An Investigation of the Electrophilic Active Site of Histidine and Phenylalanine Ammonia-lyase

JASON D. GALPIN

B.Sc. (1992), M.Sc. (1996), Simon Fraser University

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY in

THE FACULTY OF GRADUATE STUDIES

Department of Chemistry

We accept this thesis as conforming to the required standard:

THE UNIVERSITY OF BRITISH COLUMBIA

April, 2001

© Jason D. Galpin, 2001
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Chemistry

The University of British Columbia
Vancouver, Canada

Date April 24, 2001
Abstract

Histidine ammonia-lyase from *Pseudomonas putida* (hutH, EC 4.3.1.3) contains a unique active site electrophile that forms autocatalytically via a dehydrating backbone cyclization of the three amino acids alanine 142, serine 143 and glycine 144 and a second dehydration of the serine 143 side chain. The electrophile is a 4-methylidene-imidazole-5-one (MIO) that, following enzyme inactivation in the presence of cysteine and oxygen at pH 10.5, has been isolated as a component of a chromophoric octapeptide. The modified MIO is found to be an oxidized adduct of cysteine to which another cysteine or β-mercaptoethanol is linked through a disulfide bridge. One and two-dimensional NMR experiments as well as high-resolution mass and UV absorbance spectra indicate that each MIO-containing octapeptide consists of at least three interconverting isomers, with a covalent linkage between the exocyclic MIO methyldiene carbon and the amine of cysteine. This information and data from inactivation studies employing analogues of cysteine lacking an amino functionality suggest an inactivation mechanism that strongly supports the Retey mechanism of HAL with histidine. It is likely that the closely related enzyme phenylalanine ammonia-lyase (PAL, EC 4.3.1.3) shares this mechanism.

A fluorine-containing analogue of L-phenylalanine, *erythro* 3-fluorophenylalanine, was synthesized and was found to act as a competitive inhibition of PAL from poplar, with a *K*\textsubscript{i} value of 45±5 μM. This is significantly smaller than the *K*\textsubscript{M} value of 250 μM for L-phenylalanine, indicating that *erythro* 3-fluorophenylalanine is a reasonably good competitive inhibitor of PAL.

A variety of analogues of L-histidine were synthesized, including (S)-2-halo-3-(imidazol-4-yl) propionic acid, where halo is fluoro, chloro or bromo, and β-(pyrazol-1-yl)-L-alanine. These compounds were found to be neither substrates nor inhibitors of HAL. Additional experiments were performed on 4-methylimidazole/glycine and imidazole/L-alanine to test for isotope
exchange in the imidazole ring but none were observed. 2-methylimidazole was found to be a competitive inhibitor of HAL with a $K_i$ value of 20 mM.
# Table of Contents

Abstract.................................................................................................................................ii  
Table of Contents..................................................................................................................iv  
List of Figures..........................................................................................................................viii  
List of Tables..........................................................................................................................xiii  
Abbreviations and Symbols.....................................................................................................xiv  
Acknowledgments...................................................................................................................xvii  
Dedication...............................................................................................................................xviii  

Chapter One: The Ammonia-lyases HAL and PAL.................................................................1  
  1.1 Introduction......................................................................................................................2  
  1.2 The Chemical Elimination of Ammonia...........................................................................3  
  1.3 The Enzymatic Elimination of Ammonia.........................................................................7  
    1.3.1 Aspartate Ammonia-lyase.......................................................................................8  
    1.3.2 Phenylalanine Ammonia-lyase.............................................................................11  
    1.3.3 Histidine Ammonia-lyase.....................................................................................12  
    1.3.4 Evidence for a Shared Mechanism.......................................................................13  
  1.4 Proposals for an Active Site Electrophile.........................................................................13  
  1.5 Evidence for Dehydroalanine as the Electrophile.........................................................16  
  1.6 The Hanson and Havir Mechanism for HAL.................................................................19  
  1.7 Possible Origins of Dehydroalanine..............................................................................21  
  1.8 Sequence Comparison and Mutants of HAL and PAL.................................................22  
  1.9 The Retey Mechanism for HAL....................................................................................23  
  1.10 The Exchangeability of Imidazole Protons...................................................................25  
  1.11 Thesis Aims...................................................................................................................27
Chapter Two: Inhibition of Phenylalanine Ammonia-lyase

2.1 Introduction

2.2 Results and Discussion

2.2.1 Isolation and Purification of PAL

2.2.2 Synthesis of 3-Fluorophenylalanine

2.2.3 Inhibition of PAL by 3-Fluorophenylalanine

2.3 Materials and Methods

2.3.1 General Information

2.3.2 Purification of PAL

2.3.3 Activity Assay for PAL

2.3.4 Synthesis of 3-Fluorophenylalanine

2.3.5 Inhibition Assays for PAL

Chapter Three: Structural Determination of the HAL Electrophile

3.1 Introduction

3.2 Results and Discussion

3.2.1 Isolation and Purification of HAL

3.2.2 Inactivation of HAL with L-Cysteine and Peptide Isolation

3.2.3 Characterization of the Octapeptides

3.2.3.1 N-Terminal Analysis

3.2.3.2 High Resolution Mass Spectrometry

3.2.3.3 Amino Acid Analysis

3.2.3.4 One and Two-Dimensional NMR Spectroscopy

3.2.3.4.1 One-Dimensional Proton NMR
4.3 Overall Conclusions.................................................................133

4.4 Materials and Methods.........................................................134

4.4.1 General Information.........................................................134

4.4.2 Synthesis of (S)-2-Chloro-3-(imidazol-4-yl) Propionic Acid.....134

4.4.3 Synthesis of (S)-2-Fluoro-3-(imidazol-4-yl) Propionic Acid.....134

4.4.4 Spectrophotometric Assay of Halogen Analogues of Histidine.....135

4.4.5 Fluoride Ion Concentration Assay for (S)-2-Fluoro-3-(imidazol-4-yl)
        Propionic Acid.........................................................................136

4.4.6 Inhibition Assay of HAL with (S)-2-Chloro-3-(imidazol-4-yl)
        Propionic Acid.........................................................................136

4.4.7 Synthesis of β-(Pyrazol-1-yl)-L-Alanine.................................136

4.4.8 Spectrophotometric Assay of β-(Pyrazol-1-yl)-L-Alanine with HAL...137

4.4.9 $^1$H NMR Assay of β-(Pyrazol-1-yl)-L-Alanine with HAL........137

4.4.10 Test for Deuterium Wash-in of 4-Methylimidazole with HAL......137

4.4.11 Inhibition of HAL with 2-Methylimidazole............................138

References.....................................................................................139
List of Figures

Figure 1.1 Stereochemistry of α,β-eliminations .................................................. 4
Figure 1.2 More O’Ferrall diagram for β-eliminations .................................. 4
Figure 1.3 pK_a values for neutral and protonated leaving groups ................. 6
Figure 1.4 Hofmann elimination of trimethylamine from ergothioneine ........ 6
Figure 1.5 Structure of cofactors required for oxidative removal of ammonia .... 7
Figure 1.6 Anti elimination of ammonia from L-aspartate by aspartase ......... 8
Figure 1.7 Structures of 3-nitro inhibitors of aspartase and fumarase .......... 8
Figure 1.8 Possible roles for important residues in aspartase ................. 9
Figure 1.9 Anti elimination of ammonia from phenylalanine by PAL .......... 11
Figure 1.10 Compounds derived from cinnamate ........................................... 11
Figure 1.11 Anti elimination of ammonia from histidine by HAL ............... 12
Figure 1.12 Exchange of _3^H and _14^C urocanate involving an ammonia-enzyme complex .......... 14
Figure 1.13 Early model for the active site electrophile ................................... 15
Figure 1.14 Model for generating tritiated alanine from inactivated HAL .... 16
Figure 1.15 An early mechanism for HAL involving simple dehydroalanine .... 17
Figure 1.16 Evidence for iminodehydroalanine in HAL ............................... 18
Figure 1.17 Hanson and Havir HAL mechanism showing a covalent ammonia intermediate ...... 19
Figure 1.18 Generation of iminodehydroalanine from an N-terminal serine and a protein derived carbonyl ......................................................... 21
Figure 1.19 Elimination of water or hydrogen sulfide resulting in dehydroalanine .......... 23
Figure 1.20 Retey mechanism for HAL ....................................................... 24
Figure 1.21 Loss of deuterium label at C5 is explained by a covalent intermediate generated via the Retey mechanism ........................................ 26

Figure 2.1 The final step in the biosynthesis of L-phenylalanine ................. 29
Figure 2.2 The transformation of cinnamate to coniferyl alcohol .......... 30
Figure 2.3 L-(35)-3-Fluorophenylalanine with fluorine in place of the abstractable proton .......... 31
Figure 2.4 Two possible modes of inhibition of PAL by fluorophenylalanine .......... 32
Figure 2.5 Structures and K_i values of three potent PAL inhibitors .......... 32
Figure 2.6 FPLC chromatogram showing a peak corresponding to PAL and SDS PAGE gel ...... 34
Figure 3.36 HMBC spectrum of B2 showing several distant \(^1\)H/\(^{13}\)C couplings
Figure 3.37 Structure of compound similar to the MIO chromophore in HAL
Figure 3.38 Structure of Loprazolam and isomeric forms of precursor compounds showing an important H-bond in monoalkyl amines
Figure 3.39 Structures of cis and trans isomers of the B2 chromophore
Figure 3.40 Potential inactivation mechanism incorporating the Hanson and Havir model
Figure 3.41 Potential inactivation mechanism incorporating the Retey model
Figure 3.42 Graph showing oxygen requirement for HAL inactivation with 3-mercaptopropionate
Figure 3.43 Structure and formation of the MIO prosthetic group
Figure 3.44 Retey mechanism incorporating the MIO prosthetic group

Figure 4.1 A comparison showing that the chlorine analogue of histidine could be a substrate for the Retey mechanism but not for the Hanson and Havir mechanism
Figure 4.2 A comparison showing that the pyrazole analogue of histidine could be a substrate for the Hanson and Havir mechanism but not for the Retey mechanism
Figure 4.3 Numbering convention and pH behavior of the histidine imidazole
Figure 4.4 Synthesis of (S)-2-chloro-3-(imidazol-4-yl) propanoic acid
Figure 4.5 Proposed mechanism for the formation of the rearrangement product observed in fluorodeazotization of L-phenylalanine
Figure 4.6 Changes in fluoride ion concentration of 50 mM
(S)-2-fluoro-3-(imidazole-4-yl) propionate in 800 \(\mu\)L of 5 mM potassium fluoride
Figure 4.7 Synthesis of \(\beta\)-(pyrazol-1-yl)-L-alanine
Figure 4.8 Mass spectrum of products from a reaction of imidazole and 11 showing masses rationalized by oligomeric species
Figure 4.9 Glycine and \(m\)-hydroxytoluene synergistically inhibit PAL
Figure 4.10 Isotope exchange at the 5-position of the histidine imidazole is possible to rationalize using the Hanson and Havir mechanism
Figure 4.11 Structures of the pairs of components of two histidine analogues
Figure 4.12 Justification for C5 proton exchange in 4-methylimidazole and imidazole using the Retey mechanism
Figure 4.13 Justification for C5 proton exchange in 4-methylimidazole only, using the Hanson and Havir mechanism
Figure 4.14 Scheme illustrating route for deuterium incorporation into 4-methylimidazole through a possible covalent intermediate.........................130

