer with zinc [39].

#### *1.2.6 Zn(II) induced dimerization of Tat protein*

One pathway to discovering a therapeutic solution in the treatment of AIDS consists of synthesizing inhibitors that block the interaction of Tat with various cellular co-factors. However, another plausible approach to drug design is the metal-induced formation of Tat homodimers that might lead to inactivation of HIV-1 transcription due to blocking of binding sites [21]. Tat has been proposed to form a Zn(II)-induced dimer that is stabilized through interaction between Zn(II) and the

cysteine-rich region based on evidence from ultra-violet (UV) absorption spectroscopy, circular dichroism (CD) spectropolarimetry, and sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) [21]. In particular, it was shown that Tat binds to two mole equivalents of Zn(II) and a model for the dimer proposed an anti-parallel arrangement of the two monomers. The attractiveness of Tat dimerization can be seen in development of anti-viral drugs for inhibiting binding with important cellular cofactors. The dimerization may cause overcrowding of the arginine-rich region, inhibiting the interaction with the U-rich bulge of TAR and cyclin T1. Interestingly, *in vivo* dimerization of the Tat protein was reported by Bogerd et al. by using a genetic approach in yeast cells [56]. However, the oligomerization reported was independent of the cysteine-rich region of Tat. Furthermore, it has been widely reported that Tat may also exist as a monomer during the transactivation of HIV-1 [95]. The fact that there are contradicting reports of the Tat protein existing as a monomer or dimer in nature calls for the need of further structural research into the Tat folding mechanism and the role of zinc.

### *1.2.7 Zinc Coordination Geometry*

In biological cells, zinc functions in regulating cellular processes through binding to motifs within six different classes of enzymes. As previously mentioned, zinc plays a pivotal role in nearly 300 proteins. There are three types of zinc binding sites: structural, catalytic, and co-catalytic (Figure 1.6). Zinc has a tendency to bind

specifically to cysteine and histidine residues accompanied by stabilizing interactions from the carboxylic groups of glutamic and aspartic acids [85].

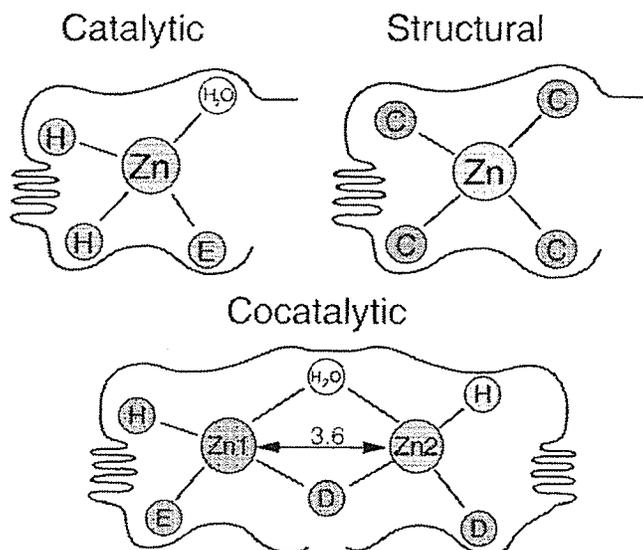


Figure 1.6: Illustration of the zinc coordination geometry for the catalytic, structural, and co-catalytic sites of zinc binding. The H and C denote histidine and cysteine residues, respectively. Reprinted with permission of Auld et al. [85]

In a catalytic zinc site, the zinc is usually coordinated predominately with histidine residues. In particular, the N $\epsilon$ 2 nitrogen specifically binds to the zinc due to its ability for charge dispersion. It is noteworthy that water is always coordinated to zinc. Typically, the coordination number of the catalytic site is 4 or 5 with the distorted tetrahedral or trigonal-bipyramidal geometry. In the structural zinc site, the preferred amino acid is cysteine while no water molecule coordinates with the

metal. In cases where zinc coordinates with a histidine residue, the site is generally termed a “zinc finger”. Similar to catalytic zinc sites, structural zinc sites generally have coordination numbers of 4 to 5. In the event where a protein contains 2 or more zinc binding sites that are close in proximity, a co-catalytic zinc site is found. Generally, the binding of 2 zincs within the protein is stabilized by the interaction between either a water molecule or the side chain of Asp. Typically, the distance between zinc ions is between 3 to 4 Å [85].

One of the factors that makes zinc an attractive metal for protein analysis is its ability to form different coordination geometry complexes. In the zinc-binding protein, alcohol dehydrogenase [58], the zinc is located in the bottom of the active-site and is coordinated to thiol groups belonging to Cys-46 and Cys-174, the N-atom of His-67 and an O-atom of either water or a hydroxide molecule. The coordination geometry is that of a distorted tetrahedron. However, the existence of five- and six-coordinated zinc-complexes in nature adds a new level of complexity in the characterization of zinc-binding proteins. In a study of the zinc-substituted cytochrome c protein reported by Qian et al., it was determined that the coordination geometry of the zinc complex is 6 (distorted octahedron) using a comparison between computational calculations and distance constraints derived from NOE measurements [59].

### 1.3 SDS-Polyacrylamide Gel Electrophoresis

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) is a well-established technique for the separation of proteins based on their molecular mass ( $M_r$ ). It is a commonly used method for the identification and confirmation of the purity of a protein [60]. The protein sample is treated with an anionic detergent, SDS, which binds to most proteins at a constant ratio of approximately 1.4 g SDS/g protein. This applies a negative charge to all the proteins in a sample, which overwhelms the inherent charge on the protein and leads to all the proteins achieving the same charge-to-mass ratio. The protein solution is then added to the polyacrylamide gel that acts as the sieving component. Since the proteins have the same charge-to-mass ratio, the mobility of the proteins through the gel is based on their molecular weight. One feature of SDS-PAGE that is very attractive is its ability to determine the oligomeric state of a protein. In the SDS-PAGE analysis of HIV-1 Tat protein (86 amino acids) reported by Frankel et al. it was discovered that the protein migrated as a monomer (15 kDa) in the absence of metal and as a dimer (34 kDa) in the presence of Cd(II) [61]. The high apparent mass of the monomer (15 kDa vs. 9 kDa) is owing to the high net positive charge on Tat which decreases its electrophoretic mobility.

## 1.4 Circular Dichroism Spectropolarimetry

Circular dichroism (CD) spectropolarimetry is an excellent method for investigating the conformation of a protein in solution. In particular, it can provide useful information about the protein's secondary structure for the study of protein-protein or protein-ligand interactions. The basic phenomenon of CD is based on the differential absorption of right- and left-circularly polarized light ( $\Delta A = A_r - A_l$ ) [64]. In general, only chromophores that are asymmetric (chiral) or chromophores that are located in an asymmetric environment produce a CD signal. To study the tertiary or quaternary structure of the protein, one monitors the near-UV (310-255 nm) CD region where the contributing chromophores correspond to Trp, Tyr, and Phe. In the far-UV (180-250 nm) CD regime, the peptide bond is the main contributing amide chromophore that reports on the secondary structure of the protein. Analysis of the far-UV CD spectrum can elucidate the secondary structural class of the protein. For proteins high in  $\alpha$ -helical content, intense negative bands at 208 nm and 222 nm accompanied by a strong positive band at 192 nm are observed in the CD spectrum [Figure 1.7, yellow]. The CD spectrum of proteins high in  $\beta$ -sheet content contain a negative band at 210-225 nm and a strong positive band at 190-200 nm [Figure 1.7, blue]. Disordered proteins are sometimes called random coils and are typically observed to contain a strong negative band at 195-200 nm and either a weak negative or positive band at 210-230 nm in the CD spectrum [Figure 1.7, red]. Distinguishing between the random coil and left-handed

polyproline helix has been difficult. The consensus seems to be that a weak positive band between 215-220 nm accompanied by a strong negative band at 200 nm is diagnostic of the polyProII conformation. The CD spectrum of the  $\beta$ -turn resembles that of the  $\alpha$ -helix but the bands are weaker [216]. However, it must be remembered that aromatic residues such as Trp, Tyr, and Phe absorb strongly in the far-UV region and their presence can complicate the analysis of secondary structure.

The observed ellipticity,  $\theta_{obs}$ , which is proportional to the difference in absorbance between the right- and left-circularly polarized light, is measured in mdeg. However, measured ellipticity is usually expressed as mean residue ellipticity (MRE) using formula shown in Equation 1.

$$MRE = \frac{\theta_{obs}(mdeg)}{10 \times n \times C_p \times l} \quad [1]$$

The  $n$  denotes the number of amino acid residues in the protein while  $C_p$  is the molar concentration of the protein. Furthermore,  $l$  is the path length of the cell in centimeters and the units of MRE are commonly expressed as  $\text{deg}\cdot\text{cm}^2\cdot\text{dmol}^{-1}$  [211, 212].

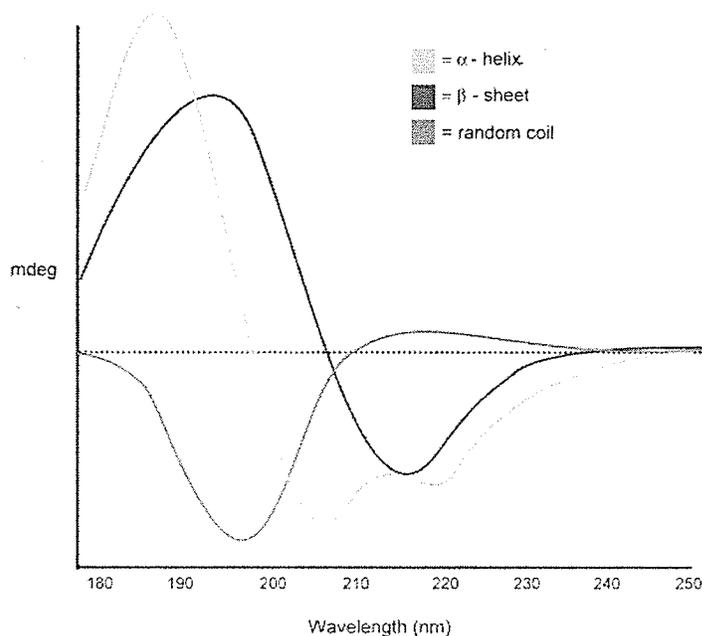


Figure 1.7: CD spectra illustrating the key characteristics of random coil (red),  $\beta$ -sheet (blue), and  $\alpha$ -helix (yellow) circular dichroism [30].

## 1.5 Nuclear Magnetic Resonance Spectroscopy

Nuclear magnetic resonance (NMR) spectroscopy has been an essential tool in acquiring data for determination of protein structures at the atomic level since the late 1980's and is complementary to X-ray crystallography [65]. In the late 1950's, the very first NMR experiment was conducted on a biological macromolecule, Ribonuclease [163]. With the evolution of high-field NMR magnets, Fourier Transformation, powerful and inexpensive computers, and multidimensional NMR techniques, the number of NMR studies on proteins has substantially increased.

Although 1D  $^1\text{H}$  NMR has been extensively used for the characterization of organic molecules, it is inefficient for the study of larger biomolecules due to the presence of significantly more hydrogen atoms that leads to overcrowding of resonance lines in the NMR spectrum. Here, high-resolution two-dimensional heteronuclear and homonuclear NMR are extensively used for the structural characterization of the Tat protein. To obtain even higher structural resolution, isotopic  $^{15}\text{N}$ - and  $^{13}\text{C}$ -labeling of the Tat protein is routinely used along with multidimensional NMR spectroscopy.

#### *1.5.1 Heteronuclear Single Quantum Correlation Spectroscopy*

Heteronuclear single quantum correlation spectroscopy (HSQC) is a commonly used 2D NMR technique for structural determination of proteins. Through coherence transfer steps, the experiment correlates a heteronucleus ( $^{15}\text{N}$  or  $^{13}\text{C}$ ) with a proton resonance [66]. One of the main advantages of HSQC NMR is the ability to study the protein in its native conformation and monitor the changes during protein-ligand interactions. Figure 1.8 shows the pulse sequence for a standard HSQC sequence. The first insensitive nucleus enhanced by polarization transfer (INEPT) step produces an antiphase magnetization for the proton nucleus (I) along the x-axis. It is noteworthy that the application of two  $180^\circ$  pulses following the first  $\tau/2$  period refocuses the chemical shift modulation. Coherence transfer from the proton to the directly attached heteronucleus S ( $^{15}\text{N}$  or  $^{13}\text{C}$ ) is accomplished by simultaneous  $90^\circ$  pulses on both spins just before b in Figure 1.8. Following the first

INEPT step, the heteronucleus magnetization is left to evolve with its chemical shift during the  $t_1$ -evolution time. The application of the  $180^\circ$  pulse on the proton in the middle of the  $t_1$  period refocuses the evolution of heteronuclear  $J_{IS}$  coupling. The simultaneous application of  $90^\circ$  pulses on both spins results in transfer of magnetization from the S-spin back to the I-spin as antiphase magnetization. The proton antiphase magnetization is refocused into observable in-phase magnetization by simultaneous application of  $180^\circ$  pulses on both spins following the third  $\tau/2$  period. The heteronucleus magnetization is decoupled using the Waltz-16 [57] sequence during the detection period of the proton magnetization.

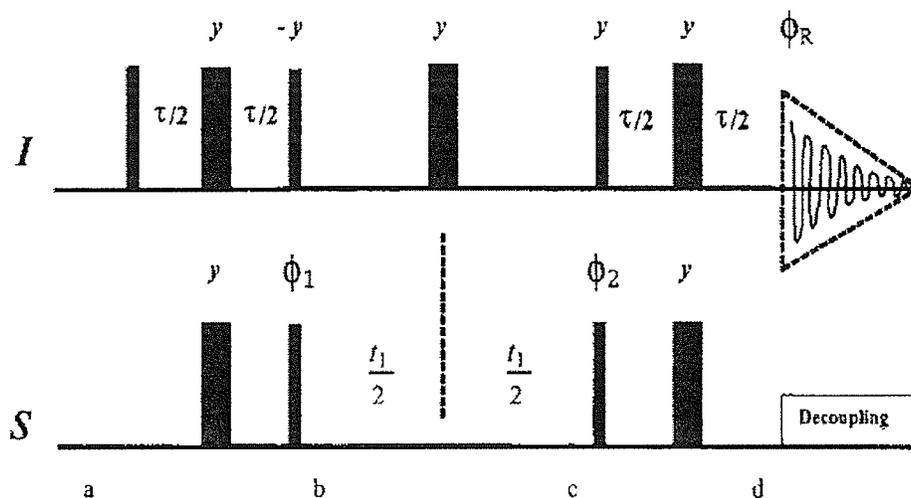


Figure 1.8: Pulse sequence for a standard HSQC sequence [106]. Thick solid bars indicate  $180^\circ$  pulses while the thin solid bars indicate  $90^\circ$  pulses. I and S are symbols to represent the nuclei involved (i.e. I =  $^1\text{H}$ , S =  $^{15}\text{N}$ ). The Waltz-16 sequence is used to decouple the S-spins during acquisition. The delay  $\tau$  is set to  $1/(2J_{\text{IS}})$ . The phase cycling  $\Phi_1$ ,  $\Phi_2$ , and  $\Phi_R$ , corresponds to application of (x,-x, x,-x), (x, x, -x, -x), and (x, -x, -x, x), respectively.

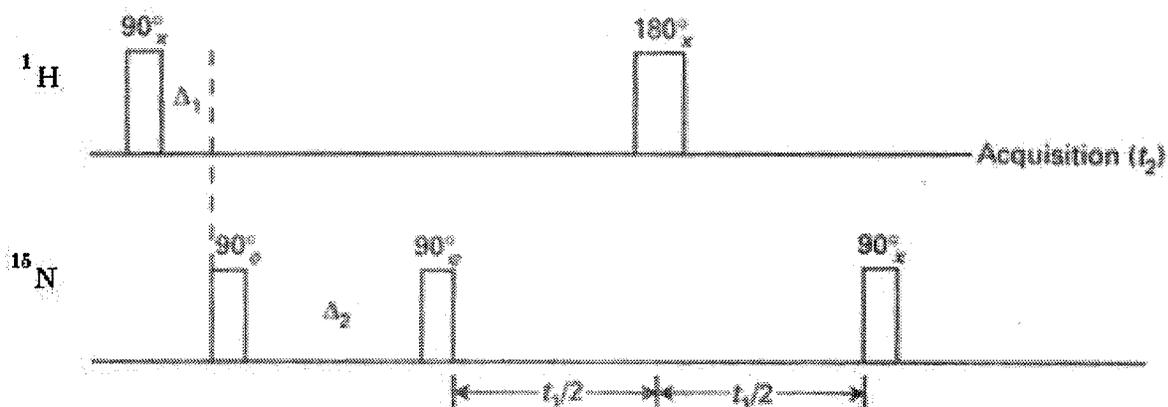
### 1.5.2 Heteronuclear Multiple Bond Correlation Spectroscopy

Long-range coupling NMR spectroscopy is a complementary technique to HSQC in the structural determination of proteins. It has the advantage of providing structural information about the connectivities of heteronuclei to the observed resonance (usually a proton) that is 2 or more bonds away. Figure 1.9A shows the pulse sequence of a general heteronuclear multiple-bond correlation (HMBC) experiment [106]. The first  $90^\circ$  pulse creates  $^1\text{H}$  magnetization in the xy plane. The

first  $90^\circ$  pulse on the  $^{15}\text{N}$  nucleus suppresses one-bond correlations and permits observation of long-range correlations with smaller coupling constants. Heteronuclear multiple-quantum coherence for protons that are directly coupled to  $^{15}\text{N}$  nuclei are removed by phase cycling. The second  $90^\circ$  pulse along the  $^{15}\text{N}$  dimension generates multiple-quantum coherence for long-range connectivities. The effects of proton chemical shifts during  $t_1$  are removed by the  $180^\circ$  proton pulse. It is noteworthy that phase cycling of the second  $90^\circ$  nitrogen pulse removes proton signals that do not contain long-range coupling information.

HMBC spectroscopy was an important technique in the investigation of the tautomeric states of histidine residues in the signal-transduction protein, III(Glc) [68]. Using the HMBC pulse sequence, the cross-peaks for  $\text{N}_{\epsilon 2}\text{-H}_{\epsilon 1}$  and  $\text{N}_{\epsilon 2}\text{-H}_{\delta 2}$  of the  $\text{N}_{\epsilon 2}\text{-H}$  tautomeric or  $\epsilon$ -tautomer state (Figure 1.9B) were observed to undergo significant chemical shift changes during the phosphorylation of III(Glc). Figure 1.9B shows the 3 possible tautomeric states of histidine along with the corresponding cross-peaks obtained from a  $^1\text{H}\text{-}^{15}\text{N}$  HMBC experiment. This suggests that it might be possible to monitor changes in the histidine residues of the Tat protein due to the binding of  $\text{Zn}(\text{II})$ , which may provide information on the role of histidine residues in metal-binding.

A



B

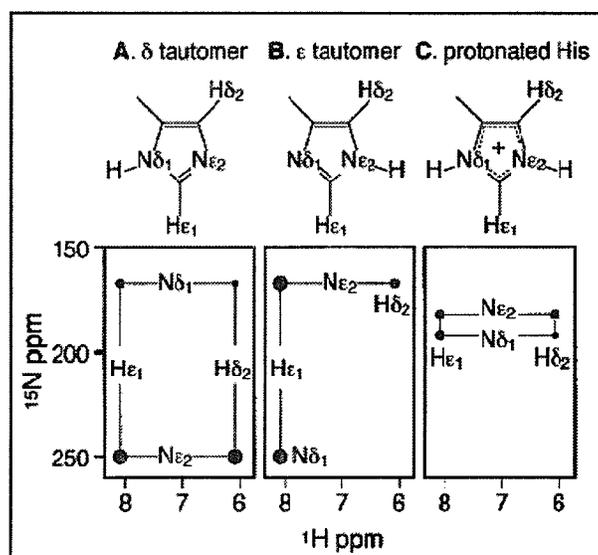


Figure 1.9: (A) Pulse sequence for Heteronuclear Multiple-Bond Correlation (HMBC) spectroscopy. Illustration was reprinted with permission from Homans SW [106]. (B) A schematic diagram of the three possible tautomeric states of Histidine and the HN cross-peaks in HMBC spectra. Illustration was reprinted with permission from Pelton et al. [117].

### 1.5.3 Nuclear Overhauser Effect Spectroscopy

Nuclear Overhauser effect spectroscopy (NOESY) is a 2D method for measuring interatomic distances by NMR and is used as the primary method for solving the complete solution structure of a molecule [106]. This experiment is specifically applied to protons whereby connectivities through space can be elucidated if the distance between protons is within 5 Angstroms ( $\text{\AA}$ ). Subsequently, NOE cross-peak intensities can be translated into distance constraints for structural determination. Figure 1.10 shows the pulse sequence for a general NOESY experiment. In the NOESY pulse sequence, three  $90^\circ$  radiofrequency (r.f.) pulses are preceded by a relaxation delay, and separated by an evolution ( $t_1$ ) and a mixing time ( $\Delta$ ). In general, the relaxation delay (preparation period or  $t_D$ ), denotes the time that should be left between passes through the pulse sequence for relaxation to restore the spins to thermal equilibrium [107]. Ideally, the total time of  $t_2$  and  $t_D$  should be in the range of 5 times the longitudinal relaxation time ( $T_1$ ). However, for macromolecules at higher field strength, the  $T_1$  may become longer requiring experimental testing to optimize  $t_D$ . In order to explain the NOESY pulse sequence, we must assume that there are two close-lying protons A and S that are not scalar-coupled (through space coupled or dipole-dipole coupled). The first  $90^\circ$  pulse creates xy magnetization by forcing the longitudinal z-magnetization of nuclei A and S towards the y-axis. During the subsequent  $t_1$  time period, the magnetization precesses in the  $x'y'$ -plane (rotating frame) to a point where after time  $t_1$  the angles traveled by the magnetizations for A and S are  $\Omega_{At_1}$  and  $\Omega_{St_1}$ , respectively ( $\Omega_A$  and

$\Omega_S$  are the precession frequencies of nucleus A and S, respectively) [109]. The magnetization on the  $x'y'$ -plane is forced to the z-axis by a second  $90^\circ$  pulse. Magnetization transfer between spins A and S occurs during the mixing period at a rate determined by cross-relaxation or chemical exchange between the two spins. The mixing period is important to NOE intensity because it allows NOE build-up. It has been shown by Kay et al. that the mixing time dependence of the peak intensities is described as an initial build-up of magnetization followed by a transition to the maximum and ending with the decay of the magnetization [111]. Furthermore, reliable internuclear distances are obtained from the rate of the initial NOE build-up. A third  $90^\circ$  pulse forces the  $x'z$  precessing magnetization to return to the  $x'y'$ -plane. After the third  $90^\circ$  pulse, each vector then precesses in the  $x'y'$  plane according to their characteristic Larmor frequency during  $t_2$ . A 2D Fourier transformation with respect to  $t_1$  and  $t_2$  will yield a general 2D NOESY spectrum. Only magnetizations that were transferred between the two spins during  $\Delta$  will generate cross-peaks while those that fail to migrate during  $\Delta$  will produce diagonal peaks. It is noteworthy that some "COSY peaks" may appear in the NOESY spectrum and to avoid this situation, one has to carry out phase cycling of the first two pulses [109].

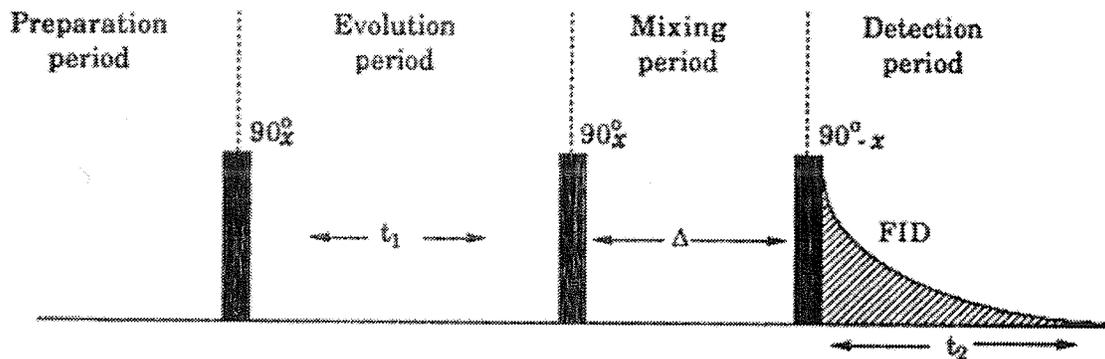


Figure 1.10: Pulse sequence of a NOESY experiment. Solid rectangular bars represent  $90^\circ$  radiofrequency pulses. The mixing period is denoted by  $\Delta$  while  $t_1$  and  $t_2$  represent the evolution (indirect) and detection (direct) periods, respectively. It is noteworthy that a preparation period precedes the first  $90^\circ$  pulse. Illustration was reprinted with permission of Rahman et al [109].

In theory, the NOE intensity can be associated with the distance ( $r$ ) between the two  $^1\text{H}$  spins by Equation 2 [105].

$$\text{NOE} \propto 1/r^6 \quad [2]$$

In reality, quantitative determination of the distance measurements is inadvisable because spin diffusion and intramolecular mobility within the macromolecule complicate the distance measurements. Currently, the appropriate method for

quantitative distance measurement consists of recording a series of NOESY spectra at short mixing times in order to measure the rates of cross relaxation ( $\sigma_a$ ,  $\sigma_b$ ; NOE build up) between two spins separated by an unknown distance  $r_a$  and spins separated by a known distance  $r_b$  (e.g. methylene protons) (Equation 3). The calculated distances are then classified in ranges (1.8 – 2.5 Å, 1.8 – 3.7 Å, 1.8 – 5.0 Å) for use in structure-determining algorithms.

$$\sigma_a/\sigma_b = r_b^6/r_a^6 \quad [3]$$

Rotating frame NOESY (ROESY) [130] has been implemented as an alternative to NOESY in cases where the protein has a rotational correlation time ( $\tau_c$ ) such that  $\omega\tau_c = 1$  ( $\omega$  is the Larmor precession frequency). In such a case, the NOE intensity is zero. Fortunately, the ROESY is non-zero throughout the entire range of the rotational correlation times. ROESY is an experiment in which homonuclear NOE effects are measured under spin-locked conditions. Figure 1.11 illustrates the relationship between NOE intensity and the protein's rotational correlation time for both rotational NOESY (ROESY) (short dashed line, 500 MHz) and NOESY (dotted line, 500 MHz; long dashed line, 300 MHz) experiments [69]. For large molecules with slow rotational correlation times, the NOE intensities in NOESY and ROESY are close to -1.0 and 0.7, respectively. For smaller molecules with fast rotational correlation times, both experiments yield NOE intensities of approximately 0.4. The significant difference between the NOESY and ROESY

experiments occurs when the protein of interest falls in between the range of large and small molecular size or alternatively, slow and fast rotational correlation times. For the NOESY experiment, the NOE intensities are very close to zero when the protein falls within this molecular mass regime ( $\tau_c = 10^{-9}$  to  $10^{-10}$  sec;  $M_r = 2$  kDa). In contrast, for the ROESY experiment the NOE intensities are approximately 0.7. It is speculated that the Tat protein might fall within the intermediate rotational correlation time regime because the protein is unfolded and its dynamics may reduce the NOE intensity making the ROESY experiment a more attractive method for structural determination. Thus, NOESY and ROESY experiments were conducted to determine which pulse sequence would yield the most intense NOE cross-peaks.

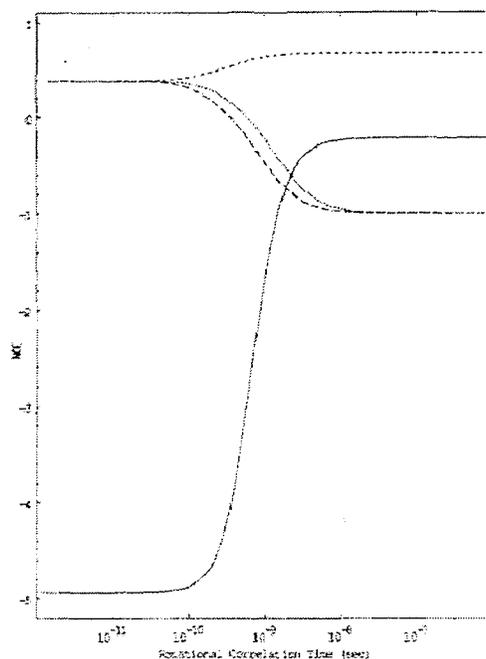


Figure 1.11: Effect of rotational correlation time on NOE. The NOESY calculations were carried out at 500.13 (---) and 300.13 MHz (●●●). The maximum heteronuclear  $^1\text{H}$ - $^{15}\text{N}$  NOE at 500.13 MHz is shown by a solid line. The main advantage of the ROESY experiment (---; 500.13 MHz) is that the NOE signal is positive and non-zero for all rotational correlation times whereas for the NOESY experiment the NOE may become zero at intermediate rotational correlation times. Illustration obtained from [36] with the permission of J. O'Neil.

#### 1.5.4 $^{113}\text{Cd}$ Cadmium NMR Spectroscopy

The naturally occurring isotope  $^{113}\text{Cd}$  (nuclear spin =  $\frac{1}{2}$ ) is a useful metal for probing the metal-binding properties in biomolecules [112]. In particular,  $^{113}\text{Cd}$  NMR is a powerful method in the investigation of the functional roles of metal ions

in metalloproteins. The fact that cadmium can be a substitute of the more biologically common zinc in metallo-enzymes and still maintain biological activity makes it an attractive tool for probing the structure and function of metalloproteins. The relative sensitivity of  $^{113}\text{Cd}$  is comparable to that of  $^{13}\text{C}$  ( $1.59 \times 10^{-2}$ ) with a natural abundance of 12.26% [113]. In general, it is possible to obtain information about changes in metal coordination, dissociation constant, and ligand exchange dynamics through metal titration experiments with cadmium. The main advantage of cadmium over other metals is the wide chemical shift dispersion of  $^{113}\text{Cd}$  that reflects the shielding of  $^{113}\text{Cd}$  by its coordination ligands and spans from -100 ppm to 850 ppm relative to the accepted reference material, 0.1 M  $\text{Cd}(\text{ClO}_4)_2$  [70]. It is noteworthy that fast chemical exchange among different species of  $^{113}\text{Cd}$ -complex is common due to formations of multiple interactions between ligands that are often hydrated. The dynamic exchange between species with different coordination geometries often leads to chemical shift averaging and line broadening. Unfortunately, in cases where chemical exchange broadening is dominant due to intermediate-exchange, interpretation of observed resonances becomes very difficult.

However, there are many examples in the literature showing that  $^{113}\text{Cd}$  NMR is a very sensitive technique for determination of the coordination environment that often leads to structural information regarding the coordination geometry of the binding sites. The chemical shift of the  $^{113}\text{Cd}$  nucleus shows large changes that reflect fluctuations in the ligands and coordination geometry during metal-binding

as a consequence of the contribution from the large electron cloud of the metal. Since upon binding to Cd(II), sulfur atoms are the most deshielded compared to other ligands. The  $^{113}\text{Cd}$  NMR shifts are particularly sensitive for the determination of the number of sulfur donor atoms. An example of the usefulness of  $^{113}\text{Cd}$  NMR is shown in the study of the human liver protein, metallothionein [114]. The uniqueness of this protein stems from the fact that it contains 20 cysteine residues out of 61 residues.  $^{113}\text{Cd}$  NMR spectral results show a multiplet of peaks around 600-680 ppm that originates from a four-metal cluster found within the metallothionein- $^{113}\text{Cd}$  complex.  $^{113}\text{Cd}$  NMR can also distinguish between the number of thiol ligands in a protein. In a study of plastocyanin (a blue-copper binding protein) by Engeseth et al. [115], it was shown that the observed  $^{113}\text{Cd}$  chemical shift was found to be in the vicinity of 400 ppm indicative of a single Cd-thiolate binding site. Proteins that contain 3 or more Cd(II) binding sites exhibit  $^{113}\text{Cd}$  NMR chemical shifts more downfield relative to the standard, 0.1 M  $\text{Cd}(\text{ClO}_4)$ . Thus,  $^{113}\text{Cd}$  NMR chemical shifts were observed near 500 ppm in a study of horse liver alcohol dehydrogenase, in which Bobsein determined that the enzyme contains two metal-binding sites [116]. Perhaps the factor that contributes significantly to the sensitivity of  $^{113}\text{Cd}(\text{II})$  NMR is the type of ligands involved in cadmium-binding rather than the number of coordinated cadmium. For instance, the more thiols coordinated to cadmium may shift the observed peak downfield towards 600-700 ppm due to extensive deshielding of the Cd(II) resonance (see Figure 1.12). The utility of  $^{113}\text{Cd}$  NMR suggests that one could use it to study the metal-binding

properties of the HIV-1 Tat protein and to investigate whether the cysteine residues within the cysteine-rich region are in fact playing a vital role as previously determined by Frankel et al. [39].

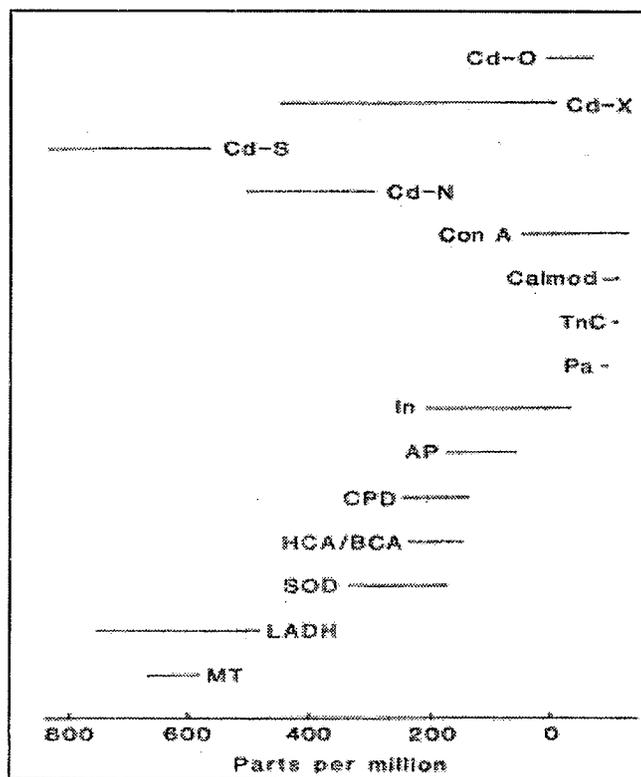


Figure 1.12:  $^{113}\text{Cd}$  Chemical shift ranges for cadmium metallothioneins and cadmium complexes. Illustration obtained from [219] with permission from Ellis P.D.

## 1.6 Ultraviolet Absorption Spectroscopy using Cd(II)

To confirm the interaction of the cysteine residues toward cadmium, the use of ultraviolet (UV) absorption spectroscopy is conducted to provide direct evidence of the existence of a thiolate-Cd(II) binding. In a UV absorption spectroscopic study of the cysteine-rich protein, Metallothionein, by Vasak et al. [207] it was determined that an increase in the absorption at 250 nm of a metallothionein-Cd(II) complex is due to the coordination of thiolate towards Cd(II). It is noteworthy that the UV range of interest (220-350 nm) is usually masked by the presence of aromatic groups. The absorption increase is most likely due to ligand-metal charge transfer (LMCT) between the sulfur atom of the cysteine (thiolate) and the Cd(II) [208]. Although, UV absorption spectroscopy can provide essential information about the thiolate-Cd(II) environment, such as metal to protein stoichiometries [209], it is an inefficient method for the determination of metal-sulfur bond length, coordination geometry, and overall structure. Thus, further investigation of the protein-Cd(II) interaction with NMR and CD spectroscopy will be required. Here, UV spectroscopy is used to investigate the interaction of the cysteine-rich region of HIV-1 Tat to Cd(II) in hopes of retrieving information about the protein's metal-binding environment.

## 1.7 Goals of the Research

One of the goals of this research was to characterize the structural changes of a recombinant HIV-1 Tat protein in the presence of Zn(II) and Cd(II) using CD spectroscopy. Another goal of this research was to obtain structural information regarding which regions within the Tat protein were essential for metal coordination using multinuclear NMR spectroscopy. The application of  $^1\text{H}$ - $^{15}\text{N}$  HSQC NMR allows one to probe the change in Tat's backbone conformation as a direct influence of the coordination with Zn(II). Using  $^{113}\text{Cd}$  NMR, we hope to determine the binding affinity of the Tat-Cd(II) complex, as well as, to characterize regions of the protein that were directly involved in metal-binding. The goal of conducting UV absorption spectroscopy on the Tat-Cd(II) complex was to determine if the Cys-rich region of the HIV-1 Tat protein was essential for metal-binding. Ultimately, the main goal of this research was to investigate the metal-binding interaction of the HIV-1 Tat protein in order to determine the coordination geometry, binding affinity, overall structure, and whether or not the Tat protein forms a dimer in the presence of Zn(II) or Cd(II).

## 2. Materials and Methods

### 2.1 Plasmid Construction

The Tat expression system was developed and optimized by Dr. Gillian Henry using the plasmid pSV2tat72 donated by Dr. Alan Frankel through the AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH [118]. Using the polymerase chain reaction (PCR), the plasmid pSV2Tat72 was used as a template for the amplified DNA sequence containing the *Tat* exon 1 gene. The gene was then removed and ligated into the expression vector pET28b(+). The expression vector contains an N-terminal His<sub>6</sub> plus T7 tag. The pET28Tat plasmid was subsequently incorporated into *E. coli* BL21(DE3)pLysS cells for expression of the 72 amino acid His-tag Tat protein.

### 2.2 Expression of Unlabeled Tat in *E. coli*

An overnight culture (50 ml) of *E. coli* was diluted 1:100 into a 1 L baffled flask containing modified Terrific Broth (TB) (Sigma, St. Louis, MO) medium supplemented with glycerol (0.8%), chloramphenicol (34 µg/ml), and kanamycin (30 µg/ml). Cells were grown at 37 °C until an OD at 600 nm of 0.6-0.9 and induced with 60 mg of isopropyl-β-D-thiogalactopyranoside (IPTG) (Sigma, St. Louis, MO). Following induction, the cells were grown for 5 hours, harvested by centrifugation

(2,607 g) (GSA-Sorvall RC-5B, Du Pont Instruments, Waltham MA), resuspended into an extraction buffer (100 mM sodium phosphate, pH 7.2, 10 mM Tris-HCl, and 10 mM TCEP), and stored at -70 °C until needed.

### **2.3 Expression of Labeled $^{15}\text{N}$ -Tat in *E. coli***

A 50 ml *E.coli* culture grown overnight at 37 °C in TB medium supplemented with chloramphenicol (34 µg/L) and kanamycin (30 µg/L) was evenly divided into 4-1 L baffled flasks containing TB media. Each 1 L culture was grown at 37 °C to an OD of 0.6-0.9, cooled in an ice-bath for 15 minutes, and the cells harvested by centrifugation at 2,607 g for 15 minutes. The cell pellets were washed twice with 100 ml of phosphate buffer (42.2 mM sodium phosphate, 22.0 mM potassium phosphate, pH 7.5, 86 mM sodium chloride), pooled, and transferred into a 1 L flask containing M9 minimal medium (Table 2.1) supplemented with 34 µg/L chloramphenicol and 30 µg/L kanamycin. The  $^{15}\text{N}$ -labeled ammonium chloride [ $^{15}\text{NH}_4\text{Cl}$  99%, Cambridge Isotope Laboratories, Inc.] was incorporated into the M9 minimal medium by dissolving 1 g with 100 ml double-deionized water (ddH<sub>2</sub>O) followed by the addition of the remaining ingredients. Double-deionized water was obtained by passing the Parker building reverse osmosis water through a Barnstead Nanopure II system for removal of minerals through deionization. The 1 L M9-medium culture was induced with 240 mg of IPTG and grown for 5 hours. The cell pellets were harvested by centrifugation at 2,607 g (GSA-Sorvall RC-5B, Du Pont

Instruments, Waltham MA) for 15 minutes, resuspended in extraction buffer, and stored at -70 °C until needed.