Figure 4.15 $^1$H NMR spectra of 4-methylimidazole and HAL in buffered D$_2$O showing imidazole C2 and C5 protons with deuterium wash-in at C2 only.........................131

Figure 4.16 Graph showing competitive inhibition of HAL with 2-methylimidazole...............132
List of Tables

Table 1.1  Equilibrium constants for the elimination of ammonia from amino acids; 
  $pK_a$ values for the abstractable proton.................................................................10
Table 1.2  Protein sequence homology of three HAL and four PAL enzymes..............22

Table 3.1  $K_M$ or $K_I$ values for various amino acids and HAL..................................48
Table 3.2  Effect of sulfhydryl compounds on the activity of HAL at pH 10.5.................50
Table 3.3  Summary of the purification of HAL............................................................56
Table 3.4  Summary of octapeptides with basic model structures, 
  absorbance maxima and neutral masses.................................................................64
Table 3.5  High-resolution neutral masses of octapeptides and 
  differences from the MIO octapeptide.................................................................67
Table 3.6  Summary of empirical formulae that match the observed 
  mass differences between A1 and B2 and the MIO octapeptide..............................68
Table 3.7  Summary of data from all NMR experiments on 
  octapeptide B2 in DMSO-$d_6$..............................................................................77
## Abbreviations and Symbols

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^2$H, D</td>
<td>deuterium</td>
</tr>
<tr>
<td>$^3$H, T</td>
<td>tritium</td>
</tr>
<tr>
<td>$\delta$</td>
<td>chemical shift (ppm)</td>
</tr>
<tr>
<td>$\varepsilon$</td>
<td>extinction coefficient ($M^{-1} \text{ cm}$)</td>
</tr>
<tr>
<td>$\lambda_{\text{max}}$</td>
<td>wavelength for largest absorbance peak</td>
</tr>
<tr>
<td>m</td>
<td>multiplet</td>
</tr>
<tr>
<td>d</td>
<td>doublet</td>
</tr>
<tr>
<td>dt</td>
<td>doublet of triplets</td>
</tr>
<tr>
<td>s</td>
<td>singlet</td>
</tr>
<tr>
<td>ACN</td>
<td>acetonitrile</td>
</tr>
<tr>
<td>AIP</td>
<td>2-aminoindan-2-phosphonic acid</td>
</tr>
<tr>
<td>AOPP</td>
<td>(S)-2-aminoxy-3-phenyl propionic acid</td>
</tr>
<tr>
<td>APEP</td>
<td>(R)-(1-amino-2-phenylethyl) phosphonic acid</td>
</tr>
<tr>
<td>CD</td>
<td>circular dichroism</td>
</tr>
<tr>
<td>COSY</td>
<td>CORrelation SpectroscopY</td>
</tr>
<tr>
<td>Da</td>
<td>dalton</td>
</tr>
<tr>
<td>DAST</td>
<td>diethylaminosulfur trifluoride</td>
</tr>
<tr>
<td>DHA</td>
<td>dehydroalanine</td>
</tr>
<tr>
<td>DMeAD</td>
<td>dimethylazodicarboxylate</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethylsulfoxide</td>
</tr>
<tr>
<td>DMSO-d$_6$</td>
<td>per-deuterated dimethylsulfoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>EC</td>
<td>Enzyme Commission</td>
</tr>
<tr>
<td>ESIMS</td>
<td>electrospray ionization mass spectrometry</td>
</tr>
<tr>
<td>GFP</td>
<td>green fluorescent protein</td>
</tr>
<tr>
<td>HAL</td>
<td>histidine ammonia-lyase</td>
</tr>
<tr>
<td>HMBC</td>
<td>Heteronuclear Multiple Bond multiple quantum Coherence</td>
</tr>
<tr>
<td>HMQC</td>
<td>Heteronuclear Multiple Quantum Coherence</td>
</tr>
<tr>
<td>HPLC</td>
<td>high pressure liquid chromatography</td>
</tr>
<tr>
<td>HRMS</td>
<td>high resolution mass spectrometry</td>
</tr>
<tr>
<td>hutH</td>
<td>HAL gene</td>
</tr>
<tr>
<td>$K_{eq}$</td>
<td>equilibrium constant</td>
</tr>
<tr>
<td>$K_i$</td>
<td>dissociation constant for an enzyme-inhibitor complex</td>
</tr>
<tr>
<td>$K_M$</td>
<td>Michaelis constant</td>
</tr>
<tr>
<td>MIO</td>
<td>4-methylidene-imidazole-5-one</td>
</tr>
<tr>
<td>NBS</td>
<td>N-bromosuccinimide</td>
</tr>
<tr>
<td>NMR</td>
<td>nuclear magnetic resonance</td>
</tr>
<tr>
<td>PAL</td>
<td>phenylalanine ammonia-lyase</td>
</tr>
<tr>
<td>PAGE</td>
<td>polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PEG</td>
<td>polyethylene glycol</td>
</tr>
<tr>
<td>PITC</td>
<td>phenylisothiocyanate</td>
</tr>
<tr>
<td>PKU</td>
<td>phenylketonuria</td>
</tr>
<tr>
<td>PTC</td>
<td>phenylthiocarbamyl</td>
</tr>
<tr>
<td>PTH</td>
<td>phenylthiohydantoin</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulfate</td>
</tr>
<tr>
<td>THF</td>
<td>tetrahydrofuran</td>
</tr>
<tr>
<td>TMSI</td>
<td>trimethylsilyl iodide</td>
</tr>
</tbody>
</table>
Tris  
tris(hydroxymethyl)amino methane

TsOH  
toluenesulfonic acid

UV  
ultraviolet

**Standard Abbreviations for Amino Acids**

<table>
<thead>
<tr>
<th>Single Letter</th>
<th>Three Letter</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Ala</td>
<td>alanine</td>
</tr>
<tr>
<td>C</td>
<td>Cys</td>
<td>cysteine</td>
</tr>
<tr>
<td>D</td>
<td>Asp</td>
<td>aspartate</td>
</tr>
<tr>
<td>E</td>
<td>Glu</td>
<td>glutamate</td>
</tr>
<tr>
<td>F</td>
<td>Phe</td>
<td>phenylalanine</td>
</tr>
<tr>
<td>G</td>
<td>Gly</td>
<td>glycine</td>
</tr>
<tr>
<td>H</td>
<td>His</td>
<td>histidine</td>
</tr>
<tr>
<td>I</td>
<td>Ile</td>
<td>isoleucine</td>
</tr>
<tr>
<td>K</td>
<td>Lys</td>
<td>lysine</td>
</tr>
<tr>
<td>L</td>
<td>Leu</td>
<td>leucine</td>
</tr>
<tr>
<td>M</td>
<td>Met</td>
<td>methionine</td>
</tr>
<tr>
<td>N</td>
<td>Asn</td>
<td>asparagine</td>
</tr>
<tr>
<td>P</td>
<td>Pro</td>
<td>proline</td>
</tr>
<tr>
<td>Q</td>
<td>Gln</td>
<td>glutamine</td>
</tr>
<tr>
<td>R</td>
<td>Arg</td>
<td>arginine</td>
</tr>
<tr>
<td>S</td>
<td>Ser</td>
<td>serine</td>
</tr>
<tr>
<td>T</td>
<td>Thr</td>
<td>threonine</td>
</tr>
<tr>
<td>V</td>
<td>Val</td>
<td>valine</td>
</tr>
<tr>
<td>W</td>
<td>Trp</td>
<td>tryptophan</td>
</tr>
<tr>
<td>Y</td>
<td>Tyr</td>
<td>tyrosine</td>
</tr>
</tbody>
</table>
Acknowledgements

A good deal of credit for the insight leading to the overall findings expressed herein should go to Martin Tanner, whose timely and helpful (sometimes crucially so) suggestions and interpretations greatly aided my investigations, both chemical and enzymological. A debt of gratitude is equally owed to Brian Ellis for his original idea about the GFP-MIO link, for his help with the problems we encountered and for ushering a synthetic organic chemist into the world of microbiology and biotechnology. The two of them also deserve thanks for actually paying me to do this challenging and fun stuff. I thank Professor Ray Andersen for his advice on $^1$H NMR and structure-based searching.

I acknowledge the assistance of those members of the Tannerite and Ellisian clans who basically taught me protein chemistry and enzymology along the way. In particular, Robert Campbell and Stephanie Butland greatly assisted me in my many struggles, scientific and otherwise.

Lastly, I would like to express my appreciation to the many, many members of the UBC research community who ran assays and spectra for me. Their contributions also made this possible.
Dedication

This thesis is dedicated to

The fair and lovely

Samantha Julia

Monckton

♥
Chapter One

The Ammonia Lyases

HAL and PAL
1.1 Introduction

DNA and protein are arguably the two most widely studied macromolecules in history, as both are intimately involved with life processes in all organisms and accordingly each holds a special place in our imaginations. The former is commonly known as the molecule of heredity and information, as coded within its elegantly wound double helix are the blueprints for each and every protein that make up an organism. Protein then must surely rank as the molecule of action, with enzymes taking the role of factory to DNA’s family library, turning out useful molecules and regulating a dizzying array of biochemical processes that make life possible. It follows that the role of the budding enzymologist is akin to that of a novice mechanic, puzzling out how the levers and tools of these complex, mysterious and beautiful factories work.

It was (and still is) the ambition of this apprentice to study a family of enzymes that are involved in the transformation of particular aromatic amino acids. The special features of these enzymes have led to some enriching discoveries and it is my goal to share this experience with the reader. I further hope that the brief recounting here of earlier investigations in this area will illustrate how a scientific model can be slowly built up, typically gaining complexity and nuance with each finding until a sound, defensible theory is put forward.

All organisms maintain concentrations of important molecules by balancing the processes of anabolism (building up) and catabolism (degradation). For example, a ready pool of amino acids is required for the optimal assembly of protein. Occasionally an excess of one or more types of amino acid will build up and therefore degradation must be increased in order to restore a balance. In most cases the first step of enzymatic amino acid degradation involves the oxidative removal of the α-amino group to produce an α-keto acid, which can then be shuttled into a variety of biochemical pathways. Histidine and phenylalanine ammonia-lyase (HAL and PAL) however
constitute a unique class of enzymes that catalyze a difficult, non-oxidative elimination of ammonia from their substrates resulting in conjugated unsaturated products. They not only remove very non-acidic protons but also expel a very poor leaving group from their aromatic amino acid substrates. Intriguingly, both HAL and PAL are known to contain a highly unusual and reactive electrophilic moiety, frequently referred to as dehydroalanine. Furthermore, there exists a great controversy over the true mechanism of these enzymes, with two very different proposals currently under consideration. All of these features make HAL and PAL compelling targets of research. It is the aim of this chapter to chronicle some of the published results in terms of our current understanding of the active site of HAL and PAL, and illustrate two mechanisms that have been proposed for the enzymatic elimination of ammonia. A strategy for determining the real mechanism is presented at the end of this chapter.