Table 2.1: M9 Minimal Medium ingredients obtained from [119]

| Component                        | mM                   |
|----------------------------------|----------------------|
| Na <sub>2</sub> HPO <sub>4</sub> | 42.3                 |
| KH <sub>2</sub> PO <sub>4</sub>  | 22                   |
| NaCl                             | 8.5                  |
| NH <sub>4</sub> Cl               | 12.8                 |
| EDTA                             | 1.3x10 <sup>-4</sup> |
| FeCl <sub>3</sub>                | 5.1x10 <sup>-2</sup> |
| ZnCl <sub>2</sub>                | 6.1x10 <sup>-3</sup> |
| CuCl <sub>2</sub>                | 9x10 <sup>-3</sup>   |
| CoCl <sub>2</sub>                | 6.0x10 <sup>-3</sup> |
| H <sub>3</sub> BO <sub>3</sub>   | 1.6x10 <sup>-3</sup> |
| MnCl <sub>2</sub>                | 1.2x10 <sup>-4</sup> |
| CaCl <sub>2</sub>                | 1                    |
| MgSO <sub>4</sub>                | 0.3                  |
| Glucose                          | 11.1                 |
| Biotin                           | 4.1x10 <sup>-3</sup> |
| Thiamin                          | 3.7x10 <sup>-3</sup> |

## 2.4 Purification of His-tagged Tat

The bacteria were lysed by applying 2 freeze-thaw cycles and sonication. Small amounts of ribonuclease I (RNase I) (Sigma, St. Louis, MO) and deoxyribonuclease I (DNase I) (Sigma, St. Louis, MO) were added to reduce the viscosity of the extract. The supernatant was collected after centrifugation at 10000 g for 30 minutes (SS34-Sorvall RC-5B, Du Pont Instruments, Waltham MA) and the protein was purified by metal affinity chromatography. Chromatography was performed with a 4 ml TALON Superflow metal affinity resin (CLONETECH) column stored in 20 % ethanol. Column preparation was initiated by washing with five column volumes (CV) of ddH<sub>2</sub>O followed with 10 ml of 300 mM sodium chloride. This step was required to remove any loosely bound metal ions. The column was subsequently washed with five CV of ddH<sub>2</sub>O and then equilibrated with seven CV of equilibration buffer (100 mM sodium phosphate, 10 mM Tris-HCl, pH 7.2, 6 M guanidine-HCl, 10 mM TCEP). The pH of the supernatant containing Tat was adjusted to 7.2 with ammonium hydroxide and directly loaded onto the column. After loading the sample, the column was washed with five CV of equilibration buffer. Seven CV of washing buffer (50 mM sodium phosphate, pH 6.4, 6 M guanidine-HCl, 10 mM TCEP) was then applied to the column to remove any loosely bound protein. Finally, bound Tat was eluted by addition of four CV of elution buffer (50 mM sodium acetate, pH 4.0, 6 M guanidine-HCl, 10 mM TCEP). Guanidine-HCl and TCEP were removed by dialysis at room temperature. Standard acetic acid solutions for dialysis

were prepared by dilution of stock glacial acetic acid into 3-1 L volumes of water to produce three buffers with concentrations of 0.1 M, 0.05 M, and 0.01 M acetic acid. All dialysis buffers were degassed and purged with Argon. The eluted Tat solution was collected in 1.5 ml eppendorf tubes for a total of 10 fractions. Only tubes that showed brown coloration were pooled, transferred into 30 cm of cellulose dialysis tubing (MWCO = 1000 Dalton) (Fisher), and allowed to equilibrate with each acetic acid buffer for 4 hours. A final dialysis with 1 L of ddH<sub>2</sub>O for 4 hours was used to remove any residual salt and most of the acetic acid. 100 µl of the final purified Tat solution was diluted with 900 µl of ddH<sub>2</sub>O, which was subsequently used to determine protein concentration by UV absorption spectroscopy ( $\epsilon = 9090 \text{ M}^{-1}\text{cm}^{-1}$ ). The remaining Tat solution was freeze-dried (Virtis Freezemobile 5SL).

## **2.5 Alkylation of the Cysteine-rich Region of Tat**

Blocking of thiols by alkylation was carried out with iodoacetamide. Tat protein (1 mg) was dissolved in extraction buffer (10 mM Tris-HCl, 10 mM TCEP, 6 M Guanidine-HCl) at pH 8.5. Iodoacetamide (20 mM) was added to the protein solution and allowed to react for 1 hr. Alkylated Tat protein was purified by metal-affinity chromatography followed by dialysis to remove excess iodoacetamide, guanidine-HCl, and TCEP, as described above. Alkylated Tat protein was freeze-dried (Virtis Freezemobile 5SL).

## 2.6 Circular Dichroism Spectropolarimetry

Circular dichroism (CD) spectra were recorded on samples in a 100  $\mu\text{m}$  path length cylindrical cuvette from 240 to 180 nm on a Jasco model J-810 spectropolarimeter at room temperature. Test samples with Tat protein concentration of 0.2 mM were prepared in a solution containing 10 mM acetate and 10 mM Tris-HCl at pH 4.0. Data were collected at a rate of 2 nm per min. The CD spectra were analyzed to determine if Tat possesses any intrinsic secondary structure and to determine the effects of pH, Zn(II), and Cd(II) on Tat's conformation. The CD instrument's intensity and wavelength were calibrated with (+)-10-camphorsulphonic acid (4.3 mM) [120]. All CD spectra were subtracted from the baseline containing 10 mM acetate/Tris-HCl buffer.

### *2.6.1 pH Titration of Tat by CD*

Briefly, 400-500  $\mu\text{M}$  of Tat protein were dissolved in 0.5 ml of 10 mM acetate/Tris-HCl buffer at pH 4.0 and titrated with increasing aliquots of diluted ammonium hydroxide. The solution was allowed to equilibrate for 15 minutes prior to CD spectrum acquisition. The spectral range was from 185 to 240 nm at room temperature.

## 2.7 SDS Polyacrylamide Gel Electrophoresis

SDS-PAGE was performed with the Laemmli System [121] on a SE 260 Mini-Vertical Unit (Hoefer Pharmacia Biotech). Stacking and separating gels were prepared with 125 mM Tris-HCl (4% w/v polyacrylamide, pH 6.8) and 375 mM Tris-HCl (12.5% w/v polyacrylamide, pH 8.8) buffers, respectively, both containing 0.1 % w/v SDS. The electrophoresis running buffer was prepared with 0.025 M Tris-HCl, 0.192 M glycine, 0.1% w/v SDS, at pH 8.3. Freeze-dried Tat protein (2 mg) was solubilized with 500  $\mu$ l of acetate buffer at pH 4.0, mixed with an equal volume of 2X sample treatment buffer, and heated for 30 minutes in boiling water at 100 °C. Pre-Stained Protein Markers (10  $\mu$ l) (*Fisher BioReagents*) were used as standards for determination of molecular weight for separated protein ( $\beta$ -galactosidase, 118 kDa; Bovine serum albumin, 85 kDa; Ovalbumin, 47 kDa; Lactate dehydrogenase, 36 kDa;  $\beta$ -lactoglobulin, 26 kDa; Lysozyme, 20 kDa). After electrophoresis, proteins were stained with Coomassie stain solution (0.025% Coomassie Brilliant Blue R-250, 40% methanol, 7% acetic acid) for 4 hours, and the gel was subsequently transferred into a destaining solution (40% methanol, 7% acetic acid), and left overnight. Initially, protein solutions were boiled to approximately 100 °C in a water bath heated with a Bunsen burner. However, realization that the Tat protein is disordered led to future avoidance of heat denaturation prior to loading onto the gels. The low pH prevents the protein solution from aggregation. 10 mM of tris(2-carboxyethyl)phosphine was subsequently added to the protein solution.

## 2.8 Nuclear Magnetic Resonance (NMR) Spectroscopy

### 2.8.1 Heteronuclear Single Quantum Correlation Spectroscopy

Standard 2D gradient sensitivity-enhanced  $^1\text{H}$ - $^{15}\text{N}$  Heteronuclear Single Quantum Correlation (HSQC) (obtained from the Varian BioPack pulse sequence database and originally from [122]) spectra of uniformly  $^{15}\text{N}$ -labeled Tat were acquired on a 600 MHz Varian INOVA spectrometer equipped with a triple resonance probehead. NMR samples were prepared with 0.5-1.0 mM  $^{15}\text{N}$  Tat in 600  $\mu\text{l}$  of NMR buffer (10 mM acetate- $d_4$ , pH 4.0, 5%  $\text{D}_2\text{O}$ , 80  $\mu\text{M}$  sodium sulfite, 0.02% sodium azide) at 293 K. Sample preparation and subsequent Zn(II) additions were conducted under an anaerobic atmosphere by degassing and then purging the NMR tube with argon prior to NMR spectrum acquisition. 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) was used as an internal standard for chemical shift referencing. The Wishart method was used for referencing to water [123]. A total of 128 scans were signal averaged for every 256 complex  $t_1$  points with acquisition times of 107 ms. The transmitter frequencies for  $^1\text{H}$  and  $^{15}\text{N}$  were set to 599.8 and 150.8 MHz, respectively. Sweep widths for the direct and indirect dimensions were set to 7196 Hz and 2187 Hz, respectively. Prior to Fourier transformation in the indirect dimension, spectra were processed with the following parameters: they were apodized using a squared cosine bell function; zero filled to twice the data set size, and Zhu-Bax linear predicted [131]. The data were processed and analyzed with

Spin Works 3.0 (Marat, K. University of Manitoba, Canada), NMRPipe [124], and Sparky [125].

For the zinc titrations, it is noteworthy that the dilution effect can be ignored due to the fact that 5  $\mu\text{L}$  of the  $\text{ZnSO}_4$  solution was added to 600  $\mu\text{L}$  of Tat sample during each titration point prior to NMR data acquisition. Thus, dilution of the Tat protein per titration step reduces the concentration to 0.9917 of its original value.

### *2.8.2 Heteronuclear Multiple-Bond Correlation Spectroscopy*

$^1\text{H}$ - $^{15}\text{N}$  Heteronuclear Multiple-Bond Correlation (HMBC) [128] spectra were measured to detect the correlation between  $^1\text{H}$  and  $^{15}\text{N}$  through  $^2\text{J}_{\text{NH}}$  coupling within the imidazole ring of the histidine residues. Two-dimensional  $^1\text{H}$ - $^{15}\text{N}$  HMBC spectra were acquired on  $^{15}\text{N}$ -labeled Tat at concentrations of 0.5-1.0 mM in 600  $\mu\text{l}$  of NMR buffer in the presence and absence of Zn(II) at pH 4.0. The relaxation delay and  $^{15}\text{N}$  spectral width were set to 1.0 s and 140 ppm, respectively. A total of 128 scans were averaged for every 256  $t_1$  complex points with acquisition times of 106.5 ms. The sweep widths for the direct and indirect dimensions were set to 5996 Hz (10 ppm) and 85071 Hz (140 ppm), respectively. Prior to Fourier transformation in the indirect dimension, spectra were processed with the following parameters: they were apodized using a squared cosine bell function and zero filled to twice the data set size. In addition, a long-range J-coupling constant of 6.0 Hz was used as an average for setting the refocusing delay time of 41.6 milliseconds. The data were

processed and analyzed with Spin Works 3.0 (Marat, K. University of Manitoba, Canada), NMRPipe [124], and Sparky [125].

### *2.8.3 Nuclear Overhauser Effect Spectroscopy*

Two-dimensional NOESY [127] spectra were acquired on an unlabeled Tat (0.5-1.0 mM) sample in 600  $\mu$ l of NMR buffer. Signal optimization was conducted by acquiring NOESY spectra with varying mixing times of 50, 100, 150, 200, 300, and 400 ms. Rotating frame NOESY (ROESY) [130] spectra were recorded with a mixing time of 200 ms to investigate the rotational correlation time of Tat through peak intensity comparison with NOESY results. The Watergate [129] pulse sequence was applied to both NOESY and ROESY experiments to suppress the solvent signal. Zero-filling to twice the data set size and apodization of the spectra using a squared cosine bell function were carried out prior to Fourier transformation. The spectral width was set to 5996 Hz (10 ppm). The data were processed and analyzed with Spin Works 3.0 (Marat, K. University of Manitoba, Canada), NMRPipe [124], and Sparky [125].

### *2.8.4 $^{113}\text{Cd}$ NMR Spectroscopy*

$^{113}\text{Cd}$  NMR spectra were obtained on a Varian 600 MHz spectrometer equipped with a tunable broadband probehead with the resonance frequency set to 133 MHz. The NMR sample was prepared by dissolving 1.82 mg of lyophilized Tat in 700  $\mu$ l of  $\text{D}_2\text{O}$

NMR buffer (10 mM acetate-d<sub>4</sub>, 80 μM sodium sulfite, 0.02% sodium azide, 100% D<sub>2</sub>O) at pD 5.0 followed by titration with 1, 2, 4, 6, 8, 10, and 20 mole equivalents of <sup>113</sup>CdCl<sub>2</sub> (<sup>113</sup>CdCl<sub>2</sub>, 93.53 % Cambridge Isotope Laboratories, Inc.). The pD of the <sup>113</sup>Cd NMR solution was determined by using the equation pD = pH + 0.40 [126]. The NMR protein sample was prepared under anaerobic conditions by degassing and purging the NMR tube prior to data acquisition. A small amount of Tat NMR solution [100 μl] was set aside for determination of protein concentration by UV absorption spectroscopy. All <sup>113</sup>Cd NMR spectra were referenced relative to 0.1 M Cd(ClO<sub>4</sub>)<sub>2</sub>. Acquisition parameters for direct detection of <sup>113</sup>Cd were as follows: a relaxation delay of 0.5 s was used, the sweep width was set to 100,000 Hz (751.9 ppm), and 40,000 free induction decays were acquired. The acquisition times were 2 hrs and 20 seconds.

#### 2.8.4.1 Mathematica Calculations for Binding Affinity

Using a well established software program, Mathematica (2008, Wolfram Research, Inc.), the binding affinity of the Tat protein towards Cd(II) can be determined by non-linear least-squares fitting to a binding equation. The underlying assumption that has to be made in order to accurately determine the metal dissociation constant ( $k_D$ ) is that the concentration of the protein has to be significantly less than the  $k_D$  [213]. Under this condition, the concentration of the Cd(II) bound to the protein will be only a small fraction of the total Cd(II) concentration. Thus, the concentration of free Cd(II) is approximately equal to the

total concentration of Cd(II) added ( $[Cd]_{total} = [Cd]_{free} + [Cd]_{bound}$ ). It is noteworthy that the concentration of the protein (300  $\mu$ M) is kept constant while the metal is titrated over a range (1-20) of mole equivalents. This permits the use of a simplified equation for determining the binding affinity (Equation 4)

$$\frac{[AB]}{[A]_{Total}} = \frac{[B]}{[B] + K_D} \quad [4]$$

The  $[AB]$  denotes the concentration of Tat-Cd(II) complex, while  $[A]$  and  $[B]$  denote the concentration of the Tat protein and Cd(II)<sub>free</sub>, respectively. Under this equation one can plot the fraction bound ( $[AB]/[A]_{Total}$ ) against the concentration of  $[B]$  and determine the  $k_D$ . Figure 2.1 shows a model hyperbolic curve defined by Equation 4 that indicates that the  $k_D$  is the concentration of  $[B]$  at exactly half the fraction bound (0.5). The assumption here is that the change in  $^{113}\text{Cd}$  NMR chemical shift at each metal titration point is an indication of protein-metal binding. Furthermore, the change in chemical shift increases at each point suggesting more protein is binding to the metal. For the  $^{113}\text{Cd}$  NMR data, we fit the total metal concentration against the change in chemical shift to a single binding equation (Equation 5):

$$y = C_i - (r[C_i - C_o]) \quad [5]$$

where,

$$r = \frac{[B]_{Total}}{[B]_{Total} + K_D}$$

The  $C_i$  and  $C_o$  denote the  $^{113}\text{Cd}$  chemical shift at saturation of the binding sites and at zero metal concentration, respectively. It is noteworthy that Equation 5 is derived from a study by Battle et al. [214] on the binding affinity of a protein-RNA interaction.

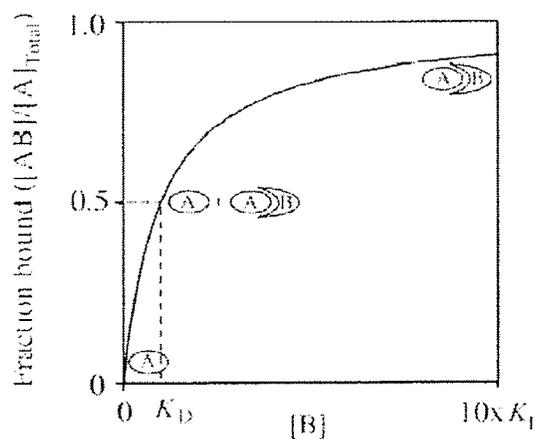


Figure 2.1: Model for the binding curve showing the  $K_D$  is defined as the concentration of B at which the fraction bound is 0.5. Illustration was reprinted with permission of Goodrich et al. [165].

## 2.9 Cadmium Titration by Ultraviolet Absorption Spectroscopy

Metal-binding experiments of apo-Tat with 0.5-2.0 molar equivalent additions of Cd(II) were recorded at room temperature on a Varian CARY13E UV-Visible spectrophotometer using a 1 cm quartz cuvette. All UV spectra were recorded with a scan rate of 30 nm min<sup>-1</sup> from 330 to 230 nm. Tat protein [16 μM] was dissolved in 1 ml of degassed acetate buffer [10 mM] at pH 4.5. For Cd(II) titrations, aliquots of 10 μl of Cd(II) [800 μM] were added prior to each acquisition. The UV absorption spectrometer was blanked with 10 mM acetate buffer at pH 4.5.

## 3. Results

### 3.1 Protein Expression, Purification, and Protein Concentration

The growth of *E. coli* cells containing recombinant Tat plasmid was initially monitored by measuring the OD<sub>600nm</sub> using a Klett-Summerson Photoelectric, which correlates the turbidity of the sample to the cell concentration.

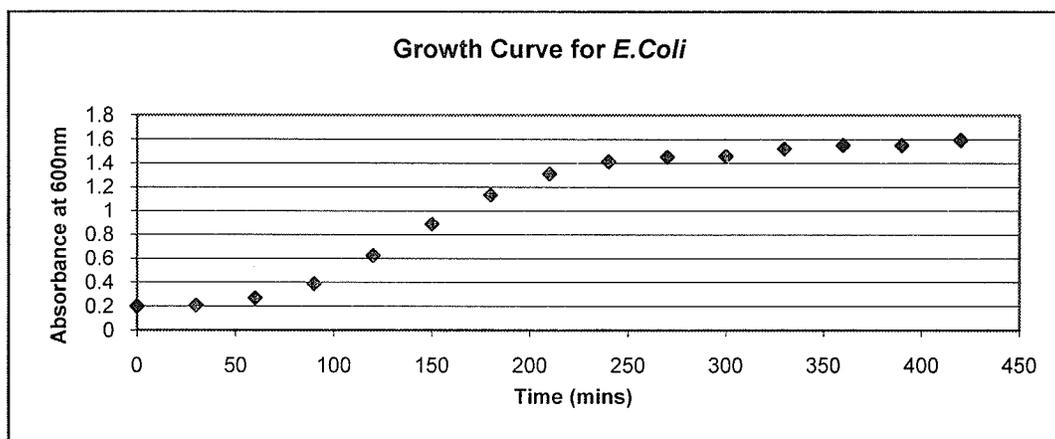


Figure 3.1: Growth curve for *E. coli* cells measured at 600 nm in 30 min intervals at 37 °C.

In Figure 3.1, the observed time course of bacterial growth follows a sigmoidal curve, which consists of three phases: lag, exponential, and stationary. The lag phase takes place from 0 to 60 minutes, during which only a small increase in the optical density is observed. Between 60 and 240 minutes, the population enters the log phase whereby the cell number increases in a logarithmic fashion. The

stationary phase occurs from 240 to 440 minutes, during which the rate of bacterial growth plateaus. Tat protein synthesis is usually induced with isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG) in early to mid-log phase corresponding to an optical density (OD) at 600 nm of 0.6 [140]. Determination of an *E. coli* growth curve, and the rate of bacterial growth during the log phase, is helpful for determining when a given culture should be induced based on OD<sub>600nm</sub> measurements of cultures. Typically, 10 grams of *E. coli* cells are harvested per litre of TB medium.

### *3.1.1 Protein Purification via Cobalt Metal-affinity Chromatography*

The recombinant Tat protein contains an N-terminal His-tag that allows purification via cobalt metal-affinity chromatography [Figure 3.2]. Following cell lysis and separation of cell membrane components from the cytoplasmic components, cell lysate is added to the metal-affinity column. Proteins that contain the His-tag will bind to the cobalt at pH 7.2 and subsequently elute with elution buffer at pH 4.1. Based on experience, fractions that contain Tat protein are brown in color and are pooled and dialyzed to remove TCEP and guanidine-HCl. The appearance of the brown coloration is likely due to the presence of loosely bound cobalt to the Tat protein. The freeze-dried Tat protein appears white suggesting that the bound cobalt is substantially removed during dialysis.

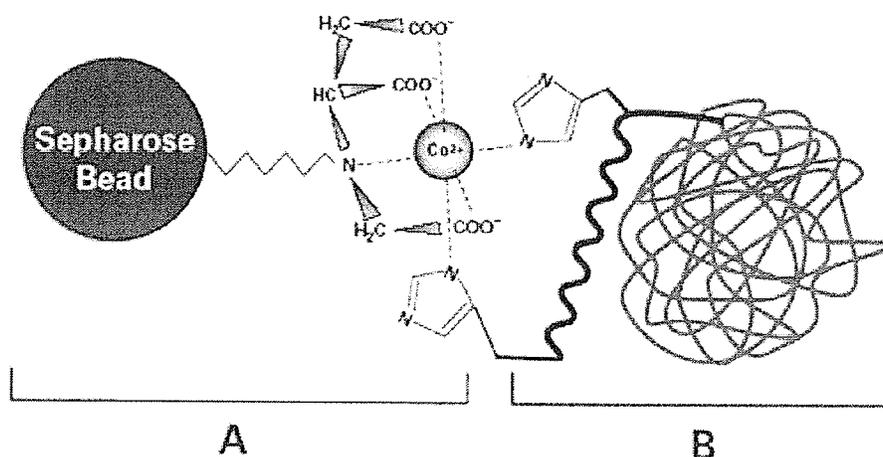


Figure 3.2: Interaction between the cobalt metal-affinity column (A) to the imidazole ring of the His<sub>6</sub>-tag (B) [210].

Despite the positive effects of the His-tag as previously mentioned, removal of the affinity tag is desirable to prevent unpredicted structural changes in order to study the protein in its physiological conformation. However, attempts at removal of the His-tag using thrombin have been unsuccessful possibly due to the presence of an internal thrombin cleavage site between Lys-61 and Ala-62 [141].

### 3.1.2 Ultraviolet Absorption Spectroscopy of HIV-1 Tat

The measurement of protein yield is carried out using ultraviolet (UV) absorption spectroscopy at 280 nm and the Beer-Lambert Law. Following dialysis of purified Tat protein, 100  $\mu$ l of the protein solution is diluted with 900  $\mu$ l of water and a UV absorption spectrum is collected. In Figure 3.3, the spectrum of Tat shows the characteristic shoulder band at 290 nm and a strong band at 280 nm corresponding

to tryptophan absorption. Furthermore, the absence of light scattering in the spectrum suggests that the Tat protein is in a reduced monomeric form. The Tat protein concentration is determined by using the absorbance measurement at 280 nm and the molar extinction coefficient of the protein ( $\epsilon = 9090 \text{ M}^{-1}\text{cm}^{-1}$ ). Typically, 5 mg of the Tat protein is purified per litre of *E. coli* cell culture.

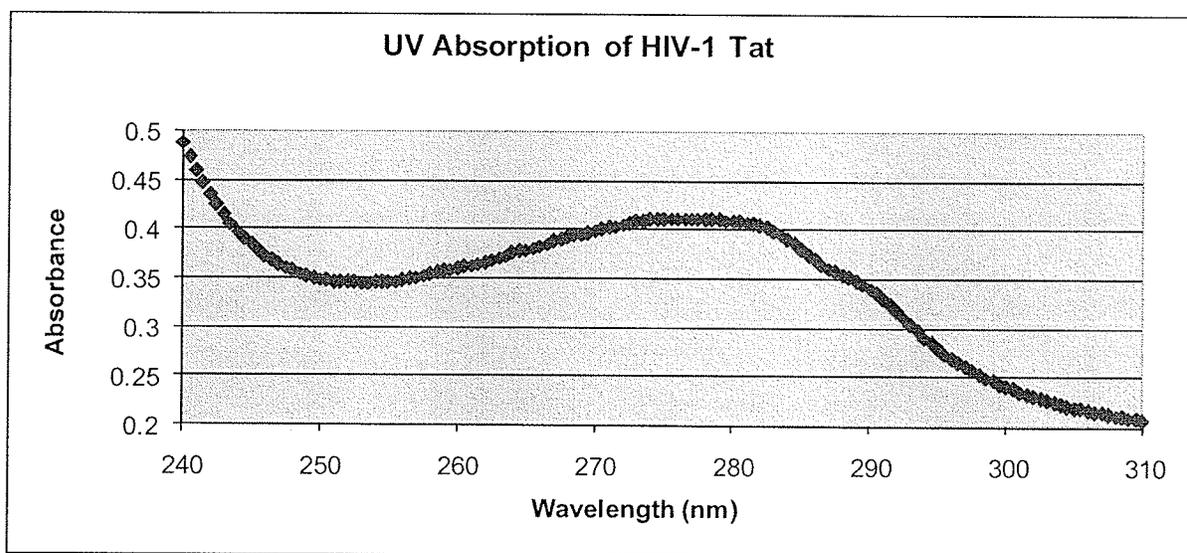


Figure 3.3: UV Absorption spectrum of HIV-1 Tat (0.473 mg/ml) taken from 240 nm to 310 nm. The solvent blank is water.

## 3.2 SDS-PAGE Analysis

### *3.2.1 Analysis of Tat Fractions from Cobalt Metal-affinity Column.*

SDS-PAGE is a well-established technique for investigating the cross-linking properties of proteins containing multiple cysteine residues. Although the Tat protein is in a reduced monomeric form at pH 4.1, it is very likely to form oligomers via intermolecular disulfide-bond formation at pH above 6 or in aerobic or non-reducing conditions. Tat protein fractions from cobalt metal-affinity chromatography were analyzed by SDS-PAGE. Figure 3.4 shows an SDS-PAGE electrophoregram of fractions 1-4 eluted from the resin. Lane 1 shows the protein molecular weight markers. In lane 2, no bands are observed in fraction 1 indicating that the Tat protein has not eluted yet. In lanes 3 (fraction 2) and 4 (fractions 3 and 4), an intense band with an apparent molecular mass of 17 kDa is observed suggesting the Tat protein is in a predominantly reduced monomeric form. The discrepancy between the theoretical molecular mass of 10 kDa and an apparent molecular mass of 17 kDa observed on the SDS gel is likely due to the Tat protein having an overall high positive charge of +25.2 at pH 4.0 [142]. This high net positive charge leads to the retardation of the protein's electrophoretic mobility. Interestingly, additional fainter bands are observed in lanes 3 and 4 of Figure 3.4 indicating proteins with higher apparent molecular mass than the Tat monomer suggesting the presence of small amounts of Tat oligomers or possibly protein

impurities. Because of the observation of similar bands, SDS-PAGE has not been used routinely to establish the purity of Tat protein preparations [141].

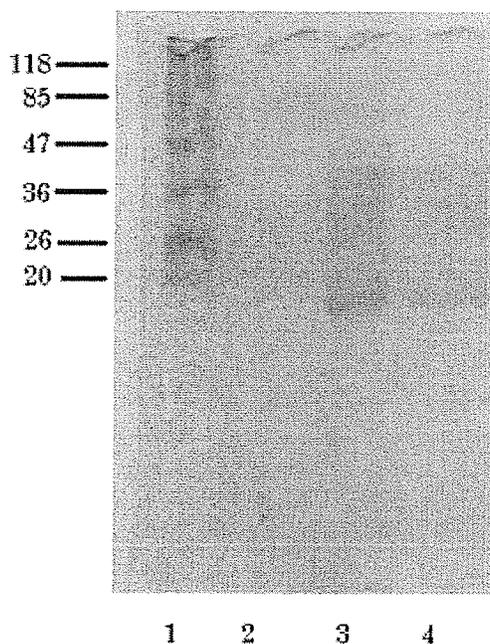


Figure 3.4: Coomassie Blue-stained SDS-PAGE electrophoregram identifying Tat protein from each fraction collected from the cobalt metal-affinity column shown in Figure 3.2. Standard protein marker (10  $\mu$ l) is added to lane 1. Lanes 2 and 3 show fractions 1 and 2, respectively. Lane 4 shows fractions 3 and 4 pooled.

### *3.2.2 Analysis of the Effect of TCEP on Tat*

To determine the nature of the multiple bands observed on the polyacrylamide gel shown in Figure 3.4, SDS-PAGE of the same samples was done in the presence of TCEP. TCEP is a powerful reducing agent used in the purification of Tat [132] that allows the formation of reduced monomeric Tat by preventing disulphide cross-linking. Although TCEP has been shown to be an excellent reducing agent in SDS-PAGE sample preparation buffer at low pH, at neutral pH it requires the presence of 6 M guanidine-HCl to prevent the formation of a Tat-TCEP precipitate. In Figure 3.5, lane 1 shows the protein molecular weight markers. In lane 2, corresponding to fraction 1, no bands are observed. In lane 3, column fraction 2 shows the monomeric Tat band at 17 kDa along with multiple higher molecular weight bands indicating formation of Tat oligomers or protein impurities. Furthermore, an intense band is observed at the top of the stacking gel for lane 3. Lane 4 contains fractions 3 and 4 and is very similar in appearance to lane 3. Lanes 5 and 6 contain the same material as lanes 3 and 4, respectively, except that 10 mM TCEP was added. TCEP nearly eliminates all of the high molecular weight bands in lanes 3 and 4 leaving a predominant band at 17 kDa indicating the reduction of oligomeric Tat to the monomeric form and confirming the high purity of the protein preparation. This experiment also shows that SDS-PAGE is a useful tool for the analysis of Tat protein as long as a strong reducing agent such as TCEP is included in the protein solution.

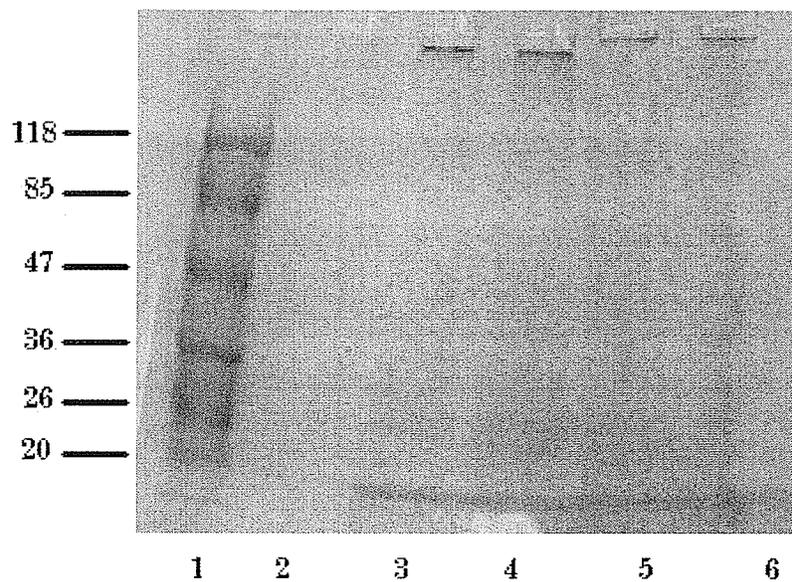


Figure 3.5: SDS-PAGE results of Tat protein fractions collected from a cobalt metal-affinity column. Fractions 1 and 2 from the column are present in lanes 2 and 3, respectively. Fractions 3 and 4 were pooled and added to lane 4. Lanes 5 and 6 are duplicates of lanes 3 and 4, respectively, except for the addition of TCEP (10 mM). Lane 1 contains standard protein markers (10  $\mu$ l).

### *3.2.3 Determining the Optimal Concentration of TCEP*

Figure 3.5 shows that 10 mM TCEP affords nearly complete reduction of Tat disulphides. To determine if higher concentrations of TCEP can eliminate the remaining oligomerized protein observable on the electrophoregram, electrophoresis of Tat protein was done on samples containing a range of concentrations of TCEP. Figure 3.6 illustrates an SDS-PAGE electrophoregram of Tat protein in the presence of TCEP (added to the sample prior to electrophoresis) at concentrations ranging from 1 to 100 mM. In lane 3, corresponding to the Tat protein in the absence of TCEP, multiple bands are observed suggesting the formation of Tat oligomers. In the presence of 1 mM TCEP (lane 4), the multiple bands corresponding to Tat oligomers are observed to become less intense relative to those in lane 3. In lane 5, a further decrease in intensity of the Tat oligomers is observed upon addition of 5 mM TCEP, however it must be noted that the total amount of protein appears to be lower than in lanes 3 and 4. At a TCEP concentration of 10 mM [lane 6], a predominant band at 17 kDa is observed suggesting a pronounced reduction of oligomeric Tat to the monomeric form. Interestingly, in lane 7 (at a TCEP concentration of 12.5 mM) a fainter monomeric Tat band is observed with apparent molecular weight of 17 kDa but all the higher molecular weight oligomers have been eliminated. Doubling the TCEP concentration to 25 mM (lane 8) shows the monomeric Tat band migrating at a slightly higher MW relative to the monomeric Tat bands in lanes 3 – 7. In lanes 9 and 10, the monomeric Tat band is observed to be less intense and migrates at a significantly higher MW. The decrease in intensity

of the monomeric Tat band at TCEP concentrations above 10 mM might be due to precipitation of the protein by TCEP. A plausible explanation for the observation of higher MW monomeric Tat band is that a solution containing a very high concentration of acidic TCEP may cause a decrease in the pH of the stacking gel from pH 6.8. The rationalization of the pH effect of TCEP is substantiated by measuring the pH of a Tat protein solution before and after the addition of TCEP. After addition of 100 mM TCEP to a Tat protein solution [0.475 mM], the pH is observed to be 1.91, which is a significant decrease from pH 4.0. This may affect the ability of the stacking gel to concentrate the protein solution prior to running through the resolving gel as compared to the Tat solution without TCEP. The inability of the Tat solution to be stacked likely explains the observation of a broadened, higher MW monomeric Tat band.

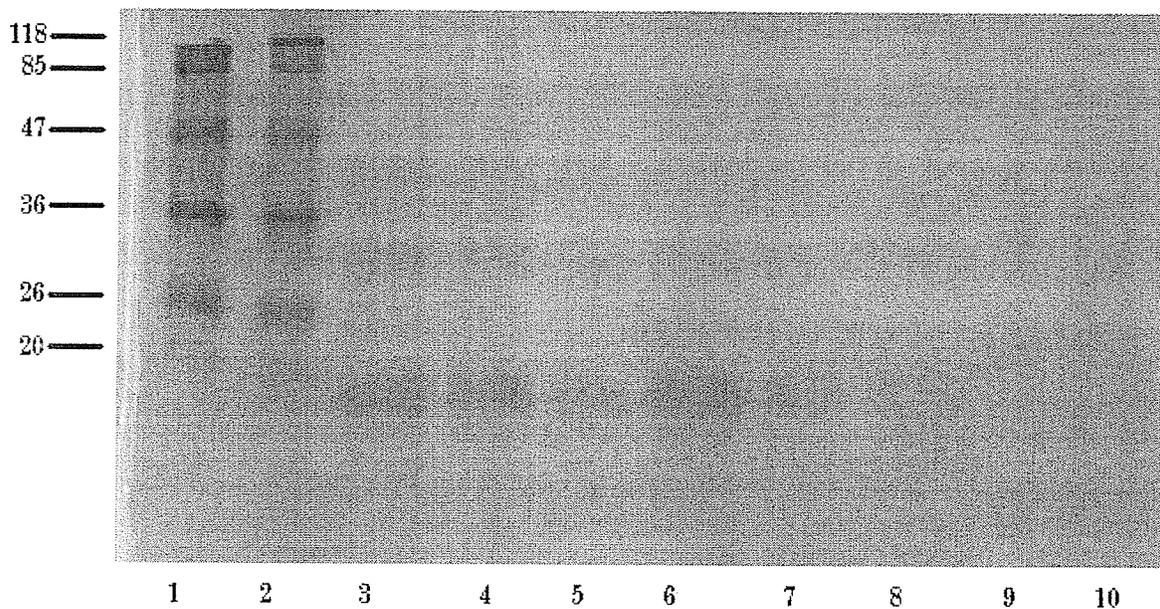


Figure 3.6: SDS-PAGE analysis of Tat in the presence of increasing concentrations of TCEP. Lanes 1 and 2 contain standard protein markers (10  $\mu$ l). Lane 3 contains Tat protein (0.475 mM) in the absence of TCEP and serves as a control. TCEP was added to Tat protein prior to electrophoresis at concentrations of 1 mM (lane 4), 5 mM (lane 5), 10 mM (lane 6), 12.5 mM (lane 7), 25 mM (lane 8), 50 mM (lane 9), and 100 mM (lane 10). Note that the stacking gel is missing from the above SDS gel.

### *3.2.4 Analysis of Alkylated Tat*

Above pH 4.5, the Tat protein precipitates from solution forming disulphide cross-linked oligomers. In order to understand the folding and dimerization of Tat above pH 4.5, a study of the redox reactivity of the cysteine residues was initiated. Initially, the reactivity of the protein towards iodoacetamide was examined as a mechanism for quenching the oxidation of the cysteine residues. The ultimate goal was to study the reactivity of disulfide bond formation in hope of revealing a possible folding of the protein in an oxidizing environment. Alkylation of the cysteine residues was performed using 100 mM iodoacetamide for 6 hrs. For these studies, SDS-PAGE is used as the primary tool for the analysis of the Tat oligomeric state. Figure 3.7 is an electrophoregram of the same amounts of untreated and alkylated Tat. In lane 2, a predominant band at 17 kDa is observed accompanied by a fainter band at 28 kDa indicating the presence of Tat dimers. Faint bands at higher  $M_r$  are indicative of higher oligomers of the protein. Lane 3 shows the result of alkylating Tat protein, in which a single intense band at 17 kDa is observed with very little oligomer. Comparing the monomeric band of Tat in lane 2 and that of alkylated Tat in lane 3 reveals a significant difference in intensity suggesting that the iodoacetamide prevents the formation of Tat oligomers during electrophoresis and is evidence of successful protein alkylation.

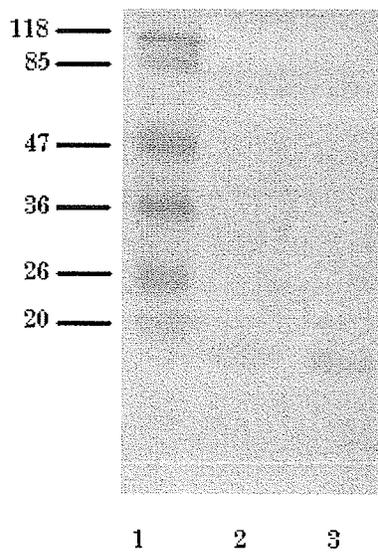


Figure 3.7: SDS-PAGE analysis of Tat and alkylated Tat protein. Lane 1 contains the standard protein markers (10  $\mu$ l). Lane 2 contains Tat protein and serves as a control. Lane 3 contains alkylated Tat protein. Alkylation of the Tat protein was carried out with 100 mM iodoacetamide for 6 hrs.

### 3.3 Circular Dichroism

#### *3.3.1 Deconvolution of the CD Spectrum of HIV-1 Tat*

CD is an excellent method for determining protein conformation in solution by measuring the difference in absorbance for left- and right-handed circularly polarized light [133]. Typically, measurements are carried out in the far-UV region of the spectrum (180-240 nm), as the amide chromophore (peptide bond) absorption probes the protein secondary structure. In Figure 3.8, the far-UV CD spectrum of the Tat protein at pH 4.17 shows a strong negative band at 199 nm and a much weaker negative band at 225 nm, suggestive of a random coil conformation [133, 143]. It is noted that strong light absorption is observed at wavelengths below 183 nm (not shown) and giving rise to noise in the CD signal.

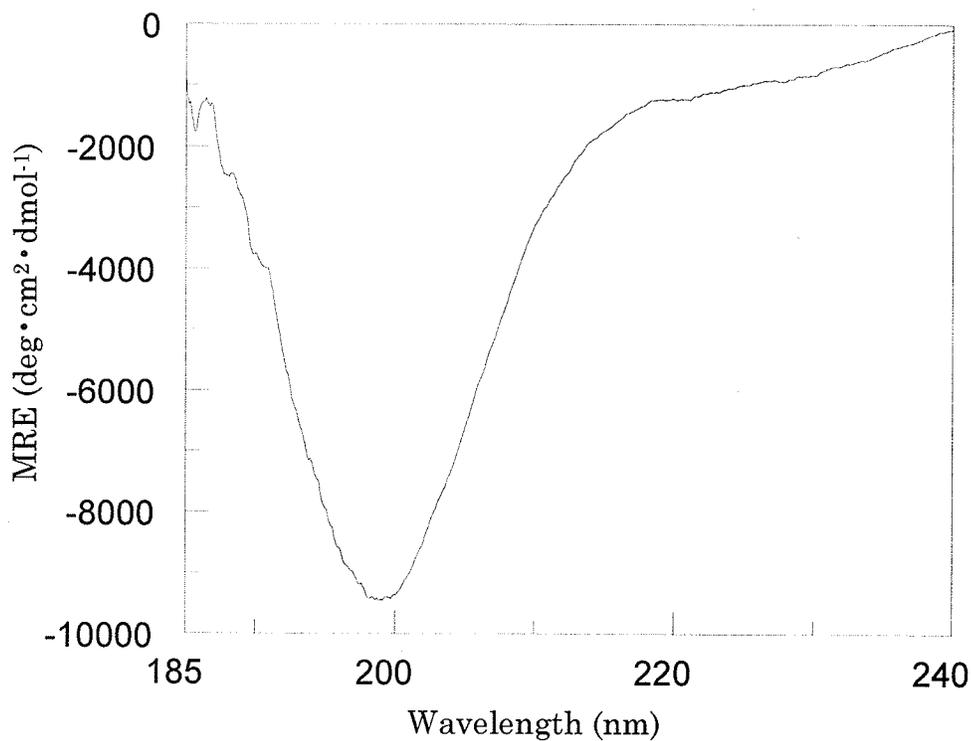


Figure 3.8: Far-UV CD spectrum of HIV-1 Tat at room temperature. Tat protein concentration was 173  $\mu\text{M}$  dissolved in 10 mM acetate buffer at pH 4.17. An algorithm taken from the web-based program Dichroweb was used for deconvolution of the Tat CD spectrum [145]. The CD spectrum was baseline subtracted using a blank solution containing 10 mM acetate/Tris-HCl buffer.