1.2 The Chemical Elimination of Ammonia

Prior to delving into the enzymology of HAL and PAL, the non-enzymatic elimination of ammonia should be examined. To this end, the nomenclature and stereochemistry of ammonia elimination will be briefly reviewed. In general, elimination reactions involve a deprotonation and a loss of a leaving group, generating an element of unsaturation. Throughout this work the terms α and β will be used to describe the carbons on which the leaving group amine and proton are situated respectively. A β-elimination proceeds with the loss of two σ-bonds from adjacent carbon atoms (the β C-H and α C-N of an amino acid in the case of HAL and PAL) and the formation of a new C-C double bond between them. A syn or anti elimination occurs when the hydrogen and amine depart from the same or opposite faces of the incipient double bond, respectively (figure 1.1) and a cis or trans olefin may be generated depending on the nature of R1–R4.
Figure 1.1 Stereochemistry of α,β eliminations

A useful visual tool for discussing the timing of the C-H and C-N bond cleavage is a More O’Ferrall diagram, represented here as a two-dimensional reaction coordinate where the bond order for the C-H bond is shown on the X-axis and that of the C-N bond is shown on the Y-axis (figure 1.2). A pathway from the bottom left corner to the upper right represents the sequence of

Figure 1.2 More O’Ferrall diagram for β eliminations
bond cleavages as starting material goes to products. It is instructive to examine the boundary conditions in the diagram, and for this purpose three variants are illustrated from a continuum of mechanistic possibilities. Where C-N cleavage fully precedes that of C-H, a carbocation intermediate results and the mechanism is termed E1. The opposite sequence involves a carbanion intermediate and in this case the reaction proceeds via an E1cb mechanism. For the E1 and E1cb paths the open circles in the figure are used to denote the intermediate ion formed along each route. The concerted case, where C-H and C-N cleave simultaneously in a single transition state (denoted by the dark circle), is known as an E2 elimination.

The great utility of the diagram is that it provides a simple rationale for the effects that substituent groups on the α and β carbons have on the actual path a reaction may go through. For example, electron-withdrawing groups (R₁ or R₂) on the β carbon could stabilize a growing negative charge there and would facilitate an E1cb mechanism. Similarly, electron donating groups (R₃ or R₄) on the α carbon could stabilize a positive charge and would favour an E1 mechanism. However, the actual pathway a compound will follow is established by more than simply the presence of charge stabilizing constituents. In particular, the ease or difficulty of leaving group departure can have a major determining effect.

Compounds with very poor leaving groups tend to exhibit the characteristics of E1cb mechanisms. Obviously, the more difficult it is for a leaving group to depart, the greater the likelihood that negative charge will 'build up' in the molecule as the proton is removed. If indeed an E1cb mechanism is operative it may be possible to observe the reversible exchange of protons between the substrate and base. For instance, if exchangeable deuterium is present in the medium then recovered starting material should contain deuterium. E1 or E2 mechanisms don’t readily provide the opportunities for isotope exchange because of the lack of any stabilized anion. These concepts will be revisited in the following section on enzymatic elimination of ammonia, as there is good evidence for an E1cb mechanism in the examples presented.
Figure 1.3 Approximate pKa values for neutral and protonated leaving groups

It is interesting to contrast the elimination of ammonia with a comparable dehydration reaction. Both involve the removal of a proton and the departure of a leaving group. But while generating the same unsaturated product, it is more difficult to eliminate ammonia than water (figure 1.3) in that the amine anion, being much more basic than hydroxide, is a much poorer leaving group. In order for the amine to have a reasonable chance of departing it needs to be positively charged. However, an ammonium ion would largely be deprotonated in the alkaline environment required for the removal of the vicinal proton.

One way out of this chemical dilemma is to peralkylate the amine to effect a Hofmann elimination (figure 1.4). The trimethylamine is much more readily eliminated than ammonia but still requires harsh conditions. Of course, one of the advantages that enzymes offer is an environment where it is possible to have acidic and basic functional groups in close proximity without neutralization, and the pK<sub>a</sub>s of bound ionizable groups can be altered by several units.

Figure 1.4 Hofmann elimination of trimethylamine from ergothioneine
1.3 The Enzymatic Elimination of Ammonia

![Chemical Structures]

There are several enzymes that remove ammonia from substrates, including those that require external cofactors like pyridoxal phosphate (aminotransferases) or NAD$^+$ (glutamate dehydrogenase) or FAD (D-amino acid oxidase) (figure 1.5). These enzymes catalyze the oxidative removal of ammonia and give ketones as products. The enzymes under scrutiny in this study, however, eliminate ammonia in a redox neutral process. Aspartate ammonia-lyase (aspartase; EC 4.3.1.1), histidine ammonia-lyase (HAL; EC 4.3.1.3) and phenylalanine ammonia-lyase (PAL; EC 4.3.1.3) are the best understood of this class and shall be examined in detail to introduce concepts that form the basis for the majority of this work. Each of these three enzymes is thought to process its substrate through an E1cb or E1cb-like mechanism with substantial carbanionic character, consistent with the presence of the very poor leaving group ammonia. The acidity of the abstractable proton is thus a major determining factor in the overall rate of the elimination reaction. The substrate specific and enzymatic factors governing anion formation and stability will be given specific attention in the following sections.
1.3.1 Aspartate Ammonia-lyase

Aspartase eliminates ammonia from L-aspartate by the removal of the pro-R hydrogen in an anti fashion, producing fumarate (figure 1.6). It is closely related to another fumarate producing enzyme, namely fumarase (EC 4.2.1.2), which makes its product through the dehydration of malate. Fumarate is an important intermediate in the citric acid cycle and as such fumarase has been extensively studied. Not surprisingly, there is strong evidence that aspartase and fumarase employ very similar mechanisms and are thought to process their substrates through an E1cb pathway.

An investigation by D. J. T. Porter and H. J. Bright demonstrated that both enzymes are inhibited by carbanionic substrate mimics (figure 1.7). The authors prepared 3-nitro analogues of aspartate and malate because of the stabilizing effect the nitro group has on the anionic forms of A and also because the nitro group is structurally and electronically similar enough to the carboxylate it replaces.

![Figure 1.6 Anti elimination of ammonia from L-aspartate by aspartase](image1)

Aspartase eliminates ammonia from L-aspartate by the removal of the pro-R hydrogen in an anti fashion, producing fumarate (figure 1.6). It is closely related to another fumarate producing enzyme, namely fumarase (EC 4.2.1.2), which makes its product through the dehydration of malate. Fumarate is an important intermediate in the citric acid cycle and as such fumarase has been extensively studied. Not surprisingly, there is strong evidence that aspartase and fumarase employ very similar mechanisms and are thought to process their substrates through an E1cb pathway.

An investigation by D. J. T. Porter and H. J. Bright demonstrated that both enzymes are inhibited by carbanionic substrate mimics (figure 1.7). The authors prepared 3-nitro analogues of aspartate and malate because of the stabilizing effect the nitro group has on the anionic forms of A and also because the nitro group is structurally and electronically similar enough to the carboxylate it replaces.

![Figure 1.7 Structures of 3-nitro inhibitors of aspartase and fumarase](image2)
For fumarase \((X = \text{OH})\) Porter et al. found that the anionic form \(A\) competitively bound 18,000-fold more tightly than the neutral form \(B\), and the ratio of the \(K_M\) value for malate and \(K_i\) value for the anionic inhibitor \(A\) was 900. Similarly, for aspartase \((X = \text{NH}_2)\) the anionic form \(A\) bound 1,630-fold better than the neutral \(B\), and \(K_M(\text{Asp})/K_i(A)\) was also coincidentally 1,630. The results suggest that the anionic compounds are transition state, or transient intermediate analogues and that the mechanisms of both enzymes involve 3-carbanions.

There is also evidence for a bound carbanion from X-ray structures of aspartase. The enzyme does not require any cofactors or prosthetic groups and certain key residues (arginine 29 and lysine 327) are essential for enzyme activity (figure 1.8).\(^9\) One model suggests that arginine 29 forms a stabilizing hydrogen bond to the enolate and is involved in lowering the energy required for its formation, however the actual residue involved in deprotonating the substrate is not known.

![Figure 1.8 Possible roles for important residues in aspartase](image)

Of course, from the structure of aspartate it is clear that the proton to be removed is already relatively acidic due to the presence of an adjacent carbonyl functionality. Estimates of the \(pK_a\) of this proton are in the range of 23 to 26. Combined with the data from the 3-nitro inhibition study it is therefore very reasonable to suspect that an E1cb mechanism is employed by aspartase with an enolate intermediate.

In terms of \(\beta\) proton activation, the fundamental difference between aspartate and histidine or phenylalanine is the lack of a greatly acidifying function in either of the latter. The \(pK_a\) value of
the β protons in these substrates has been estimated to be at least 40,¹⁰ about fifteen orders of magnitude larger than that of aspartate! Kinetically it would seem that HAL and PAL have a tremendously difficult job to do. Thermodynamically, however, the elimination is more favoured for the aromatic substrates, owing to the benefit of extended conjugation (table 1.1).¹⁰ Regardless, the fact that ammonia-lyases for amino acids lacking β-carboxyl or aromatic groups have not been reported suggests that the conjugation such groups provide is essential to bring the activation energy down to a reasonable level for catalysis.

\[
\text{R} \quad \begin{array}{c}
\alpha \quad \text{CO}_2^-
\end{array} \quad \text{NH}_3 \quad \begin{array}{c}
\beta
\end{array}
\]

\[
\underset{K_{eq}}{\rightleftharpoons} \quad \text{R} \quad \begin{array}{c}
\text{CO}_2^-
\end{array} \quad + \quad \text{NH}_3
\]

<table>
<thead>
<tr>
<th>L-Amino acid</th>
<th>R</th>
<th>(K_{eq} ) (M)</th>
<th>(pK_a (\beta \ H))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspartate</td>
<td>Carboxyl</td>
<td>(4.6 \times 10^{-3})</td>
<td>25</td>
</tr>
<tr>
<td>Histidine</td>
<td>Imidazolyl</td>
<td>3.0</td>
<td>(~40)</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>Phenyl</td>
<td>4.3</td>
<td>(~40)</td>
</tr>
</tbody>
</table>

*Table 1.1 Equilibrium constants for the elimination of ammonia from amino acids; \(pK_a\) values for the abstractable proton (from reference 10)*
1.3.2 Phenylalanine Ammonia-lyase

PAL eliminates ammonia from L-phenylalanine by removal of the pro-S hydrogen in an anti fashion to give the conjugated product trans-cinnamic acid (figure 1.9). It is considered an important plant enzyme because of its role as a branch point for phenylpropanoid biosynthesis, diverting phenylalanine from protein synthesis. Compounds that are ultimately generated from cinnamate include monolignols, the flavonoids and coumarins (figure 1.10). PAL from a variety of sources has been studied, including potato, maize, soybeans, tobacco and poplar. In microorganisms the enzyme catabolizes exogenous phenylalanine for use as a carbon source. In some cases oligosaccharide side chains are attached to the enzyme but are not required for activity.

The regulation and expression patterns of the enzyme from trees are being studied as well. It is hoped that from a general understanding of the modes of expression of PAL in timber...
that the level of lignification in wood might be controlled or at least diminished. There is great interest in managing the degree and type of lignification in trees, as a major part of the cost of pulping wood is the separation of lignin, the second most abundant biomolecule on earth, from the most abundant, namely cellulose.

An additional interesting feature of PAL is its potential as a therapeutic agent. The enzyme has not been detected in mammals, but could conceivably be administered to reduce excess levels of phenylalanine in the blood of patients suffering from the disorder phenylketonuria (PKU). This classic genetic disease is associated with hyperphenylalaninemia and mental retardation, and arises from a defect in the enzyme phenylalanine monooxygenase (EC 1.14.16.1). PKU is currently treated from birth with a low phenylalanine diet, the lifelong compliance with becomes very difficult. Recently, experiments on a mouse PKU model used PAL as a substitute for the monooxygenase; the cinnamate produced was found to be a harmless metabolite.

1.3.3 Histidine Ammonia-lyase

\[
\begin{align*}
\text{L-Histidine} & \quad \text{HAL} \quad \text{trans Urocanate} \\
\text{\textit{trans} Urocanate} & \quad + \quad \text{NH}_3
\end{align*}
\]

*Figure 1.11 Anti elimination of ammonia from histidine by HAL*

Histidine ammonia-lyase eliminates ammonia from L-histidine by removal of the pro-R hydrogen in an anti fashion to give \textit{trans}-urocanic acid (figure 1.11). It has been found in many organisms, from fungi and bacteria to mammals, but not in plants. Its major function seems to be the catabolism of excess histidine, although elevated levels of HAL in skin tissues have suggested that urocanic acid may also act as a mediator of ultraviolet light induced immunosuppression.16
The jaundice-like condition resulting from defective HAL, known as histidinemia, is not considered a life-threatening ailment.\textsuperscript{17}

HAL from \textit{Pseudomonas putida} is the most frequently studied form of this enzyme, in part because it is easily overexpressed as a recombinant enzyme in \textit{E. coli}. The enzyme functions optimally in the presence of divalent cations but does not absolutely require them.\textsuperscript{18} Similarly, thiols have been shown to stimulate HAL's activity but their role seems to be one of keeping the enzyme in its more active reduced state (the oxidized and reduced forms are separable).

\subsection*{1.3.4 Evidence for a Shared Mechanism}

There is a great deal of evidence for a shared mechanism for HAL and PAL. Both eliminate free ammonia from aromatic amino acids via removal of a non-acidic hydrogen to give \textit{trans} acid products and both have similar pH/rate profiles. Both enzymes are homotetramers and are inactivated upon treatment with reactive carbonyl-directed reagents like sodium borohydride and cyanide ion. Probably the most compelling evidence for a common mechanism (and ancestry) is the level of sequence homology between the two, which will be discussed in more detail in section 1.7. Consequently, information from the study of one enzyme is often directly compared to that from the other with little distinction.

\subsection*{1.4 Proposals for an Active Site Electrophile}

Some of the earliest work on HAL and PAL showed that they could be permanently inactivated with nucleophilic reagents. One explanation for their rapid inactivation with borohydride was that the enzymes contained an active site electrophile. In 1962 Peterkofsky\textsuperscript{19} developed a model for electrophilic HAL that involved a covalent enzyme-ammonia complex, although he never speculated what the actual electrophile in the enzyme might be. He found that
while $^{14}$C-labeled urocanic acid and solvent-derived tritium can be exchanged into histidine, $^{15}$N-ammonia cannot and that both exchanges require the presence of histidine (figure 1.12). This was interpreted as evidence for the ordered release of products from HAL, with urocanate departing first and ammonia, being covalently attached, departing second. Several years later it was shown that ammonia could in fact be incorporated back into histidine if it is present in very high concentrations. Presumably this occurs via the reverse reaction as opposed to ammonia intercepting a bound product prior to release. Peterkofsky also proposed a concerted elimination

![Diagram of molecular reactions](image)

$^{15}$NH$_3$ does not exchange into histidine

Figure 1.12 Exchange of $^3$H and $^{14}$C urocanate involving an ammonia-enzyme complex
because he could find no difference in the rates of exchange of tritium and urocanic acid from labeled histidine. This belief also turned out to be incorrect, as there is now strong evidence for a carbanionic intermediate.\textsuperscript{21}

Five years later T. A. Smith, F. H. Cordelle and R. H. Abeles\textsuperscript{22} published a communication on the inactivation of HAL by nucleophilic reagents that were known to attack carbonyl groups. Following up the earlier work of H. Tabor and A. H. Mehler\textsuperscript{23} they permanently inactivated the enzyme with sodium borohydride, phenylhydrazine or potassium cyanide and demonstrated that the attack was occurring at the active site by showing protection by substrate. When sodium borotritide was employed the protein became permanently labeled. They interpreted these results in terms of an active site carbonyl that could form either a protonated carbinolamine or an imine with the amine of histidine (figure 1.13), supposedly creating a better leaving group.\textsuperscript{24} D. S. Hodgins came to the same conclusions by studying the inactivation of PAL.\textsuperscript{25}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.13.png}
\caption{Early model for the active site electrophile (from reference 22)}
\end{figure}
1.5 Evidence for Dehydroalanine as the Electrophile

The aforementioned role for a carbonyl residue in an enzyme was new, and in 1969 R. B. Wickner\textsuperscript{26} attempted to isolate the labeled species from inactivated HAL. Upon treatment with sodium borotritide and acid hydrolysis of the inactivated enzyme he was able to isolate the majority of radioactivity in a single amino acid, alanine. This finding resulted in a new model for a reactive electrophilic center in HAL involving dehydroalanine (DHA), previously observed only in microbial peptides.\textsuperscript{27} Presumably the inactivation involves nucleophilic attack of the tritide at the $\beta$ carbon of dehydroalanine, a process of conjugate addition (figure 1.14). The presence of tritiated alanine was confirmed by demonstrating co-migration with externally added $^{14}$C-alanine. This result led other investigators to try different radiolabeled inactivators, and a mechanism emerged involving nucleophilic attack of substrate amine on dehydroalanine (figure 1.15). This rather spartan proposal seems to be shifting responsibility for explaining the difficulty inherent in eliminating ammonia from the substrate to that of removing ammonia from the enzyme dehydroalanine. However, a more sophisticated model was on the horizon.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure114.png}
\caption{Model for generating tritiated alanine from inactivated HAL}
\end{figure}
During that same year I. L. Givot, T. A. Smith and R. H. Abeles\textsuperscript{28} published results that greatly expanded on Wickner's findings. They found evidence for further unsaturation in the dehydroalanine moiety and elaborated a new mechanism to include this feature. Givot et al. used nucleophilic $^{14}$C-nitromethane to permanently inactivate and label HAL, and demonstrated substrate protection. Depending on what order the subsequent proteolysis, reduction and acid hydrolysis were run, three different radioactive species were isolated. They proposed that the one common precursor for 4-amino-2-hydroxybutyrate, 2,4-diaminobutyrate and $\beta$-alanine was iminodehydroalanine (figure 1.16). One rationalization for the presence of the imine is that it prevents isomerization of the $\alpha,\beta$ double bond of dehydroalanine. This new unsaturation would turn out to play a key role in the first major proposal for the mechanism of HAL.
Figure 1.16 Evidence for iminodehydroalanine in HAL
1.6 The Hanson and Havir Mechanism for HAL

The discovery of the previously unknown electrophilic group at the active site of HAL and PAL soon led to proposals for how it interacts with the substrate. Kenneth Hanson and Evelyn Havir\(^{29}\) developed a model for the catalytic role of the iminodehydroalanine involving a nucleophilic addition of the histidine amino group at the methyldiene carbon, based in part on the inactivation products previously isolated (figure 1.17). The former imine carbon is protonated in 1 and an imine-to-enamine tautomerization occurs from 2 to 3.

\[
\text{Figure 1.17 Hanson and Havir HAL mechanism showing a covalent ammonia intermediate}
\]
At this point the covalent adduct can be thought of as a vinylogous amide that is now set up for the elimination of ammonia. The authors propose that as 3 is converted to 4 an internal carbonyl-amine hydrogen bond lowers the energy barrier to elimination. The overall effect is to enhance the leaving-group-ability of the amine. Reformation of the imine in 5 sets up the complex for the expulsion of ammonia and regeneration of the iminodehydroalanine in 6.

This rather elaborate proposal attempts to address the problem of the poor leaving group but says nothing about facilitating the difficult removal of the β proton. In 1970 Mildvan et al.\textsuperscript{30} offered an explanation when they showed via proton relaxation rates that the imidazole of the substrate histidine was coordinated to Mn\textsuperscript{2+} in HAL. The metal is therefore implicated in activating the β proton by electron withdrawal from the imidazole ring. This finding would also help explain why L-cysteine, containing a thiol capable of coordinating metals, is the most potent amino acid inhibitor of HAL, having a $K_i$ of 1 mM compared to a $K_M$ of 5 mM for L-histidine. This crucial role for the metal, however, seems somewhat untenable, as HAL’s activity is only diminished by a factor of between five and ten in the complete absence of divalent metals.\textsuperscript{19}

The Hanson and Havir mechanism at least rationalizes some of the findings from previous studies on HAL and PAL; it explains the outcome of labeling experiments using radioactive nucleophiles and theorizes a covalent ammonia-adduct that is released second, after urocanate. Its major shortcoming, however, is that it doesn’t satisfactorily explain the problem of hydrogen activation on the substrate β carbon. It would be another 25 years before an alternative mechanism attempted to address this failure.
1.7 Possible Origins of Iminodehydroalanine

The iminodehydroalanine postulated by Abeles et al. is the result of a post-translational modification of normal amino acids in HAL. Determining the nature of the precursor residues is an important step in understanding how this unusual electrophile forms. To this end the HAL gene from *Pseudomonas putida* HAL was cloned and sequenced by M. W. Consevage and A. T. Phillips in 1990. The DNA sequence revealed an open reading frame of 1,530 base pairs, corresponding to a protein subunit of molecular weight 53 574 Da. The N-terminal sequence as determined by Edman degradation of HAL exactly matched that predicted by the gene sequence (minus an initiating formyl methionine). The first five residues were determined to be Thr-Glu-Leu-Thr-Glu. This finding was important, because it eliminated one possible explanation for

![Diagram of protein carbonyl and iminodehydroalanine](image)

Figure 1.18 Generation of iminodehydroalanine from an N-terminal serine and a protein derived carbonyl

iminodehydroalanine attachment in HAL (figure 1.18). The protein carbonyl in the figure could originate from the peptide backbone or from a serine side-chain oxidized to an aldehyde. The presence of the N-terminal threonine immediately following the initiator amino acid meant that dehydroalanine could not be at the N-terminal position, since it can only be derived from a serine or cysteine residue. Additionally, the fact that the HAL molecule is identical in size to that predicted from the gene sequence excludes a second possibility, that HAL starts out as a much
larger progenitor with some internal excision leading to iminodehydroalanine. The remaining possibilities are that this electrophilic moiety is present as part of the main peptide chain or is somehow attached to a side chain.

1.8 Sequence Comparison and Mutants of HAL and PAL

In 1990, Taylor et al.\textsuperscript{32} published a protein sequence comparison of three HAL and four PAL genes that revealed sufficient amino acid homology for them to suggest that HAL and PAL utilize a common mechanism. There are several sections of absolutely conserved residues among all seven species, the longest of which is the stretch of five amino acids alanine, serine, glycine, aspartate and leucine, or ASGDL (table 1.2). For reasons of clarity these are numbered 142-146, according to their position in \textit{P. putida} HAL. Obviously, the more conserved a region of amino acids is between enzymes from different species, the more critical their role in enzyme function or structure. Serine 143 in particular became the focus of much attention; it became a favorite candidate for the precursor of dehydroalanine.