A large number of computer-based algorithms are widely used today to predict the secondary structural content of a protein based on comparison of experimental CD data to well established reference databases. A web-based program, Dichroweb, is used here to determine the percentages of alpha-helix, beta-sheet, and unordered structure of the Tat protein [136]. In Table 3.1, deconvolution of the CD spectrum of apo-Tat at pH 4.17 using three different algorithms show that on average, the protein is mainly composed of disordered secondary structure (35.3 %), followed by 22.6 % of  $\beta$ -turns, and the remainder is made up of  $\beta$ -strand 1 (13.0%) and 2 (11.2%). Interestingly, all three algorithms agree that the Tat protein is predominantly disordered and does not contain a significant amount of  $\alpha$ -helical structure, which contradicts some of the literature NMR data of the HIV-1 Tat protein [144]. The high content of  $\beta$ -turns and  $\beta$ -sheet is not surprising. At low pH the highly basic protein may contain regions of polypeptide with an extended backbone conformation resembling that of a  $\beta$ -sheet. In addition,  $\beta$ -turns are often observed in disordered proteins and are thought to be in equilibrium with  $\beta$ -strand, disordered, and polyproline II helix (135).

Table 3.1: Predicted secondary structure content of HIV-1 Tat through deconvolution using the web-based program Dichroweb.

| Algorithm | Helix 1 | Helix 2 | Strand 1 | Strand 2 | Turns | Disordered | Total |
|-----------|---------|---------|----------|----------|-------|------------|-------|
| CDSSTR    | 0.000   | 0.030   | 0.230    | 0.120    | 0.230 | 0.360      | 0.970 |
| SELCON3   | 0.024   | 0.047   | 0.140    | 0.100    | 0.209 | 0.346      | 0.866 |
| ContnLL   | 0.014   | 0.072   | 0.021    | 0.117    | 0.238 | 0.350      | 1.000 |
| Average   | 0.013   | 0.050   | 0.130    | 0.112    | 0.226 | 0.352      | 0.945 |

### 3.3.2 Effects of pH on Tat CD Spectra

The HIV-1 Tat protein structure was shown above to adopt a predominantly random coil conformation (with 35.2 % disordered structure) in 10 mM acetate buffer at pH 4.17 by CD spectropolarimetry and this agrees with the results of an earlier study that used multinuclear NMR spectroscopy to study the protein at the same pH [26]. However, several NMR analyses of the protein at higher pH have suggested that Tat has a condensed globular structure with significant alpha helix content [54, 144]. To determine the conformation of HIV-1 Tat at physiological pH, a pH titration experiment with ammonium hydroxide was conducted. Figure 3.9A shows CD spectra of Tat protein from pH 4.17 to 8.06. At low pH Tat exhibits a typical random coil conformation, as evidenced by weak and strong minima at 220 nm and 198 nm, respectively [133]. The band at 199 nm is significantly diminished in intensity as a function of increasing pH between pH 4.51 and 8.06, whereas the band at 220 nm is only slightly diminished. Figure 3.9B, shows a linear decrease in intensity measured at 199 nm from pH 4.51 to 7.51. However, the loss in ellipticity

is likely due to Tat protein dilution at higher pH values. Figure 3.9C shows that as the pH values increase from 4.17 to 8.06, the absorbance measured at 199 nm decreases linearly, suggestive of protein dilution from the addition of ammonium hydroxide. Precipitation of the Tat protein is ruled out due to the absence of a scattered light spectrum. Another consideration is that during the pH measurement of the Tat protein solution, small amounts of the positively charged protein may stick to the negatively charged pH glass electrode. Figure 3.9D shows the CD spectrum of Tat at different pH values corrected for dilution making use of the measured absorbance changes. Clearly, raising the pH of the Tat solution imposes a dramatic secondary structural change as evidenced by the significant decrease in ellipticity. Especially noteworthy is the observation of a lack of any isodichroic point in the overlaid spectra.

To determine the meaning of the decrease in ellipticity at higher pH values, deconvolution of the Tat protein at pH 8.04 was carried out to characterize the change in conformation as a function of pH. Table 3.2 shows the deconvolution results of the Tat protein at pH 8.04 using the web-based Dichroweb. The average of all three different algorithms shows that the Tat protein is primarily made up of disordered secondary structure (32.0%), followed by 25.5%  $\beta$ -turns. The remaining structural composition of the protein consists of 25.7% of  $\beta$ -sheet and 17.2% of  $\alpha$ -helical. In Table 3.3, the comparison between the secondary structure percentage of Tat at pH 4.0 and 8.04 shows a decrease in disordered structure by 3.2% followed by an increase in  $\alpha$ -helical and  $\beta$ -content ( $\beta$ -turns +  $\beta$ -sheets) structures by 10.9% and

4.4%, respectively. Based on the analysis of the secondary structural content of HIV-1 Tat, the protein is becoming more ordered at higher pH values.

Table 3.2: Predicted secondary structure content of HIV-1 Tat at pH 8.04 through deconvolution using the web-based program Dichroweb.

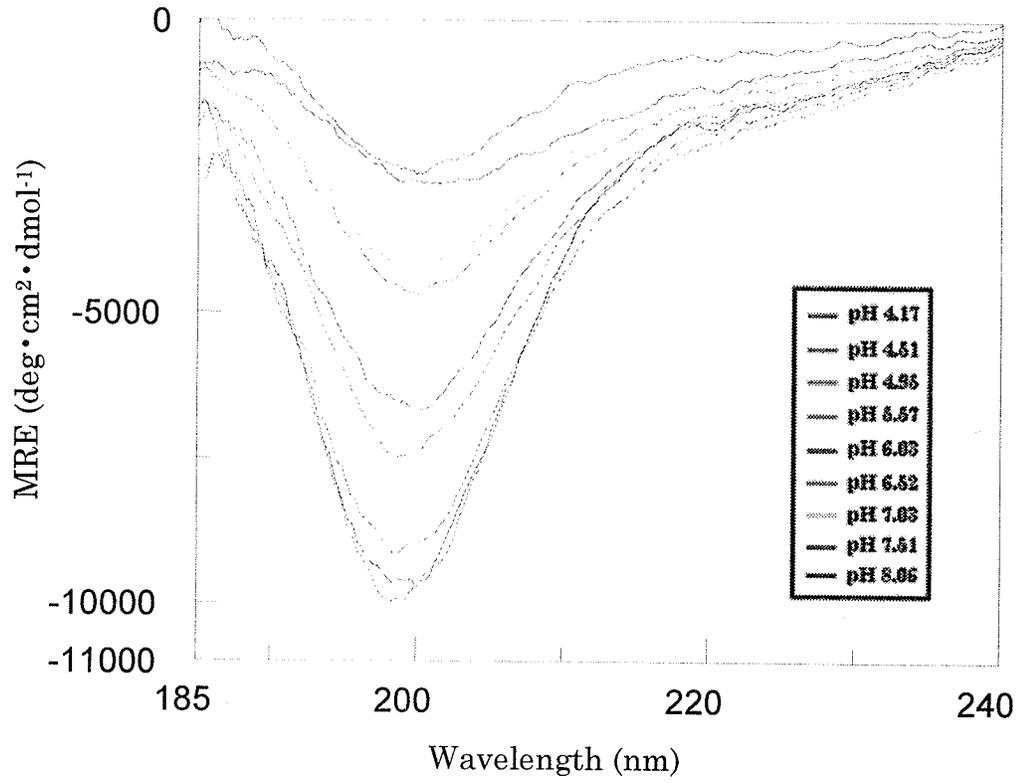
| Algorithm | Helix 1 | Helix 2 | Strand 1 | Strand 2 | Turns | Disordered | Total |
|-----------|---------|---------|----------|----------|-------|------------|-------|
| CDSSTR    | 0.090   | 0.110   | 0.180    | 0.110    | 0.240 | 0.280      | 1.010 |
| SELCON3   | 0.058   | 0.104   | 0.170    | 0.116    | 0.240 | 0.312      | 1.000 |
| ContinLL  | 0.053   | 0.100   | 0.088    | 0.108    | 0.285 | 0.367      | 1.001 |
| Average   | 0.067   | 0.105   | 0.146    | 0.111    | 0.255 | 0.320      | 1.004 |

Table 3.3: Comparison between secondary structural content of HIV-1 Tat at pH 4.0 and 8.04

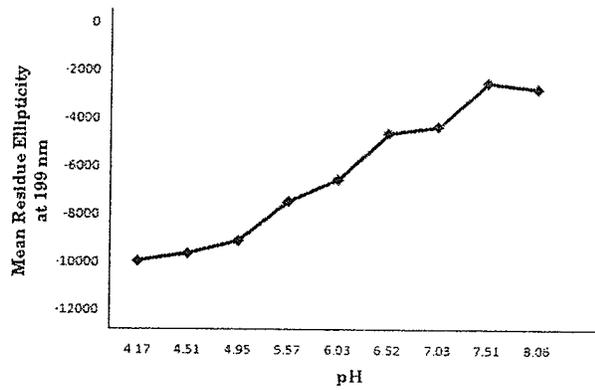
|            | Helix 1 | Helix 2 | Strand 1 | Strand 2 | Turns | Disordered | Total |
|------------|---------|---------|----------|----------|-------|------------|-------|
| Difference | 0.054   | 0.055   | 0.016    | -0.001   | 0.029 | -0.032     | 0.058 |

Figure 3.9: (A) CD spectra of Tat (190  $\mu\text{M}$ ) at different pH conditions in 10 mM acetate buffer at room temperature. (B) Plot of uncorrected CD intensity at 199 nm as a function of pH. (C) Changes in High-Tension Voltage (absorption) at 199 nm as a function of pH. (D) CD spectra of Tat (190  $\mu\text{M}$ ) at different pH values corrected for dilution effects. (E) Mean residue ellipticity at 199 nm of Tat protein at different pH values corrected for dilution. All CD spectra were baseline subtracted using a blank solution containing 10 mM acetate/Tris-HCl buffer.

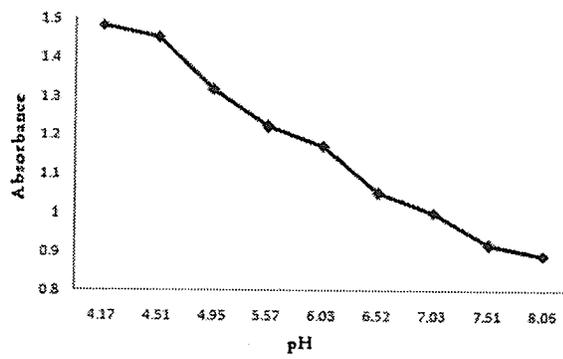
A



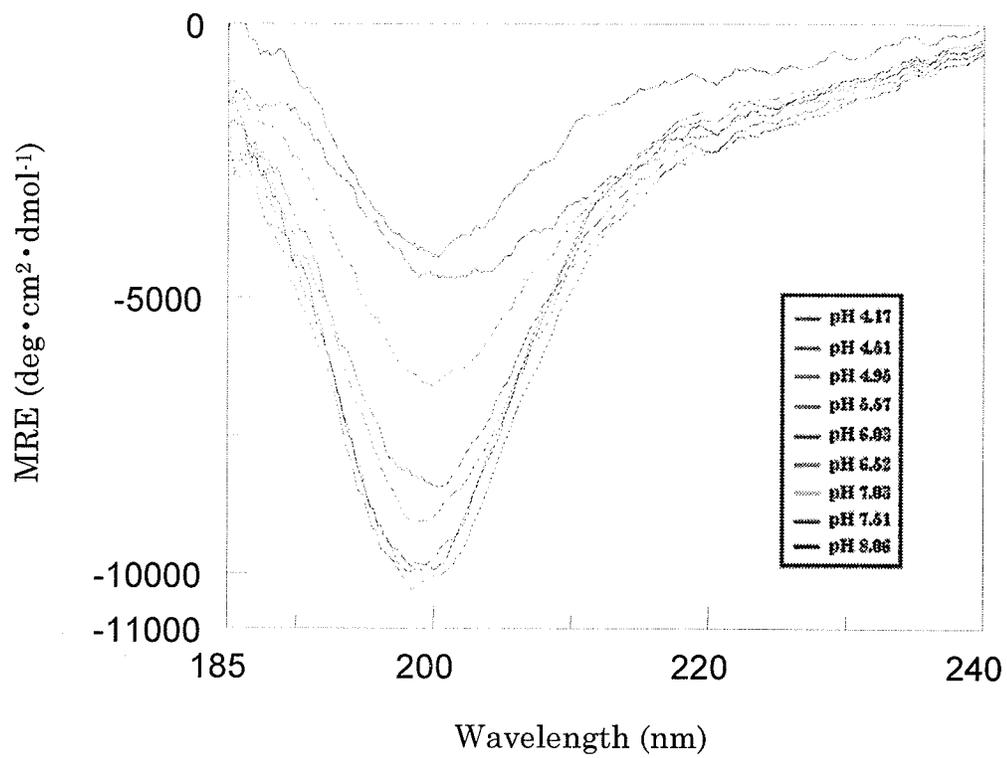
B



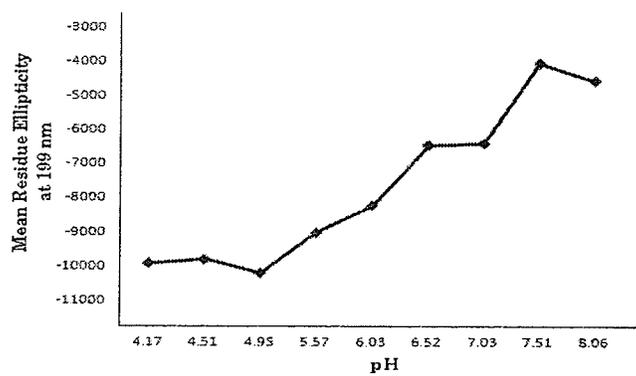
C



D



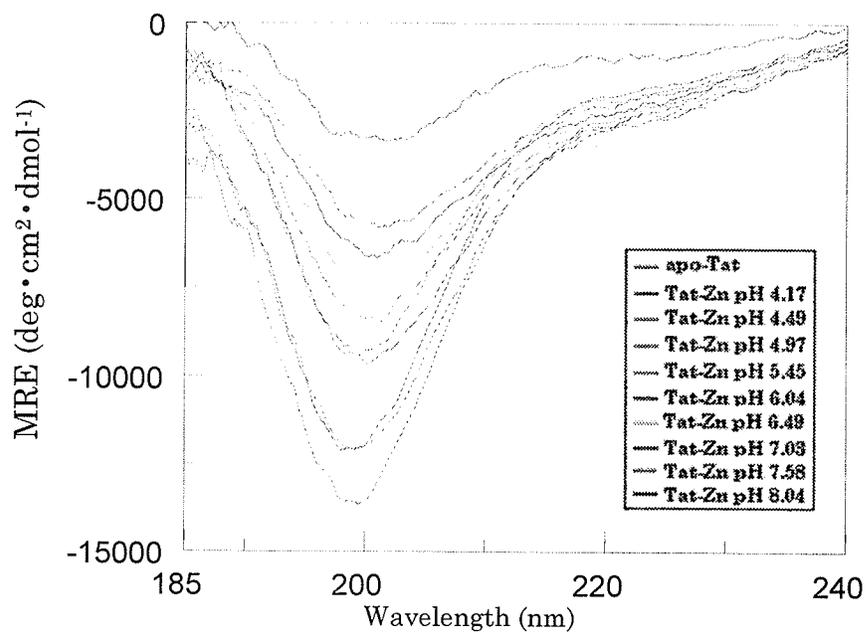
E



### 3.3.3 Effects of pH on Tat-Zn(II) Complex

The study of metal-induced protein folding has been a popular field for many years, particularly in the study of metallothionein proteins [136]. In the presence of Zn(II), it has been suggested that the Tat protein forms a dimer [39] whereby the cysteine residues from each monomer are coordinating with the metal. By coordinating the cysteine residues via metal-linked dimerization, it might be possible to raise the pH of the solution and prevent or delay Tat precipitation. To test this hypothesis and to explore the possibility of the formation of a metal-linked dimer, Tat's conformation in the presence of Zn(II) was analyzed with CD spectropolarimetry. Similar to the spectrum of apo-Tat, Figure 3.10A shows that the spectrum of Tat in the presence of 0.5 mole equivalents of Zn(II) at low pH appears to be that of a random coil. The addition of 0.5 mole equivalents of Zn(II) causes a small decrease in ellipticity suggestive of a change in backbone conformation upon binding to the metal. However, because this small change was not reproducible (see below) it could also have been the result of a small increase in the pH of the solution upon addition of the zinc. As the pH increases from 4.17 to 8.04, the intensity at 199 nm decreases in a linear fashion (Figure 3.10B) similar to what was observed in the absence of metal (Figure 3.9). It may be significant however, that the pH 8.04 form of the protein has significantly less ellipticity in the presence of metal than in its absence. Thus, the lower random coil ellipticity may suggest that partial metal-induced folding has taken place.

A



B

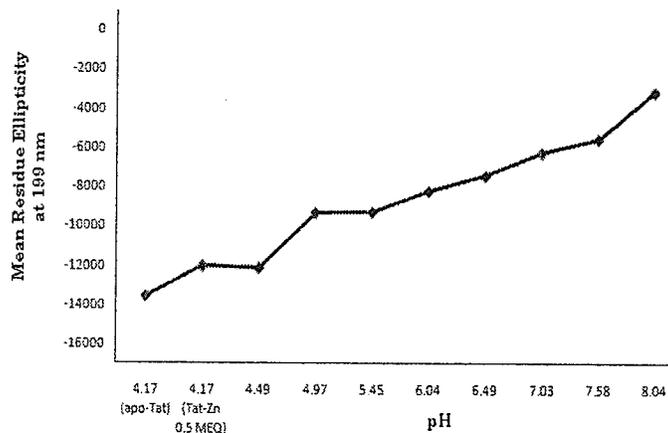
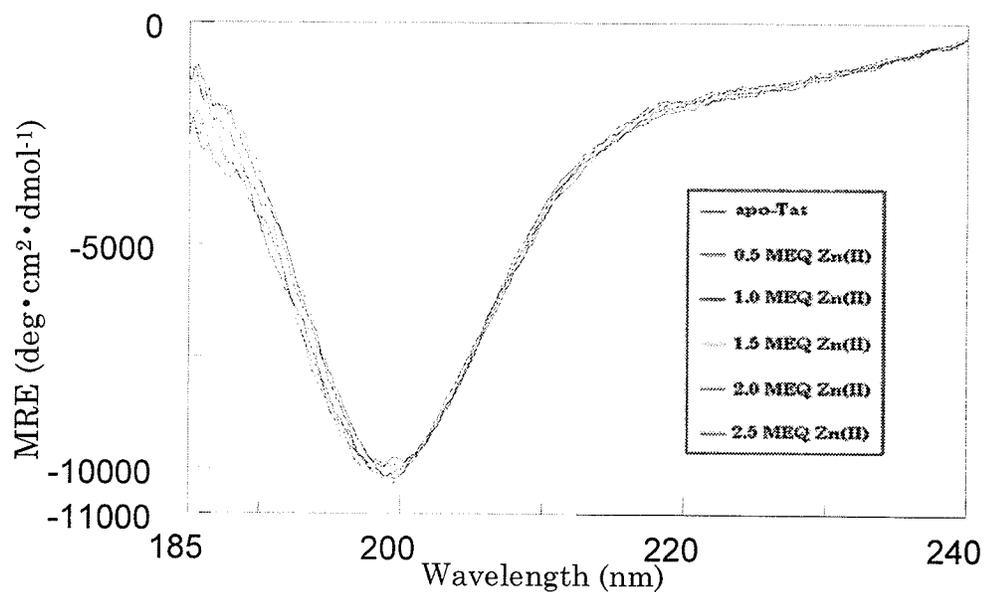


Figure 3.10: CD spectra of Tat [190  $\mu\text{M}$ ] in the absence and presence of 0.5 MEQ Zn (II) at various pH conditions corrected for dilution effects. B) Analysis of pH effect using mean residue ellipticity measured at 199 nm. All CD spectra were baseline subtracted using a blank solution containing 10 mM acetate/Tris-HCl buffer.

### *3.3.4 Zn(II) Titration of Tat at pH 4.0*

Tat's conformation in the presence of 0.5 to 2.5 mole equivalents of Zn(II) was studied by CD spectropolarimetry at pH 4.0. The CD spectra of Tat in Figure 3.11A show very small change in ellipticity at 199 nm in the presence of Zn(II) suggesting that Zn(II) has little effect on the conformation of the protein at pH 4.0. This is likely due to Tat having a low binding affinity for Zn(II) at low pH. A closer examination of the ellipticity from 0 to 0.5 mole equivalents of Zn(II) shows a change of approximately 137 deg cm<sup>2</sup>/dmole, which is substantially lower than the result of Tat in the presence of 0.5 mole equivalents of Zn(II) during the pH dependency experiment shown previously (Figure 3.10; 1439 deg·cm<sup>2</sup>/dmole). Since the same concentration of Tat was used for both experiments (190 μM), it seems likely that the change in ellipticity observed in Figure 3.10 upon addition of zinc arose from a change in pH.

A



B

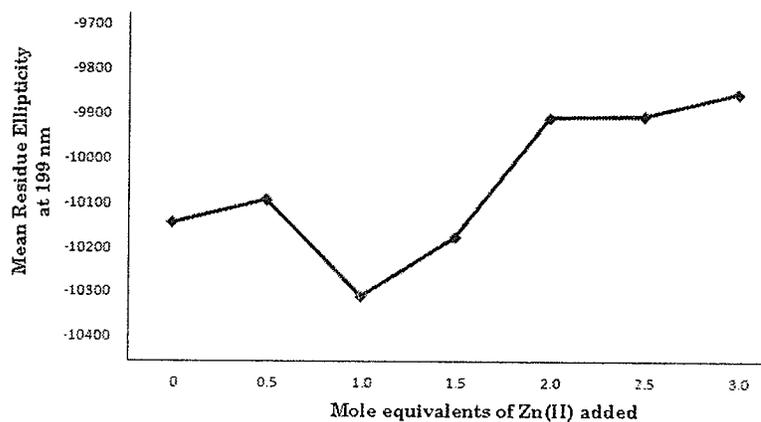
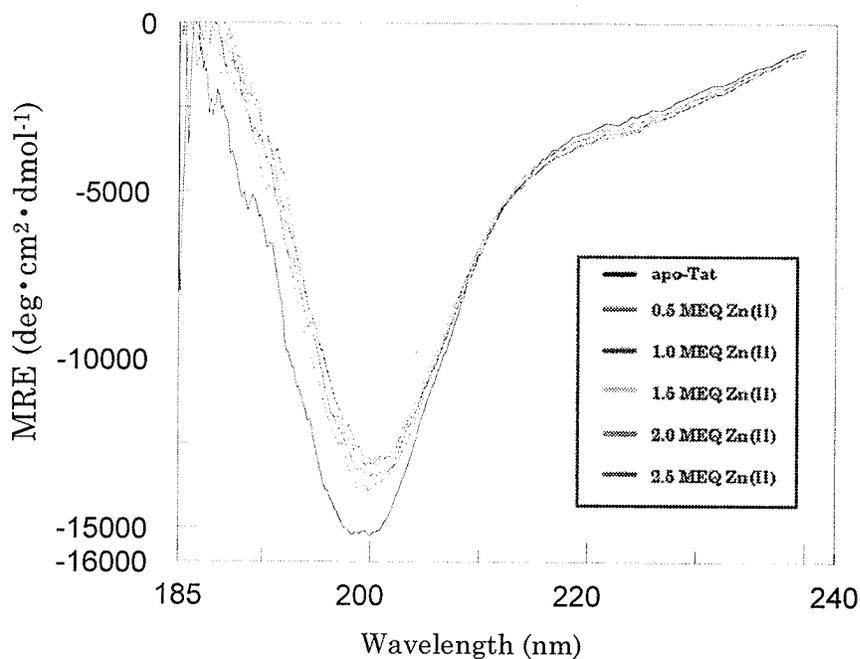


Figure 3.11: (A) CD spectra of apo-Tat (190  $\mu\text{M}$ ) in the presence of increasing concentrations of Zn(II) at pH 4.0 corrected for dilution effects. (B) Analysis of the mean residue ellipticity results at 199 nm as a function of the mole of equivalents of Zn(II) added. All CD spectra were baseline subtracted using a blank solution containing 10 mM acetate/Tris-HCl buffer.

### 3.3.5 Zn(II) Titration of Tat at pH 5.0

The effects of Zn(II) on Tat's conformation were analyzed at pH 5.0 using CD spectropolarimetry. In Figure 3.12A, the CD spectra show that in the presence of 0.5 mole equivalents of Zn(II) at pH 5.0, a significant loss in the ellipticity at 199 nm is observed but there is very little change at longer wavelengths. Further additions of Zn(II) at pH 5.0 caused smaller decreases in intensity. Figure 3.12B shows that the change in ellipticity between apo-Tat and Tat in the presence of 0.5 mole equivalent Zn(II) is approximately 1436 deg·cm<sup>2</sup>/dmole at 199 nm. The significant loss in ellipticity is followed by a plateau at higher Zn(II) concentrations where the overall change in ellipticity is found to be 213 deg·cm<sup>2</sup>/dmole at 199 nm. The significant change in Tat's conformation in the presence of 0.5 mole equivalents of Zn(II) at pH 5.0, is likely due to the change in ionization of the side chains of residues that are important to metal-binding. In particular, raising the pH from 4.0 to 5.0 may induce the deprotonation of several aspartic, glutamic, histidine or, less likely, cysteine residues. That the ellipticity loss is greater up to 0.5 mole equivalents of Zn(II) supports the idea of the formation of a metal-linked dimer.

A



B

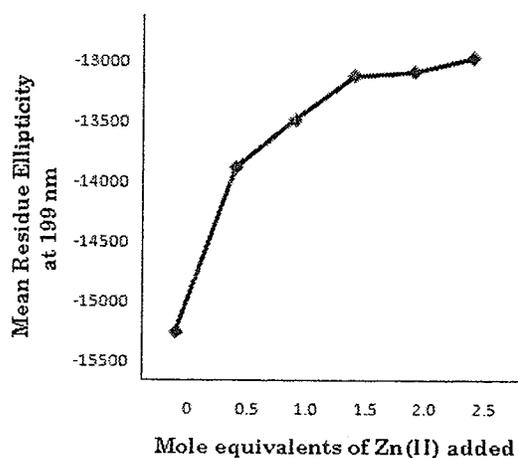
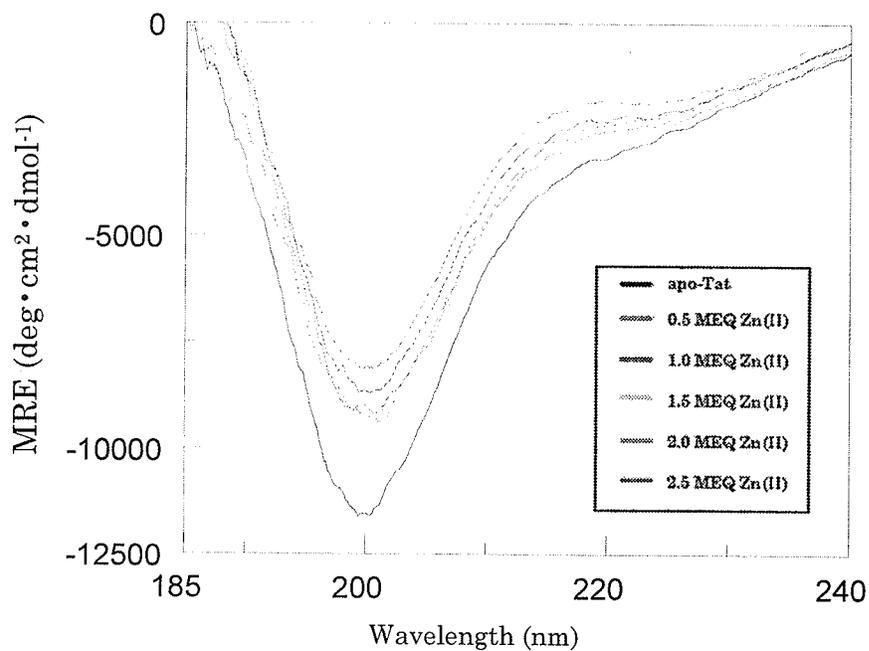


Figure 3.12: (A) CD spectra of apo-Tat (190  $\mu\text{M}$ ) in the presence of increasing concentrations of Zn(II) at pH 5.0 corrected for dilution effects. (B) Analysis of the mean residue ellipticity at 199 nm as a function of the mole of equivalents Zn(II). All CD spectra were baseline subtracted using a blank solution containing 10 mM acetate/Tris-HCl buffer.

### 3.3.6 Zn(II) Titration of Tat at pH 7.2

In order to study metal-protein interaction at physiological conditions, Zn(II) titrations of apo-Tat were conducted at pH 7.2 using CD. In Figure 3.13A, the CD results show that at 0.5 mole equivalents of Zn(II), a significant loss in ellipticity at 199 nm is observed and smaller decreases are observed at 220 nm. Smaller decreases in intensity occur at all wavelengths at higher mole equivalents of metal. Mean residue ellipticity measurements at 199 nm confirm that the greatest decrease in intensity occurred between 0 and 0.5 mole equivalents of Zn(II) with a change of  $2439 \text{ deg}\cdot\text{cm}^2/\text{dmole}$  (Figure 3.13B). Between 0.5 and 2.0 mole equivalents of Zn(II) the change in ellipticity is  $682 \text{ deg}\cdot\text{cm}^2/\text{dmole}$ . During the last Zn(II) addition another significant loss of ellipticity is observed with a change of  $1090 \text{ deg}\cdot\text{cm}^2/\text{dmole}$ . The larger effect of Zn(II) on the conformation of Tat at pH 7 compared to pH 5 is likely due to further ionization of His and Cys side-chains leading to higher affinity for the metal. The initial change in Tat's conformation followed by a gradual but smaller change at higher Zn(II) concentrations is likely due to the Tat protein containing high and low affinity binding sites and the inflection point at 0.5 mole equivalents is further support for the formation of a metal-linked dimer [39]. The spectra are also notable for their lack of an isodichroic point.

A



B

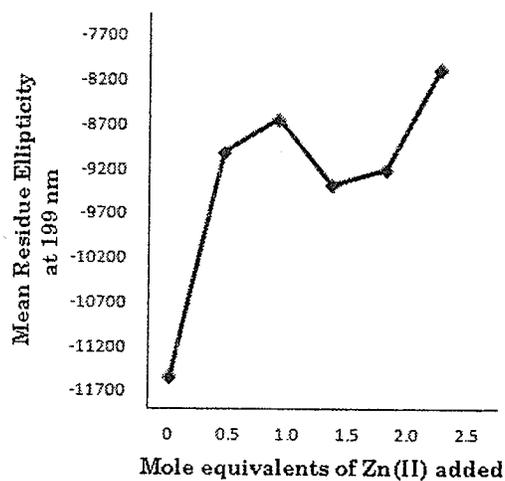


Figure 3.13: (A) CD spectra of apo-Tat (190  $\mu$ M) in the presence of increasing concentrations of Zn(II) at pH 7.2 corrected for dilution effects. (B) Analysis of the mean residue ellipticity results at 199 nm as a function of the mole of equivalents Zn(II). All CD spectra were baseline subtracted using a blank solution containing 10 mM acetate/Tris-HCl buffer.

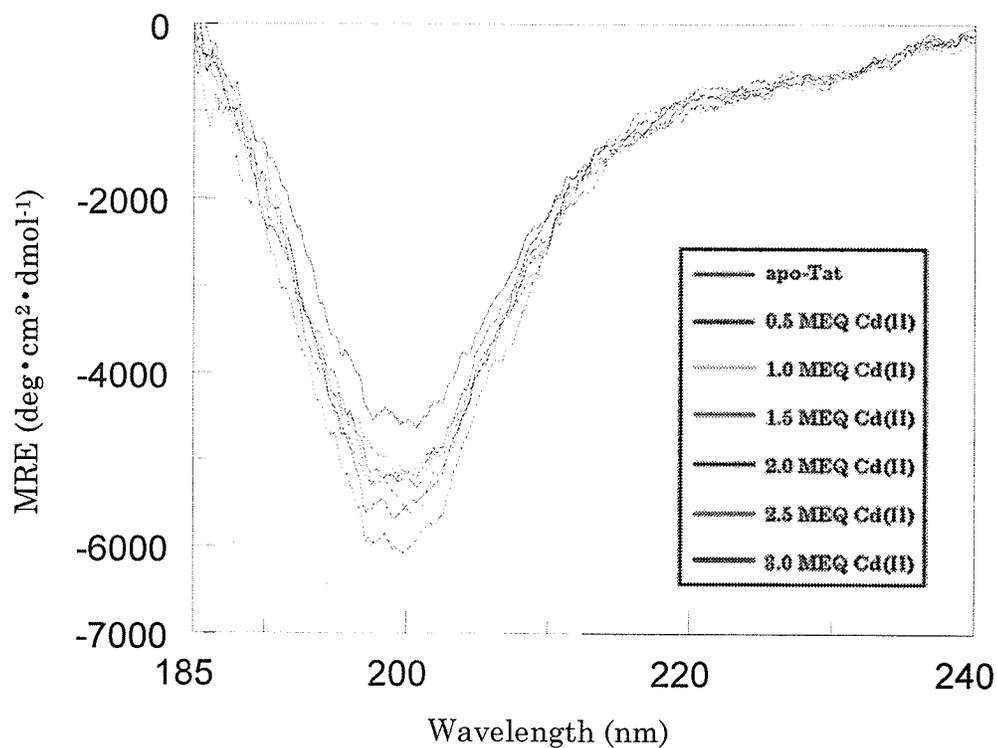
Zn(II). All CD spectra were baseline subtracted using a blank solution containing 10 mM acetate/Tris-HCl buffer.

### *3.3.7 Cd(II) Titration of Tat at pH 4.0*

Competition experiments with various divalent metals showed that the Tat protein has a higher binding-affinity for Cd(II) than Zn(II) where metal binding was measured by UV absorption spectroscopy [39]. It was shown by UV absorption spectroscopy that upon addition of Cd(II), the Tat-Zn(II) complex dissociates and a new complex, Tat-Cd(II), is formed. To explore the use of CD spectropolarimetry to monitor Cd(II) binding to Tat, a Cd(II) titration of Tat was performed. In Figure 3.14A, the CD spectra of apo-Tat and Tat at pH 3.95 in the presence of 0.5 to 3.0 mole equivalents of Cd(II) show significant losses in ellipticity at 199 nm only. The change throughout the entire Cd(II) titration is measured to be 1000 deg cm<sup>2</sup>/dmole (Figure 3.14B). Furthermore, the overall change in absorbance at 199 nm throughout the entire titration increases from 1.21 to 1.29 [data not shown] indicating that neither dilution effects nor precipitation of the Tat protein influenced the CD results. A possible explanation for the slight increase in absorption may be an indirect side effect of using Cd(II) whereby the metal may be absorbing in the wavelength range that is of interest. However, based on CD absorption measurements of Cd(II) in the absence of Tat, results show no significant increase in absorption as a function of increasing concentration of cadmium [data not shown]. The observed binding of Cd(II) to Tat at low pH confirms the higher affinity of Cd(II) for Tat compared to Zn(II) that does not bind at all (see Figure

3.11). The efficacy of Tat's ability to bind metal may be hindered due to the protonation of the histidine and cysteine residues at pH 3.95. The titration curve does not show an inflection point at 0.5 mole equivalents of metal added and thus does not confirm the formation of a metal-linked dimer.

A



B

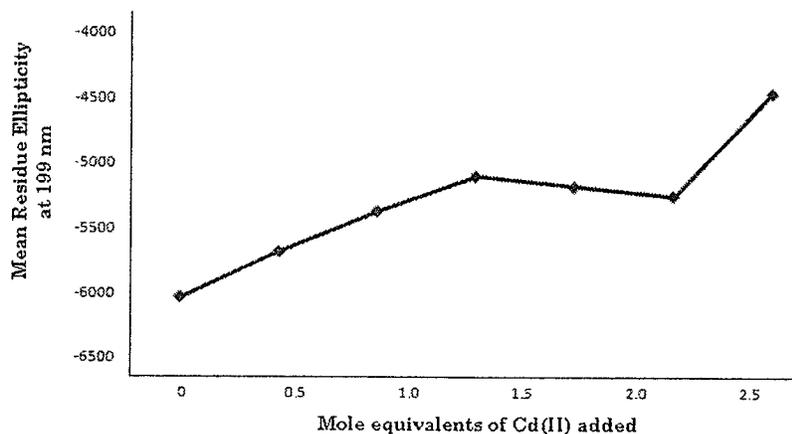
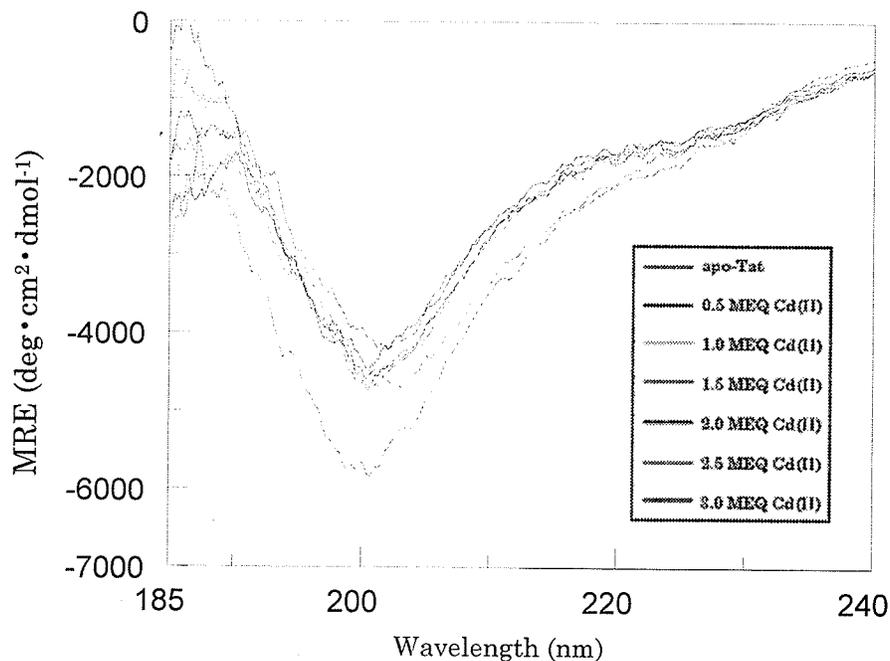


Figure 3.14: (A) CD spectra of apo-Tat (135  $\mu$ M) in the presence of increasing concentrations of Cd(II) at pH 3.95. (B) Analysis of the mean residue ellipticity at 199 nm as a function of the mole of equivalents Cd(II). All CD spectra were baseline subtracted using a blank solution containing 10 mM acetate/Tris-HCl buffer.

### 3.3.8 Cd(II) Titration of Tat at pH 5.0

Figure 3.15A shows the CD spectra of apo-Tat and Tat with increasing concentrations of Cd(II). Clearly, a significant loss of ellipticity is observed at all wavelengths upon the addition of 0.5 mole equivalents of Cd(II) as shown in Figure 3.15B (1675 deg·cm<sup>2</sup>/dmole). Further additions of Cd(II) result in much smaller changes in ellipticity. The fact that 0.5 mole equivalents of cadmium cause a much larger loss in ellipticity is further support for the suggestion of a metal-linked dimer. These CD results agree with those of the Tat-Zn(II) CD titration experiment at pH 5.0 [see Figure 3.12] and provide further evidence that as the pH rises, Tat's binding affinity increases due to changes in the ionization of the histidine and cysteine side chains. Similar to the zinc experiments, the spectra in Figure 3.15 show little evidence of an isodichroic point. Interestingly, the change in absorption throughout the entire titration increases from 1.33 to 1.40 [data not shown] suggesting that neither dilution nor precipitation influenced the CD results. Furthermore, the increase in absorption may indicate a slight contribution from Cd(II) absorption.