In 1994 Retey et al.\textsuperscript{33} published a key paper which showed that serine 143 dehydrates to form dehydroalanine in \textit{P. putida}. Using the technique of site-directed mutagenesis they

<table>
<thead>
<tr>
<th></th>
<th>138</th>
<th>139</th>
<th>140</th>
<th>142</th>
<th>143</th>
<th>144</th>
<th>145</th>
<th>146</th>
<th>147</th>
<th>148</th>
<th>149</th>
<th>150</th>
<th>151</th>
<th>152</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rat HAL</td>
<td>T</td>
<td>V</td>
<td>G</td>
<td>A</td>
<td>S</td>
<td>G</td>
<td>D</td>
<td>L</td>
<td>A</td>
<td>P</td>
<td>L</td>
<td>S</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>\textit{B. subtilis} HAL</td>
<td>S</td>
<td>L</td>
<td>G</td>
<td>A</td>
<td>S</td>
<td>G</td>
<td>D</td>
<td>L</td>
<td>A</td>
<td>P</td>
<td>L</td>
<td>S</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>\textit{P. putida} HAL</td>
<td>S</td>
<td>V</td>
<td>G</td>
<td>A</td>
<td>S</td>
<td>G</td>
<td>D</td>
<td>L</td>
<td>A</td>
<td>P</td>
<td>L</td>
<td>S</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>\textit{R. toruloides} PAL</td>
<td>T</td>
<td>I</td>
<td>S</td>
<td>A</td>
<td>S</td>
<td>G</td>
<td>D</td>
<td>L</td>
<td>S</td>
<td>P</td>
<td>L</td>
<td>S</td>
<td>Y</td>
<td>I</td>
</tr>
<tr>
<td>\textit{R. rubra} PAL</td>
<td>T</td>
<td>I</td>
<td>S</td>
<td>A</td>
<td>S</td>
<td>G</td>
<td>D</td>
<td>L</td>
<td>S</td>
<td>P</td>
<td>L</td>
<td>S</td>
<td>Y</td>
<td>I</td>
</tr>
<tr>
<td>\textit{P. crispum} PAL</td>
<td>T</td>
<td>I</td>
<td>T</td>
<td>A</td>
<td>S</td>
<td>G</td>
<td>D</td>
<td>L</td>
<td>V</td>
<td>P</td>
<td>L</td>
<td>S</td>
<td>Y</td>
<td>I</td>
</tr>
<tr>
<td>\textit{P. vulgaris} PAL</td>
<td>T</td>
<td>I</td>
<td>T</td>
<td>A</td>
<td>S</td>
<td>G</td>
<td>D</td>
<td>L</td>
<td>V</td>
<td>P</td>
<td>L</td>
<td>S</td>
<td>Y</td>
<td>I</td>
</tr>
</tbody>
</table>

\textit{Table 1.2} Protein sequence homology of three HAL and four PAL enzymes with bold conservations
exchanged this serine for an alanine (S143A) and found the correctly folded mutant to be devoid of catalytic activity. The obvious conclusion is that the alanine mutant lacks a hydroxyl group and is therefore incapable of dehydrating. A later result\textsuperscript{34} helped to confirm the fate of serine 143, whereby it was mutated into cysteine (S143C) to give an enzyme that was indistinguishable from the wild type. Chemical tests were performed to quantitate the number of cysteine residues present per HAL monomer and the S143C mutant and wild type enzymes gave the same result. The new cysteine residue seemed to have ‘lost’ its thiol functional group. This implied that hydrogen sulfide could be eliminated in the place of water from serine to generate dehydroalanine (figure 1.19).

![Figure 1.19 Elimination of water or hydrogen sulfide resulting in dehydroalanine](image)

**Figure 1.19** Elimination of water or hydrogen sulfide resulting in dehydroalanine

1.9 The Retey Mechanism for HAL

In 1995 M. Langer, A. Pauling and J. Retey introduced a radical new mechanism for the elimination of ammonia from histidine.\textsuperscript{35} Their model eschews the iminodehydroalanine of Abeles et al. and considers the electrophile to be a simple dehydroalanine. However, they propose a dramatically different role for it. The authors believe the true function of the dehydroalanine is not to bond to the departing amine but rather to activate the β proton for removal via an initial Friedel-Crafts-type electrophilic addition to the aromatic ring of histidine.\textsuperscript{36} Specifically, attack from the 5-imidazole position is believed to lead to an immonium intermediate in which the β-hydrogen is activated for removal by a catalytic base from the enzyme (figure 1.20).
Figure 1.20 Retey mechanism for HAL with L-histidine or L-5-nitrohistidine. \(X=H\), no reactivity without DHA; \(X=NO_2\), good reactivity without DHA

The original impetus for Retey’s new mechanism was the behavior of the S143A dehydroalanine-less mutant HAL with a histidine analogue, 5-nitrohistidine, which is a moderately good substrate of HAL. Those findings supported a model in which the nitro group on the substrate was acting as a ‘surrogate dehydroalanine’ in those cases where the dehydroalanine had been eliminated through either mutagenesis or borohydride reduction. With 5-nitrohistidine as a substrate the \(K_M\) and \(V_{max}\) values were essentially identical among wild type, mutant and borohydride inactivated HAL. Earlier it had been shown that the kinetic deuterium isotope effect observed with \([\beta^{-2}H_2]\) histidine disappears when 5-nitro-[\(\beta^{-2}H_2\)] histidine is the substrate, indicating that the (at least partially) rate determining step of proton abstraction is made much more favorable by the presence of the nitro group. It therefore seemed reasonable that the role of dehydroalanine also serves to activate the substrate for proton removal by forming a sigma complex with the histidine imidazole at the C5 position.
There are several serious problems with the Retey mechanism as it is formulated in figure 1.20. Firstly, it completely ignores the previously established imine component of the electrophile. Secondly, there is a loss of aromaticity in the imidazole after it attacks to form the sigma complex, which would not be offset through generation of the enol form of dehydroalanine. Thirdly, it depicts the opposite order of product release as determined by earlier investigators. However, the bold proposal for β proton activation is the only mechanism to seriously address this somewhat overlooked aspect of the elimination of ammonia from histidine. It also helps explain the exchangeability of the C5 imidazole proton noted several years earlier.37

1.10 The Exchangeability of Imidazole Protons

In section 1.2 it was mentioned that there was evidence for an E1cb pathway in HAL and PAL. A key finding was the exchangeability of the imidazole C5 protons with solvent by T. Furata et al.37 They demonstrated this by preparing L-[5-²H] histidine and monitoring urocanate (and histidine) at C5 for deuterium washout in the presence of HAL. They observed a steady increase in the incorporation of solvent derived protons in urocanate at C5, consistent with a reversible stepwise mechanism involving a carbanion intermediate. This finding supports the Retey mechanism, as a reversible deprotonation at C5 is possible from the sigma complex formed in the initial nucleophilic attack of histidine imidazole (figure 1.21).
Figure 1.21 Loss of deuterium label at C5 is explained by a covalent intermediate generated via the Retey mechanism.
1.11 Thesis Aims

Two mechanisms have been proposed for the elimination of ammonia from histidine and phenylalanine that employ an electrophile at the enzyme active site. However, the catalytic function of this crucial element is in dispute with radically different roles proposed in each case. Each of the mechanisms may be used to explain particular observations of enzyme/substrate behavior, but neither can rationalize all of them. In order to have a sound foundation from which to investigate the mechanism of HAL it is essential to ascertain the structure of the enzyme electrophile, which is variously proposed as dehydroalanine or iminodehydroalanine. Only then will we be confident enough to design experiments that address the ambiguity of the role of the electrophile in the ammonia-lyases HAL and PAL.

The first goal of this thesis is to firmly establish the structure of the catalytic electrophile in HAL and PAL by recovery of labeled peptides containing the electrophile. Physical methods will be employed to measure peptide masses, nuclear resonances and UV absorbances. Once this goal has been achieved the second can be addressed; determining which of the two mechanistic proposals is correct, if either.

A series of histidine substrate analogues will be synthesized and tested for the elimination of ammonia with HAL. These compounds will contain structural features that may allow them to be processed by the enzyme to generate ammonia, but only by utilizing one of the two possible mechanisms. For example, one compound might yield urocanate with HAL utilizing the Retey mechanism but would not be expected to do so on mechanistic grounds if the Hanson and Havir model is correct. A second compound would be designed to test the opposite hypothesis.

Additional compounds will be synthesized in an attempt to inhibit the enzymes with the consideration of electrophile participation. It is our belief that, taken together, these data will allow a sound model for the mechanism of HAL and PAL to emerge.
Chapter Two

Inhibition of Phenylalanine Ammonia-lyase
2.1 Introduction

The term ‘primary metabolite’ is used in the classification of biological molecules to denote a component of an organism that is involved in essential life processes and has a broad distribution in all living things. A complex web of enzyme-catalyzed reactions, often cyclical in nature, determines an array of these products that are indispensable for the maintenance of core cellular functions. The amino acids are good examples of primary metabolites, and their biosynthesis and ultimate usage has been of interest from the earliest days of biochemistry. Not surprisingly, there are separate biosynthetic pathways for amino acids with distinctive side-chains, such as the aliphatic and aromatic amino acids.

The first step in the synthesis of the aromatic amino acids tryptophan, tyrosine and phenylalanine is the coupling of phosphoenol pyruvate and erythrose-4-phosphate at the beginning of the Shikimate (or Chorismate) pathway. In the assembly of phenylalanine a further nine steps are required, each employing a different enzyme. In most microbes the last step is the transfer of an amino group from glutamate to phenylpyruvate by an amino transferase (figure 2.1). At this

![Chemical reaction](image_url)

*Figure 2.1 The final step in the biosynthesis of L-phenylalanine*
point there can be several fates for phenylalanine; hydroxylation to tyrosine (in animals and some microbes), incorporation into proteins or peptides, degradation to starting materials or, in the case of plants, shuttling into the phenylpropanoid pathway via PAL.

PAL is an important enzyme in plants because it is a branch point from primary to secondary metabolism. As the name suggests, a secondary metabolite is not necessarily an element that is essential for the basic survival of an organism, but a compound that imparts some benefit to the host, such as antiparasitism or the ability to withstand drought. There are many examples of secondary metabolites from plants and microorganisms that also provide benefits to humans in the form of drugs, including the alkaloids, penicillins and polyketides.

The enzymes responsible for secondary metabolism usually become active under conditions in which their products become most beneficial to organisms. For instance, PAL’s regulation is linked to physical wounding, pathogen infection, exposure to UV radiation and seasonal plant growth cycles. A variety of downstream products (coumarins, flavonoids) help the plant repair, protect and fortify itself. PAL’s direct product, trans cinnamate, may eventually become oxidized in the aromatic ring, methylated to various degrees and reduced to coniferyl alcohol, a monolignol (figure 2.2). At least three different types of monolignols are subsequently joined together to form the cross-linked polymer lignin, imparting rigidity and water impermeability to plant tissues and helping the plant withstand external stresses.

\[
\begin{align*}
\text{trans Cinnamate} & \rightarrow \text{Various enzymes} \rightarrow \text{Coniferyl alcohol} \\
\end{align*}
\]

Figure 2.2 The transformation of cinnamate to coniferyl alcohol
Arising largely from the commercial issues surrounding removal of lignin from pulped wood, researchers are interested in the regulation of lignin biosynthesis via control of PAL activity. One approach is the development of inhibitors of the enzyme, as the diminished activity of PAL could eventually lead to an altered lignin profile in vivo. Thus, one of the earliest aims of this project was to synthesize a fluorine-substituted analogue of phenylalanine and evaluate it as an inhibitor of PAL from poplar (figure 2.3). By replacing the abstractable pro-S hydrogen with fluorine it was hoped that a tight binding substrate analogue might be realized. There is precedence for fluorinated amino acids inhibiting enzymes. At this point in the project we were considering the validity of the Hanson and Havir mechanism, so that this compound was designed to take advantage of the proposed amine-enzyme linkage to produce a covalently bound inhibitor.