A



B

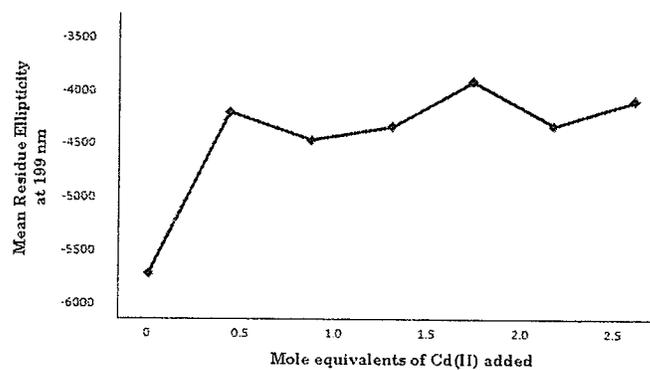


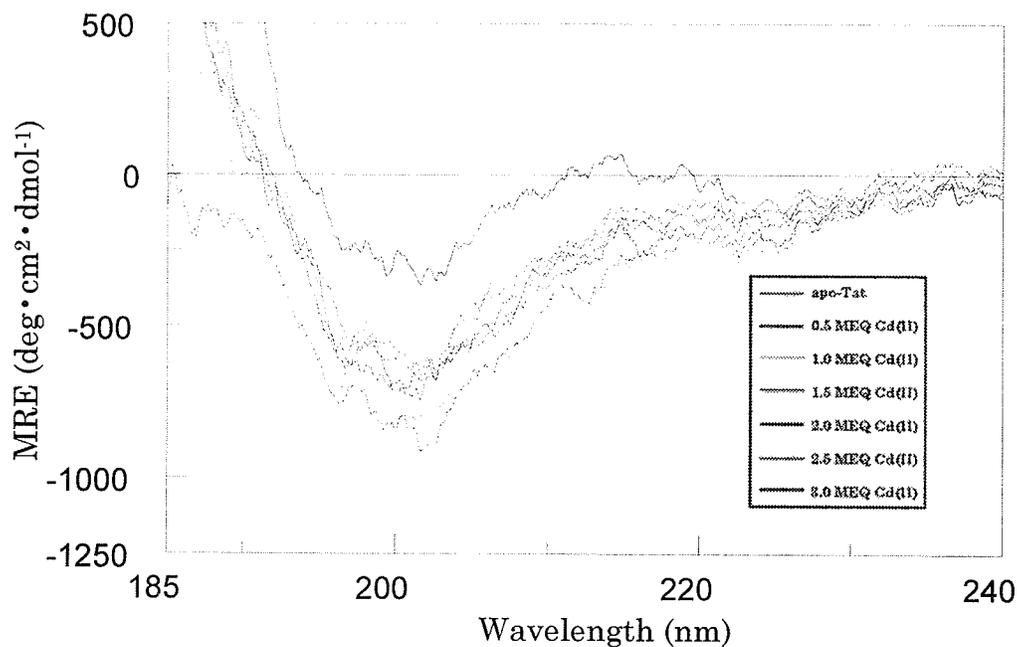
Figure 3.15: (A) CD spectra of apo-Tat [176  $\mu$ M] in the presence of increasing concentrations of Cd(II) at pH 5.08. (B) Analysis of the mean residue ellipticity results at 199 nm as a function of the mole of equivalents Cd(II). All CD spectra were baseline subtracted using a blank solution containing 10 mM acetate/Tris-HCl buffer.

### 3.3.9 Cd(II) Titration of Tat at pH 7.1

In Figure 3.16, the CD spectra of apo-Tat and Tat in the presence of increasing concentrations of Cd(II) show small losses in ellipticity at low metal concentrations followed by a larger loss in ellipticity at higher metal concentrations confirming that the binding interaction with Cd(II) induces a change in the protein's conformation. It further demonstrates that the pH of the Tat's environment plays a vital role in metal binding. One would expect that the change in ellipticity at each Cd(II) titration point would be greater than the corresponding Zn(II) titration point due to the fact that Tat is expected to have a higher binding affinity towards Cd(II) than Zn(II) [61]. However, one must take the protein concentration into account. Since, the Tat concentration of the Cd(II) experiment is substantially less than that of the Zn(II) titration experiment, it suggests that protein concentration may be a factor in metal-binding. It is noteworthy that attempts to conduct CD analysis of Tat at pH 7.0 with higher protein concentrations were unsuccessful due to precipitation of protein. This experiment was quite informative because the samples in Figure 3.16 were prepared by dissolving up Tat in buffer at pH 7.12 followed by addition of metal. The protein concentration achieved is much lower (30  $\mu$ M) than that achieved for the zinc binding experiments at pH 7.2 (190  $\mu$ M, Figure 3.13). In those experiments the Tat solution was prepared at pH 4.0 and then titrated to pH 7.2 followed by addition of metal which resulted in much higher protein concentration. Clearly, it would be worthwhile to repeat the cadmium titration by

preparing the protein as was done for Figure 3.13 which would hopefully achieve a higher protein concentration.

A



B

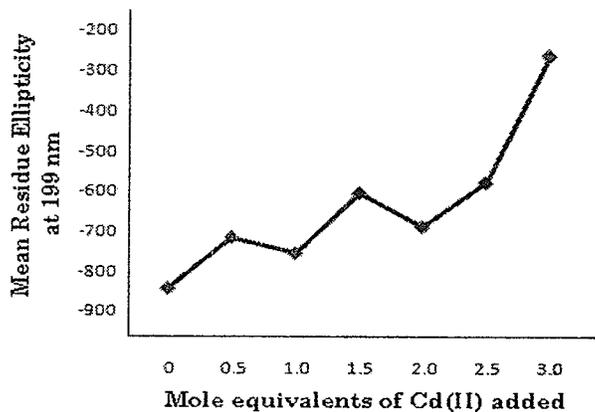


Figure 3.16: (A) CD spectra of apo-Tat (30  $\mu$ M) in the presence of increasing concentrations of Cd(II) at pH 7.12. (B) Analysis of the mean residue ellipticity results at 199 nm as a function of the mole of equivalents Cd(II). All CD spectra were baseline subtracted using a blank solution containing 10 mM acetate/Tris-HCl buffer.

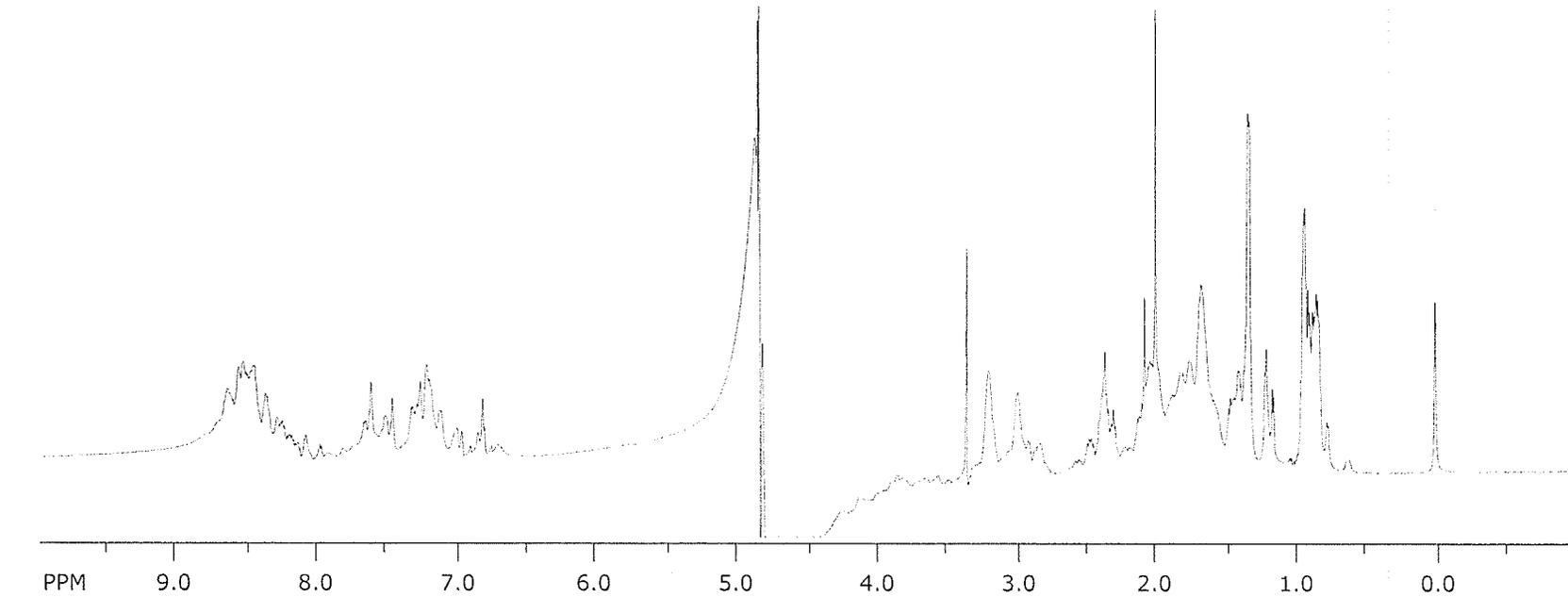
### 3.4 Nuclear Magnetic Resonance Spectroscopy (NMR)

#### 3.4.1 Natural Abundance 2D $^1\text{H}$ - $^{15}\text{N}$ HSQC and 1D $^1\text{H}$ NMR of Unlabeled HIV-1 Tat

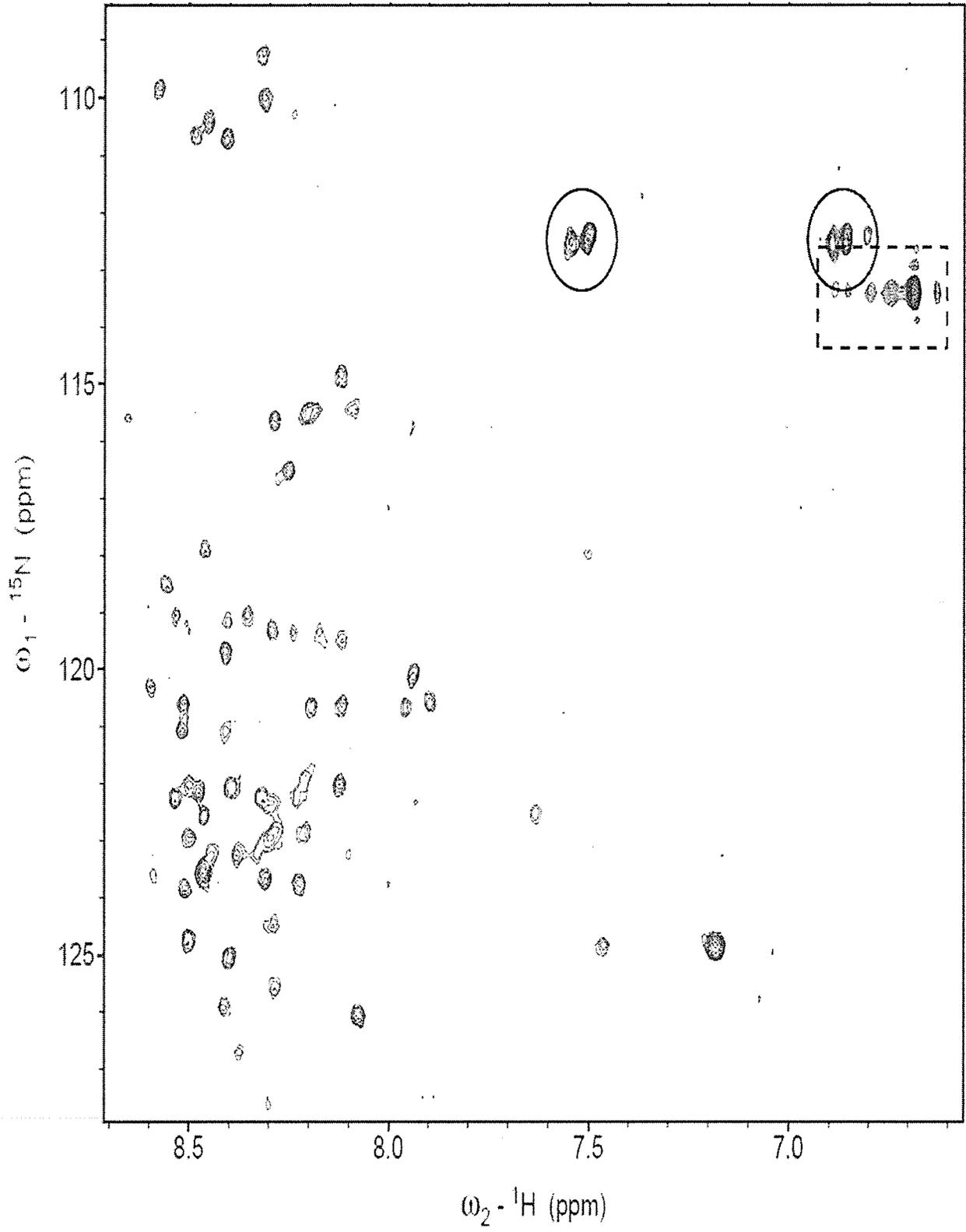
Figure 3.17A shows a 1D  $^1\text{H}$  NMR spectrum of unlabeled Tat protein with the proton chemical shifts of the backbone NH resonances observed in the range of 6.5 to 9.0 ppm relative to the internal standard, DSS at 0 ppm. The distortion in the centre of the spectrum at 5 ppm arises from the water resonance; the aliphatic protons resonate in the range of 0-5 ppm.

Overcrowding of the peaks in the presence of a large water signal makes assignment of the 1D  $^1\text{H}$  NMR spectrum of Tat protein very difficult. In Figure 3.17B, the 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum of unlabeled Tat protein shows 62 cross-peaks out of a possible 91 backbone amides accompanied by the  $\text{NH}_2$  resonances of the side chains of Gln and Asn. The close similarity of the spectrum to the published assigned spectrum of HIV-1 Tat [137] confirms the identity of the protein. Previous published NMR data of HIV-1 Tat showed that the protein's conformation exists in a disordered state through comparison of the observed chemical shifts to standardized random coil chemical shifts with corrections for local sequence effects [138, 141]. The narrow dispersion of the backbone amide resonances in Figure 3.17B confirms that the protein is substantially disordered at pH 3.5.

Figure 3.17: (A) One-dimensional  $^1\text{H}$  NMR and (B) Two-dimensional natural abundance  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of apo-Tat (300  $\mu\text{M}$ ) in NMR buffer at pH 3.5 at 293 K. Side chain  $\text{NH}_2$  cross-peaks corresponding to Gln and Asn are outlined in solid circles. Cross-peaks found in the dashed rectangle correspond to impurities. Both spectra were recorded on a Varian INOVA 600 MHz spectrometer equipped with a triple resonance probehead. Acquisition time in (A) was 34 seconds and in (B) was 24 hrs.



B



### *3.4.2 2D $^1\text{H}$ - $^{15}\text{N}$ HSQC NMR on $^{15}\text{N}$ -labeled HIV-1 Tat and Sequence Assignments*

A more sensitive and less time consuming approach to acquiring NMR data for protein structure determination is the use of isotopic labeling of proteins. In Figure 3.18, the  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum of HIV-1 Tat shows the amide cross-peaks of residues together with their assignments. The assignments were made by comparing the positions of the observed peaks to the resonance assignments published by Dr. Shojania [141] for the protein in similar conditions of temperature and pH. Excellent agreement was observed between the two sets of resonances.

Figure 3.18:  $^1\text{H}$ - $^{15}\text{N}$  HSQC NMR spectrum of isotopically-labeled HIV-1 Tat protein (635  $\mu\text{M}$ ) acquired on a Varian INOVA 600 MHz spectrometer equipped with a triple resonance probehead at pH 4.0 at 293 K. Acquisition time was 2 hrs. Backbone amide cross-peaks were assigned using previously published NMR data of HIV-1 Tat [137].



### 3.4.3 2D $^1\text{H}$ - $^{15}\text{N}$ HSQC NMR on $^{15}\text{N}$ -Labeled Tat in the Absence and Presence of Zn(II)

#### 3.4.3.1 Addition of 0.5 Mole Equivalents of Zn(II)

One of the advantages of solution NMR spectroscopy over absorption and circular dichroism spectroscopies is the ability to study atomic level structural perturbations in proteins in the presence of cellular co-factors such as metals. Figure 3.19 shows  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of  $^{15}\text{N}$ -labelled HIV-1 Tat in the absence (red) and presence (blue) of 0.5 mole equivalents of Zn(II) at pH 4.02. Addition of zinc causes significant line broadening and chemical shift changes in the spectrum. Figures 3.20 and 3.21 show the changes in amide cross-peak intensity and chemical shift, respectively, that is induced by the addition of Zn(II). Careful inspection of the resonance intensities shows that the majority of the residues that lose significant intensity upon addition of Zn(II) are found within the Cys-rich region (Figure 3.20). Cross-peaks for residues Cys-42, Asp-44, Cys-45, Cys-47, Cys-51, Cys-57 and Lys-70 are observed to disappear from the spectrum in the presence of 0.5 mole equivalents of Zn(II). Furthermore, the decrease in intensity of these residues relative to other cross-peaks suggests that the changes in the Cys-rich region are in addition to the line broadening observed throughout the protein and likely due to dynamic exchange between Zn(II) bound and unbound Tat protein. This supports the suggestion that the cysteine residues are in fact directly binding to Zn(II). Furthermore, cross-peaks corresponding to residues Ala-1, Ala-41, Thr-40, Ala-62, Lys-71, and Lys-91, are observed to decrease in intensity. It is important to note

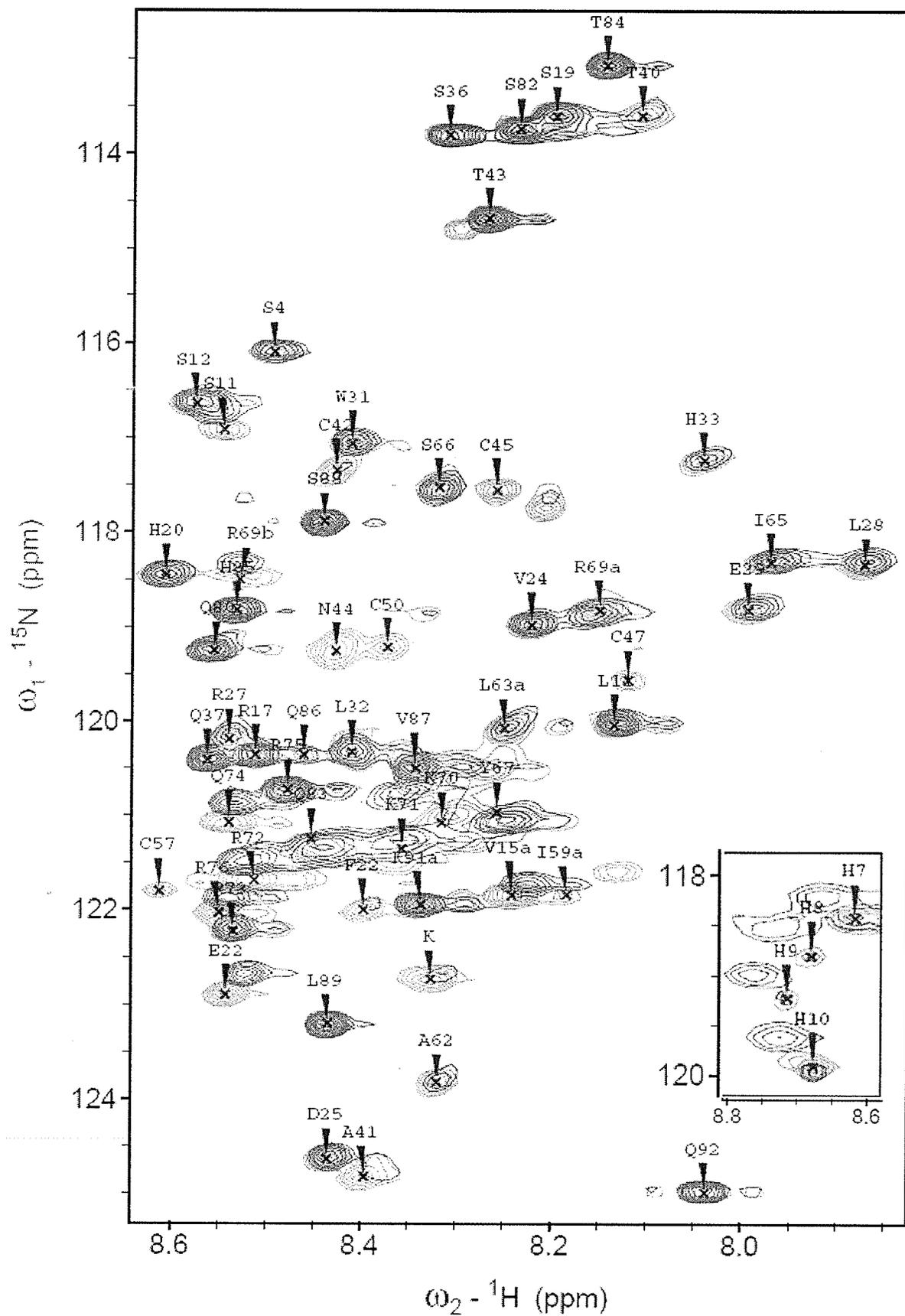
that the chemical shifts corresponding to the asparagine and glutamine NH<sub>2</sub> side chains remain unchanged for both apo-Tat and the Tat-Zn complex spectra [data not shown]. However, significant line broadening is observed of all the backbone amide cross-peaks for the Tat protein in the presence of 0.5 mole equivalents of Zn(II) suggesting that the protein backbone is undergoing intermediate exchange between unbound and bound metal-protein conformations. Quantitative analysis of the line broadening effects on a few arbitrary cross-peaks shows that the intensities decrease on average by 13.4% in the presence of 0.5 mole equivalents of Zn(II) (data not shown).

Analysis of the chemical shift changes in the direct dimension (<sup>1</sup>H) in the presence of Zn(II) in Figure 3.21A shows changes for residues His-7, His-8, His-9, His10, Glu-22, Ala-41, and Phe-52. Figure 3.21B shows that the Histidine tag residues His-7, His-8, His-9, and His10 and Glu-22 undergo chemical shift changes in the nitrogen dimension. Furthermore, out of the 11 residues within the Arg-rich motif (Arg-69, Lys-71, Arg-72, Arg-73, Gln-74, and Arg-76) are observed to undergo chemical shift changes suggesting that this region either changes conformation [147] due to Zn(II) binding or is directly involved in the binding [148, 149]. Note that the cysteine residues within the Cys-rich region do not appear to undergo any chemical shift changes because the resonances disappear from the spectrum. It is also worth noting that the chemical shift changes are relatively small. That is, most of the resonances move considerably less than 0.2 ppm for <sup>1</sup>H and 2 ppm for <sup>15</sup>N,

which are the ranges (2 standard deviations) for random coil chemical shifts in the NMR Biological Magnetic Resonance data bank.

More specifically, the average chemical shift changes of the histidine residues along the  $^1\text{H}$  and  $^{15}\text{N}$  dimension in the presence of Zn(II) is approximately -0.03 and 0.24 ppm, respectively, which is well below two standard deviations ( $2\sigma = 1.36$  ( $^1\text{H}$ ); 8.2 ppm ( $^{15}\text{N}$ )) from Histidine backbone data from the chemical shift tables in the Biological Magnetic Resonance Bank data base [157]. This suggests that the Zn(II) interaction with the histidine residues within the His-tag region is weak.

Figure 3.19: 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of apo-Tat (red) and Tat in the presence of 0.5 mole equivalents of Zn(II) (blue) at pH 4.02 at 293 K. The Tat protein concentration in both spectra is 517  $\mu\text{M}$ . Spectral data were recorded on a Varian INOVA 600 MHz NMR spectrometer equipped with a triple resonance probehead. The acquisition time of each spectrum was 2 hrs and 40 min. Cross-peaks corresponding to the His-tag are shown in the inset box.



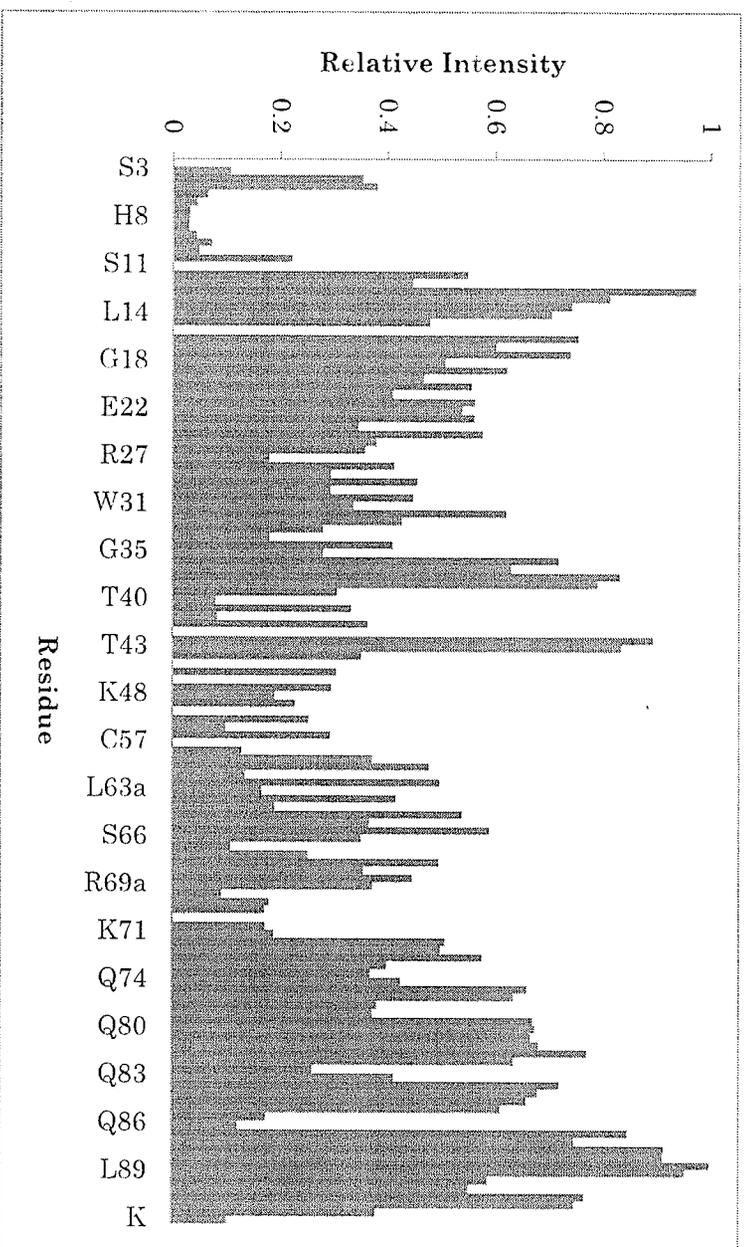
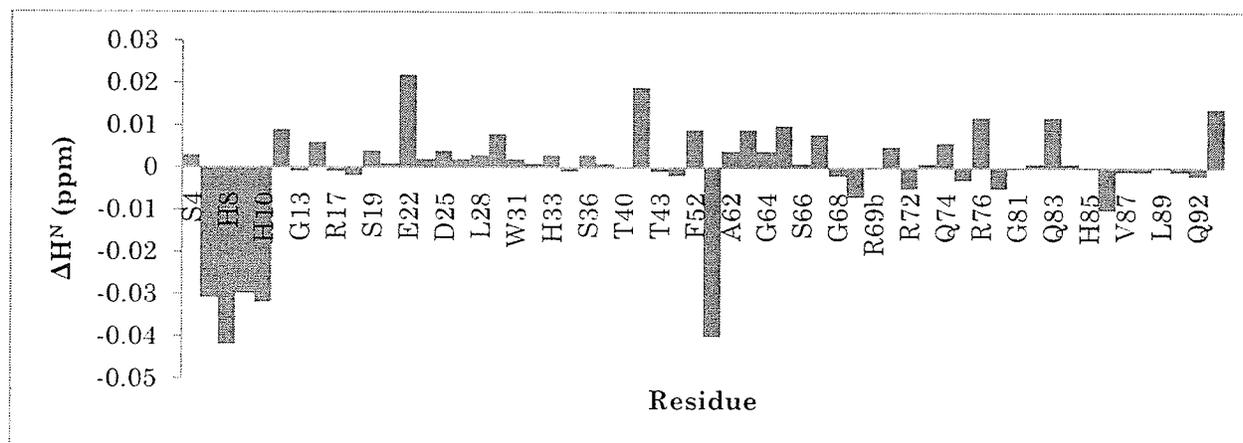


Figure 3.20:  $^1\text{H}$ - $^{15}\text{N}$  HSQC cross peak intensities for residues of apo-Tat (red) and Tat in the presence of 0.5 mole equivalents of Zn(II) (blue).

A



B

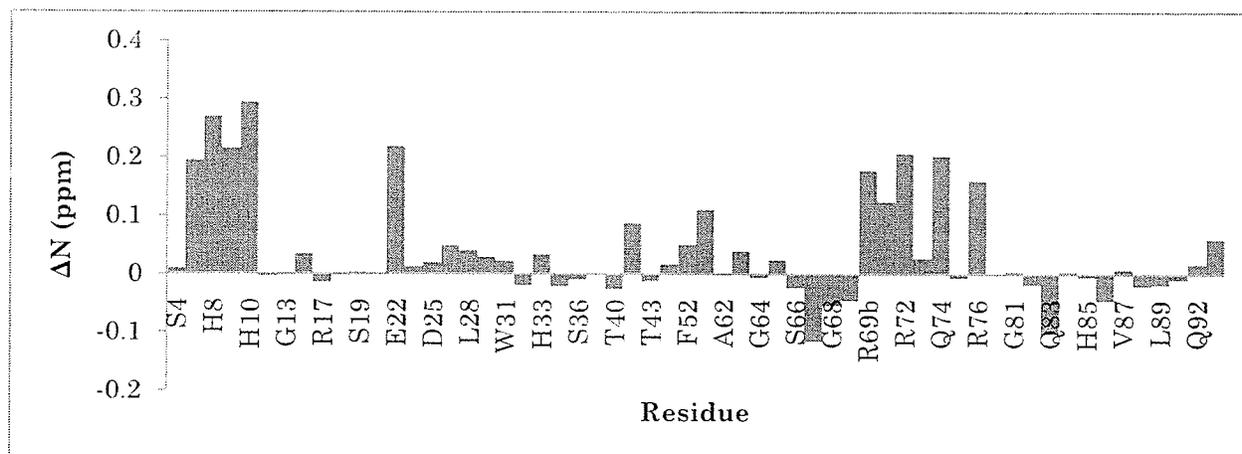
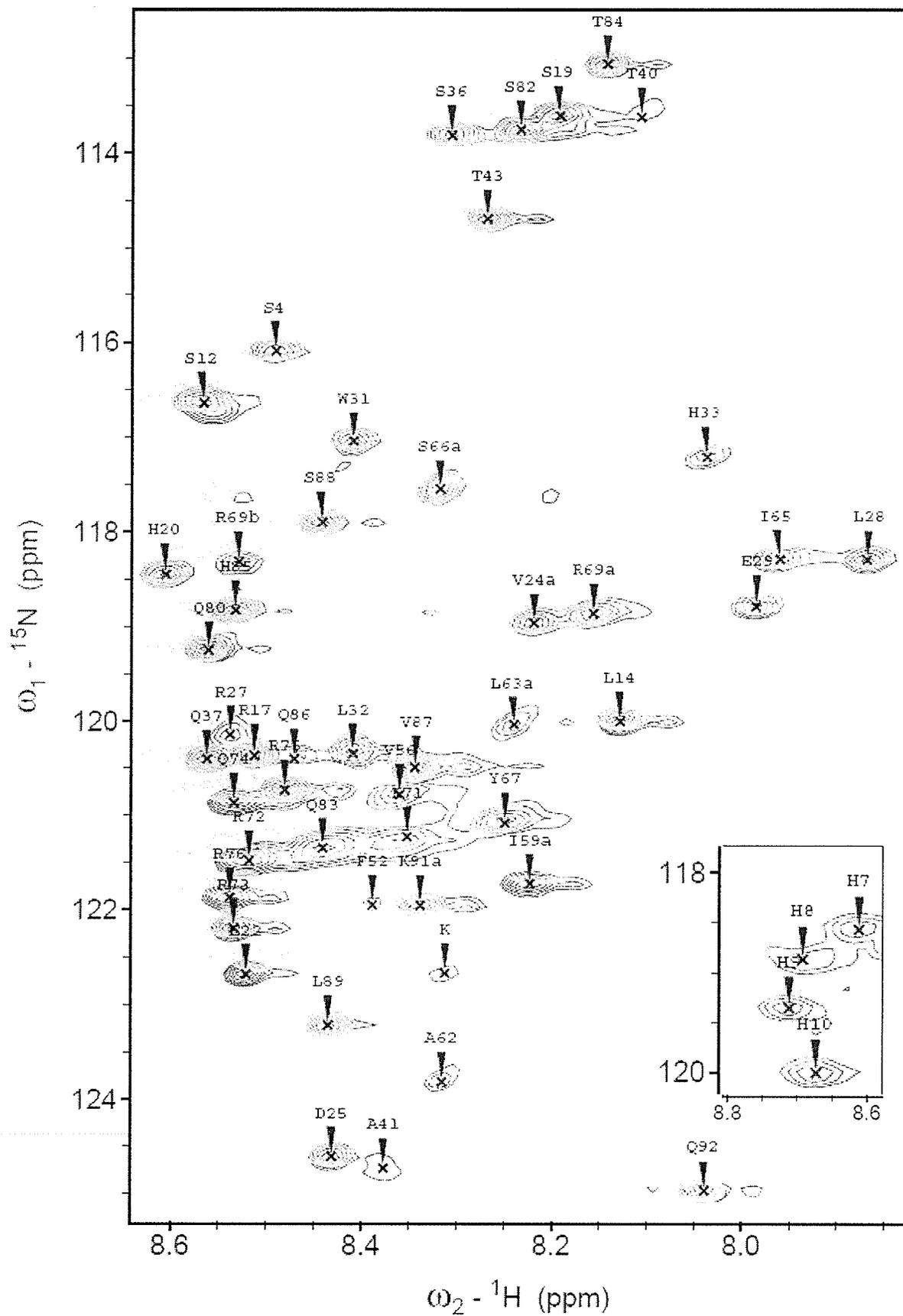


Figure 3.21: Chemical shift difference plots for (A)  $^1\text{H}^{\text{N}}$  and (B)  $^{15}\text{N}$ . The difference plots were obtained by subtracting the chemical shifts obtained from the  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of apo-Tat from Tat in the presence of 0.5 mole equivalents of Zn(II).

### 3.4.3.2 Addition of 1.0 Mole Equivalent of Zn(II)

Figure 3.22 shows the  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of Tat in the presence of 0.5 (blue) and 1.0 (green) mole equivalents of Zn(II) whereby chemical shift changes and the increase in intensity for residues within the His-tag region are observed. In particular, intensities from residues His-7, His-8, His-9, and His-10 are greater in the  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum of Tat in the presence of 1.0 mole equivalent Zn(II) than in the 0.5 mole equivalent spectrum suggesting that this region may be involved in metal-binding. Figure 3.23 shows the changes in HSQC resonance intensity induced by the addition of 0.5 and 1.0 mole equivalents of Zn(II) in which further line broadening of the resonances is observed. Figure 3.24 shows the  $^1\text{H}$  and  $^{15}\text{N}$  chemical shift differences between Tat containing 0.5 and 1.0 mole equivalents of Zn(II). Most of the residues that are observed to undergo significant changes are found within the Arg-rich region suggesting further conformational changes in the basic domain (residues Arg-69, Lys-71, Arg-72, Arg-73, Glu-74, Arg-75, and Arg-76) of the protein upon binding of Zn(II). Figure 3.24A shows chemical shift changes along the proton dimension for residues His-7, His-8, His-9, His-10, Glu-22, Ile-59, leu-63, and lys-71. For the indirect dimension ( $^{15}\text{N}$ ), chemical shift changes (Figure 3.24B) are observed for residues His-7, His-8, His-9, His-10, Lys-14, Glu-22, Lys-32, Ile-59, Ala-62, Lys-3, Arg-72, Gln-74, and Arg-76.

Figure 3.22: 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of Tat in the presence of 0.5 (blue) and 1.0 (green) mole equivalents of Zn(II) at pH 4.02 at 293 K. Tat protein concentration in both spectra is 517  $\mu\text{M}$ . Spectral data were recorded on a Varian INOVA 600 MHz spectrometer equipped with a triple resonance probehead. The acquisition time of each spectrum was 2 hrs and 40 min. Cross-peaks corresponding to the His-tag are shown in the inset box.



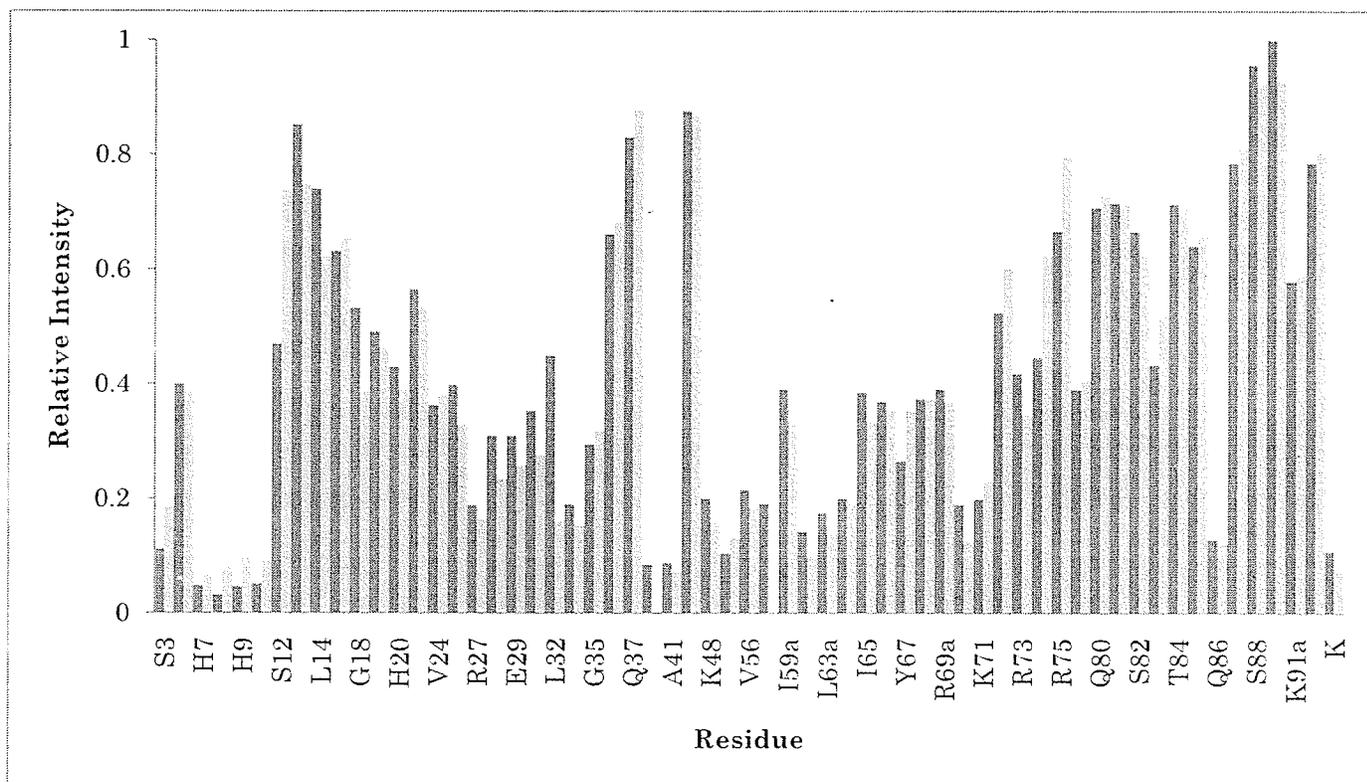
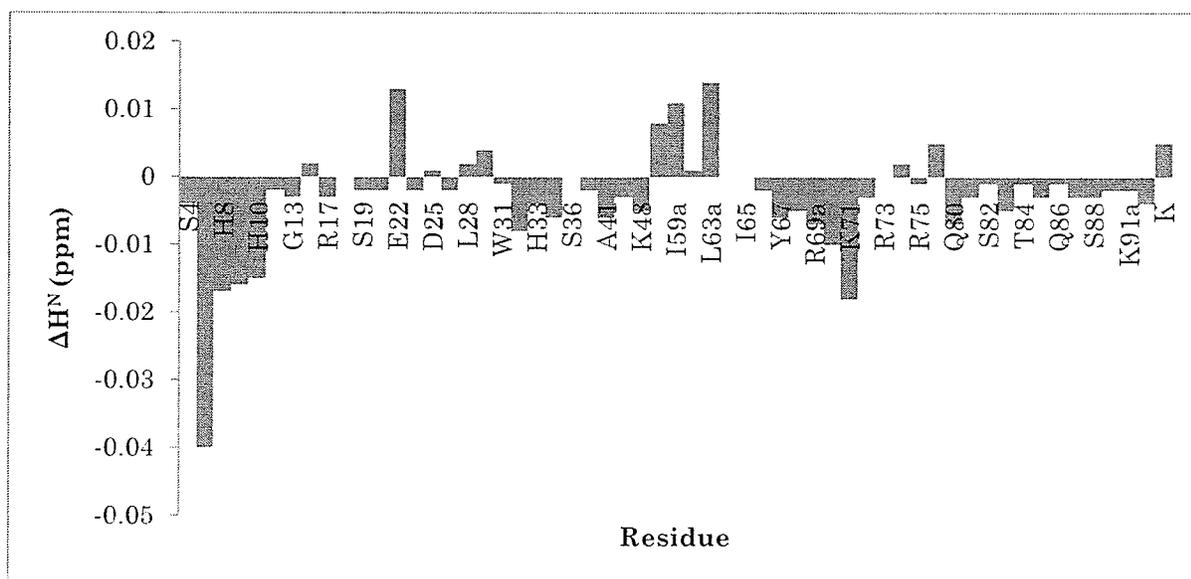


Figure 3.23:  $^1\text{H}$ - $^{15}\text{N}$  HSQC cross peak intensities for residues of Tat in the presence of 0.5 (blue) and 1.0 (green) mole equivalents of Zn(II).

A



B

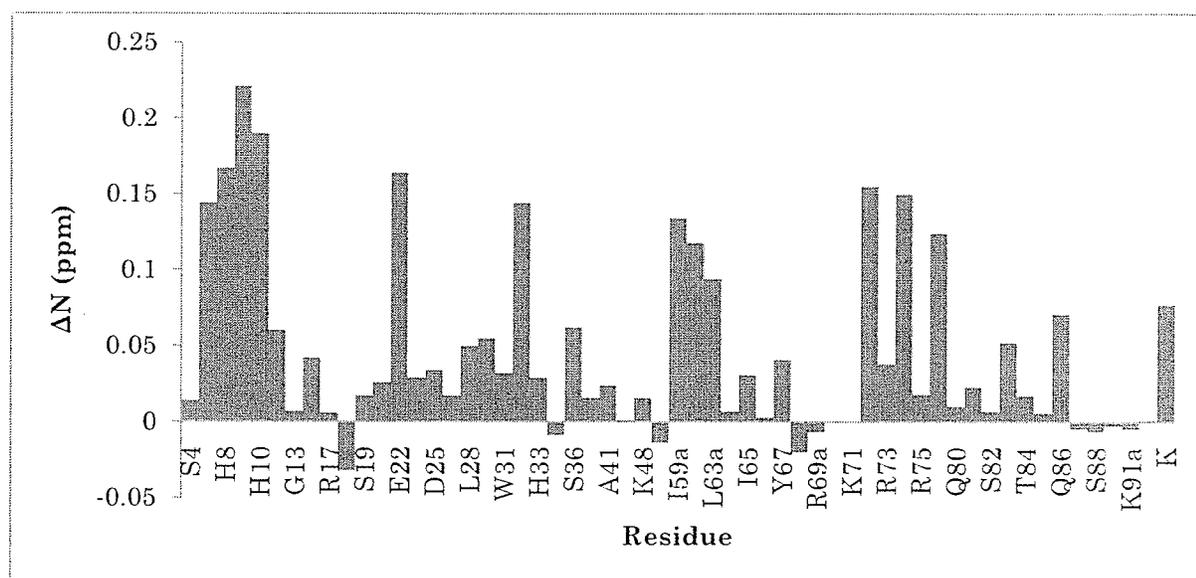
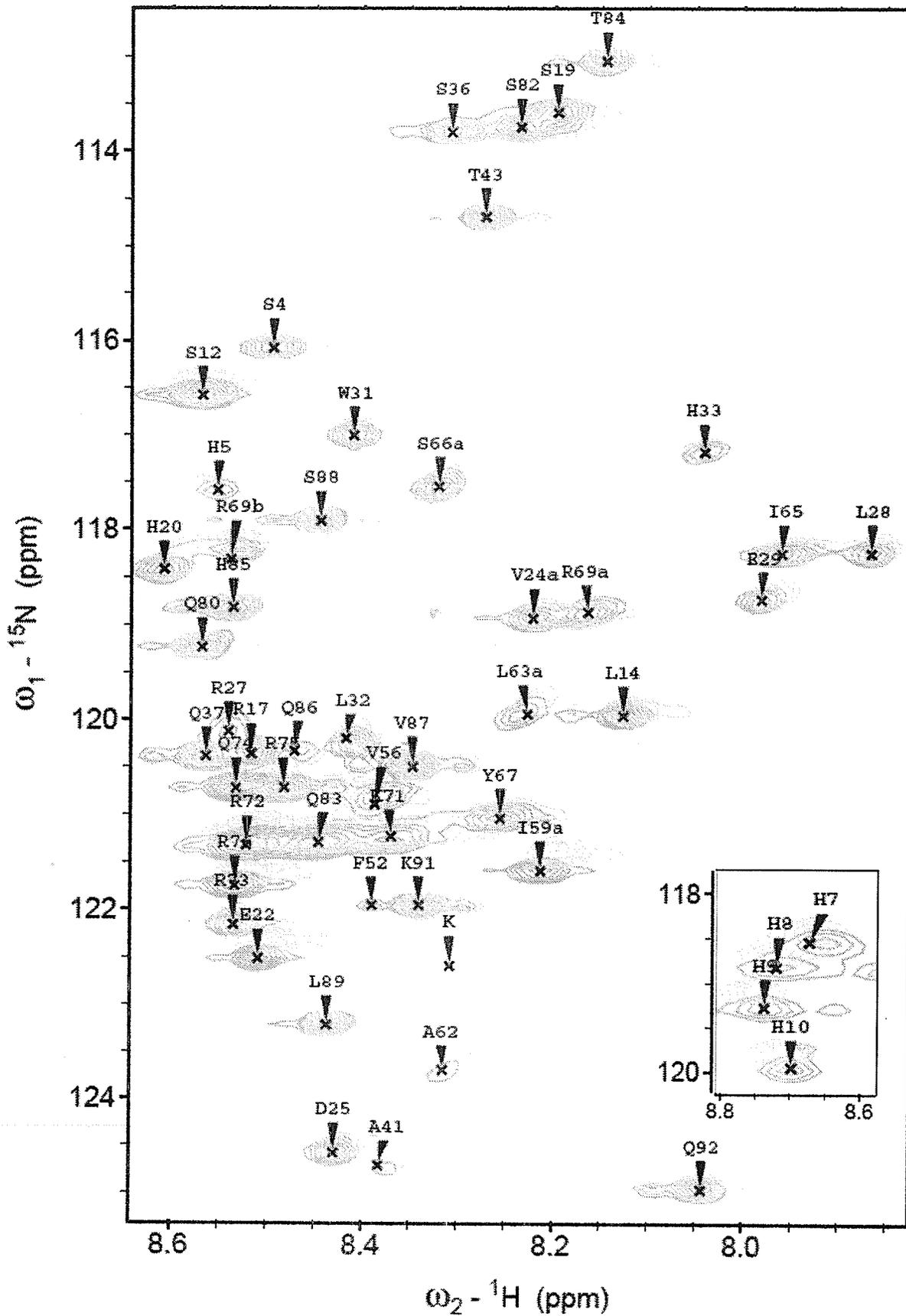


Figure 3.24: Chemical shift difference plots for (A)  $^1H^N$  and (B)  $^{15}N$ . The difference plots were obtained by subtracting the chemical shifts obtained from the  $^1H$ - $^{15}N$  HSQC spectra of Tat in the presence of 1.0 mole equivalents of Zn(II) from Tat in the presence of 0.5 mole equivalents of Zn(II).

#### 3.4.3.3 Addition of 2.0 Mole Equivalents of Zn(II)

Figure 3.25 shows the  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of Tat in the presence of 1.0 (green) and 2.0 (yellow) mole equivalents of Zn(II). Small chemical shift changes are observed upon addition of a second mole equivalent of zinc for residues located within the Arg-rich and His-tag regions. Residues His-5, His-7, His-8, His-9, and His-10 are observed to undergo chemical shift changes suggestive of Zn(II) binding. Furthermore, residues within the Arg-rich motif: Arg-69, Arg-72, Arg-73, Arg-74, and Arg-75 are also observed to undergo significant chemical shift changes. Interestingly, the negatively charged residues (found within the N-terminus) Glu-22, Asp-25, and Glu-29 also undergo chemical shift changes at each Zn(II) addition. This might indicate a direct interaction with Zn(II) but could also indicate that the metal-binding is stabilized through hydrogen-bond formation of the imidazole nitrogen of the histidine to the oxygen of the carboxyl groups of Asp and Glu [85]. Figures 3.26 and 3.27 show the changes in resonance intensity and chemical shift, respectively, induced by the addition of 1.0 and 2.0 mole equivalents of Zn(II). Careful inspection of the residue intensity comparison between 1.0 and 2.0 mole equivalents of Zn(II) shows a slight decrease in nearly all residues in the latter due to line broadening, which is a direct effect of metal binding. Similar to the analysis of the comparison between residue intensities between 0-0.5 and 0.5-1.0 mole equivalents of Zn(II), Figure 3.27 shows residues corresponding to the His-tag and the Arg-rich regions undergo changes between 1.0 to 2.0 mole equivalents of Zn(II) in both the indirect ( $^{15}\text{N}$ ) and direct ( $^1\text{H}$ ) dimensions.

Figure 3.25: 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of Tat in the presence of 1.0 (green) and 2.0 (yellow) mole equivalents of Zn(II) at pH 4.02 at 293 K. Tat protein concentration in both spectra is 517  $\mu\text{M}$ . Spectral data were recorded on a Varian INOVA 600 MHz NMR spectrometer equipped with a triple resonance probehead. The acquisition time of each spectrum was 2 hrs and 40 min. Cross-peaks corresponding to the His-tag are shown in the inset box.



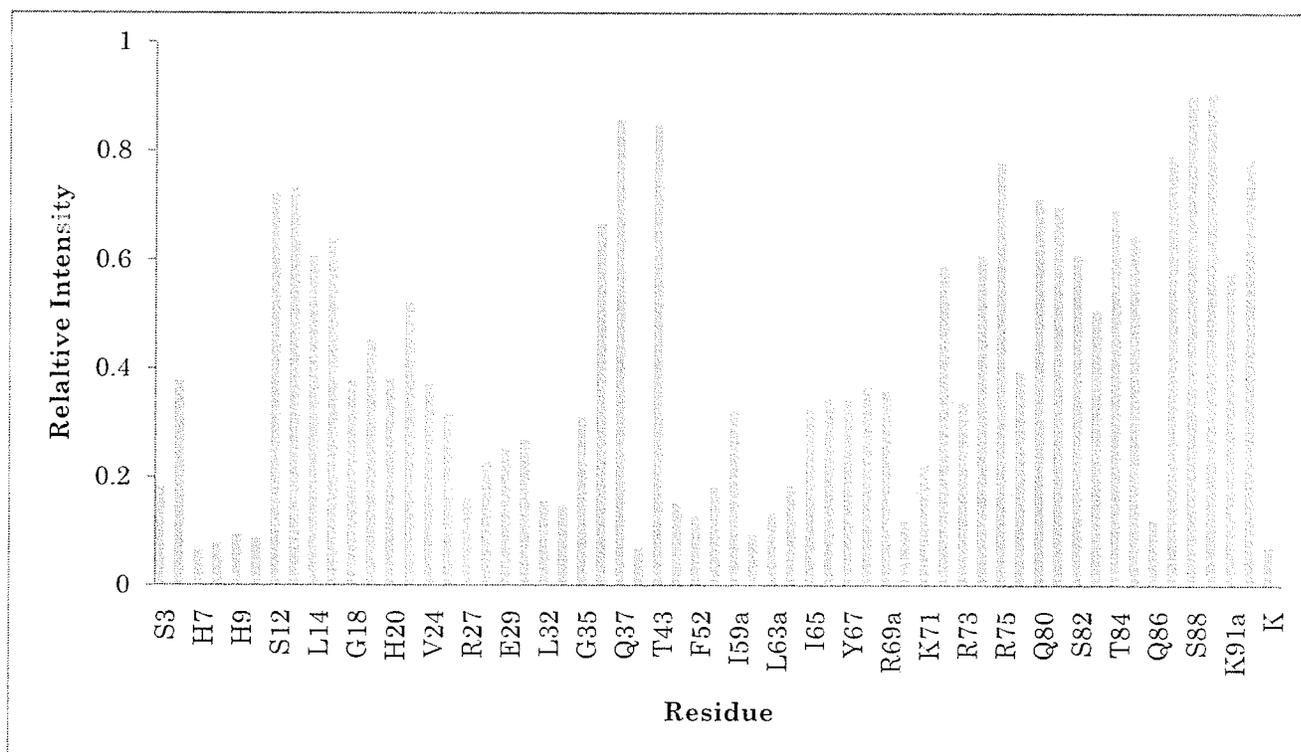
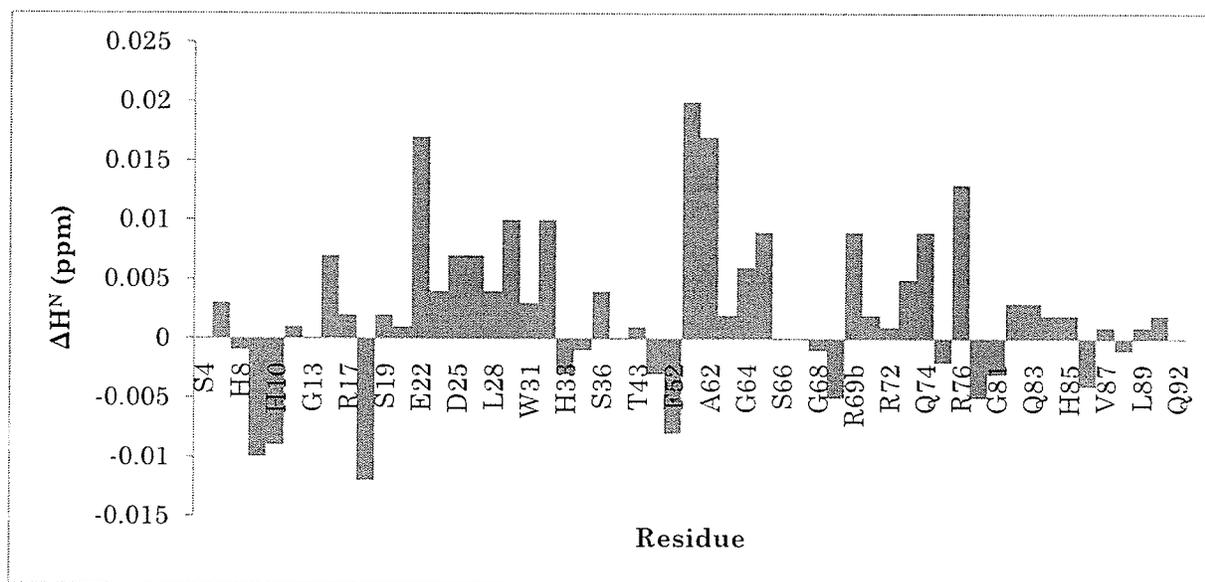


Figure 3.26:  $^1\text{H}$ - $^{15}\text{N}$  HSQC cross peak intensities for residues of Tat in the presence of 1.0 (green) and 2.0 (yellow) mole equivalents of Zn(II).

A



B

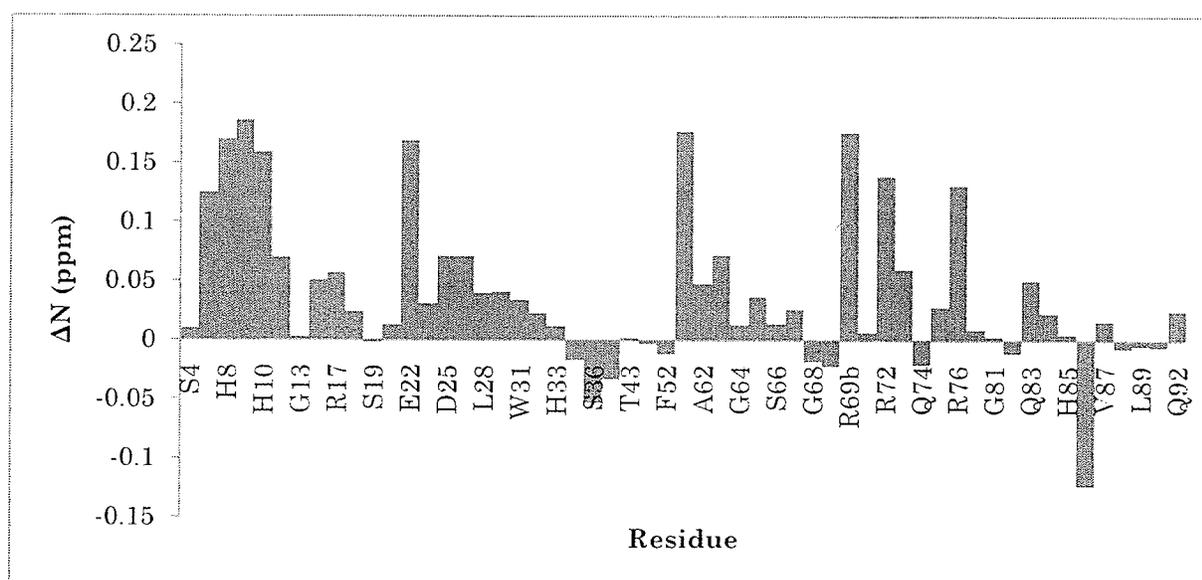


Figure 3.27: Chemical shift difference plots for (A)  $^1\text{H}^N$  and (B)  $^{15}\text{N}$ . The difference plots were obtained by subtracting the chemical shifts obtained from the HSQC spectra of Tat in the presence of 1.0 mole equivalents of Zn(II) from Tat in the presence of 2.0 mole equivalents of Zn(II).

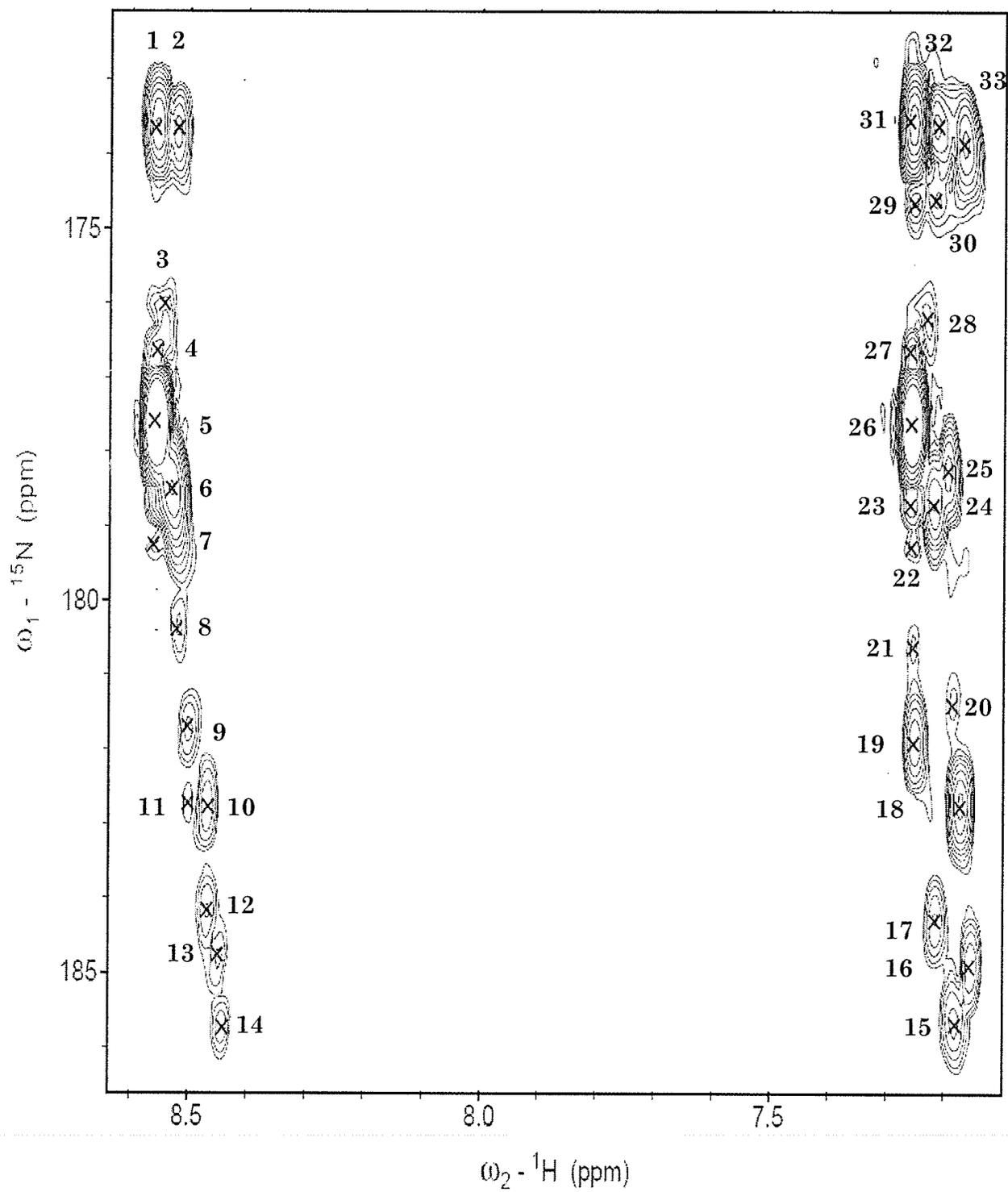
#### 3.4.4 $^1\text{H}$ - $^{15}\text{N}$ HMBC Spectra of Apo-Tat and the Tat-Zn Complex

A direct interaction between the histidine imidazole side-chains and Zn(II) is somewhat surprising at pH 4.02 since the  $\text{pK}_a$  of His is 5.86 and the side-chains will be substantially protonated at pH 4.02. To further investigate the contribution of the 10 histidine residues to Zn(II) binding,  $^1\text{H}$ - $^{15}\text{N}$  HMBC experiments of HIV-1 Tat protein in the presence of increasing Zn(II) concentrations were carried out. In particular, the  $^1\text{H}$ - $^{15}\text{N}$  HMBC experiment allows the detection, *via* the heteronuclear two-bond J-coupling, of the cross-peaks between the  $\text{N}_{\epsilon 2}$  and  $\text{N}_{\delta 1}$  atoms and the  $\text{H}_{\beta 2}$  and  $\text{H}_{\epsilon 1}$  atoms of the imidazole ring. Figure 3.28A shows the  $^1\text{H}$ - $^{15}\text{N}$  HMBC spectrum of apo-Tat at pH 4.0 with the observation of 33 cross-peaks in 2 narrow regions in the proton dimension: 8.56-8.44 ppm and 7.25-7.15 ppm [Appendix A]. Furthermore, the  $^{15}\text{N}$  chemical shifts are observed to be dispersed in the region of 186.01 to 173.61 ppm. Figure 3.28B shows the  $^1\text{H}$ - $^{15}\text{N}$  HMBC spectral overlay of apo-Tat and Tat in the presence of increasing Zn(II) concentrations: 0 (red), 0.5 (blue), 1.0 (green), and 2.0 (yellow) mole equivalents. It is noteworthy that the addition of 0.5 mole equivalents of Zn(II) causes the disappearance of 15 cross-peaks suggestive of a direct interaction of Zn(II) with the His-tag. Comparison of the  $^1\text{H}$ - $^{15}\text{N}$  HMBC spectra of Tat with 0.5 and 1.0 molar equivalents of Zn(II) shows a further decrease in the number of cross-peaks from 18 to 16 and the movement of peaks downfield in the  $^{15}\text{N}$  dimension. Finally, addition of 2.0 mole equivalents of Zn(II) to the Tat protein results in a spectrum with only 12 cross-peaks with  $^{15}\text{N}$  chemical shifts observed in the narrow region of 179.07-173.52 ppm. Based on

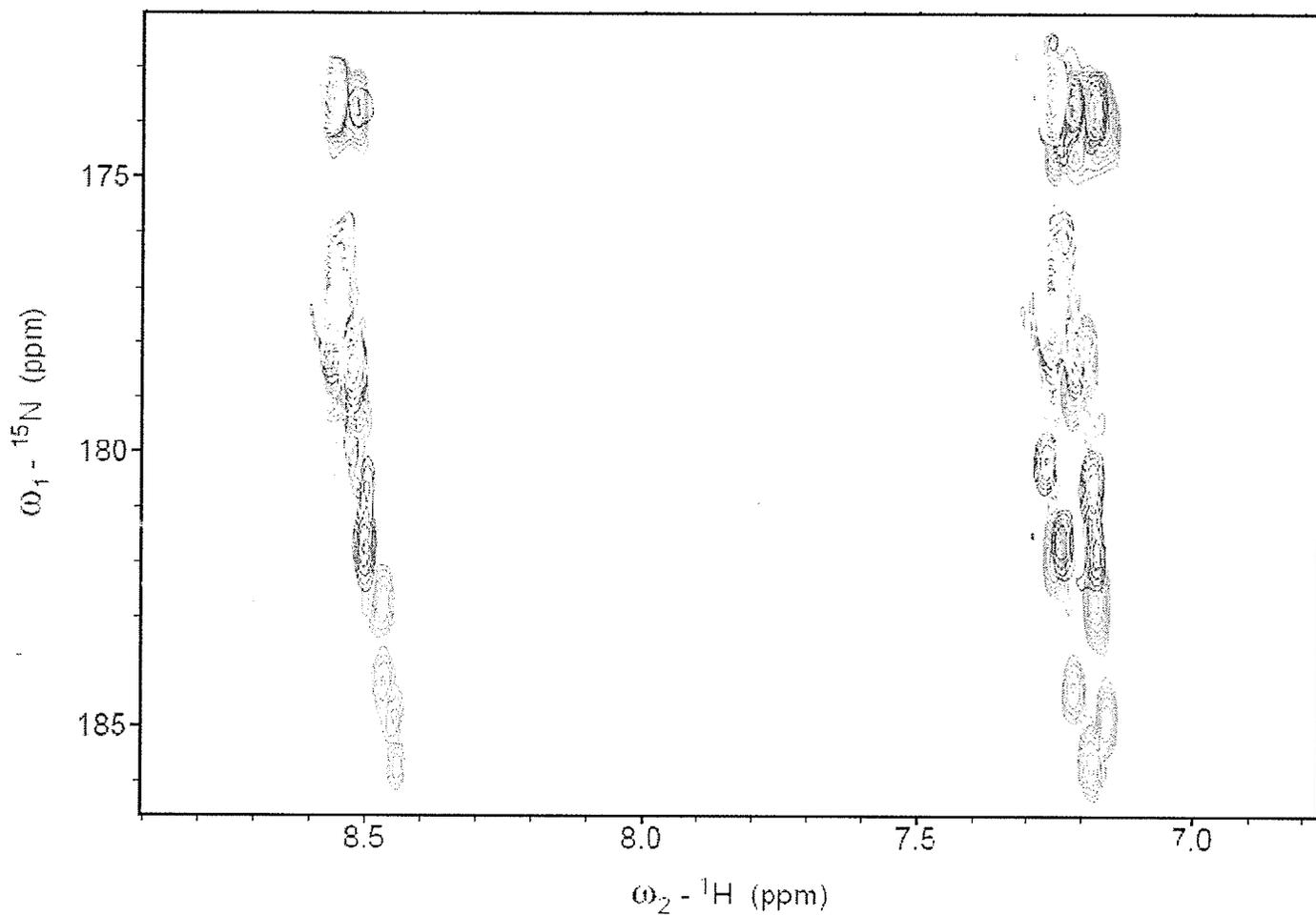
comparison of the observed imidazole chemical shifts with the literature [150] the imidazole rings in Tat are determined to be in a cationic ionization state whereby the positive charge is delocalized over the  $N_{\delta 1}$  and  $N_{\epsilon 2}$  atoms [Figure 3.28C]. The disappearance of cross-peaks in the presence of Zn(II) is likely due to the Tat-Zn(II) complex exhibiting dynamic exchange between bound and free states and possibly between four- and five-metal coordination geometries [138].

Figure 3.28: (A)  $^{15}\text{N}$ - $^1\text{H}$  HMBC spectrum of apo-Tat showing the cross peaks corresponding to  $^2J_{\text{NH}}$  coupling within the imidazole rings of the histidine residues at pH 4 and 293 K. (B) An overlay of the  $^{15}\text{N}$  - $^1\text{H}$  HMBC spectra of apo-Tat with increasing concentrations of Zn(II): 0 (red) , 0.5 (blue) , 1.0 (green), and 2.0 (yellow) mole equivalents. (C) A schematic diagram of the three possible tautomeric states of histidine, reprinted from [150] with the permission of Pelton et al. All spectra were recorded on a Varian INOVA 600 MHz spectrometer. A J-coupling constant of 6 Hz was used to set the refocusing with a 41.6 ms delay (4J coupling) for the experiment [151]. Due to the inability to assign the cross-peaks, numbers will be used to label each cross-peak and their chemical shifts are listed in Appendix A.

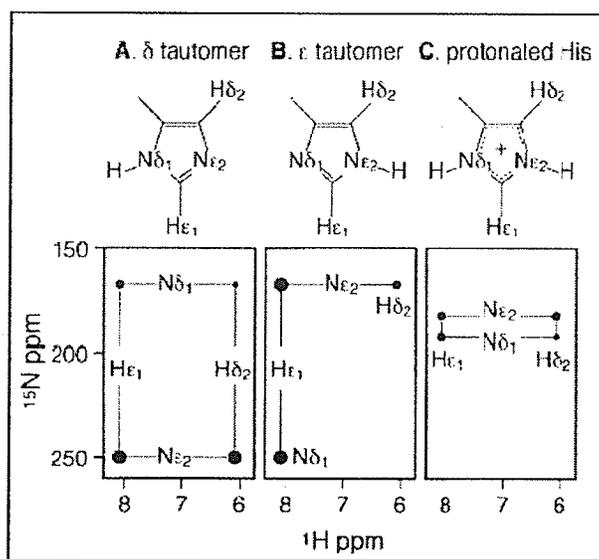
A



B



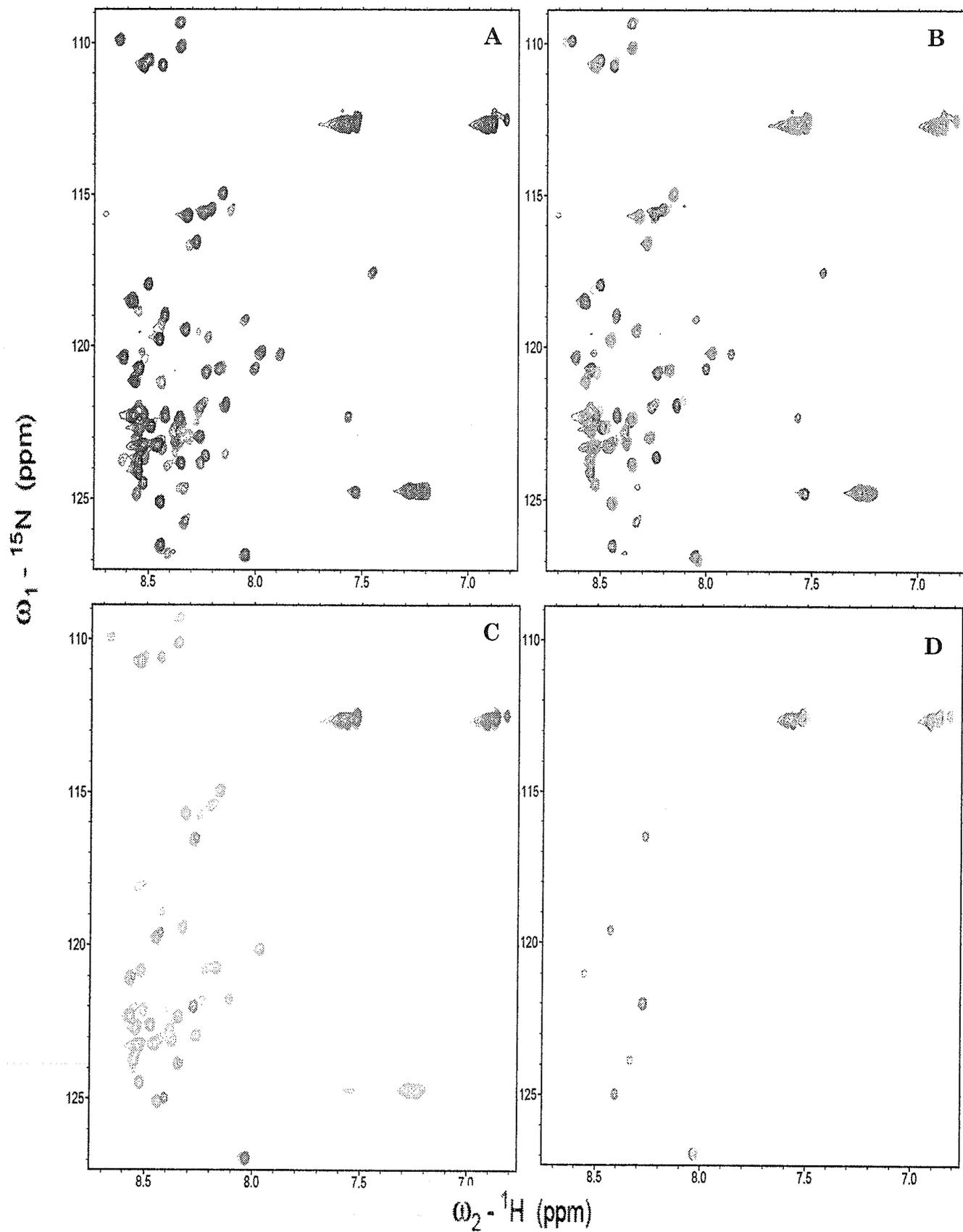
C



### *3.4.5 2D $^1\text{H}$ - $^{15}\text{N}$ HSQC NMR on $^{15}\text{N}$ -Labeled Tat in the presence of 0.5 Mole Equivalents of Zn(II) at different pH values*

The effects of pH on the Tat-Zn(II) complex were probed by 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC NMR to determine the stability of the metal-protein interaction and whether or not it can prevent aggregation at higher pH. The NMR tube containing the Tat sample was degassed and then purged with Ar prior to data acquisition. In Figure 3.29, the  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of Tat in the presence of increasing Zn(II) concentrations show a gradual reduction in residue cross-peaks as the pH value increases. In addition, the overall intensity of all the cross peaks decrease with increasing pH. With the increase from pH 4.02 to 5.11, comes a dramatic disappearance of cross-peaks from 56 to 41. Further decreases in cross-peaks are also observed as the pH is raised from 5.11 to 6.12 (41 to 7). Finally, at pH 7.03, only one residue cross-peak is observed. The significant reduction of cross-peaks upon titration with ammonium hydroxide is likely due to precipitation of the Tat protein upon oxidation of the cysteine side chains that results in the formation of many disulphide cross-linked multimers observable on SDS-PAGE (see Figure 3.4). The observation of white precipitate in the Tat solution at pH above 6 adds further evidence to the oxidation of the protein and the formation of multimers. It is noteworthy that the dilution of the protein with ammonium hydroxide has no influence on the NMR results due to the fact that only 10  $\mu\text{l}$  of diluted base is added per pH titration to an initial volume of 600  $\mu\text{l}$ .

Figure 3.29: (A)  $^1\text{H}$ - $^{15}\text{N}$  HSQC overlay of apo-Tat (red) and Tat in the presence of 0.5 (blue) MEQ of Zn at pH 4.42. (B)  $^1\text{H}$ - $^{15}\text{N}$  HSQC overlay of Tat in the presence of 0.5 MEQ of Zn at pH 4.42 (blue) and 5.11 (green). (C)  $^1\text{H}$ - $^{15}\text{N}$  HSQC overlay of Tat in the presence of 0.5 MEQ of Zn at pH 5.11 (green) and 6.12 (magenta). (D)  $^1\text{H}$ - $^{15}\text{N}$  HSQC overlay of Tat in the presence of 0.5 MEQ of Zn at pH 6.12 (magenta) and 7.03 (yellow). Tat protein (600  $\mu\text{M}$ ) was dissolved in 700  $\mu\text{l}$  of NMR buffer at pH 4.02 and data were acquired at 293 K. All spectra were recorded on a Varian INOVA 600 MHz spectrometer. The acquisition time was 2 hrs and 40 min.



### 3.4.6 2D $^1\text{H}$ - $^{13}\text{C}$ HMQC NMR on $^{13}\text{C}$ -Labeled Tat and Tat-Zn(II) Complex

To get a better understanding of the binding of Tat to Zn(II),  $^1\text{H}$ - $^{13}\text{C}$  HMQC NMR analysis was carried out to investigate the involvement of the cysteine and histidine side-chains. By examining the  $^{13}\text{C}^\beta$  chemical shifts of cysteine residues it is possible to get insight into the effects of thiolate-metal interactions [152, 153, 154]. Aromatic resonances are more likely to report effects of the binding of metals to histidine [155, 156]. It is noted that the sequential assignment of the  $^1\text{H}$ - $^{13}\text{C}$  HMQC data has not been carried out leaving the cross-peaks to be identified using arbitrary numbers. Assignment of the amino acid side-chain resonances would require further information obtained from 3D NMR experiments that can correlate the  $^{13}\text{C}^\beta$  cross-peaks with the corresponding  $\text{H}^\beta$  such as a HSQC-TOCSY experiment. Based on the measured chemical shifts from hundreds of proteins stored in a well-established chemical shift database [157], cross-peaks corresponding to cysteine residues in the aliphatic region of the HMQC spectrum (Figure 3.30A) are predicted to be observed within the region along the  $^{13}\text{C}$  dimension between 28 to 33 ppm. It is noteworthy that more than one residue contains  $^{13}\text{C}^\beta$  resonances (other than cysteine) and are also found in this region such as Arg, Glu, Gln, Lys, Met, Pro, and Val. Any significant chemical shift changes of cross-peaks within this region in the presence of 0.5 mole equivalents of Zn(II) can provide evidence to suggest that the metal may be coordinating with the cysteine residues but this would have to be followed up with further experiments involving resonance assignments. In the

aliphatic region of the HMQC spectrum, shown in Figure 3.30B, one would predict that the  $^{13}\text{C}^\beta$  of the histidine residue would be observed in the region of 28-34 ppm. Out of the four possible aromatic residues, Tryptophan has a  $^{13}\text{C}^\beta$  chemical shift within the range of interest. Since Tat contains only 1 Trp, there is a good probability that cross-peaks found within this region correspond to histidine.

In Figure 3.30, the  $^1\text{H}$ - $^{13}\text{C}$  HMQC spectra of apo-Tat (red) and Tat in the presence of 0.5 mole equivalents of Zn (blue) show only minor chemical shift changes in both the aliphatic (A) and aromatic regions (B). Figure 3.31A shows an expanded view of the aliphatic  $^1\text{H}$ - $^{13}\text{C}$  HMQC spectrum highlighting the resonances that undergo small changes. It is noted that cross-peaks 2, 3, and 4 undergo small chemical shift changes while cross-peak 5 disappears in the presence of Zn(II) possibly owing to interactions with cysteine residues. In Figure 3.31B and 3.31C, the aromatic regions of the  $^1\text{H}$ - $^{13}\text{C}$  HMQC spectrum show small chemical shift changes of 5 cross-peaks upon addition of Zn(II). It is likely that these changes reflect interactions between Zn(II) and the side-chains of histidine.

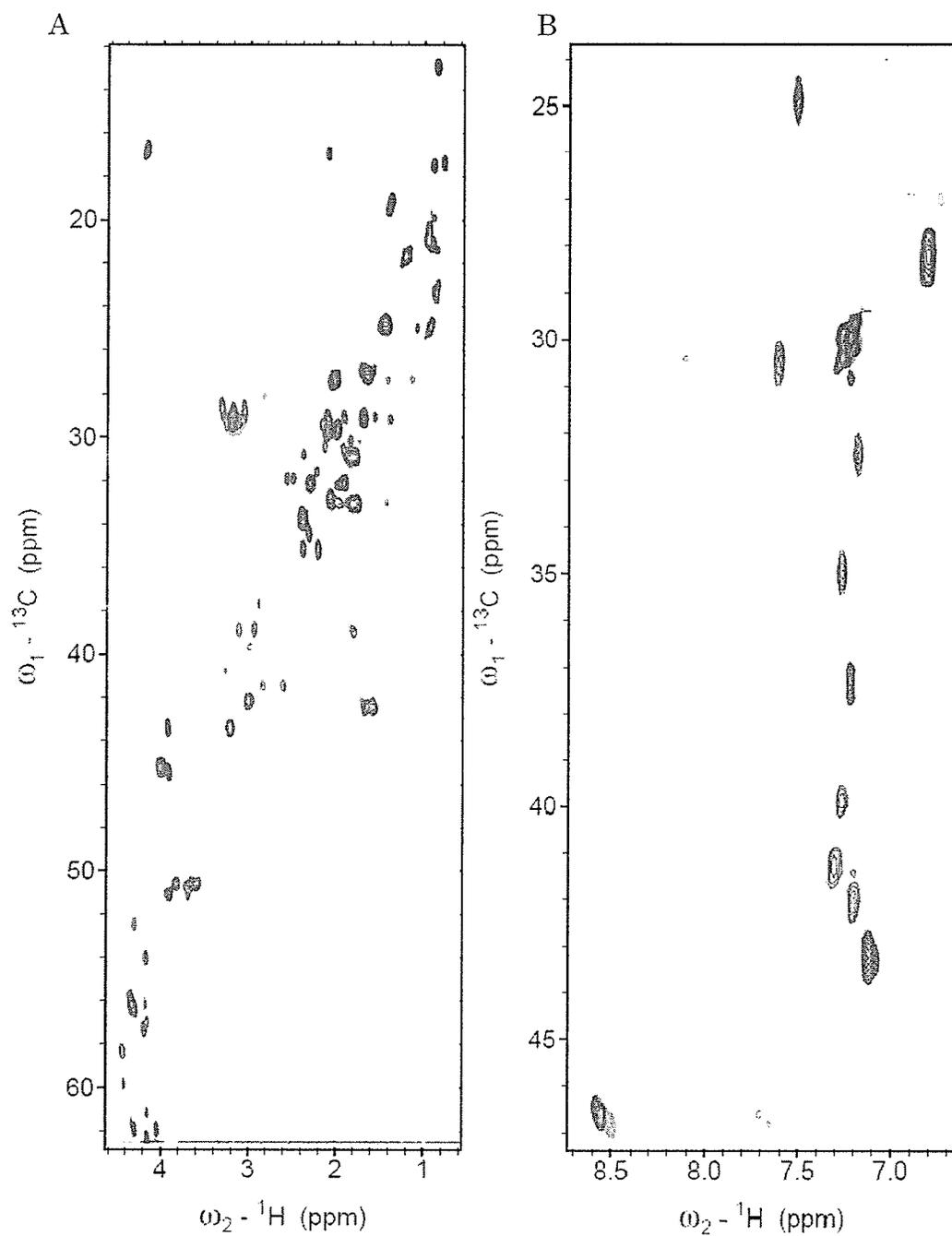
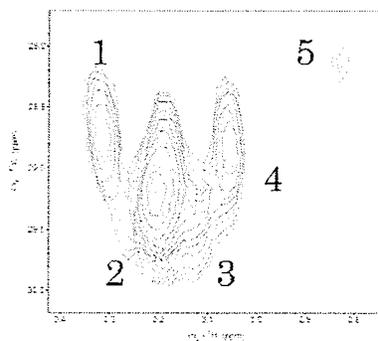
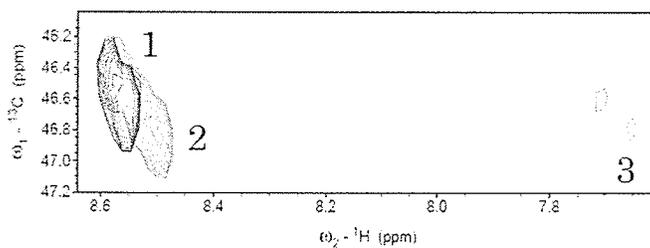


Figure 3.30: Aliphatic (A) and aromatic (B) regions of the  ${}^1\text{H}$ - ${}^{13}\text{C}$  HMQC spectra of apo-Tat (red) and Tat in the presence of 0.5 mole equivalents of Zn(II) (blue) at pH 4.0 at 293 K. All spectra were recorded on a Varian INOVA 600 MHz spectrometer.