![L-Phenylalanine and L-(3S)-3-Fluorophenylalanine](image)

*Figure 2.3 L-(3S)-3-Fluorophenylalanine with fluorine in place of the abstractable proton*

There are several modes by which 3-fluorophenylalanine might inhibit enzyme activity. Conceivably, once in the active site of PAL, the 3-fluorophenylalanine could accumulate as an enzyme adduct through the amine functional group, shown in the bottom of figure 2.4 employing the Hanson and Havir mechanism. It was believed that this covalent bond between the amine and the prosthetic group might be detected, perhaps through mass analysis of proteolytic fragments of PAL treated with the analogue. In addition, a properly positioned nucleophilic residue could directly displace the fluorine, leading to a second mode of attachment (figure 2.4, upper figure). Or the inhibitor may simply occupy the binding site with tenacity.
Three of the most potent inhibitors of PAL discovered so far are (S)-2-aminooxy-3-phenyl propionic acid (AOPP), (R)-(1-amino-2-phenylethyl) phosphonic acid (APEP) and 2-aminoundan-2-phosphonic acid (AIP) (figure 2.5).\textsuperscript{41} Each of these involves a modification in the amine or carboxylate of phenylalanine, and AIP introduces a ring. It has not been determined how the structural changes in these inhibitors relate to the ordering of their $K_i$ values. The best of the three, AOPP, simply contains an oxygen atom inserted between the $\beta$ carbon and amine.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure2_4}
\caption{Two possible modes of inhibition of PAL by fluorophenylalanine}
\end{figure}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure2_5}
\caption{Structures and $K_i$ values of three potent PAL inhibitors}
\end{figure}
Our strategy of fluorine for hydrogen replacement assumes that the basic residue responsible for removal of the β proton would be incapable of removing the fluorine atom, and that the relatively minor change to the steric profile of the molecule would allow it access to the binding domain of the enzyme. Depending on how well the compound performs against PAL in vitro, it may be tested in poplar cell cultures as a general inhibitor of phenylpropanoid biosynthesis.

2.2 Results and Discussion

2.2.1 Isolation and Purification of PAL

Recombinant PAL from Populus trichocarpa x deltoides was obtained from heterologous expression of the corresponding cDNA in insect cell cultures via a baculovirus expression vector. Cells were lysed, filtered and subjected to linear gradient anion exchange FPLC. Fractions containing PAL activity were combined and stored at −20°C (figure 2.6). An SDS PAGE gel of active fractions showed a major band corresponding to a 78 kDa PAL standard. Further purification was not attempted partly because of the sensitivity of the enzyme and because our goal of $K_i$ determination does not require extreme purity. The $K_M$ for phenylalanine was determined to be 250 μM (figure 2.7), which was in agreement with that determined elsewhere.
Figure 2.6 FPLC chromatogram showing major peak corresponding to PAL and SDS PAGE gel of FPLC fractions indicating PAL in #23, 24

Mar = Broad range molecular weight marker
Std = PAL standard 78 kDa

Figure 2.7 Reciprocal plot for PAL with L-phenylalanine as substrate
2.2.2 Synthesis of 3-Fluorophenylalanine

![Erythro enantiomers](image1)

L-(3S)-3-Fluorophenylalanine 1a  D-(3R)-3-Fluorophenylalanine 1b

![Threeo enantiomers](image2)

D-(3S)-3-Fluorophenylalanine 1c  L-(3R)-3-Fluorophenylalanine 1d

Figure 2.8 Structures of the erythro and threeo isomers of 3-fluorophenylalanine

Racemic 3-fluorophenylalanine 1 was synthesized as erythro and threeo stereoisomers according to the method of M. J. O’Donnell, C. L. Barney and J. R. McCarthy\(^4\) (figure 2.8). The primary target was one of the erythro enantiomers 1a, as it contains both the amino group in the proper L configuration and a fluorine atom in place of the abstractable pro-S hydrogen. Erythro and threeo 3-fluorophenylalanine were separated as their benzophenone imine benzyl esters. The retrosynthesis indicates a key step is the alkylation of protected glycine by a fluorinated benzyl synthon (figure 2.9).

![Retrosynthesis of 3-fluorophenylalanine](image3)

Figure 2.9 Retrosynthesis of 3-fluorophenylalanine
The benzyl ester of glycine 2 was first prepared by Fischer esterification with the azeotropic removal of water from glycine and benzyl alcohol. A trans-imination followed, using the benzophenone imine 3 prepared from phenylmagnesium bromide and benzonitrile, leading to fully protected glycine 4 (figure 2.11).

The traditional approach to generating an imine is to condense an amine with an aldehyde or ketone while removing the water produced. However, the desire for chromatographic separation of the erythro and threo stereoisomers required the use of the stable yet bulky benzophenone imine of glycine, which proved difficult for O'Donnell et al. to synthesize. The major problem was that the dehydrating conditions typically employed lead to cyclization between two glyclyl esters (figure 2.10). By employing a trans-imination procedure the desired product (4) was produced in high yield.

\[
\begin{align*}
2 \text{H}_2\text{N} & \text{CO}_2\text{Et} \quad \text{-} 2 \text{EtOH} \\
\end{align*}
\]

\[\text{Figure 2.10 Cyclization of glyclyl esters occurs under traditional imination conditions}\]

Benzyl fluoride 5 was prepared from benzyl bromide and potassium fluoride by nucleophilic substitution, utilizing the potassium coordinating polyether PEG 200 to enhance the nucleophilicity of the fluoride ion. Bromine was re-introduced in the following step with N-bromosuccinamide to give \(\alpha\)-bromo-\(\alpha\)-fluorotoluene 6. This very unstable compound was quickly coupled to 4 to give 7, which was resolved into erythro and threo stereoisomers through several rounds of flash column chromatography. Treatment of each with trimethylsilyl iodide followed by an acidic work-up gave the hydrochloride of 1. The free acids were obtained by treatment with propylene oxide.
Figure 2.11 Synthesis of erythro and threo 3-fluorophenylalanine
The alkylation of the protected glycine equivalent was carried out via phase transfer catalysis to minimize dialkylation, as the introduction of a fluoroalkyl group has precedence for activating the second $\alpha$ hydrogen by induction.\textsuperscript{49} This strategy of using a pre-fluorinated alkylating agent eliminates the need for such harsh reagents as fluorine gas or sulfur tetrafluoride in hydrofluoric acid.\textsuperscript{50} However, the milder fluorinating reagent diethylaminoisulfur trifluoride (DAST) has successfully been employed in fluorodehydroxylation reactions leading to similar types of erythro and threo compounds.\textsuperscript{51}

The presence of a benzyl ester protecting group allowed for easy deprotection with trimethylsilyl iodide, thus avoiding problematic hydrolytic conditions where fluorine can be lost.\textsuperscript{52} The characteristic splitting of the $\alpha$ and $\beta$ $^1$H NMR signal due to coupling to the $\beta$ $^{19}$F were observed for both erythro and threo products (figure 2.12).

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure212.png}
\caption{Characteristic proton splitting in protected erythro 3-fluorophenylalanine}
\end{figure}
2.2.3 Inhibition of PAL by 3-Fluorophenylalanine

_Erythro_ and _threo_ 3-fluorophenylalanine were tested as inhibitors for PAL at concentrations from 30 to 120 mM. Solutions of higher concentration were not prepared due to the limited amount of I available. It was found that the pair of _erythro_ enantiomers (1a and 1b) inhibited PAL competitively, with a $K_i$ value of 45±5 μM (figure 2.13), while the _threo_ pair (1c and 1d) had no detectable effect on the reaction rate (figure 2.14).

The contribution to the inhibition from the D- (3R) isomer 1b of the _erythro_ pair is questionable, as D-phenylalanine is itself neither a substrate nor an inhibitor for PAL. If that is true the $K_i$ value for the L- (3S) 1a is actually closer to 22 μM, half of that determined for the equimolar mixture of 1a and 1b. This value is an order of magnitude smaller than the $K_M$ for the natural substrate L-phenylalanine, but much larger than those of the inhibitors mentioned in figure 2.5. One explanation for the tighter binding of 1a is that the introduction of the fluorine atom in phenylalanine allows for the possibility of a new dipolar interaction in the active site. Fluorine is the most electronegative atom and it is conceivable that it could be interacting with a positively charged residue in PAL. Unfortunately, the relatively poor performance of these fluorinated inhibitors compared to those in figure 2.5 meant that they would not be tested for phenylpropanoid inhibition _in vivo_.

39
Figure 2.13 Erythro isomers of 3-fluorophenylalanine exhibit competitive inhibition

Figure 2.14 Threo isomers of 3-fluorophenylalanine exhibit no inhibition
The observed lack of inhibition with the *threo* isomers is interesting, as 1d has the proper L configuration and was expected to be similar enough to L-phenylalanine to bind to the active site of PAL. In fact, the effect of electron withdrawal due to the geminal fluorine atom could be expected to enhance the acidity of the abstractable proton, allowing 1d to be a viable substrate. Apparently though, the only effect of the fluorine atom is to prevent binding, perhaps by forcing the adoption of a stable conformation. T. Tsushima *et al.* have published a study on rotamer populations for both L-*erythro* and L-*threo* 3-fluorophenylalanine based on analysis of coupling constants at various pH and temperatures (figure 2.15). The report demonstrates the stabilizing hydrogen bond that exists between the fluorine and ammonium groups in each case. It seems likely that the best candidates for inhibiting PAL would have a conformation with *anti* carboxylate and phenyl groups in order to mimic the incipient *trans* double bond, which 1d would seem to display according to the top box in the figure. Their findings, however, do not show that one rotamer is dramatically favoured over the others and therefore do not help to explain why 1a is an inhibitor while 1d is not. Perhaps the stabilizing hydrogen bond effect is retarded in the region of

![Chemical structures](image)

*Figure 2.15 Rotamer populations for 1a and 1d from Tsushima *et al.* The most populated conformer is boxed.*
PAL's binding pocket. Alternative explanations could include unfavorable dipole-dipole or steric interactions with active site residues that lower the binding constant of 1d.

The initial testing of the isomers of 1 with PAL was optimized to detect reversible inhibition on a time scale of a few minutes. As alluded to in figure 2.4, the possibility remained that erythro and threo 3-fluorophenylalanine were either very slowly permanently inactivating PAL (perhaps through a covalent adduct) or were being processed as alternate substrates, leading to 3-fluorocinnamate (figure 2.16). The production of fluorocinnamate might not have been noticed during the normal PAL assay if this compound doesn't absorb at 290 nm. An alternative assay was therefore developed.

Both erythro and threo 1 were pre-incubated with PAL for several hours and at regular intervals aliquots were removed and assayed for activity. In each case no decrease in PAL activity was observed indicating a lack of slow, irreversible inhibition. Additionally, the incubation mixtures showed no change in absorbance between 200 and 500 nm, indicating that 3-fluorocinnamate was not being slowly produced. It was therefore concluded that the only discernable effect of 3-fluorophenylalanine on PAL was that the erythro stereoisomers (and presumably the L-(3S) enantiomer 1a) were inhibiting PAL competitively with a $K_i$ of 45 µM.