A



B



C

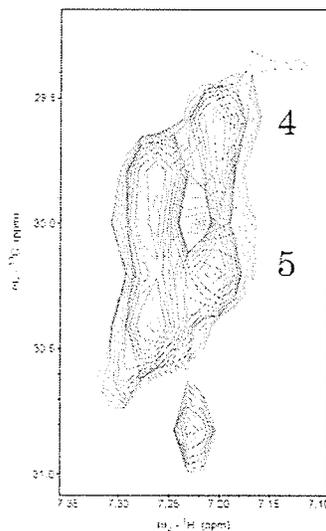


Figure 3.31: Expanded regions of the  $^1\text{H}$ - $^{13}\text{C}$  HMQC spectra of the (A) aliphatic and (B and C) aromatic regions of apo-Tat (red) and Tat in the presence of 0.5 mole equivalents of Zn(II) (blue). See Appendix B for tabulation of the chemical shifts of the observed cross-peaks.

### 3.4.7 NOESY/ROESY NMR

2D  $^1\text{H}$ - $^1\text{H}$  NOESY NMR experiments can provide information for the determination of protein structure and structural changes due to the binding of metals [158]. Furthermore, NOESY spectroscopy is useful for the assignment of aliphatic and aromatic side-chain resonances [159]. However, Tat is a 10 kDa protein that is found to be unfolded [55] and its dynamics may reduce the NOE cross-peaks to zero [see Figure 1.11 in **Introduction**]. A solution to this problem is the use of ROESY, which generates a non-zero NOE value regardless of size of the protein. The purpose of acquiring NOESY and ROESY spectra is to determine if the ROESY experiment improves the sensitivity of NOE measurement in Tat over the more common NOESY experiment.

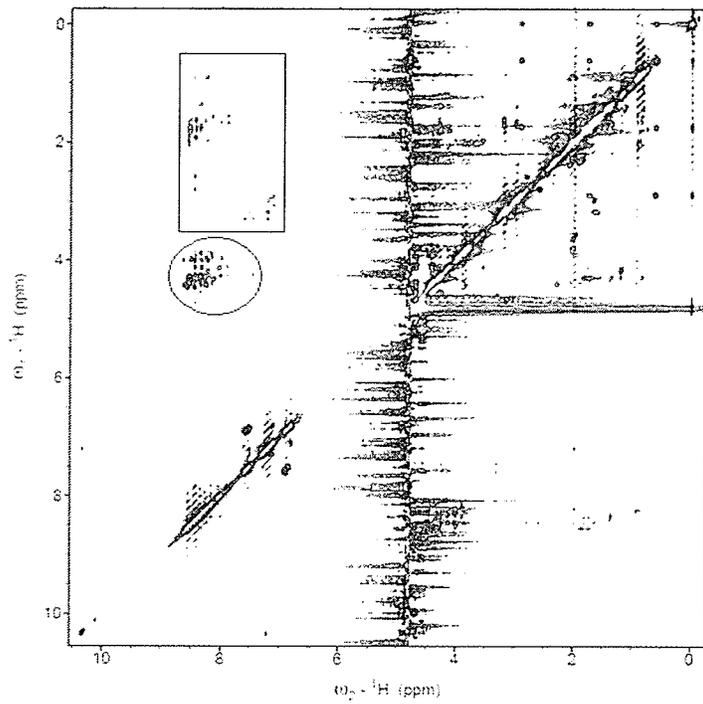
#### 3.4.7.1 NOESY and ROESY Experiments of Unlabeled Tat Protein

Figure 3.32 shows the 2D  $^1\text{H}$ - $^1\text{H}$  NOESY and ROESY spectra of unlabeled Tat at pH 4.0. It is noteworthy that all parameters were kept constant between the two experiments including those used for data processing. Clearly, the spectrum corresponding to the NOESY experiment contains more NOE cross-peaks than the ROESY experiment indicating that the former is a more appropriate experiment than the latter. Furthermore, the NOESY result suggests that the Tat protein's dynamics (rotational correlation time) do not fall within the regime that reduces the NOE signal to zero. While both NOESY and ROESY data are processed identically,

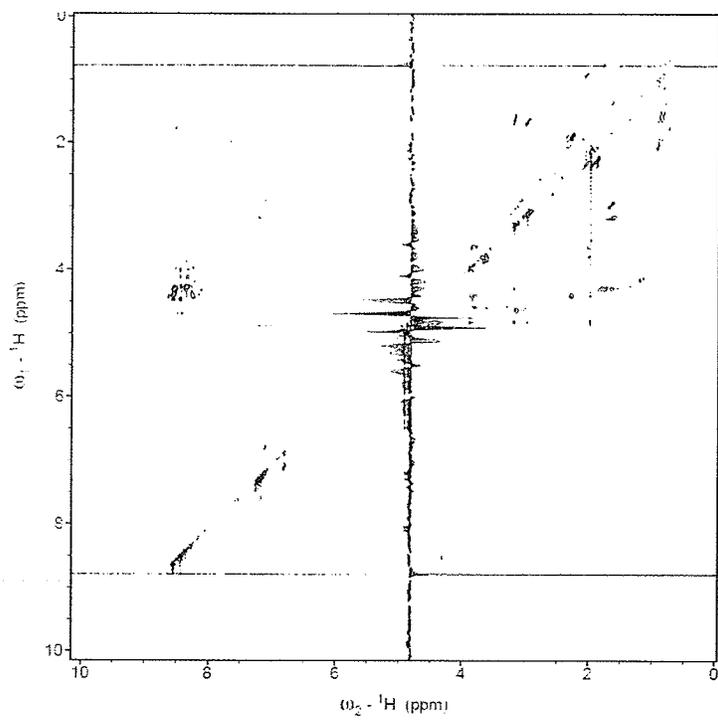
the former contains more noise (background) than the latter. Based on analysis of 1D NOESY and ROESY spectra, the signal-to-noise ratio (determined by comparing the diagonal cross-peak intensity to the baseline) of the NOESY spectrum is greater than the ROESY experiment [data not shown]. It is noteworthy that initial attempts for acquiring data of Tat using the ROESY were carried out with a WET water suppression pulse sequence. Unfortunately, the WET-ROESY experiment yielded very poor signal-to-noise ratio. In general, the maximum NOESY intensity is in the range of 1.0 whereas for the ROESY experiment, a maximum of 0.7 can be obtained [66] which must also be kept in mind when comparing the two experiments.

Figure 3.32:  $^1\text{H}$ - $^1\text{H}$  NOESY (A) and ROESY (B) spectra of apo-Tat acquired in 12 hrs at pH 4.0 at 293 K using a Varian INOVA 600 MHz NMR spectrometer. Both spectra were processed with Sparky using identical contour level settings [160]. All spectra were recorded on a Varian INOVA 600 MHz spectrometer. Both spectra were calibrated with DSS at 0 ppm. Cross-peaks representing NOE connectivities between amide protons of aromatic residues and either  $\alpha$ -protons of all residues or  $\delta$ -protons of Pro or  $\beta$ -protons of Ser and Thr are shown in the inset circle. Cross-peaks representing NOE connectivities between amide protons of aromatic residues and aliphatic side chain protons are shown in the inset box [110].

A



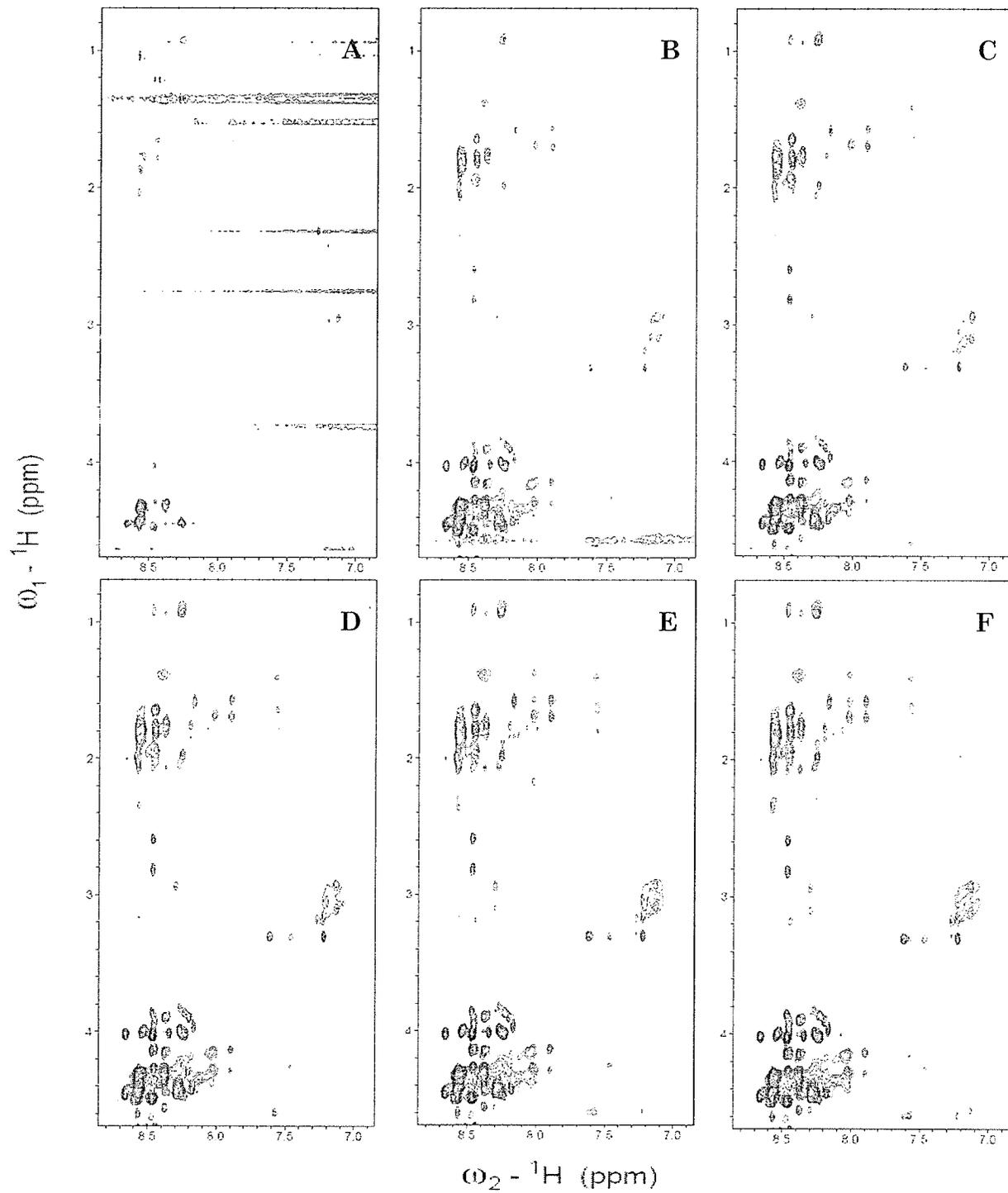
B



### 3.4.7.2 Unlabeled Tat Protein with Different Mixing Times

Experimenting with a range of NOESY mixing times may give an indication of what settings can give the optimal signal for NOE cross-peak observations. Figure 3.33 shows  $^1\text{H}$ - $^1\text{H}$  NOESY spectra of apo-Tat acquired with identical NMR parameters except for different mixing times. Clearly, the  $^1\text{H}$ - $^1\text{H}$  NOESY spectrum of apo-Tat with a mixing time of 300 ms shows the most intense NOE cross-peaks. The optimum mixing time has to be selected in order to allow the maximum amount of NOE build up to take place which increases sensitivity [110]. It is noteworthy that we are not interested in the chemical exchange rates but would like to determine what NOESY parameters are suitable for the Tat protein

Figure 3.33:  $^1\text{H}$ - $^1\text{H}$  NOESY spectra of Tat at pH 4.0 and at 293 K with different mixing times: 50 (A), 100 (B), 150 (C), 200 (D), 300 (E), and 400 ms (F). All spectra were recorded on a Varian INOVA 600 MHz spectrometer. All spectra were processed using Sparky with identical parameters and contour level settings.

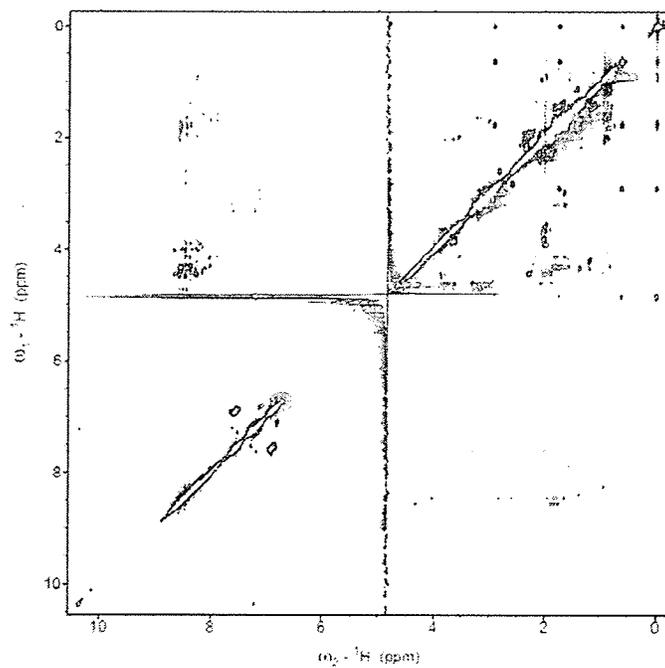


### 3.4.7.3 NOESY measurements of Apo-Tat and Tat-Zn(II)

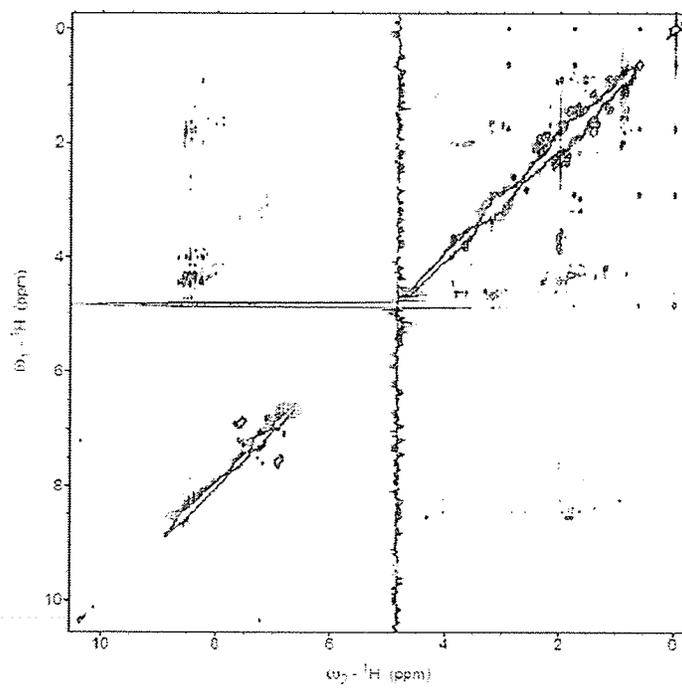
Using 2D  $^1\text{H}$ - $^1\text{H}$  NOESY NMR, it is possible to correlate two protons that are very close in proximity to one another and in this way obtain structural information about a protein. In the presence of Zn(II), it has been proposed that Tat may form a dimer [39] that may bring protons of opposite monomers close together and within range of NOESY detection that were out of range in the monomer. In Figure 3.34, the  $^1\text{H}$ - $^1\text{H}$  NOESY spectra of apo-Tat (A) and Tat in the presence of 0.5 mole equivalents Zn(II) (B) at pH 4.0 are compared. The spectra show no significant difference in the NOE cross-peaks suggesting no dramatic structural change in Tat's conformation upon addition of Zn(II). Thus, the indications of weak binding of zinc observed in the circular dichroism and heteronuclear NMR experiments (see sections 3.3 and 3.4) have little effect on the overall conformation of the protein at pH 4.0.

Figure 3.34:  $^1\text{H}$ - $^1\text{H}$  NOESY spectra of apo-Tat (A) and Tat in the presence of 0.5 mole equivalents of Zn(II) (B) at pH 4.0 at 293 K. The acquisition time for both spectra was 24 hrs. NMR data were acquired on a Varian INOVA 600 MHz NMR spectrometer. Concentration of the Tat protein was 524 mM. The mixing time for both spectra was 200 ms.

A



B



#### 3.4.7.4 Unlabeled Tat Dialyzed against Ni-free NTA (immobilized nickel nitrolotriacetic acid)

Another explanation for why no change in NOE cross-peaks were observed upon addition of zinc is the possibility of residual Co(II) bound to the Tat protein after dialysis. If, prior to the addition of Zn(II), the Co(II) is loosely coordinating with the Tat protein, the addition of Zn(II) may have displaced the Co(II) while leaving Tat's conformation unchanged. To determine whether residual Co(II) may be weakly attached to the Tat protein after dialysis with acetate buffer, 2D  $^1\text{H}$ - $^1\text{H}$  NOESY NMR was carried out on a Tat protein solution that had been dialyzed with a divalent chelator, Ni-free NTA, prior to freeze-drying. In theory, the Ni-free NTA resin will bind the Co(II) that may be attached to Tat via the His-tag or Cys-rich regions. In Figure 3.35, the  $^1\text{H}$ - $^1\text{H}$  NOESY spectrum of apo-Tat that underwent dialysis with Ni-free NTA resin shows no significant NOE cross-peak changes relative to apo-Tat that is purified under the standard protocol [see Figure 3.34A]. This suggests that no residual Co(II) remains weakly bound to apo-Tat after dialysis with glacial acetic acid or that metal exchange has no effect on the conformation of the protein. However, the concentrations of the Tat solutions that were dialyzed against NTA resin and the control were found to be 644  $\mu\text{M}$  and 524  $\mu\text{M}$ , respectively. It is unlikely that this difference in concentration could mask a major conformational difference between the two samples.

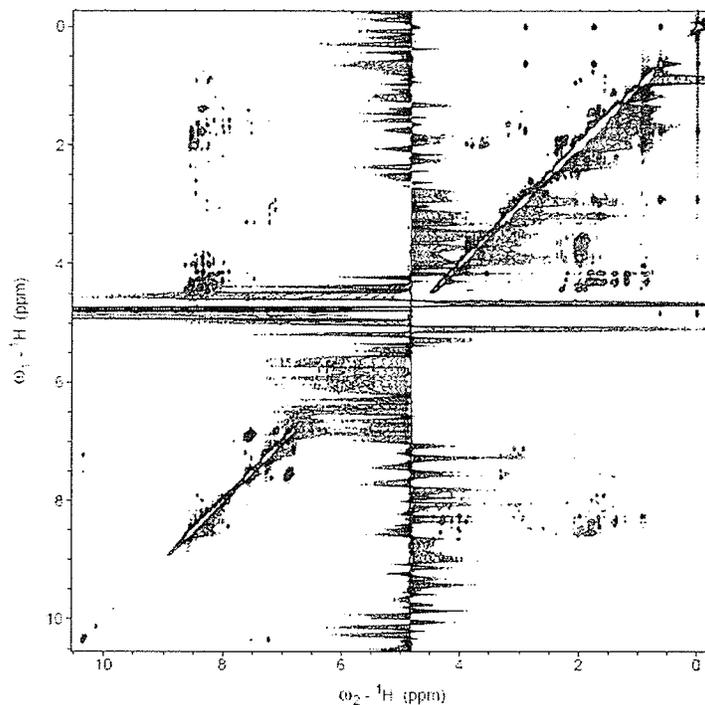


Figure 3.35:  $^1\text{H}$ - $^1\text{H}$  NOESY spectrum of apo-Tat ( $644\ \mu\text{M}$ ) that was dialyzed against Ni-free NTA resin for 24 hrs, freeze-dried and dissolved in  $700\ \mu\text{l}$  of NMR buffer at pH 4.0 at 293 K. The spectrum was collected on a Varian INOVA 600 MHz spectrometer. The acquisition time was 12 hrs.

An alternative approach in determining the presence of residual cobalt following column purification is to acquire a  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum of Tat protein that was dialyzed against Ni-free NTA and examine the relative peak intensities in regions believed to be coordinating with Zn(II) based on comparison with a control Tat solution. In Figure 3.36, the relative peak intensities of HSQC data from Tat dialyzed against Ni-free NTA and a control Tat sample are compared. The dialyzed protein shows significant line broadening within the cysteine-rich region accompanied by a dramatic increase in intensity for residues found near the His-tag compared to the control Tat protein purified under standard conditions.

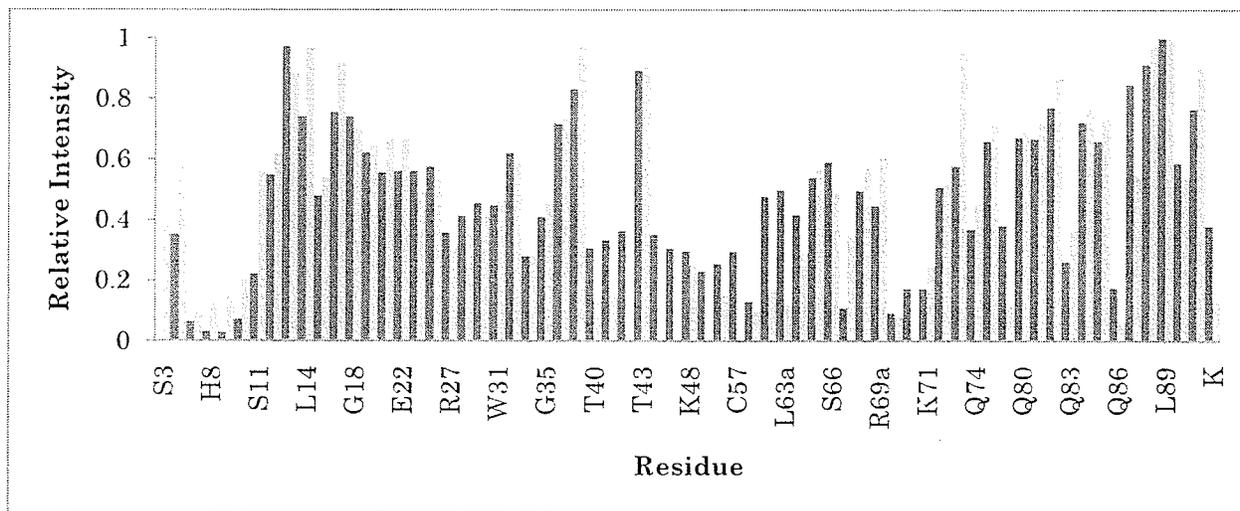


Figure 3.36:  $^1\text{H}$ - $^{15}\text{N}$  HSQC cross peak intensities for residues of Tat (538  $\mu\text{M}$ ) (orange) dialyzed against Ni-free NTA resin and control Tat (300  $\mu\text{M}$ ) purified under standard conditions (red). Both spectra were normalized using peak intensity from Lys-89.

Figure 3.37 shows the relative peak intensities of dialyzed apo-Tat and the same solution in the presence of 0.5 mole equivalents of Zn(II). Comparisons between the two samples show a small increase in the intensities of histidine residues within the His-tag accompanied by overall line broadening or a decrease in intensity of the remaining resonances. Interestingly, all seven cysteine residues in the Cys-rich region are not observed in either the apo-Tat or Tat-Zn(II) complex  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra. If any residual cobalt were indeed associating with the Cys-rich region, then the removal of the metal would show an increase in the intensity of the cysteine residues due to the absence of paramagnetic relaxation.

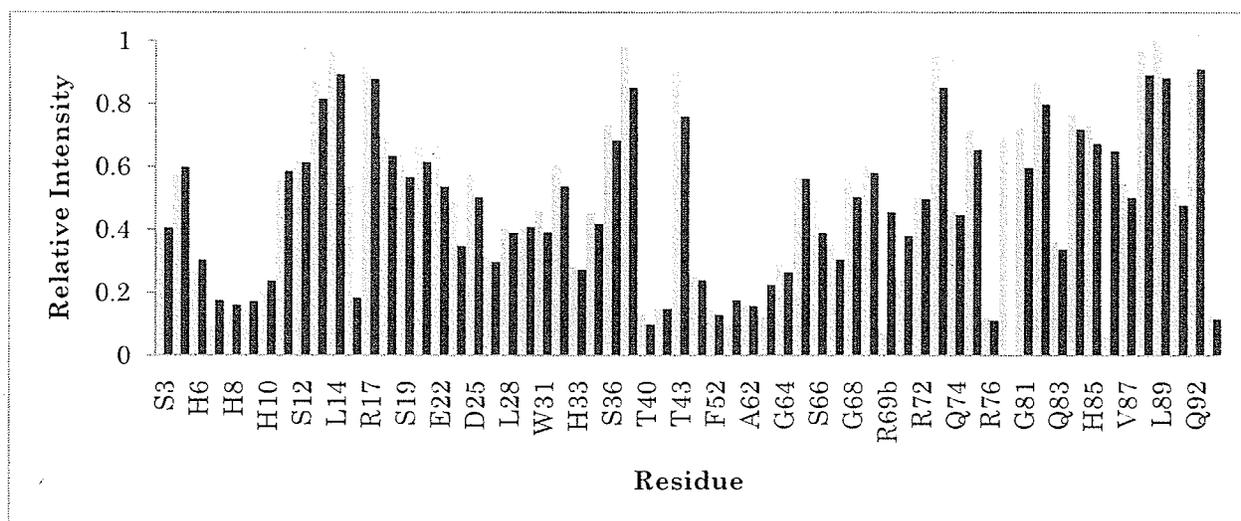


Figure 3.37:  $^1\text{H}$ - $^{15}\text{N}$  HSQC cross peak intensities for residues of apo-Tat (orange) and Tat in the presence of 0.5 mole equivalents of Zn(II) (blue) whereby the former sample has been dialyzed against Ni-free NTA to remove any residual cobalt.

### 3.4.8 $^{113}\text{Cd}$ NMR Titration Experiment

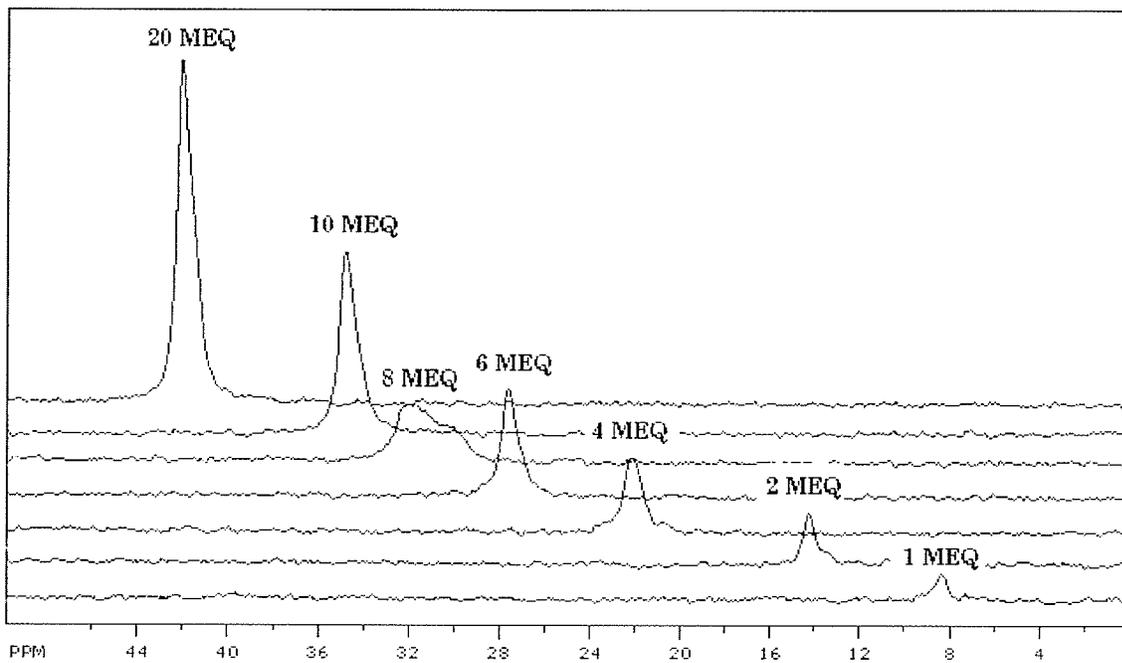
The chemical shift of  $^{113}\text{Cd}$  is particularly sensitive to the nature and number of ligands in its metal coordinating environment and cadmium substitutes well for calcium, magnesium, copper, manganese, and zinc in a number of biomolecules [162]. To further investigate the metal-binding properties of the cysteine-rich Tat protein,  $^{113}\text{Cd}$  NMR spectra were acquired. In Figure 3.38A, the  $^{113}\text{Cd}$  NMR spectra of  $\text{CdCl}_2$  titrations of the Tat protein at pD 5.0 show significant chemical shift changes upon addition of metal. Interestingly, the chemical shift of  $^{113}\text{Cd}$  at 1 mole equivalents (observed at 8.7 ppm) is substantially lower than the literature NMR values of  $^{113}\text{Cd}$  coordination to sulfur atoms that are typically located near 600 to 700 ppm [139]. There are two possible explanations for this observation. One is that cadmium exists in a coordinating environment that contains a mix of nitrogen and oxygen ligands. The other possibility is that the cadmium is in rapid exchange between oxygen-binding ligands at -100 ppm and sulphur ligands at 600-750 ppm. In either case, the shift from shielded to deshielded environments as  $^{113}\text{Cd}$  is added to Tat suggests an increasing influence from sulphur and nitrogen ligands. This suggests that the  $^{113}\text{Cd}$  is likely to undergo dynamic exchange between bound and unbound Tat-Cd(II) complex with preference for the unbound form. Further addition of  $^{113}\text{CdCl}_2$  to the Tat protein causes gradual downfield chemical shift changes suggestive of increased metal-binding. In the 8 mole equivalents of Cd(II) spectrum, it is observed that the peak broadens and decreases in intensity. This is

most likely due to a number of factors such as poor shimming, formation of bubbles, or rapid transient formation of insoluble Tat-Cd(II) complex.

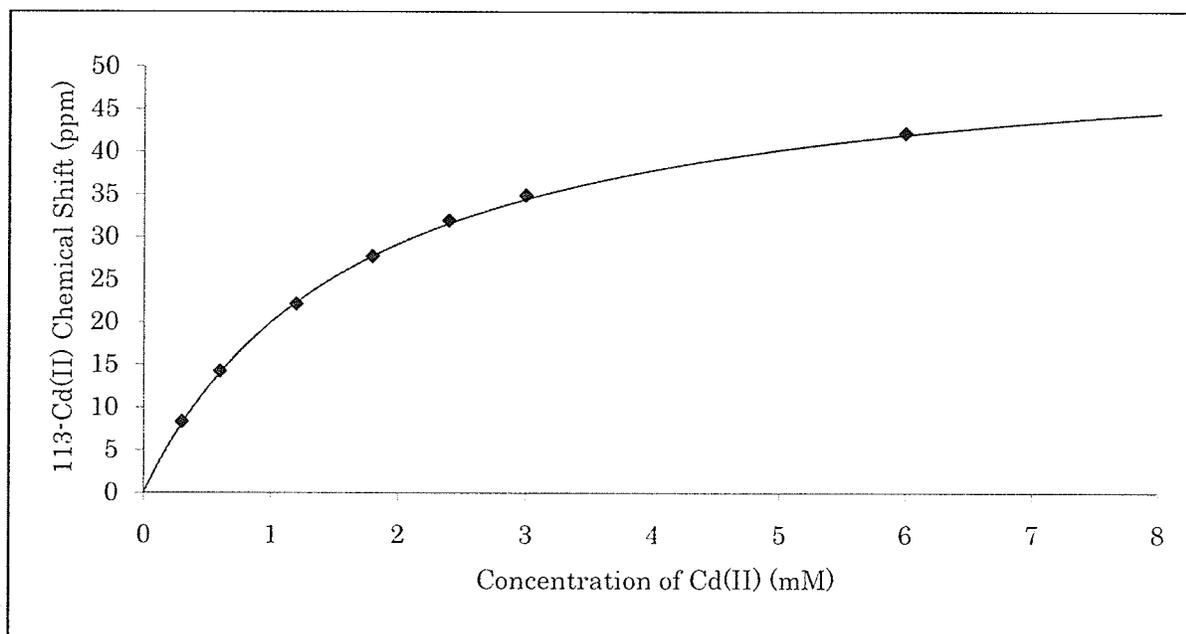
Figure 3.38B shows a plot of the  $^{113}\text{Cd}$  chemical shift versus the concentration of Cd(II) added to the protein. The line through the points is a proper fit to the binding equation given in the Methods section. It is likely that the metal-binding site of the Tat protein is saturated at metal concentration above 10 mM. Results from the Mathematica calculations show that the chemical shift at saturation of the binding site is 52.98 ppm while the chemical shift at zero metal concentration is -1.98 ppm. Careful inspection of Figure 3.38B shows that at half the chemical shift for saturation (26.49 ppm), the concentration of Cd(II) is 1.8 mM with the Tat concentration at 300  $\mu\text{M}$ . This matches perfectly with the Mathematica curve fitting data showing that the  $k_D$  is 1.81 mM.

Figure 3.38: (A)  $^{113}\text{Cd}$  NMR spectra of  $\text{CdCl}_2$  titrations with HIV-1 Tat protein at pD 5.0 at 293 K.  $\text{Cd}(\text{ClO}_4)_2$  (0.1 M) was used as an external standard with chemical shift at 0 ppm. All spectra were recorded on a Varian INOVA 600 MHz spectrometer equipped with a broadband probehead. The Tat protein was at a constant concentration of 300  $\mu\text{M}$  and the cadmium concentration was in the range of 1 to 20 mole equivalents. (B) Mathematica fit of  $^{113}\text{Cd}$  chemical shift against concentration of metal.

A



B

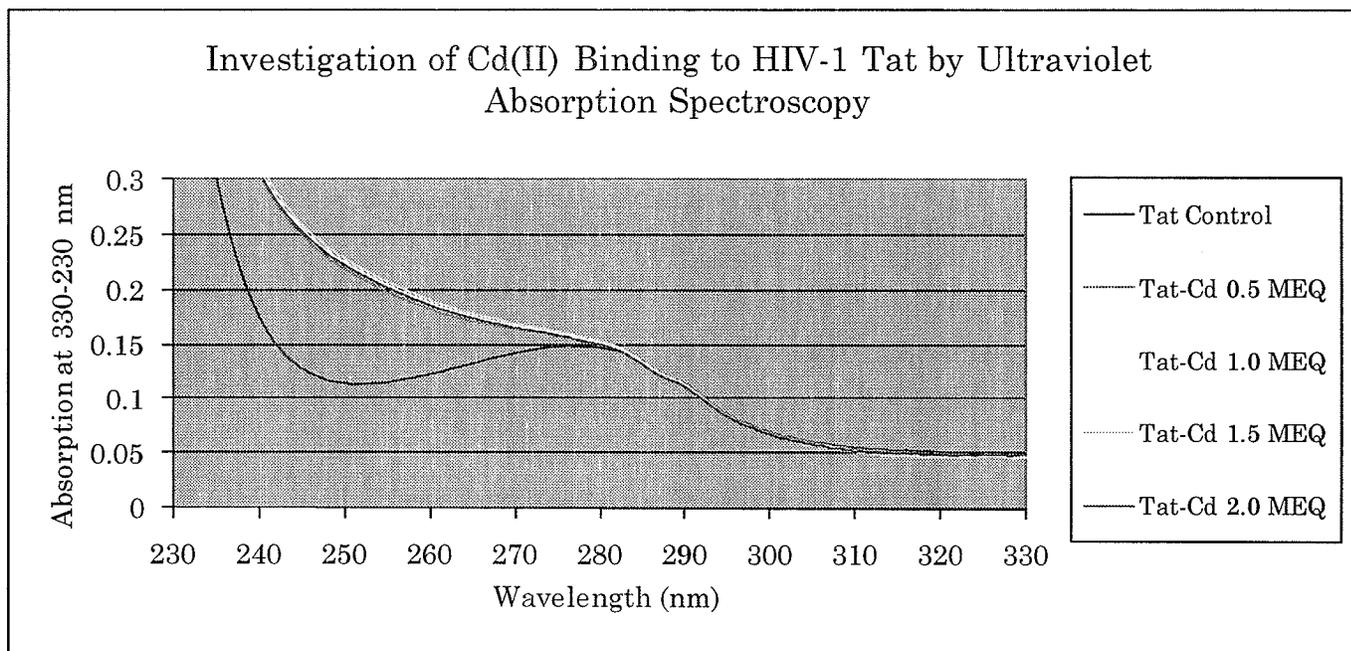


### 3.5 UV Absorption Analysis of Tat in the presence of Cadmium

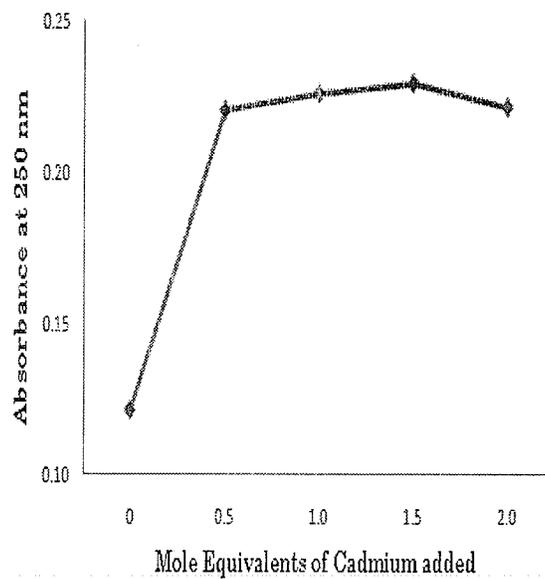
In Figure 3.39A, the UV absorption spectra of apo-Tat and Tat at pH 4.5 in the presence of increasing concentrations of Cd(II) show a pronounced absorption increase at 250 nm in the presence of 0.5 mole equivalents Cd(II) followed by minor changes with further Cd(II) titration. This suggests that the binding of Cd(II) to the Tat protein saturates at metal concentrations above 0.5 mole equivalents. The increase in absorption at 250 nm is likely due to thiolate-to-Cd(II) electron transfer [161]. Figure 3.39B shows a graph of the absorbance at 250 nm with respect to the mole equivalents of cadmium added. Clearly, saturation occurs after the initial addition of 0.5 mole equivalents of cadmium lending further support to the formation of a metal-linked dimer.

Figure 3.39: (A) Ultraviolet absorption spectra of apo-Tat with 0.5-2.0 molar equivalents of cadmium using a 1 cm cuvette at pH 4.5. The UV spectrometer is blanked with acetate buffer at pH 4.5. Protein concentration was 16  $\mu\text{M}$ . (B) Graph of absorption at 250 nm vs. mole equivalents of cadmium added.

A



B

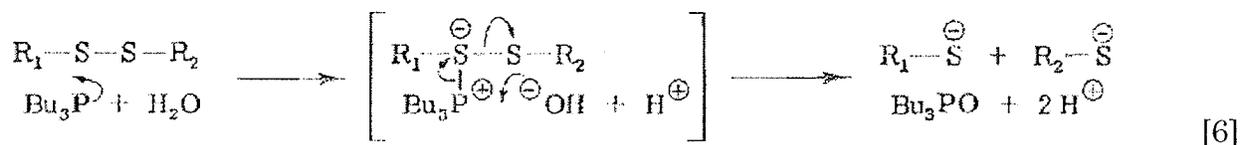


## 4. Discussion

### 4.1 Protein Purification and Expression

Isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) was used to induce the transcription of the *Tat exon 1* (amino acids 1-72) gene within the expression vector pET-28c(+) which contains a commonly used His-affinity tag for purification via cobalt metal-affinity chromatography. Purification of the protein using this method is based on the change in ionization of the imidazole ring of the histidine residues controlled by varying the pH of the buffers. Binding buffer, at pH 7.2, causes the imidazole ring of the histidine residues within the affinity tag to become deprotonated permitting the interaction between the  $N_{\epsilon 1}$  atom of His ( $pK_a = 5.86$ ) and the immobilized cobalt. The elution of the Tat protein occurs when the pH of the buffer drops below 5.86 after which the  $N_{\epsilon 1}$  atoms become protonated and the protein dissociates from the column.

To prevent oxidation of the cysteine residues at neutral pH, tris(2-carboxyethyl)phosphine (TCEP) was added as a sulfhydryl reducing agent [132]. In terms of stoichiometry, one mole of TCEP reduces one mole of disulfide [Equation 6] [167]. The mechanism of TCEP reduction of disulfide bonds has the cleavage of the disulphide as the rate-determining step in Equation 6.



In practice, all disulfide bonds can be reduced with a 5-20% excess of TCEP [169]. No denaturing agent is required to increase efficacy of TCEP reduction. However, it is noted that in the case where the reduced protein precipitates the presence of denaturing agent may be necessary. The efficacy of reduction by TCEP has been shown to cover a wider range of pH values as compared to dithiothreitol (DTT) and  $\beta$ -mercaptoethanol ( $\beta$ ME) [167]. DTT is prone to oxidation in the presence of metal ions and is not stable in reduced form for a long period of time. In contrast, TCEP is much more stable and has the added advantage of being a weaker chelator of biologically relevant metal ions (such as Zn(II)) than DTT [168] permitting the studies of protein metal-binding interactions in reduced conditions. In the case of the Tat protein however, metal titrations in the presence of TCEP were not feasible as TCEP precipitates the protein in the absence of guanidine-HCl.

Although guanidine-HCl has been commonly used as a denaturing agent in protein biochemistry [197], it has been proven to be beneficial in stabilizing the Tat protein during purification. The purpose of adding 6 M guanidine-HCl to all buffers involved in the Tat protein purification is to maximize the yield of purified Tat protein. It has been shown that in the presence of 6 M guanidine-HCl the Tat protein elutes from the metal affinity resins in a volume of 10 ml rather than 40 ml [137]. It is likely that high guanidine-HCl concentrations prevent interactions

between the cationic protein and the chromatographic resin permitting elution in a minimum volume. Since the Tat protein has been shown to be a disordered protein [26], the denaturing effect of guanidine can be ignored. Another advantage of 6 M guanidine-HCl is that it prevents Tat precipitation in the presence of TCEP at neutral pH presumably by blocking interactions between the Tat protein's basic region and the carboxylic groups from TCEP. Several steps have been carefully placed within the purification protocol to minimize the risk of oxidation of the Tat protein. In particular, the dialysis of the Tat protein was carried out in a reducing environment where the buffers were at low pH (3.50), degassed, and then purged with argon.

## 4.2 SDS-PAGE

On SDS-polyacrylamide gel electrophoresis (SDS-PAGE), the 92-residue monomeric Tat protein migrates at an apparent molecular size of 17-kDa, which is comparable to literature values showing an 86 amino acid Tat protein migrating at 15-kDa [39]. It is possible that the presence of the His-tag accounts for the difference of 2-kDa in apparent molecular size. The significant difference in the apparent molecular size of Tat (17-kDa) and the theoretical mass (10,509 Da) [137] is due to the protein's high net positive charge that retards its electrophoretic mobility.

In SDS-PAGE, the Tat protein has been shown to be susceptible to oligomerization due to the oxidation of the cysteine residues as it migrates through

the resolving gel at pH 8.8 even in the presence of DTT [see Figure 3.2.2 in **Results**]. TCEP has been commonly used as a reducing agent prior to SDS-PAGE analysis [132]. In the presence of 10 mM TCEP, a predominantly monomeric Tat band is observed due to the more complete reduction of disulfide bonds afforded by the more powerful oxidizing agent. Since the protein is prepared at low pH, TCEP is ideally suited to preventing the cross-linking between the cysteine residues. The fact that TCEP reduces the Tat oligomers to the monomeric form via disruption of disulfide bond formation provides indirect evidence that the cysteine residues play a significant role in Tat protein oligomerization [198]. Results in this thesis show also that caution must be exercised in the use of TCEP as it will readily reduce the pH of the sample and in high concentrations alter the electrophoretic mobility of proteins.

#### *4.2.1 Alkylation of Tat*

The purpose of alkylating the Tat protein was based on a previously published study of the Bovine pancreatic trypsin inhibitor (BPTI) [63] whereby the folding mechanism was investigated through measuring the rate constants of disulfide formation with glutathione using strong acid as a quenching agent and HPLC to separate the folding intermediates.

The chemical modification of cysteine residues within the Cys-rich region of the Tat protein was carried out using the uncharged alkylation reagent, iodoacetamide. It is noteworthy that initial attempts at alkylation of the cysteine residues were carried out using iodoacetic acid. Iodoacetic acid has been shown to be

an excellent modifier of sulfhydryl groups of glutathione and cysteine that led to the inactivation of glyoxalase [180]. Alkylation attempts with iodoacetic acid were unsuccessful due to frequent precipitation of the Tat protein in the presence of the reagent. A plausible explanation for the observation of protein precipitate is the interaction of the basic region of one monomer to the negatively charged carboxyl groups of the alkylated cysteines on another monomer. The use of iodoacetamide introduces an uncharged amide on the cysteine side chain and prevents precipitation.

The identification of alkylated Tat protein was conducted using SDS-PAGE whereby a monomeric band at 17-kDa is expected. In SDS-PAGE, the modified protein showed a lower abundance of oligomers compared to the un-alkylated control suggesting the presence of largely alkylated monomeric Tat protein. Confirmation of the extent of alkylation by mass spectrometry was not done as this part of the project was discontinued. In a study of the human extracellular superoxide dismutase (EC-SOD) [199], the protein was allowed to react with iodoacetamide and subsequently applied onto the SDS-PAGE. Results of the study showed a band at 28 kDa corresponding to monomeric EC-SOD. It is noteworthy that EC-SOD readily undergoes homodimerization via disulfide bond linkage of Cys-219.

HPLC analysis of alkylated Tat protein was not successful due to problems with protein detection and solubility. Many factors might have caused this. For example, it is possible that the pH of the buffers within the column may not have been suitable for the Tat protein leading to a possibility of the protein precipitating during chromatography. The observation of multiple peaks on the chromatograms may indicate a source of contamination in the guard column or solvent buffer. It is noteworthy that little information on cysteine alkylation of the Tat protein is found in the literature.

### 4.3 Circular Dichroism

#### *4.3.1. Secondary Structural Analysis of Apo-Tat and Content predicted by Dichroweb*

Circular dichroism (CD) is a well-established spectroscopic technique that predicts the secondary structure of the protein of interest. The wavelength range of interest was chosen to be 240 to 180 nm (far-UV band) because it contains contributions from the peptide bonds (amide chromophores) that in turn can provide secondary structural information on the Tat protein. The CD spectrum of HIV-1 Tat protein at pH 4 reveals that the secondary structure of the protein is likely composed of a random coil as evidenced by the strong negative band at 199 nm and a weak negative band at 220 nm [62]. This result agrees with the literature on CD of the Tat protein. Far-UV CD studies of five synthesized HIV-1 Tat variants (86-101 residues) [171] all showed distinct strong negative bands at 199 nm accompanied by a weaker negative band at 230 nm, thus supporting our conclusion that the Tat

protein is mainly disordered. Furthermore, the literature suggests little secondary structural differences among these Tat variants.

Dichroic deconvolution of the Tat CD spectrum suggests that the protein is comprised of 35.3 % disordered structure, 22.6%  $\beta$ -turns, 24.3%  $\beta$ -sheet, and 6.3 %  $\alpha$ -helix. In the secondary structural analysis of CD data of the Tat Mal variant (87 amino acid residues) [172], isolated from a subtype D HIV-1 strain, it was revealed that the protein is composed of 43% unordered structure, 30% beta-turns, 22% beta-sheet, and 5%  $\alpha$ -helix. Comparison of the secondary structural percentages of Tat<sub>1-72</sub> with the Tat Mal variant shows an agreement among all four types of secondary structure. These results are also in general agreement with the view that true random coil conformations exist in equilibrium with  $\beta$ -turns,  $\beta$ -strand, and polyproline II helix (Figure 4.1) [135]. Furthermore, the basic segment of the protein is likely to exist in an extended  $\beta$ -sheet-like conformation owing to the concentration of positive charge in that segment at pH 4.0. It seems likely as well that the N-terminus may sample the polyproline helix conformation owing to the presence of five Pro in the first 18 residues and a nearby sixth Pro in the N-terminal affinity tag [see Figure 1.4 in **Introductions**]. Note that deconvolution of polyproline II helix and random coil structures are not usually attempted owing to the similarity of the CD spectra. These results also agree with NMR studies on the conformation of the Tat protein that describes the protein as a random coil [55].

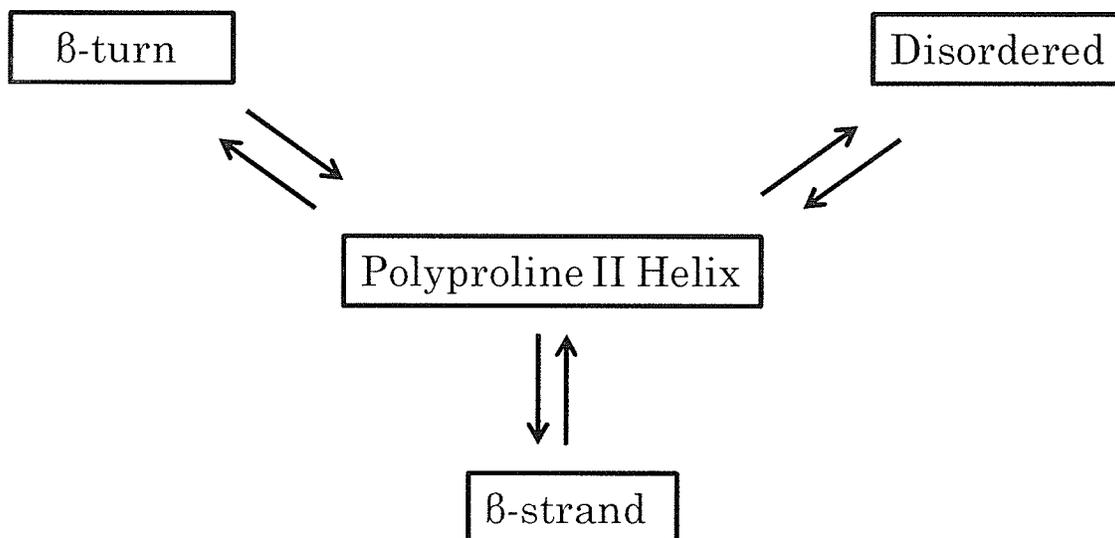


Figure 4.1: Illustration of the equilibrium between multiple conformations: Polyproline II Helix, Disordered,  $\beta$ -strand, and  $\beta$ -turn [135].

A possible explanation for the slight deviation of Tat's secondary structure percentage from the literature values is that different algorithms were used for the calculations and that Dichroweb [145] may contain a significantly more updated reference dataset. In addition, our version of the protein contains the His affinity tag.

#### 4.3.2 Effects of pH Titration

A pH titration on apo-Tat was conducted to determine whether the protein's secondary structure is influenced by the environment's pH. As the pH increases, the ellipticity at 199 nm significantly decreases suggestive of a pH-dependent structural change that results in a substantial loss of random coil structure. Raising the pH of the solution will change the ionization of the residue's side chains. In

particular, the ionization of Asp, Glu, and His will reduce the net positive charge on the protein which may directly change the structure of the protein by decreasing the amount of extended structure. In addition, at neutral pH the acidic N-terminus may directly interact with the basic segment reducing the random coil content. Reducing the net charge on the protein may also permit some helical conformation to form; recall that the  $\alpha$ -helix is the most compact of the common conformations of proteins.

The pH titrations of Tat did not exhibit an isodichroic point. In general, when a protein exists in only two main conformations a pH or temperature change shifts the equilibrium populations and the two-state transition is diagnosed by an isodichroic point in the spectrum. An isodichroic point is observed when the CD curves obtained under different conditions intersect. For example, an isodichroic point is observed at 208 nm during the pH titration study the XAO peptide [215], whereby a gradual decrease in the ellipticity at 199 nm and a gradual increase in ellipticity at 215 nm are observed as a function of pH indicating equilibrium between the polyproline conformation and  $\beta$ -strand. Since the CD spectra of Tat at different pHs do not show a gradual increase at 215 nm, no isodichroic point is observed. The lack of an isodichroic point in the CD spectra suggests two possible interpretations. One is that the protein exists in more than two conformations in equilibrium. This is not surprising in view of the Dichroweb secondary structure analysis indicating 3 well-populated conformations. The second is that changes in the far UV ellipticity of aromatic groups mask a two-state transition.

In a CD study of a Tat fragment (47-72) by Loret et al. [182], it was shown that the negative peak at 198 nm decreased in intensity as the pH of the Tat solution was increased from 7 to 11 indicating that further decreases in random coil conformation can occur at higher pH's than measured here. Furthermore, analysis of the CD spectra of the Tat fragment showed that as the pH was raised, the percentage of Tat's  $\alpha$ -helical structure increased by 2 % while the percentage of other structure decreased by 2 %. The importance of the environment's pH towards Tat's biological function was studied by Rayne et al. whereby they determined that exogenous HIV-1 Tat protein requires low endosomal pH (5 or 6) in order to reach the cytosol of T-cells [174]. Besides the effects of pH on the structure of the protein described above it should also be recalled that the Tat protein is highly reactive at pH values higher than 6.0 due to formation of disulfide cross-linking and this may contribute to the increased activity at lower pHs.

To further investigate the effects of increasing the pH on Tat's conformation, a deconvolution of the CD spectrum of Tat at pH 8.04 was carried out using Dichroweb. Results show that as the pH is raised from 4.17 to 8.04, the percentage of random coil decreased by 3.2 % while the percentage of  $\alpha$ -helix increased by 10.9 %. Thus, the increase in pH caused the conformation of the Tat protein to become more ordered due to ionization of side chains corresponding to key residues involved in stabilizing the protein's secondary structure.

In a CD study of the Tat peptide that consists of the cysteine-rich region (21-38) by Huang et al. [41], it was shown that in the presence of 0.5 mole equivalents of

Zn(II), a significant loss of ellipticity at 199 nm was observed indicative of a conformational change induced by Zn(II) through chelation with the thiols of the cysteine residues at pH 7.0. Unfortunately, it does not provide any information about the pH-dependency of the Tat-Zn(II) complex.

#### *4.3.3 Effects of Zn(II) at pH 4.0, 5.0, and 7.2*

Examination of CD spectra of the apo-Tat protein in the presence of increasing concentrations of Zn(II) at pH 4.0 show no significant change in ellipticity at 199 nm indicative of little change in the protein's conformation upon binding with the metal. Interestingly, at pH 5.0, a significant reduction in ellipticity at 0.5 mole equivalents of Zn(II) is observed suggesting that a change in the ionization states of certain residues is significantly influencing the metal-binding properties of Tat. The  $pK_a$ 's of Asp, Glu, His, and Cys are 3.7, 4.3, 6.0, and 8.2, respectively suggesting an order of likelihood of interaction with the metal as the pH is increased. Ordinarily, one would not expect cysteine thiols to interact with Zn(II) at pH 5.0 however, it is possible that one or more cysteines have a lower  $pK_a$  and could be ionized at pH 5.0. This is suggested by the poor solubility of the protein above pH 4.5. The loss in ellipticity at 199 nm suggests a loss of random coil structure upon the binding of Zn(II). The inflection point in the titration curve at 0.5 mole equivalents of metal added supports the suggestion of the formation of a metal-linked dimer. The ellipticity changes at higher mole equivalents show that additional Zn(II) binding is possible.

The CD spectra of the Tat protein in the presence of increasing concentrations of Zn(II) at pH 7.2 show significant losses in ellipticity at each metal addition suggestive of metal-induced reduction in random coil content. The fact that the ellipticity losses are greatest after addition of 0.5 mole equivalents of Zn(II) also supports the formation of a Zn(II)-linked dimer at pH 7.2. The larger magnitude of the ellipticity change at pH 7.2 compared to pH 5.0, suggests that the binding affinity is higher resulting in a larger conformational change at higher pH. In view of the known high affinity of histidine and cysteine side chains for Zn(II) it seems likely that the His tag and Cys-rich regions are the sites of Zn(II) binding. Previously published CD spectra of HIV-1 Tat in the presence of divalent cations such as Zn(II) or Cd(II) agree with our results. In particular, a CD study of the Tat peptide comprised of the cysteine-rich region (Tat<sub>21-38</sub>) showed distinct losses in the ellipticity at 199 nm upon addition of 1 and 2 mole equivalents of Zn(II) [41]. It is noteworthy that the CD spectrum of apo-Tat<sub>21-38</sub> showed a very similar band at 199 and 230 nm to our Tat protein.

#### *4.3.4 Effects of Cd(II) at pH 4.0, 5.0, and 7.1*

The Tat protein's conformation in the presence of Cd(II) was examined using far-UV CD spectropolarimetry. Based on competition experiments, Tat has been shown to have a higher binding affinity towards Cd(II) than Zn(II) using UV absorption spectroscopy [39]. Examination of the CD spectra of apo-Tat and the Tat-Cd(II) complex at pH 3.95 shows only very small decreases in ellipticity suggesting that a

very weak metal-binding interaction has only slightly changed the conformation of the protein. This agrees with the CD results of apo-Tat in the presence of Zn(II) at pH 4.0 which also showed no change in ellipticity. A previously published CD study on Tat in the absence and presence of CdCl<sub>2</sub> reported by Frankel et al. showed that the addition of Cd(II) did not lead to a noticeable change in the ellipticity of the Tat spectrum at pH 7.2. However, the addition of 6 M urea to the apo-Tat prior to the metal titration experiment may have prevented binding of metal [39].

Addition of Cd(II) to Tat at pH 5.0 resulted in a major loss of ellipticity at 0.5 mole equivalents of metal added and only minor changes at higher pHs in reasonably good agreement with the Zn(II) titration at the same pH. A CD study of the metal-binding properties of Cd(II) to peptides rich in cysteine residues showed significant structural changes as observed by the decrease in negative ellipticity at 199 nm. This supports our suggestion that decreases in ellipticity at 199 nm reflect conformational changes that convert random coil backbone into a conformation suitable for the binding of metal.

CD spectral analysis of apo-Tat in the presence of 0.5 to 3.0 mole equivalents of Cd(II) at pH 7.12 showed small but gradual changes in ellipticity at low metal concentration followed by a significant change at higher metal concentration [see Figure 3.16 in **Results**]. This experiment was hampered by the low concentrations of Tat which, in hindsight, appear to be the result of the method used to prepare the protein. In this experiment the protein was dissolved up in a pH 7.12 buffer solution and metal was then added. The result was a low concentration of protein

making it difficult to measure the effects of added metal. In contrast, the experiments with Zn(II) involved dissolving up the protein at pH 4.0, raising the pH, and then adding metal. This approach resulted in much higher concentrations of protein so it would be very worthwhile to repeat the Tat titration with cadmium at pH 7 by preparing the protein-metal mixtures at pH 4.0 followed by elevating the pH.

## 4.4 Nuclear Magnetic Resonance Spectroscopy

### 4.4.1 $^1\text{H}$ - $^{15}\text{N}$ Heteronuclear Single Quantum Correlation Spectroscopy

$^1\text{H}$ - $^{15}\text{N}$  HSQC NMR spectral analysis of unlabeled Tat protein showed clustering of similar amino acid residues within a narrow chemical shift range indicative of a disordered protein [183]. For example, resonances corresponding to glycine residues of the Tat protein were observed within the narrow range of 8.0-8.5 ppm ( $^1\text{H}$  dimension) [see Figure 3.5.2 in **Results**], which closely matches the random coil chemical shift values for glycine residues of a model peptide (Ac-G-G-X-G-G-NH<sub>2</sub>) in the presence of acidic 8 M urea reported by Schwarzsinger et al. [177]. The conclusion that apo-Tat is a disordered protein based on  $^1\text{H}$ - $^{15}\text{N}$  HSQC data is also substantiated by previous NMR and CD results showing that the Tat protein's conformation is predominantly a random coil [26].

#### 4.4.2 Effect of Zn(II)

In light of the CD experiments that showed Zn(II) and Cd(II) binding by Tat at a pH of greater than or equal to five, NMR spectroscopy was applied to determine the nature of the conformational change and the identity of the metal binding residues. Because many attempts to solubilize Tat at pH 7 at the relatively high concentrations necessary for NMR spectroscopy were unsuccessful, NMR spectra could only be measured at low pH values. Examination of the  $^1\text{H}$ - $^{15}\text{N}$  HSQC NMR spectra of apo-Tat and Tat at pH 4.02 in the presence of 0.5 mole equivalents of Zn(II) reveals significant line broadening of all cross-peaks. A possible interpretation of the observation of line broadening is a change in the kinetics of dynamic equilibrium among multiple conformations induced by the addition of metal [178]. It can be rationalized that fast dynamic inter-conversion between different conformations on the nano-second to pico-second timescale leads to an averaging of all the contributing resonances whereby a single resonance is observed in the apo-protein[12]. Addition of metal appears to slow down the averaging bringing the exchange into the microsecond to millisecond timescale where it results in appreciable line-broadening. However, analysis of the amide intensity along the entire length of the protein showed several resonances dramatically diminished in intensity as a result of Zn(II) addition [see Figure 3.20 in **Results**]. In particular, amino acid residues within the cysteine-rich region were shown to disappear. Resonance intensity comparisons between the apo-Tat and Tat-Zn(II) spectra suggest that the effect of Zn(II) on the intensities of the peaks in the Cys-rich

segment is significantly greater than the effect on the rest of the protein. This suggests a specific interaction between the cysteine side-chains and the metal. The interaction must be very weak however, as the global conformation of the protein is unchanged according to CD spectropolarimetry and indicates that at pH 4 the metal affects the dynamics of protein conformational interconversion but not the average fold measured by CD. Indeed, it is surprising that the changes in the Cys-rich region are greater than in the acidic segment considering the likely ionization states of the side-chains at pH 4.0, namely partially deprotonated for Asp and Glu and fully protonated for Cys. Another possible explanation for the overall line-broadening is metal-induced aggregation of the Tat protein [185]. However, this is unlikely because the Tat protein is known to be stable at low pH [55] and no precipitation was observed in the NMR sample.

Spectral superposition of the apo-Tat HSQC spectrum and the Tat spectrum in the presence of 0.5 mole equivalents of Zn(II) shows observable but small chemical shift changes within the histidine-tag and the Arg-rich motif indicative of the possible involvement of the histidine and arginine residues as metal ligands or of an indirect conformational effect of metal interactions with the Asp, Glu, or Cys residues. This latter rationalization is supported by the observation of chemical shift changes for residues Glu-22, Glu-29, and Asp-25. It is difficult to rule out a direct interaction with these residues and indeed the metal may be dynamically exchanging among many sites on the protein at low pH. It is noteworthy however, that these residues are only found within the N-terminus region and may function

as a stabilizing factor due to its proximity to the His-tag. In an NMR study of Bacitracin by Wasylishen et al. [184], it was shown that the binding of Cu(II) and Mg(II) by histidine residues is stabilized by the carboxylic acid groups of glutamic and aspartic acid. It is plausible that even at low pH, the histidine residues are permitted to bind to Zn(II) with the help of stabilizing interactions through residues from the N-terminus. The side chains of glutamic and aspartic acid have been shown to stabilize the histidine-Zn(II) interaction by the formation of a hydrogen-bond between the  $N_{\delta 1}-H_{\delta 1}$  and the oxygen atom of the carboxylates [85] (Figure 4.2). At pH 4, the Asp (3.8) and Glu (4.3)  $pK_a$ 's would permit partial negative charges on their side-chains that might encourage the displacement of a proton from the histidine imidazole by a zinc atom.

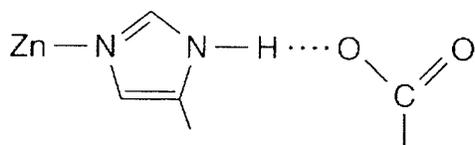


Figure 4.2: Schematic representation of a hydrogen bond between a carboxylic group of glutamic or aspartic acid with the imidazole ring of the histidine which is directly coordinated with zinc. Illustration was reprinted with permission of Auld et al. [85].

Another region of the Tat protein that may be involved in Zn(II) binding is the Arg-rich motif since chemical shift changes were observed for these residues. It is unclear how the arginine residues could play a role in the Tat protein binding to

Zn(II) but several possibilities do exist. The  $pK_a$  of the arginine side chain is 12.5, which at low pH is positively charge (protonated). It is noteworthy that the side chain of the arginine residue resembles that of guanidine, which is rarely found to function as a metal ligand. However, in a study of the stereochemistry of guanidine-metal interaction by Costanzo et al [186], it was shown by analysis of several molecule crystal structures that the neutral guanidine form is capable of metal coordination interactions with preference for in-plane coordination. Thus, it is highly unlikely that the arginine residues are directly coordinating with zinc at low pH. In Figure 4.3, a model diagram of the possible dimerization of the Tat protein in the presence of metal ions is proposed to exist in an anti-parallel orientation [61]. One can speculate that the arginine residues help stabilize this orientation through electrostatic interaction with the negative residues found within the N-terminus.

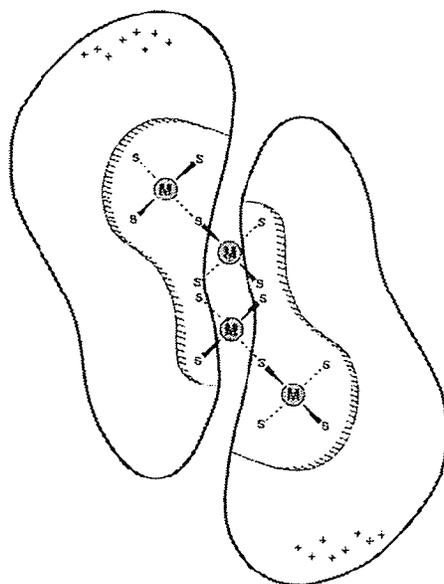


Figure 4.3: Model representation of a Tat metal-linked dimer. The M and S denote the metal ions and thiolate ligands, respectively. Each monomer contributes seven cysteine residues. The orientation of the dimer is anti-parallel with the basic region (Arg-rich) represented by positive charges. Illustration was reprinted with permission of Frankel et al. [61].

Despite observable changes in chemical shift and intensity of cross-peaks of the NMR spectra of Tat in the absence and presence of Zn(II) at pH 4.0, CD spectra showed no significant change to Tat's conformation in the presence of the metal at the same pH. A plausible explanation for the discrepancy in the two methods is that the protein concentration is much higher in the NMR experiments (on average 500  $\mu\text{M}$ ) than in the CD experiments (on average 160  $\mu\text{M}$ ). Thus, it is conceivable that the effects observed by NMR depend on the high concentrations of the protein. It

could also be that the Zn(II) is weakly coordinating with the Tat protein in a way that does not change the protein's conformation.

#### *4.4.3 $^1\text{H}$ - $^1\text{H}$ Nuclear Overhauser Effect Spectroscopy*

##### 4.4.3.1 NOESY vs. ROESY

Two-dimensional nuclear Overhauser effect (NOESY) spectroscopy is a key method in elucidating the 3D structure and folding of proteins. In general, NOESY NMR provides information for the determination of sequential connectivities and identifies long range contacts. Spectral comparison between NOESY and ROESY experiments of the Tat protein show significantly more NOE cross-peaks in the former than the latter. Comparison of peak intensity between the NOESY and ROESY experiments is valid as long as appropriate consideration is given to the fact that the maximum intensity from both experiments on the same protein sample will be different. However, based on NOE results from the two experiments, it can be concluded that the Tat protein's rotational dynamics do not fall within the regime between fast and slow rotational correlation times that often results in the NOE of zero. Furthermore, one can speculate that the Tat's rotational correlation time may fall within the regime corresponding to large molecules where the NOE intensity of the NOESY experiment is larger than the ROESY experiment [36].

#### 4.4.3.2 NOESY Analysis of Apo-Tat and Tat-Zn(II) Complex

Comparison of the NOESY spectra of apo-Tat and Tat at pH 4 in the presence of 0.5 mole equivalents of Zn(II) showed no significant difference in the number of observable cross-peaks. To understand the structural effects of Tat-Zn(II) binding, comparison of the Tat-Zn(II) NOESY results with the literature is needed. Unfortunately, very few papers have been published on NOESY NMR studies of the HIV-1 Tat protein. In a NOESY NMR study of the equine infectious anemia variant (EIAV) 1-72 Tat by Sticht et al. [187], it was shown that the protein exists in a predominantly helical conformation in a solution of 40 % trifluoroethanol/water but this is clearly not helpful to us. Interestingly, numerous publications on the use of NOESY NMR to study the Tat-TAR interactions exist in literature but the primary focus is on the TAR spectrum [188, 189]

Typically, the effect of metal-binding on a protein usually results in the appearance of additional cross-peaks in the NOESY spectrum indicating a metal-induced conformational folding [190]. However, this may not always be the case. In a NOESY NMR study of the metal-binding properties of the P5 helix of the group 1 intron [191] with  $\text{Co}(\text{NH}_3)_6^{3+}$ , the observation of significant chemical shift changes ( $^1\text{H}$ - $^{15}\text{N}$  HSQC) were accompanied by no dramatic changes in the intensity and pattern of the NOESY spectrum indicating that the binding of  $\text{Co}(\text{NH}_3)_6^{3+}$  had no influence on the conformation of the protein. The fact that significant changes in the chemical shift in the HSQC data of the Tat-Zn(II) complex were observed but no additional cross-peaks in the NOESY spectrum were observed suggests that no

dramatic conformational changes were induced upon interaction with Zn(II) and this agrees very well with the CD results. It should also be noted that the NOESY results do not support the speculation of a metal-induced Tat dimer. In theory, dimerization of two Tat monomers would yield additional NOESY cross-peaks due to the probability of two protons on each monomer to come close in proximity. However, this is not surprising given the low pH at which the experiments were conducted. Also, the lack of any difference in the NOESY spectra of apo-Tat and Tat-Zn(II) contradicts the notion of the stabilizing electrostatic interaction between the arginine-rich and the N-terminal regions.

Another possible explanation for the inconclusive results is that the dynamics of the disordered Tat protein may lead to the observation of fewer NOE cross-peaks relative to proteins that have a definitive secondary structure. It is most likely that the Tat protein is interacting with Zn(II), however, the NOESY NMR experiment is an inadequate method to detect such minute structural changes in the Tat protein. It is noteworthy that it was not practical to measure NOESY spectra of Tat bound to Zn(II) at higher pH because of protein precipitation.

#### 4.4.3.3 Analysis of NOESY Spectra at Different Mixing Times

Based on comparing the number of NOE cross-peaks and their relative intensities, the optimum mixing time was determined to be 300 ms. However, it should be pointed out that NOESY distance determination depends on the NOE intensity only on the assumption that the intensity is proportional to the rate of NOE build-up. In

practice, NOE intensities can be used once it has been determined that the measurements are in the range in which the build-up rate is linear, due to the effects of spin diffusion [110]. Spin diffusion can dramatically influence the intensity of the observed NOE through cross relaxation. Consider a system of three spins A, B, and C. A two-step cross-relaxation pathway from spin A to B followed by B to C may be more efficient than directly going from A to C and will distort the distance information in the cross-peak intensities at long mixing times. Nevertheless, subtle spin diffusion effects would not be expected to alter significantly the conclusions made here based on the similarity of the measured NOESY and ROESY experiments.

#### 4.4.3.4. NOESY and HSQC Analysis of Dialyzed Tat Sample

A NOESY analysis of a Tat protein that was dialyzed against Ni-free NTA to remove residual cobalt was carried out to confirm the absence of paramagnetic relaxation that may lead to broadening of cross-peaks. Comparison of NOESY spectra of apo-Tat dialyzed against Ni-free NTA and apo-Tat purified under standard conditions showed no significant increase in the number of observable cross-peaks suggesting that no residual cobalt remained attached to the Tat protein. Although the two samples were of different concentrations causing the comparison to be qualitative, no large effect of dialysis was evident. An alternative method to study the paramagnetic influence of residual cobalt was to conduct a  $^1\text{H}$ - $^{15}\text{N}$  HSQC experiment on the dialyzed Tat sample and compare the relative intensities of the

resonances with the Tat sample purified under standard conditions, paying particular attention to the cysteine resonances. It is noteworthy that cross-peak intensities were normalized against Leu-89. Interestingly, the cross-peaks corresponding to the cysteine residues within the Cys-rich region were absent in the HSQC spectrum of the dialyzed Tat sample, quite the opposite of what was expected. If the dialysis against NTA did remove trace metals this suggests that the intensities of resonances in the Cys-rich region are broadened by a mechanism other than paramagnetic relaxation by loosely held metals.

#### *4.4.4 $^1\text{H}$ - $^{15}\text{N}$ Heteronuclear Multiple Quantum Correlation Spectroscopy*

Based on  $^1\text{H}$ - $^{15}\text{N}$  HSQC results of the Tat-Zn(II) complex, it was suggested that the Zn(II) ion may be coordinating with the histidine residues within the His-tag region. To further investigate the role of the histidine residues in Tat-metal binding, long-range HMBC NMR was carried out to focus on the effects of Zn(II) on the imidazole rings of histidine. A binding interaction between Zn(II) and the histidine residues is suggested by the observation of chemical shift changes and a reduction in cross-peak intensities in the HMBC spectra of apo-Tat and Tat in the presence of increasing Zn(II) concentration [see Figure 3.28B in **Results**]. However, it is difficult to rule out the possibility that the effects are owing to a conformational change induced by binding at another region of the protein.

Based on comparison of the  $^1\text{H}$ - $^{15}\text{N}$  HMBC spectra of apo-Tat with the schematic diagram of the three possible imidazole tautomeric states, together with

the fact that the  $pK_a$  of the side chain is 5.86 [192], indicates that the histidine residues are in a cationic tautomeric state at low pH, as expected. A possible explanation for the disappearance of cross-peaks upon metal addition is the rapid exchange of the metal between different histidine ligands leading to dramatic line broadening. Furthermore, it appears that the cross-peaks are slowly migrating, together with reduction in intensity, towards a single histidine cross-peak pattern during each metal titration point.

#### *4.4.5 $^1H$ - $^{13}C$ HMQC Analysis of Apo-Tat and Tat-Zn(II) Complex*

Because the most dramatic effects of added Zn(II) on Tat  $^1H$ - $^{15}N$  HSQC spectra were observed in the Cys-rich region it seemed appropriate to examine the side-chain resonances of the Cys residues. Specifically, the focus was on the detection of changes in the  $^1H^\beta$ - $^{13}C^\beta$  correlation signals upon metal-binding. Unfortunately, spectral comparison between  $^1H$ - $^{13}C$  HMQC data of apo-Tat and Tat-Zn(II) showed only minor changes in the chemical shifts of cross-peaks that are found within regions where the cysteine residues would be expected to be observed and raises doubts about the possibility that the cysteine residues are directly interacting with the metal. It should be kept in mind that the cysteine thiols are expected to be protonated at pH 4 and thus are not expected to interact strongly with metal at low pH.

#### 4.4.6 $^{113}\text{Cd}$ NMR Spectroscopy

Using high-resolution 1D  $^{113}\text{Cd}$  NMR spectroscopy, one can study the binding of Cd(II) to the Tat protein, which may lead to an understanding of the protein's multiple metal-binding sites. Analysis of the  $^{113}\text{Cd}$  titration experiment of the Tat protein shows a hyperbolic relationship between  $^{113}\text{Cd}$  chemical shift and mole equivalents of metal added. This is characterized by an initial increase in chemical shift followed by a plateau at higher metal concentrations. The  $^{113}\text{Cd}$  NMR spectra in the presence of the Tat protein show significant chemical shift changes upon each titration of metal indicative of metal-binding. The cadmium chemical shift of Tat in the presence of 1 mole equivalent of Cd(II) is approximately 8.5 ppm, which is substantially different from the literature values reported by Otvos et al. for the Cd<sub>7</sub>-metallothionein-2 (Cd-MT-2) protein complex [217]. In this study,  $^{113}\text{Cd}$  NMR chemical shifts for Cd-MT-2 were found to be in the range of 600 – 680 ppm and diagnostic of cysteine thiolate binding. At 8 ppm the cadmium chemical shift suggests binding by oxygen and/or nitrogen ligands. Another possible explanation for the dramatic difference in observed chemical shift relative to the literature values is dynamic chemical exchange between the free-Cd(II) (0 ppm) and Cd(II) bound to the Tat protein with a strong shift towards the free form. This fast exchange would average the two chemical shifts. The speculation of dynamic exchange is substantiated by a study of the cadmium interaction of the extracellular organic matter (EOM) excreted by *Selenastrum capricornutum* [194] in which it was determined that the observed chemical shift for Cd(II)-EOM arises from an

equilibrium between free and bound Cd(II) species. It is noteworthy that the range in chemical shifts of Cd(II) bound to EOM is 1.3 to 6.6 ppm. Furthermore, the observed Cd(II) chemical shift is also strongly affected by coordination with N, S, and OH ligands.

Further complications in the analysis of  $^{113}\text{Cd}$  NMR results arise if the dynamic exchange is between multiple species of Cd(II) complexes with binding ligands or water [112]. A possible solution to overcome fast exchange resonance averaging is to alter the time scale of the exchange process to the slow exchange regime by modifying the pH, lowering the temperature, or changing ionic strength conditions of the sample [113].

The observation of sharp resonance lines throughout the Cd(II) titration also indicates the absence of intermediate exchange among different Tat-Cd(II) conformations. Another possible explanation for the inconsistencies of the observed chemical shift is that Cd(II) may be coordinating more strongly to the histidine residues and the carboxyl group of Asp and Glu rather than the cysteine residues. This may explain the observed  $^{113}\text{Cd}$  chemical shift between 50 – 8 ppm. The typical  $^{113}\text{Cd}$  NMR chemical shift of which the coordinating ligand is a nitrogen and oxygen atoms fall in the range of 0 to 300 ppm and -100 to 0 ppm, respectively [112]. Thus,  $^{113}\text{Cd}$  NMR results may be a product of the averaging of the chemical shift between nitrogen and oxygen rather than sulfur ligands.

Algorithms used in Mathematica, together with the chemical shift data taken from the  $^{113}\text{Cd}$  NMR experiment, led to the determination of a metal dissociation

constant of the Tat-Cd(II) complex at pH 4.6 to be 1.81 mM. The dissociation constant of Cd(II) to the Tat protein is relatively low as compared to cadmium-binding protein in rat liver ( $K_D = 1.2 \times 10^{-8} \text{ M}$ ), human serum albumin ( $K_D = 3.8 \times 10^{-5} \text{ M}$ ), and metallothionein ( $K_D < 10^{-11} \text{ M}$ ) [196]. The determined dissociation constant of Tat-Cd(II) may be indicative of a weak binding at low pH. One can assume that as the pH increases, the ionization of important residue side chains changes which may lead to an increase in Tat's metal-binding affinity. Thus, at higher pH, the dissociation constant may decrease.

#### 4.5 UV Spectroscopy of Metal-binding using Cd(II)

Using UV absorption spectroscopy, we can monitor the binding of the Cd(II) to the Tat protein. Examination of the UV absorption titration experiment of the Tat protein showed an initial increase in absorption at 250 nm at a Cd:protein ratio of 0.5:1 at pH 4.5 [see Figure 3.7 in **Results**]. Further increase in Cd(II) concentrations does not lead to significant changes in absorption at 250 nm. The initial increase in absorption at 250 nm is suggestive of a thiolate-to-Cd(II) charge transfer complex [195]. In a study of the Cd(II)-complexes with mammalian  $\beta$ -metallothionein-1 protein reported by Tio et al., it was observed that in the presence of 0.5 mole equivalents of Cd(II), the UV absorption difference underwent a significant increase at 250 nm accompanied by small changes at higher mole equivalents of Cd(II). This is indicative of the sulfur ligand binding to the Cd(II)

with saturation of metal-binding sites [161]. It also substantiates the notion that the Tat protein is capable of binding Cd(II) with low affinity, at pH below 6.

A discrepancy arises in the comparison between the UV absorption and  $^{113}\text{Cd}$  NMR results. In the UV analysis, the Tat-Cd(II) complex saturates at 0.5 mole equivalents of Cd(II) whereas the  $^{113}\text{Cd}$  NMR titration shows a gradual increase in chemical shift change at higher metal concentrations. A plausible explanation is that the UV analyses only showed the binding interaction between the Tat protein and the sulfur groups of the cysteine residues whereas the  $^{113}\text{Cd}$  NMR experiment characterized the interaction of Cd(II) to all the Tat binding sites, suggesting that in addition to binding to thiolates Cd(II) also binds weakly to carboxylates and other side-chains. This is supported by the chemical shift range of the  $^{113}\text{Cd}$  during the titration. However, the saturation of the UV absorption change at 0.5 mole equivalents of Cd(II) further support the suggestion of a metal-linker dimer complex.