![Chemical structure](image)

**Figure 2.16 3-Fluorophenylalanine as a substrate for PAL**
2.3 Materials and Methods

2.3.1 General Information

All buffers and synthetic reagents were purchased from the Aldrich or Sigma Chemical companies unless explicitly stated. Water for buffers and enzyme assays was either double distilled or filtered through a Barnstead NANOpure ultrapure water filtration system. FPLC buffers were filtered through a 0.22 μm sterilizing filter and degassed prior to use. Maintenance of *Spodoptera frugiperda* (Sf9) cells expressing PAL from recombinant PAL (pVL1393) and viral (AcNPV, BaculoGold, Pharmagen) DNA was performed by Jennifer Norton in the Department of Plant Science. Protein concentrations were determined using the Coomassie dye-binding assay with BSA as a standard.

Flash column chromatography was performed with Silica Gel 60 (230-400 mesh, E. Merck, Darmstadt). Methylene chloride, acetonitrile and isopropanol were freshly distilled from calcium hydride. Analytical thin layer chromatography was performed on aluminum-backed sheets of Silica Gel 60F254 (Merck) of thickness 0.2 mm. Visualization of synthetic compounds was accomplished by spraying plates with a mixture of sulfuric acid (30 mL), cesium sulfate (1 g) and ammonium molybdate (20 g) in water (0.5 L) followed by heat charring. $^1$H and $^{19}$F NMR spectra were recorded on a Bruker AC 200E instrument at a field strength of 200 MHz or a Bruker WH-400 instrument at 400 MHz.

2.3.2 Purification of PAL

SF9 cells were grown from a density of $5 \times 10^5$ cells mL$^{-1}$ to $2 \times 10^6$ cells mL$^{-1}$ in a 50 mL spinner flask and then infected with the recombinant virus AcNPV-PAL at a multiplicity of infection (MOI) of 1.$^{11}$ After 48 hours the cultures were centrifuged (1 500 x g, 10 min, 4 °C), the
medium was decanted and the cell pellet washed in 10 mL cold PBS. The cells were centrifuged again, resuspended in 5 mL 50 mM Tris-acetate pH 7.5 containing 5 mM β-mercaptoethanol, lysed with a Dounce homogenizer (20 strokes) and vortexed 1 min. This crude lysate was centrifuged (10 000 x g), filtered (0.45 µm) and subjected to anion exchange FPLC using an analytical HR 5/5 Mono Q column. Solvent A: 50 mM Tris-acetate, 5 mM β-mercaptoethanol, pH 7.5; solvent B: 500 mM sodium chloride in solvent A. Gradient conditions: 0% B for 10 min; 20% B for 20 min; 20 to 60% B over 60 min; 100% B for 10 min. Flow rate 0.5 mL min⁻¹. Fractions were collected at 3 min intervals and were monitored at 280 nm.

2.3.3 Activity Assay for PAL

PAL was assayed spectrophotometrically by following changes at 290 nm and 30 °C due to increase in concentration of product cinnamate. The pre-assay buffer consisted of Tris-acetate (100 mM, pH 8.8) containing 5 mM β-mercaptoethanol. The assay buffer contained L-phenylalanine (10 mM). The test for activity was run as follows: Pre-assay buffer (190 µL) was pipetted into a quartz cuvette and a constant amount of enzyme solution (10µg/10 µL) was introduced, mixed and warmed to 30 °C. Assay buffer (800 µL) was then added, mixed and monitored at 290 nm. Slopes of the absorbance versus time curves were converted into initial rates using an extinction coefficient e for cinnamate of 10 000 L mol⁻¹ cm⁻¹.

2.3.4 Synthesis of 3-Fluorophenylalanine

The synthesis of erythro and threo 3-fluorophenylalanine has been described by M. J. O'Donnell et al. Our data for 1a,b and 1c,d match that of the literature.
2.3.5 Inhibition Assays for PAL

Solutions of L-phenylalanine (1.00 to 0.05 mM) were prepared using the standard PAL assay buffer. For each of these substrate concentrations a second series of solutions were prepared of the erythro and threo 3-fluorophenylalanine (0.12 to 0.03 mM). These solutions comprised an array of substrate and inhibitor concentrations that were assayed as in section 2.3.3. The initial rates that were subsequently computed were plotted as the inverse of rate versus the inverse of substrate concentration for each concentration of inhibitor. To determine the $K_i$ value the slopes of these lines were plotted against inhibitor concentration; the intercept on the concentration axis gave the $K_i$ value.

The slow binding assay was performed as follows: PAL (100 µg/100 µL) was injected into solutions of erythro and threo 1 (1.0 mL, 200 µM) and incubated at 30 °C. At regular intervals (60 min) aliquots were removed (50 µL) and assayed for enzyme activity as in 2.3.3.

The assay for erythro and threo 1 as substrates for PAL was identical to the previous assay, except no aliquots were removed. Instead a UV spectrum was run (200-500 nm) on each solution at regular intervals (30 min) and examined for absorbance changes.
Chapter Three

Structural Determination

of the HAL Electrophile
3.1 Introduction

The presence of an unusual electrophile at the active site of HAL and PAL has been known for almost thirty years. Over that time several models have been proposed to rationalize its reactivity to nucleophilic reagents and explain the products obtained from various labeling studies. As might be expected, the earliest explanations were also the simplest, starting with one that proposed a single reactive carbonyl. A simple dehydroalanine was later invoked, followed soon after by iminodehydroalanine. Chapter One illustrated how the Retey and Hanson and Havir mechanisms incorporated these last two electrophile models. The reactive center appears to form autocatalytically, as heterologous expression results in fully active enzyme. Whatever its form, the electrophile is a crucial part of a unique system participating in the very difficult elimination of ammonia from unactivated substrates. The challenges of deciphering its structure make HAL and PAL exciting enzymes to investigate.

An interesting feature of the electrophile in HAL is the manner in which it is transformed upon treatment with cysteine. In 1974 Klee\textsuperscript{55} performed a series of experiments on \textit{Pseudomonas} HAL which showed that at high pH the enzyme could be permanently inactivated with cysteine, accompanied by the formation of an enzyme bound chromophore which absorbs at 340 nm. This finding would prove to be very important for the development of our own theories about the structure of the HAL electrophile and to ultimately promote one mechanism over the other.

Cysteine is known to interact with HAL in three distinct ways. First, cysteine contains an oxidizable thiol function and can reduce the less reactive, oxidized form of HAL to its more reactive form, presumably by the reduction of disulfide bonds. This is the chief reason that thiols are included in all HAL assay buffers. Apparently at least three stable disulfide bonds can form between the enzyme and buffer thiols, as suggested in section 3.2.1. Second, L-cysteine is a good, reversible competitive inhibitor of HAL in the normal pH range of enzyme activity. At pH above
10, however, the third way in which L-cysteine interacts with the enzyme manifests itself, through a rapid, irreversible inhibition. This last effect is accompanied by an increase in absorbance at 340 nm and is dependent on the presence of atmospheric oxygen. The last two inhibitory effects merit closer scrutiny.

<table>
<thead>
<tr>
<th>Inhibitor</th>
<th>$K_i$ or $K_M$ (mM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L-Histidine</td>
<td>7.0 ($K_M$)</td>
</tr>
<tr>
<td>D-Histidine</td>
<td>15.6</td>
</tr>
<tr>
<td>L-Cysteine</td>
<td>0.9</td>
</tr>
<tr>
<td>Glycine</td>
<td>6</td>
</tr>
<tr>
<td>L-Alanine</td>
<td>&gt;100</td>
</tr>
</tbody>
</table>

Table 3.1 $K_M$ or $K_i$ values for various amino acids and HAL (from reference 22)

Abeles et al.\textsuperscript{22} studied the inhibitory effect of a variety of amino acids on HAL (table 3.1), and found that L-cysteine was the best competitive inhibitor. Apparently the thiol of cysteine mimics the imidazole of histidine well enough to bind tightly to HAL’s active site. Closer comparison of their structures shows that the sulfur may be positioned in close proximity to a residue in HAL that normally forms a key hydrogen bond to the N3 of the imidazole (figure 3.1). The sulfur of cysteine could also be reversibly forming a covalent bond to the active site electrophile in a manner analogous to the imidazole of histidine in the Retey mechanism. A third rationalization is that the cysteine thiol is coordinating to the metal ion that other investigators have determined to be in close proximity to the imidazole of bound substrate.\textsuperscript{30} A comparison of the side chain $pK_a$s show that, in free solution at least, both the imidazole and the thiol would be neutral at the pH optimum for HAL, 9.0. However, the thiol is largely deprotonated at the higher pH conditions under which cysteine permanently inactivates HAL.
Klee performed her cysteine inactivation experiments at pH 10.5 in the presence of oxygen and cysteine. She observed the formation of the new UV absorbing species at 340 nm lagging behind the inactivation of HAL (figure 3.2, left), which lost all activity within two hours. In the absence of cysteine, HAL is stable at pH levels as high as 11.8 for hours. In addition, no permanent inactivation or chromophore formation resulted from the incubation of HAL and cysteine in an atmosphere of pure nitrogen.

It is appropriate to note here that in addition to the normal protein UV absorbances at 220 and 280 nm, HAL exhibits a small shoulder at 315 nm (figure 3.2, right). This feature becomes obscured by the much stronger absorption at 340 nm as the inactivation proceeds. The shoulder absorbance has been largely ignored in the literature but will turn out to play a role in our model for the HAL electrophile and will be elaborated at the end of this section.
Klee wanted to test other thiols for this effect and repeated the experiment with a variety of similar compounds (table 3.2). She showed that histidine and urocanate could protect HAL from inactivation, suggesting the attack occurred at the active site electrophile. Mercaptoethylamine showed the next best level of HAL inactivation after cysteine, at 13% loss of enzyme activity after two hours, while β-mercaptoethanol showed half that. Inactivation with $^{35}$S-labeled cysteine led to inactive HAL with one equivalent of $^{35}$S label incorporated per enzyme monomer. This label could only be removed under extremes of pH and denaturing conditions. Taken together, the data suggest a nucleophilic attack of cysteine thiolate at the electrophilic active site of HAL coincident with an oxygen dependent permanent inactivation and lagging chromophore generation.

<table>
<thead>
<tr>
<th>Additions (5 mM)</th>
<th>Activity T = 0</th>
<th>Activity T = 120 min</th>
<th>% Loss of Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>4.9</td>
<td>5.0</td>
<td>0</td>
</tr>
<tr>
<td>L-Cysteine</td>
<td>4.7</td>
<td>&lt;0.05</td>
<td>100</td>
</tr>
<tr>
<td>Mercaptoethylamine</td>
<td>3.8</td>
<td>3.3</td>
<td>13</td>
</tr>
<tr>
<td>β-Mercaptoethanol</td>
<td>5.0</td>
<td>4.7</td>
<td>6</td>
</tr>
<tr>
<td>L-Cystine</td>
<td>5.4</td>
<td>5.3</td>
<td>2</td>
</tr>
</tbody>
</table>

*Table 3.2 Effect of sulfhydryl compounds on the activity of HAL over time at pH 10.5. Activity in μmol per min per mL (from reference 71)*
Nearly twenty years later, D. Hernandez, J. G. Stroh and A. T. Phillips\textsuperscript{56} expanded on Klee’s findings. They attempted the inactivation with a greater range of cysteine-based inhibitors, and found that analogues with a blocked or missing α-amine could inactivate HAL but were not able to form the chromophore, and that a free thiolate was essential for enzyme inactivation (figure 3.3). The methyl ester of cysteine was also discovered to inactivate HAL and form a chromophore almost as efficiently as free cysteine. From this data one can conclude that a free thiolate is required for inactivation, a free amine is subsequently required to generate the chromophore and the carboxylate is not involved in either process. Oxygen is somehow involved in one or both steps.