### **Future Perspectives**

Based on the NMR and CD results of the Tat protein interaction with divalent cations Zn(II) and Cd(II), further experiments are required to fully understand the binding kinetics and structural changes. Using Biacore (Surface Plasmon Resonance) as an analytical technique, one could study the binding kinetics of the Tat protein and determine whether or not a dimeric Tat can be formed. One

possibility would be to attach a biotinylated Tat monomer to the Biacore sensor chip SA via interaction with the streptavidin-gold-dextran surface. Initial binding of the Tat protein onto the surface is measured in resonance units (RU). A microfluidic system containing a solution of the Tat monomer and Zn(II) or Cd(II) would be injected and flow over the solid Tat gold surface. The interaction between the Tat monomers will change the mass of the Tat monomer bound to the surface which will lead to changes in the RU over time [201]. The quantitative measurement of the dissociation and association rates of the two Tat monomers in the presence of Zn(II) or Cd(II) can then be converted into a binding affinity constant. In a Biacore study of the killer cell Ig-like receptor (KIR), it was shown that Zn(II) induces the multimerization of KIR proteins [202]. It is noteworthy that the idea of only using a solution of Zn(II) or Cd(II) as the mobile phase and studying the binding affinity of the Tat-Zn(II) or Tat-Cd(II) complex using Biacore is not feasible due to the inability to detect the low  $M_r$  of the Zn(II) and Cd(II).

The next step towards further understanding the Tat-Cd(II) complex should be to conduct  $^{113}\text{Cd}$ - $^1\text{H}$  HMQC NMR experiments that may identify ligands coordinated to Cd(II) through detection of long range couplings. This may be feasible at neutral pH if the solution is prepared by titration of the Tat-Cd(II) complex from low pH to neutrality. In an NMR study of the  $\text{Cd}_6$ -metallothionein complex [203], it was shown that the  $^{113}\text{Cd}$ - $^1\text{H}$  HMQC method can identify the Cd(II)-coupled to the  $\text{H}^\beta$  of the cysteine residues. Furthermore, the  $^{113}\text{Cd}$ - $^1\text{H}$  HMQC method can be useful in defining the geometry of the binding site, which can provide additional 3D

information. In a study of the metalloproteins by Zerbe et al. [204], it was determined that a Karplus-type correlation exists between  $H^\beta-C^\beta-S^\gamma-Cd(II)$  dihedral angle and the magnitude of the  ${}^3J({}^{113}Cd, {}^1H)$  coupling constant for the  $H^\beta$ . Thus, once the magnitude of the coupling constant is determined, one can determine the dihedral angle.

## 5. Conclusions

The focus of this research is to study the influence of metals on Tat's secondary structure and to characterize the regions within the protein that are believed to be involved in metal-binding. It is plausible that in the presence of metals such as Zn(II) and Cd(II), the Tat protein may adopt a stable conformation. However, it is more likely that the divalent cations will influence the protein's rate of conformational change.

The efficient purification of the HIV-1 Tat protein via cobalt metal-affinity chromatography depends on the presence of a reducing agent TCEP and 6 M guanidine-HCl. TCEP assists in the purification process by preventing oligomerization of the Tat protein while guanidine-HCl inhibits precipitation due to the formation of Tat-TCEP complexes. Based on SDS-PAGE analysis, the Tat protein predominantly exists in a monomeric state in the presence of 10 mM TCEP, indicating that the formation of disulfide-bonds between inter and intra-molecular cysteine residues within the Cys-rich region are responsible for oligomerization of the protein.

Results from CD conformational studies of Tat in the presence of Zn(II) and Cd(II) revealed distinct changes that may provide structural insights into the protein-metal complex. Clearly, the Tat protein has a very low binding affinity for both Zn(II) and Cd(II) at low pH. As the ionization of the side chain changes, the binding affinity of the Tat protein towards the metal increases. More specifically,

the Tat protein's conformation changes significantly in the presence of 0.5 mole equivalents of either Zn(II) or Cd(II) at pH 5.0. Furthermore, as the pH increases to 7.0, an increase in binding of the Tat protein toward Zn(II) is observed.

NMR investigations of Tat in the presence of Zn(II) and Cd(II) revealed minor structural changes in regions that are most likely to be involved in metal-binding. Results from  $^1\text{H}$ - $^{15}\text{N}$  HSQC NMR experiments of Tat in the presence of Zn(II) revealed dynamic exchange between different cysteine residues believed to be involved in metal binding. Furthermore, histidine residues within the His-tag underwent significant chemical shift changes at each addition of Zn(II). In addition to the  $^1\text{H}$ - $^{15}\text{N}$  HSQC, the  $^1\text{H}$ - $^{15}\text{N}$  HMBC experiment substantiates the rationale that the histidine residues within the His-tag are involved in metal-binding. In the  $^1\text{H}$ - $^{13}\text{C}$  HMQC experiment, small changes in chemical shifts of cross-peaks found within regions predicted to belong to cysteine residues are observed, providing additional evidence of the involvement of the Cys-rich region in metal-binding. Using  $^{113}\text{Cd}$  NMR the binding interaction of the Tat protein towards Cd(II) is characterized as a weak binding with a dissociation constant of 1.81 mM. Furthermore, the observed chemical shift range of the experiment indicates the involvement of oxygen and nitrogen ligands in metal-binding.

In summary, results from NMR, CD, and UV studies showed that the Tat protein undergoes significant structural change only in the presence of 0.5 mole equivalents of metal with the His-tag and Cys-rich region most likely to be directly involved. It is plausible that the divalent cations induce the formation of an anti-

parallel Tat dimer with a protein:metal ratio of 2:1. Unfortunately, the formation of a Tat dimer did not help stabilize the protein's conformation at higher pH. Thus, the need for further study on stabilizing the Tat protein at neutral pH is needed.  $^{113}\text{Cd}$  NMR analysis reveals a weak binding affinity in the millimolar range. This may be due to the fact that in the presence of metal, the Tat protein is undergoing conformational exchange in the intermediate regime between monomeric and dimeric Tat conformers.

## References

1. Coffin, J.; Haase, A.; Levy, J. A.; Montagnier, L.; Oroszlan, S.; Teich, N.; Temin, H.; Toyoshima, K.; Varmus, H.; Vogt, P.; Weiss, R. A. *Nature* **1986**, *321*, 10.
2. Weiss, R. A. *Science* **1993**, *260*:5112, 1273–1279.
3. Barre-Sinoussi, F.; Chermann, J. C.; Rey, F.; Nugeyre, M. T.; Chamaret, S.; Gruest, J.; Dauguet, C.; Axler-Blin, C.; Vezinet-Brun, F.; Rouzioux, C.; Rozenbaum, W.; Montagnier, L. *Science* **1983**, *220*, 868–871.
4. Reeves, J.D.; Doms, R.W. *Journal of General Virology* **2002**, *83*, 1253-1265.
5. Freed, O. S. *Cell and Molecular Genetics* **2002**, *26*:1, 13-33.
6. Harrich, D.; McMillan, N.; Munoz, L.; Apolloni, A.; Meredith, L. *Current Drug Targets* **2006**, *7*, 1595-1606.
7. Campbell, N.A.; Reece, J.B. “*Biology*” Sixth Edition, Benjamin Cummings Inc. **2002**, 328-350.
8. Weiss, R.A. *Nature* **2001**, *410*, 963-967.
9. Kulkosky, J.; Bray, S. *Current HIV Research* **2006**, *4*, 199-208.
10. Stevens, M.; Clercq, E.D.; Balzarini, *Journal of Medicinal Research Review* **2006**, *26*:5, 595- 625.
11. Wender, P.A.; Kee, J.M.; Warrington, J.M. *Science* **2008**, *320*, 649-652.
12. Karn, J. *Journal of Molecular Biology* **1999**, 235-254.
13. Cheng, B.; Price, D.H. *Journal of Biological Chemistry* **2007**, *282*:30, 21901-21912.

14. Jeang, K-T.; Xiao, H.; Rich, E.A. *Journal of Biological Chemistry* **1999**, *274:41*, 28837-28840.
15. Smith, S.M.; Pentlicky, S.; Klase, K.; Singh, M.; Neuveut, C.; Lu, C.; Rietz, M.S.; Yarchoan, R.; Marx, P.A.; Jeang, K. *Journal. Biological Chemistry* **2003**, *278:45*, 44816-44825.
16. Ping, Y-H.; Rana, T.M. *Journal of Biological Chemistry* **2001**, *276:16*, 12951-12958.
17. Bugatti, A.; Urbinati, C.; Ravelli, C.; Clercq, E.D.; Liekens, S.; Rusnati, M. *Antimicrob. Agents Chemother.* **2007**, *51:7*, 2337-2345.
18. Fulcher, A.J.; Jans, D.A. *Life* **2003**, *55:12*, 669-680.
19. Wu, Y.; Marsh, J.W. *Virology* **2003**, *77:19*, 10376-10382.
20. Frankel, A.D. *Current Opinion in Genetics and Development* **1992**, *2*, 293-298.
21. Hauber, J.; Malim, M.H.; Cullen, B.R. *Journal of Virology* **1989**, *63:3*, 1181-1187.
22. Wu, Y.; Marsh, J. W. *Microbes and Infection* **2003**, *5*, 1023-1027.
23. Wright, P.E.; Dyson, H.J. *Journal of Molecular Biology* **1999**, *293*, 321-331.
24. Weinreb, P.H.; Zhen, W.; Poon, A.W.; Conway, K.A.; Lansbury, P.T. *Biochemistry* **1996**, *35:43*, 13709-13715.
25. Bracken, C. *Journal of Molecular Graphics and Modelling* **2001**, *19*, 3-12.
26. Shojania, S.; O'Neil, J.D. *Journal of Biological Chemistry* **2006**, *281:13*, 8347-8356.
27. Lu, C.X.; Li, J.; Sun, Y.X.; Wang, Q-J.; Xin, X.L.; Geng, M.Y. *Biochemical pharmacology*, **2007**, *74*, 1330-1339.

28. Albini, A.; Barillari, G.; Benelli, R.; Gallo, R.C.; Ensoli, B. *Proc. Natl. Acad. Sci.* **1995**, *92*, 4838-4842.
29. Albini, A.; Soldi, R.; Giunciuglio, D.; Giraudo, E.; Benelli, R.; Primo, L.; Noonan, D.; Salio, M.; Camussi, G.; Rocki, W.; Bussolino, F. *Nature Medicine* **1996**, *2:12*, 1371-1375.
30. Thomas, K.A. *Journal of Biochemical Chemistry* **1996**, *2:12*, 603-606.
31. Arnau, J.; Lauritzen, C.; Petersen, G.E.; Pedersen, J. *Protein Expression and Purification* **2006**, *48*, 1-13.
32. Jenny, R.; Mann, K.G.; Lundblad, R.L. *Protein Expression and Purification* **2003**, *31*, 1-11.
33. Vendel, A.C.; Lumb, K.J. *Biochemistry* **2003**, *42*, 910-916.
34. Marzio, G.; Tyagi, M.; Gutierrez, M.I. Giacca, M. *Proc. Natl. Acad. Sci.* **1998**, *95*, 13519-13524.
35. Chaloin, O.; Peter, J.C.; Briand, J.P.; Masquida, B.; Desgranges, C.; Muller, S.; Hoebeke, J. *Cell. Mol. Life Sci.* **2005**, *62*, 355-361.
36. Ulich, C.; Dunne, A.; Parry, E.; Hooker, C.W.; Gaynor, R.B.; Harrich, D. *Journal of Virology* **1999**, *73:3*, 2499-2508.
37. Lorey, S.; Stockel-Maschek, A.; Faust, J.; Brandt, W.; Stiebitz, B.; Gorrell, M.D.; Kahne, T.; Mrestani-Klaus, C.; Wrenger, S.; Reinhold, D.; Ansorge, S.; Neubert, K. *Eur. J. Biochem.*, **2003**, *207*, 2147-2156.

- 38.Devadas, K.; Boykins, R.A.; Hardegen, N.J.; Philp, D.; Kleinman, H.K, Osa, E.O., Wang, J.; Clouse, K.A.; Wahl, L.M.; Hewlett, I.K.; Rappaport, J.; Yamada, K.M.; Dhawan, S. *Peptides* **2006**, *27*, 611-621.
- 39.Frankel, A.D.; Bredt, D.S.; Pabo, C.O. *Science* **1988**, *240*, 70-73.
- 40.Bieniasz, P.D.; Grdina, T.A.; Bogerd, H.P.; Cullen, B.R. *The EMBO Journal* **1998**, *17:23*, 7056-7065.
- 41.Huang, H-W.; Wang, K-T. *Biochemical and Biophysical Research Communications* **1996**, *227*, 615-621.
- 42.Chen, D.; Wang, M.; Zhou, S.; Zhou, Q. *The EMBO Journal* **2002**, *21:24*, 6801-6810.
- 43.Agbottah, E.; Zhang, N.; Dadgar, S.; Pumfery, A.; Wade, J.D.; Zeng, C.; Kashanchi, F. *Virology* **2006**, *345*, 373-389.
- 44.Cardarelli, F.; Serresi, M.; Bizzarri, R.; Beltram, F. *Traffic* **2008**, *9*, 528–539.
- 45.Cordingley, M.G. *Proc. Natl. Acad. Sci.* **1990**, *87*, 8985-8989.
- 46.Weeks, K.M.; Ampe, C.; Schults, S.C.; Steitz, T.A.; Crothers, D.M. *Science* **1990**, *249*, 1281-1285.
- 47.Truant, R.; Cullen, B.R. *Molecular and Cellular Biology*, **1999**, *19:2*, 1210-1217.
- 48.Campbell, G.R.; Pasquier, E.; Watkins, J.; Bourgarel-Rey, V.; Peyrot, V.; Esquieu, D.; Barbier, P.; Mareuil, J.D.; Braquer, D. *Journal of Biological Chemistry* **2004**, *46:12*, 48197-48204.
- 49.Green, M.; Loewenstein, P.M.; *Cell* **1988**, *55*, 1179-1188.

50. Xiao, H.; Neuveut, C.; Benkirane, M.; Jeang, K-T. *Biochemical and Biophysical Communications* **1998**, *244*, 384-389.
51. Mahlknecht, U.; Dichamp, I.; Varin, A.; Lint, C.V.; Herbein, G. *Journal of Leukocyte Biology* **2008**, *83*, 718-727.
52. Brake, D.A.; Debouck, C.; Biesecker, G. *Journal of Cell Biology* **1990**, *111*, 1275-1281.
53. Barillari, G.; Gendelman, R.; Gallo, R.C.; Ensoli, B. *Proc. Natl. Acad. Sci.* **1993**, *90*, 7941-7945.
54. Peloponese, J-M.; Gregoire, C.; Opi, S.; Esquieu, D.; Sturgies, L. Lebrun, E.; Meurs, Y.; Olive, D.; Aubertin, A-M.; Witvrow, M.; Pannecouque, C.; Clercq, E.D.; Bailly, C.; Lebreton, J.; Loret, E.P. *Medical Science* **2000**, *323*, 883-894.
55. Shojania, S.; O'Neil, J.D. *Journal Biological Chemistry* **2006**, *281:13*, 8347-8356.
56. Bogerd, H.P.; Fridell, R.A.; Blair, W.S.; Cullen, B.R. *J. Virol.* **1993**, *67:8*, 5030-5034.
57. Dudev, T. *Chem, Rev.* **2003**, *103*, 773-787.
58. Vallee, B.; Galdes, A. *The metallobiochemistry of Zinc Enzyme* 284-283
59. Qian, C.; Yao, Y.; Tong, Y. *J. Biol Inorg. Chem.* **2003**, *8*, 394-400.
60. Sambrook, J.; Russell, D.W.; "Molecular Cloning" **2001**, *Third Edition, Vol. 3*, Appendices A8.40
61. Frankel, A.D.; Bredt, D.; Pabo, C.O. *Science* **1988**, *240*, 70-73.
62. Martin, S.R.; Schilstra, M.J. *Methods in Cell Biology* **2008**, *44*, 263-293.
63. Weissman, J.S.; Kim, P.S. *Science*, **1991**, *253*, 1386-1393.

64. Wuthrich, K. *Jour. Biol. Chem.* **1990**, *255:36*, 22059-22062.
65. Hetzer, C.; Dormeyer, W.; Schnolzer, M.; Ott, M. *Microbes and Infection* **2005**, *7*, 1364-1369.
66. Mandal P.K. and A. Majumdar *Magnetic Resonance Part A* **2003**, *20A*, 1-23.
67. Lee, S., A.I. Arunkumar, X. Chen and D.P. Giedroc. *Journal of American Chemistry Society* **2006**, *128*, 1937-47.
68. Pelton, J.G., D.A. Torchia, N.D. Meadow, S. Roseman *Protein Sci.* **1993**, *2*, 543-558.
69. Yee, A.A.; O'Neil, J.D. *Biochemistry* **1992**, *31:12* 3135-3143.
70. Lazlo P. "NMR of Newly Accessible Nuclei" Vol 2. Academic Press **1999**, 337-363.
71. Ahmadibeni, Y., M. Hanley, M. White, M. Ayrapetov, X. Lin, G. Sun, K. P. *ChemBioChem*, **2007**, *8*, 1592-05.
72. Tubek, S.; Grzanka, P.; Tubek, I. *Biol. Trace. Elem. Res.* **2008**, *121*, 1-8.
73. Vallee, B.L.; Falchuk, K.H. *Physiological Reviews* **1993**, *79:40*.
74. Price, C.A.; Vallee, B.L. *Plant Physiol.* **1962**, *37*, 428-433.
75. Gordon, P.R.; Woodruff, C.W.; Anderson, H.L.; O'Dell, B.L. *The American Journal of Clinical Nutrition* **1982**, *35*, 113-119.
76. King, L.E.; Osati-Ashtiani, F.; Fraker, P.J. *American Society for Nutritional Sciences* **2002**, 974-979.
77. Smith, A.P.; Lee, N.M. *Amyotrophic Lateral Sclerosis*, **2007**, *8*, 131-143.
78. Petersen, L.C.; Olsen, O.H.; Nielsen, L.S.; Freskgard, P.O.; Persson, E. *Protein Science* **2000**, *9*, 859-866.

79. Falchuk, K.H.; Vallee, B.L. *Biochemistry* **1975**, *14*, 3439-3444.
80. Cousins, R.J. *Metal Metabolism and Disease* **1985**, *4*, 20-30.
81. Richards, M.P.; Cousin, R.J. *Biochemical and Biophysical Research Communications* **1977**, *75:2*, 286-294.
82. Knapp, B. "Zinc, Cadmium, and Mercury" Atlantic Europe Publishing Company Ltd. **1996**, 12-40.
83. Waisberg, M.; Joseph, P.; Hale, B.; Beyersmann, D. *Toxicology*, **2003**, *192*, 95-117.
84. Habeebu. S.S.; Liu, J.; Klaasen, C.D. *Toxicology and Applied Pharmacology* **1998**, *149*, 203-209.
85. Auld, D.S.; *Biometals* **2001**, *14*, 271-313.
86. Viscidi, R.P.; Mayur, K.; Lederman, H.M.; Frankel, A.D. *Science* **1989**, *246*, 1606-1608.
87. Li, C.J. Friedman, D.J. Wang, C.; Metelev, V.; Pardee, A.B. *Science* **1995**, *268*, 21, 429-431.
88. Westendorp, M.O.; Frank, R.; Ochsenbauer, C.; Stricker, K.; Dhein, J.; Walczak, H.; Debatin, K-M. I Krammer, P.H. *Nature* **1995**, *375*, 8, 497-500.
89. Bansal, A.K.; Mactutus, C.F.; Nath, A.; Maragos, W.; Hauser, K.F.; Booze, R.M. *Brain Research* **2000**, *879*, 42-49.
90. Cheng, J.; Nath, A.; Knudsen, B.; Hochman, S.; Geiger, J.D.; Ma, M.; Magnuson, D.S.K. *Neuroscience* **1998**, *82:1*, 97-106.

91. Nath, A.; Conant, K.; Chen, P.; Scott, C.; Major, E.O. *The Journal of Biological Chemistry* **1999**, *284:24*, 17098-17102.
92. Roshal, M.; Zhu, Y.; Planelles, V. *Apoptosis* **2001**, *6*, 103-116.
93. Bartz, S.R.; Emerman, M. *Journal of Virology* **1999**, *73:3*, 1956-1963.
94. Mann, D.A.; Frnakel, A.D. *The EMBO* **1991**, *10:7*, 1733-1739.
95. Rice A.P.; Chan, F. *Virology* **1991**, *185*, 451-545.
96. Vallee, B.L. Galdes, A. "The Metallobiochemistry of Zinc Enzymes", 284-401.
97. Alain, M.; Moulis, J-M. *Journal of Inorganic Biochemistry* **2004**, *98*, 1412-1420.
98. Pan, T.; Coleman, J.E. *Biochemistry* **1991**, *30:17*, 4212-4222.
99. Pan, T.; Coleman, J.E.; *Proc. Natl. Acad. Sci.* **1990**, *87*, 2077-2081.
100. Victory, W.W. *American Journal of Physiology* **1981**, *240:4*, 299-305.
101. Walravens, P.A. *Clinical Nutrition* **1979**, *130:2*, 134-142.
102. Yamataka, A.; Pringle, K.C.; Wyeth, J. *Journal of Pediatric Surgery* **1998**, *33:4*, 660-662.
103. Brown, R.S.; Sander, C.; Argos, P. *FEBS* **1985**, *186:2*, 271-274
104. Iuchi, S.; Kuldell, N. "Zinc Fingers Proteins: From Atomic Contact to Cellular Function" **2005**, Landes Bioscience, Eurekah.com and Kluwer Academic, Plenum Publisher.
105. Stoll, R.; Lee, B.M.; Debler, E.W.; Laity, J.H.; Wilson, I.A.; Dyson, H.J.; Wright, P.E. *Journal of Molecular Biology* **2007**, *372*, 1227-1245.
106. Homans, S.W. *A dictionary of concepts in NMR*. Clarendon Press Oxford, **1989**, 211-220.

107. Neuhaus, D.; Williamson, M.P. *The Nuclear Overhauser Effect in Structural and Conformational Analysis*. A John Wiley & Sons, Inc., Publication. **2000**, second edition, 282-321.
108. Noggle, J.H.; Schirmer, R.E. *The Nuclear Overhauser Effect: Chemical Applications*. Academic Press, New York and London, **1971**, 4-10.
109. Rahman, A-U.; Choudhary, M.I. *Solving Problems with NMR Spectroscopy*, Academic Press, Inc. **1999**, 187-193, 259.
110. Wuthrich, K. *NMR of Proteins and Nucleic Acids*. A Wiley-Interscience Publication. John Wiley & Sons. **1986**, 93-113.
111. Kay, L.E.; Torchia, D.A.; Bax, A. *Biochemistry* **1989**, *28*, 8972-8979.
112. Coleman, J.E. *Methods in Enzymology* **1993**, *227*, 16-43.
113. Vasak, M. *Biodegradation* **1998**, *9*, 501-512.
114. Boulanger, Y.; Armitage, I.M. *Journal of Inorganic Biochemistry* **1982**, *17*, 147-153.
115. Engeseth, H.R.; McMillin, D.R.; Otvos, J.D. *Journal of Biological Chemistry* **1984**, *259*:8, 4822-4826.
116. Bobsein, B.R.; Myers, R.J. *Journal of Biological Chemistry* **1981**, *256*:11, 5313-5316
117. Pelton, J., Torchia, D., Meadow, N., Roseman, S. *Protein Science*, **1999**, *2*, 543.
118. Frankel, A. D.; Pabo, C. O. *Cell* **1988**, *55*. 1198-1193.

119. Ingredients for M9 medium used for expression of isotopically labeled proteins. [http://www.embl-hiedelberg.de/nmr/sattler/lab/protocols/triple\\_6.html](http://www.embl-hiedelberg.de/nmr/sattler/lab/protocols/triple_6.html)
120. Manley, D.; O'Neil, J.D. *Methods in Enzymology* **2003**, *228*, 89-101.
121. Laemmli U. K. *Nature* **1970**, *227:5259*, 680-685.
122. Kay, L. E.; Keifer, P.; Saarinen, T. *J. Am. Chem. Soc.* **1992**, *114*, 10663-0665.
123. Wishart, D. S.; Bigam, C. G.; Yao, J.; Abildgaard, F.; Dyson, H. J.; Oldfield, E.; Markley, J.L.; Sykes, B. D. *J. Biomol. NMR* **1995**, *6*, 135-140.
124. Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol.NMR* **1995**, *6*, 227-293.
125. T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco
126. Glasoe, P., & Long, F. *J. Phys. Chem.* **1960**, *64*, 188-190.
127. Macura, S.; Ernst, R.R. *Molecular Physics* **1980**, *41:1*, 95-117.
128. Bax, A.; Summers, M. *J. Am. Chem. Soc.* **1986**, *108*, 2093-2094.
129. Hwang, T-L, Shaka, J. *Journal of Magnetic Resonance Series A.* **1995**, *112*, 275-279.
130. Bax, A.; Davis, D.G. *Journal of Magnetic Resonance* **1985**, *63*, 207-213.
131. Zhu, G.; Bax, A. *Journal of Magnetic Resonance* **1990**, *90*, 405-410.
132. Krezel, A.; Latajka, R.; Bujacz, G.D.; Bal, W. *Inorg. Chem.* **2003**, *42*, 1994-2003.
133. Martin, S.R.; Schilstra, M.J. *Methods in Cell Biology* **2008**, *84*, 263-293.

134. Arnau, J.; Lauritzen, C.; Petersen, G.E.; Pedersen, J. *Protein Expression and Purification* **2006**, *48*, 1-13.
135. Bochicchio, B.; Tamburro, A.M. *Chirality* **2002**, *14*, 782-792.
136. Ducan, K.E.; Stillman, M.J. *Journal of Inorganic Biochemistry* **2006**, *100*, 2101-2107.
137. Shojania, S. p.H.D Thesis "*Nuclear Magnetic Resonance and Dynamic Characterization of the Intrinsically Disordered HIV-1 Tat Protein*", **2007**.
138. Bertini, I.; Gray, H.B.; Lippard, S.J.; Valentine, J.S. *Bioinorganic Chemistry* **1994**, University Science Books, 37-43.
139. Ahmadibeni, Y.; Hanley, M.; White, M.; Ayrapetov, M.; Lin, X.; Sun, G.; Parang, K. *ChemBioChem* **2007**, *8*, 1592-1605.
140. Sambrook., J.; Russell, D.W. *Molecular Cloning A Laboratory Manual* Cold Spring Harbor Laboratory Press, *3*, **2001**.
141. Shajonia, S. Ph.D Thesis, *Nuclear Magnetic Resonance and Dynamic Characterization of the Intrinsically Disordered HIV-1 Tat protein*, University of Manitoba, **2007**
142. Putnam, C. "Protein Calculator Version 3.3", <http://www.scripps.edu/~cdputnam/protcalc.html>, The Scripps Research Institute, La Jolla, California, **2006**
143. Ladokhin, A.S.; Selsted, M.; White, S.H. *Biochemistry*, **1999**, *38*, 12313-12319.

144. Gregoire, C.; Peloponese, J-M.; Esquieu, D.; Opi, S.; Campbell, G.; Solomiac, M.; Lebrun, E.; Lebreton, J.; Loret, E.R. *Biopolymers (Biospectroscopy)*, **2001**, *62*, 324-335.
145. Lees, J.G.; Miles, A.J.; Wien, F.; Wallace, B.A. *Structural Bioinformatics*, **2006**, *22:16*, 1955-1962.
146. Peloponese J.M.; Gregoire, C.; Opi, S.; Esquieu, D.; Sturgis, J.; Lebrun, E.; Meur, E.; Collette, Y.; Olive, D.; Aubertin, A.M.; Witvrow, M.; Pannecoupque, C.; Clercg, E.D.; Bailly, C.; Lebreton, J.; Lorte, E.P. *Life Sciences* **2000**, *323*, 883-894
147. Blasie, C.A.; Berg, J.M. *Biochemistry* **2004**, *43*, 10600-10604.
148. Ferraroni, M.; Tilli, S.; Briganti, F.; Chegwiddden, W.R.; Supuran, C.T.; Wiebauer, K.E.; Tashian, R.E.; Scozzafava, A. *Biochemistry* **2002**, *41*, 6237-6244.
149. Aoki, S.; Iwaida, K.; Hanamoto, N.; Shiro, M.; Kimura, E.; *J. AM. CHEM. SOC.* **2002**, *124*, 5256-5257.
150. Pelton, J., Torchia, D., Meadow, N., Roseman, S. *Protein Science*, **1999**, *2*, 543.
151. Kay, L. E.; Keifer, P.; Saarinen, T. *J. Am. Chem. Soc.* **1992**, *114*, 10663-0665.
152. Krezel, A.; Hao, Q.; Maret, W. *Archives of Biochemistry and Biophysics* **2007**, *463*, 188-200.
153. Simonson, T.; Calimet, N. *Proteins: Structure, Function, and Genetics*, **2002**, *49*:37-48.
154. Roth, E.J.; Kurz, B.; Liang, L.; Hansen, C.L.; Dameron, C.T. Winge, D>; Smotkin, D. *The Journal of Biological Chemistry* **1992**, *267:23*, 16390-16395.

155. Blindauer, C. *Journal of Inorganic Biochemistry* **2008**, *102*, 507-521.
156. Vallee, B.L. *Neurochem. Int.* **1995**, *27:1*, 23-33.
157. Biological Magnetic Resonance Data Bank obtained from <http://www.bmrb.wisc.edu/> **2008**
158. Razmiafshari, M.; Kao, J.; Avignon, A.; Zawia, N.H. *Toxicology and Applied Pharmacology* **2001**, *172*, 1-10.
159. Fiorito, F.; Herrmann, Damberger, F.F.; Wuthrich, K. *Journal of Biomolecular NMR* **2008**, *42*, 23-33.
160. Goddard, T.D.; Kneller, D.G. "Sparky 3", **2008**, University of California, San Francisco
161. Kagi, J.H.; Vallee, B.L. *Journal of Biological Chemistry* **1961**, *236:9*, 2435-2442.
162. Coleman, J.E. *Methods in Enzymology* **1993**, *227*, 16-43.
163. Saunders, M.; Wishnia, A.; Kirkwood, J.G. *Journal of American Chemistry Society* **1957**, *79*, 3289-3290.
164. Farrow, M.A.; Aboul-ela, F.; Owen, D.; Karpeisky, A.; Beigelman, L.; Gait, M.J. *Biochemistry* **1998**, *37*, 3096-3108.
165. Goodrich, J.A.; Kugel, J.F. "Binding and Kinetics for Molecular Biologist" **2007**, Cold Spring Harbor Laboratory Press, pp 22-33.
166. Battle, D.J.; Doudna, J.A. *RNA* **2001**, *7*, 123-132.
167. Krezel, A.; Latajka, R.; Bujacz, G.D.; Bal, W. *Inorg. Chem.* **2003**, *42*, 1994-2003.

168. Getz, E.B.; Xiao, M.; Chakrabarty, T.; Cooke, R.; Selvin, P.R. *Analytical Biochemistry* **1999** *273*, 73-80.
169. Ruegg, U.T.; Rudinger, J. *Methods in Enzymology* **1977** *47*, 111-116.
170. Mossner, E.; Huber-Wunderlich, M.; Glockshuber, R. *Protein Science* **1998** *7*, 1233-1244
171. Campbell, G. R.; Senkaali, D.; Watkins, J. Esquieu, D.; Opi, S.; Yirrell, D.L.; Kaleebu, P.; Loret, E. *Vaccine* **2007**, *25*, 8441-8447
172. Campbell, G.R.; Pasquier, E.; Watkins, J.; Bourgaral-Rey, V.; Peyrot, V.; Esquieu, D.; Barbier, P.; Mareuil, J.D.; Braquert, D.; Kaleebul, P.; Yirrell, D.L.; Loret, E.P. *Journal of Biological Chemistry* **2004**, *279*, 48197-48204
173. Huang, H-W.; Wang, K-T. *Biochemical and Biophysical Research Communications* **1996**, *227*, 615-621
174. Rayne, F.; Vendeville, A.; Bonhoure, A.; Beaumelle, B. *Journal of Virology* **2004**, *78:21*, 12054-12057
175. Holtzhauer, M. "Basic Methods for the Biochemical Lab", Springer Lab Manual, **2006**, 23-31
176. Levison, M.E.; Josephson, A.S.; Kirchenbaum, D.M. *Experientia* **1969**, *25:2*, 126-127
177. Schwarzingler, S.; Kroon, G.; Foss, T.R.; Wright, P.E.; Dyson, H.J. *Journal of Biomolecular NMR* **2000**, *18*, 43-48
178. Hou, L.; Zagorski, M.G. *Journal of American Chemical Society*, **2006**, *128*, 9260-9261

179. Maret, W.; Larsen, K.S.; Vallee, B.L. *Proc. Natl. Acad. Sci.* **1997**, *94*, 2233-2237
180. Dickens, F. *Biochemistry* **1933**, *XXVII*:72, 1142-1151
181. Kato, T.; Hamada, D.; Fukui, T.; Hayashi, M.; Honda, T.; Murooka, Y.; Yanagihara, I. *FEBS Journal* **2005**, *272*, 2773-2783
182. Loret, E.; goergel, P.; Johnson, W.C.; Ho, P.S. *Proc. Natl. Acad. Sci.* **1992**, *89*, 9734-9738
183. Peti, W.; Smith, L.J.; Redfield, C.; Schwalbe, H. *Journal of Biomolecular NMR* **2001**, *19*, 153-165
184. Wasylishen, R.E.; Graham, M.R. *Can. J. Biochem.* **1975**, *53*, 1250-1254
185. Danielsson, J.; Pierattelli, R.; Banci, L.; Graslund, A. *FEBS Journal* **2007**, *274*, 46-59
186. Costanzo, L.D.; Flores, L.V.; Christianson, D.W. *PROTEINS: Structure, Function, and Bioinformatics* **2006**, *65*, 637-642
187. Sticht, H.; Willbold, D.; Bayer, P.; Ejchart, A.; Herrmann, F.; Rosin-Arbesfeld, R.; Gazit, A.; Yaniv, A.; Frank, R.; Rosch, P. *Eur. J. Biochem.* **1993**, *218*, 973-976
188. Du, Z.; Lind, K.E.; James, T.L. *Chemistry and Biology* **2002**, *9*, 707-712
189. Hamy, F.; Felder, E.R.; Heizmann, G.; Lazdins, J.; Aboul-ela, F.; Varani, G.; karn, J.; Klimkait, T. *Proc. Natl. Acad. Sci.* **1997**, *94*, 3548-3553
190. DeSilva, T.M.; Veglia, G.; Porcelli, F.; Prantner, A.M.; Opella, S.J. *Biopolymers*, *64*:4, 189-197
191. Colmernarejo, G.; Tinoco, I. *J. Mol. Biol.* **1999**, *290*, 119-135

192. Blomberg, F.; Maurer, W.; Ruterjans, H. *Journal of the American Chemical Society* **1977**, *99*:25, 8149-8159
193. Massiah, M.; Matts, J.A.B.; Short, K.M.; Simmons, B.N.; Singireddy, S.; Yi, Z.; Cox, T.C. *J. Mol. Biol.* **2007**, *369*, 1-10
194. Grassi, M.; Minngazzini, M. *Environ. Sci. Technol.* **2001**, *35*, 4271-4276
195. Liu, H-J.; Hupp, J.T.; Ratner, M.A. *J. Phys. Chem.* **1996**, *100*, 12204-12213
196. Aoki, Y.; Suzuki, K.T. *J. Biochem. Toxicol.* **1987**, *2*, 67-71
197. Monera, O.D.; Kay, C.M.; Hodges, R.S. *Protein Science*, **1994**, *3*, 1984-1991
198. Demasi, M.; Filho, G.M.P. Castro, L.M.; Ferreira, J.C.; Rioli, V.; Ferro, E.S. *Free Radical Biology & Medicine* **2008**, *44*, 1180-1190
199. Petersen, S.V.; Oury, T.D.; Valnickova, Z.; Thogersen, I.B.; Hojrup, P.; Crapo, J.D. *PNAS* **2003**, *100*:24, 13875-13880
200. Wishart, D.S.; Bigam, C.G.; Holm, A.; Hodges, R.S.; Sykes, B.D. *Journal of Biomolecular NMR* **1995**, *5*, 67-81
201. Lin, Y.; Pixley, R.A.; Colman, R.W. *Biochemistry* **2000**, *39*, 5104-5110
202. Vales-Gomez, M.; Erskine, R.A.; Deacon, M.P.; Strominger, J.L.; Reyburn, H.T. *PNAS* **2001**, *98*:4, 1734-1739
203. Live, D.; Armitage, I.M.; Dalgarno, D.C.; Cowburn, D. *J. Am. Chem. Soc.* **1985**, *107*, 1775-1777
204. Zerbe, O.; Pountney, D.L.; Philipsborn, W.V.; Vasak, M. *J. Am. Chem. Soc.* **1994**, *116*, 377-378

205. Sims, R.J.; belotserkovskaya, R.; Reinberg, D. *Genes and Development* **2004**, **18**, 2437-2468
206. Hetzer, C.; Dormeyer, W.; Schnolzer, M.; Ott, M. *Microbes and Infection* **2005**, **7**, 1364-1369
207. Vasak, M.; Kagi, J.H.R.; Hill, H.A.O. *J. Amer. Chem. Soc.* **1981**, *20:10*, 2852-2856
208. Turk, T.; Resch, U.; Fox, M.A.; Vogler, A. *J. Phys. Chem.* **1992**, **96**, 3818-3822
209. Stillman, M.J. *Coordination Chemistry Reviews* **1995**, **144**, 461-511
210. Talon™ Metal Affinity Resin User Manual, BD Biosciences, 2001, 5
211. Kandagal, P.B.; Seetharamappa, J.; Ashoka, S.; Shaikh, S.M.T.; Manjunatha, D.H. *International Journal of Biological Macromolecules* **2006**, **39**, 234-239
212. Calero, M.; Gasset, M. *Methods in molecular biology* **2005**, **299**, 129-151
213. Goodrich, J.A.; Kugel, J.F. "Binding and Kinetics for Molecular Biologist" **2007**, Cold Spring Harbor Laboratory Press, pp 22-33
214. Battle, D.J.; Doudna, J.A. *RNA* **2001**, **7**, 123-132
215. Makowska, J.; Rodiewicz-Motowidlo, S.; Baginska, K.; Vila, J.A.; Liwo, A.; Chmurzynski, L.; Scheraga, H.A. *PNAS* **2006**, **103:6**, 1744-1749
216. Tatham, A.S.; Drake, A.F.; Shewry, P.R. *Biochem. J.* **1989**, **259**, 471-476
217. Otvos, J.D.; Engeseth, H.R.; Wehrli, S. *Biochemistry* **1985**, **24:24**, 6735-6740.
218. Willbold, D.; Rosin-Arbesfeld, R.; Sticht, H.; Frank, R.; Rosch, P. *Science* **1994**, **264**, 1584-1587
219. Ellis, P.D. *Science* **1983**, **221**, 1141-1146

## Appendix A

List of chemical shift values for cross-peaks observed during HMBC 2-J coupling experiment of HIV-1 Tat

| Cross-Peak Position | N (ppm) | H (ppm) |
|---------------------|---------|---------|
| 1                   | 173.609 | 8.556   |
| 2                   | 173.719 | 8.521   |
| 3                   | 176.106 | 8.553   |
| 4                   | 176.554 | 8.558   |
| 5                   | 177.618 | 8.556   |
| 6                   | 178.481 | 8.529   |
| 7                   | 178.792 | 8.566   |
| 8                   | 179.278 | 8.555   |
| 9                   | 180.331 | 8.516   |
| 10                  | 181.739 | 8.499   |
| 11                  | 182.698 | 8.499   |
| 12                  | 182.780 | 8.467   |
| 13                  | 184.162 | 8.466   |
| 14                  | 184.800 | 8.446   |
| 15                  | 185.729 | 8.441   |
| 16                  | 185.647 | 7.183   |
| 17                  | 184.920 | 7.154   |
| 18                  | 184.294 | 7.214   |
| 19                  | 182.754 | 7.172   |
| 20                  | 181.897 | 7.250   |
| 21                  | 181.371 | 7.186   |
| 22                  | 180.645 | 7.254   |
| 23                  | 179.231 | 7.253   |
| 24                  | 178.726 | 7.261   |
| 25                  | 178.674 | 7.218   |
| 26                  | 177.631 | 7.257   |
| 27                  | 176.587 | 7.265   |
| 28                  | 176.228 | 7.230   |
| 29                  | 174.668 | 7.253   |
| 30                  | 174.590 | 7.216   |
| 31                  | 173.592 | 7.256   |
| 32                  | 173.615 | 7.217   |
| 33                  | 173.853 | 7.168   |

## Appendix B

Cross-peak analysis of the aliphatic and the aromatic regions of  $^1\text{H}$ - $^{13}\text{C}$  HMQC spectra of apo-Tat and Tat in the presence of 0.5 mole equivalents of Zn(II) at pH 4.0 at 293 K.

| Aliphatic Region |         |         | Aromatic Region |         |         |
|------------------|---------|---------|-----------------|---------|---------|
| Peak             | H (ppm) | C (ppm) | Peak            | H (ppm) | C (ppm) |
| 1                | 3.318   | 28.768  | 1               | 8.576   | 46.427  |
| 2                | 3.197   | 29.400  | 2               | 8.501   | 46.810  |
| 3                | 3.131   | 29.544  | 3               | 7.652   | 46.789  |
| 4                | 3.061   | 29.122  | 4               | 7.190   | 29.597  |
| 5                | 2.831   | 28.134  | 5               | 7.217   | 30.216  |