Hernandez \textit{et al.} repeated the cysteine inactivation experiment on \textit{Pseudomonas putida} HAL but went much further and isolated peptide fragments of inactivated enzyme that contained the UV absorbing species. Following reductive carboxymethylation of the cysteine inactivated enzyme using sodium borohydride and iodoacetate they embarked on two rounds of proteolytic

\textsuperscript{51}
digestion and HPLC purification. The initial chromophoric peptide obtained from treatment with the first protease, trypsin, was twenty-four amino acids long, and the second, obtained from further treatment with V8 protease, was eight. Both were isolated using reverse-phase chromatography by following their unique absorbance at 340 nm, and were found to have identical N-terminal sequences of GSVG. The fifth position in the sequence was somehow blocked. This data unambiguously identified the octapeptide fragment as GSVGASGD, or residues 138 to 145. This includes the highly conserved ASGD motif noted in table 1.2 on page 22, and contains the key ‘dehydroalanine’ precursor, serine 143. The authors then proceeded to use mass spectrometry to analyze the octapeptide.

They discovered that the mass of the isolated peptide was larger than the sequence prediction by 141 Da. In order to establish where the extra mass on the peptide was situated they employed secondary ion fragmentation to cleave the peptide bonds inside the mass spectrometer and measure the masses of the fragments produced. By comparing the masses of the known amino acids to those in the fragmentation spectrum* it was clear that the extra mass was situated on serine 143 (figure 3.4)! This clear link between the chromophore and the active site electrophile made it a very compelling target for us.

![Figure 3.4 Fragmentation pattern of the modified octapeptide from HAL showing an extra mass on serine 143 (from reference 56)](image_url)

* The only peak in the fragment profile that they could not justify was the peak of mass 611.8 [M+H]⁺. We also observed this peak in a different experiment and an explanation is offered in section 3.2.3.2.
In 1996 K. Weber and J. Retey\textsuperscript{57} performed the inactivation experiment following the procedures of Hernandez \textit{et al.} and isolated a chromophoric peptide containing an identical extra mass. They proposed a structure for their 789.4 Da peptide that involved an arginine residue attached to an oxidized dehydroalanine moiety (figure 3.5). Their proposal seems to be almost entirely based on the observation of an arginine residue in the amino acid analysis of their octapeptide and excludes any reference to iminodehydroalanine. They employed conditions that were identical to those of Hernandez \textit{et al.}, including the reductive carboxymethylation step. This interpretation will be revisited following the presentation of our findings in section 3.2.3.2.

\textbf{Figure 3.5} The Retey model for a chromophoric octapeptide (boxed) involving attachment of an arginine residue 'carved out' of HAL during the proteolysis steps (from reference 57)
Our curiosity about the structure of the chromophore from cysteine inactivated HAL led us to speculate upon the striking resemblance between a known chromophore, that of green fluorescent protein (GFP) and iminodehydroalanine (figure 3.6). We felt (especially in light of the 340 nm chromophore results) that a protein backbone cyclization analogous to that in the GFP chromophore could be operative in HAL, resulting in a five-membered ring comprising the iminodehydroalanine. The major difference between the two is that an oxidation of tyrosine 66 in GFP introduces the exocyclic unsaturation whereas a second dehydration of serine 143 accomplishes this task in HAL. The extended conjugation that the five-membered ring affords may also explain the 315 nm shoulder observed in the UV spectrum of unmodified HAL. It seemed reasonable that the oxidative attachment of some species during the cysteine inactivation of HAL could both shift the UV absorbance from 315 to 340 nm and also increase its intensity. We therefore endeavored to isolate the chromophoric octapeptide in quantities sufficient for physical characterization by one and two-dimensional NMR, high-resolution mass spectrometry and amino acid analyses. We hoped that the data we obtained might confirm our model for the electrophile in HAL and give us insight into the true mechanism of the enzyme.
Figure 3.6 Comparison of the structure of the chromophore from green fluorescent protein and a similar hypothetical cyclic iminodehydroalanine in HAL. The $\lambda_{\text{max}}$ speculation is based on the observed 315 nm shoulder absorption in unmodified HAL.

During the course of this work the crystal structure of HAL was solved by T. Schwede, J. Retey and G. E. Schulz\textsuperscript{59} that confirmed our model of the HAL electrophile. A five-membered ring containing the three amino acids ASG, which the authors name as a 4-methylidene imidazole-5-one, or MIO group, is visible at the active site in the structure from \textit{P. putida} HAL. A second publication\textsuperscript{60} demonstrated that the 315 nm shoulder in unmodified HAL is indeed due to absorbance by the MIO group. The MIO abbreviation will be used from now on to describe the cyclic iminodehydroalanine to prevent confusion. These two papers will be discussed in more detail in section 3.4.
3.2 Results and Discussion

3.2.1 Isolation and Purification of HAL

HAL from *Pseudomonas putida* was overexpressed in *Escherichia coli* from a thermally inducible pPl-Lambda expression vector. Results from a representative purification are shown in table 3.3. Bacteria were centrifuged from growth media and resuspended in phosphate buffer at 5 °C before lysis on a French press. Cellular debris was pelleted by centrifugation and the cell lysate heated to 75 °C for eight to ten minutes. The heat treatment step greatly increased the specific activity of the crude HAL and is only possible because HAL is stable at temperatures up to 80 °C. Consequently, the majority of the unwanted protein denatures and coagulates as a paste. The

<table>
<thead>
<tr>
<th>Stage of purification</th>
<th>Total protein (mg)</th>
<th>Specific activity (units/mg)</th>
<th>Total activity (units)</th>
<th>Percent recovered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell lysate</td>
<td>11 000</td>
<td>2.1</td>
<td>23 100</td>
<td>100</td>
</tr>
<tr>
<td>Heat step</td>
<td>1 700</td>
<td>10.9</td>
<td>18 500</td>
<td>80</td>
</tr>
<tr>
<td>Concentration</td>
<td>1 550</td>
<td>10.8</td>
<td>16 700</td>
<td>72</td>
</tr>
<tr>
<td>DE 52</td>
<td>400</td>
<td>33</td>
<td>13 200</td>
<td>57</td>
</tr>
</tbody>
</table>

*Table 3.3 Summary of the purification of HAL*

soluble HAL was then concentrated by centrifuge filtration to approximately 15 mg/mL and subjected to linear gradient anion exchange chromatography on a column of diethylaminoethyl beads (DE 52). At this point the enzyme was deemed sufficiently pure (~90%, figure 3.7) to proceed with the inactivation experiment because it is subjected to inactivation, digestion and chromatography in the following stage. In addition, small amounts of enzyme were subjected to a second round of purification on a MonoQ quaternary ammonium column for inhibitor studies and mass spectral analysis. The inhibition data will be discussed in Chapter Four. The activity assay
for HAL is virtually identical to that for PAL as the product urocanic acid absorbs at 277 nm and its production is easily monitored.

![Chromatogram of HAL purified from anion exchange chromatography and SDS PAGE gel of HAL during various stages of purification. Lanes 1 and 2, 100 and 20 µg protein from cell lysate; lanes 3 and 4, 20 and 4 µg from heat treatment; lanes 5 and 6, 20 and 4 µg from Mono Q anion exchange chromatography](image)

Figure 3.7 Chromatogram of HAL purified from anion exchange chromatography and SDS PAGE gel of HAL during various stages of purification. Lanes 1 and 2, 100 and 20 µg protein from cell lysate; lanes 3 and 4, 20 and 4 µg from heat treatment; lanes 5 and 6, 20 and 4 µg from Mono Q anion exchange chromatography

An unusual aspect of this part of the project was the huge scale involved. Typically, enzymes are isolated in milligram quantities from grams of bacteria and liters of media for the determination of kinetic constants and inhibition studies. We estimated that between three and five grams of enzyme (53 kDa) would be required to generate enough chromophoric octapeptide (~750 Da) for the NMR spectral studies, based on small scale experiments involving bacterial cultures of one or two liters. In order to accommodate our needs, we frequently employed the services of a 150 L fermentor in the Department of Microbiology and Immunology. We often obtained quantities of frozen bacterial paste in excess of one kilogram!* This presented challenges in tasks as mundane as resuspending the cells, as many hours of tedious pipetting were required. Additionally, French pressing a kilogram of bacteria takes an entire day. Fortunately, the

---

* On one occasion an accident occurred in which over 100 liters of *E. coli* expressing HAL was spilled over the floor of the Pilot Plant. Fortunately I was on vacation at the time and didn’t have to mop it up.
subsequent heat step greatly simplified the purification process, and in the end over three grams of purified HAL were obtained.

Globular proteins like enzymes are able to maintain their shape by such inter-residue forces as hydrogen bonding, electrostatic and hydrophobic interactions and cysteine bridges. A properly folded protein tends to maintain a hydrophobic core and displays the majority of its charged groups on the exterior, facing the aqueous environment. It is able to withstand the stress of minor temperature fluctuations. However, as the characteristic unfolding or melting temperature of the protein is approached, the thermal motions of the protein and surrounding solvent can disrupt this folded structure and the protein becomes denatured and loses function. HAL fortunately has a relatively high melting temperature\(^{62}\) of 83 °C (figure 3.8) and maintains its shape and solubility while others denature and coagulate. We expected that the tertiary and quaternary structure of HAL would support this result, and this will be examined further in the discussion of the X-ray crystal structure of HAL in section 3.4.

*Figure 3.8 Changes in the CD spectrum of HAL at 222 nm as it melts and denatures at 83 °C (from reference 62)*
Our proposal for a GFP-type backbone cyclization required properly formed HAL to have a mass that was smaller than the sequence prediction by two molecules of water. To check this we obtained a mass spectrum of MonoQ purified HAL and found a major peak that was significantly larger than our predictions. We found that we could justify this mass value by assuming both a double dehydration and the attachment of three molecules of β-mercaptoethanol through disulfide linkages (figure 3.9). Our buffer contained β-mercaptoethanol and it was possible that these could have become covalently bound to cysteine residues in HAL. Unfortunately, in attempts with dithiothreitol and other thiols we were not able to get any useful corroborating data. It is unclear why β-mercaptoethanol, intended to reduce cystine bridges, would end up covalently attached to the enzyme. Perhaps three HAL cysteine residues have unusually high oxidation potentials due to their chemical environment. This failure underscored our need to get at the electrophile in as direct a manner as possible, namely by excising it from the bulk of the enzyme.

![Mass spectrum of HAL monomer showing a mass 198 Da larger than predicted from sequence](image)

**Figure 3.9 ESIMS of HAL monomer showing a mass 198 Da larger than predicted from sequence**
3.2.2 Inactivation of HAL with Cysteine and Octapeptide Isolation

As mentioned earlier, the requirement for milligram quantities of the chromophoric octapeptide necessitates purifying gram quantities of HAL, allowing for latter losses during chromatography. The largest amount of DE 52 purified enzyme that was ever inactivated at one time was 3.5 grams. The experiment was initiated by the addition of a volume of 100 mM L-cysteine at pH 10.5 to an equal volume of HAL at the same pH. The extent of enzyme inactivation was monitored by removing aliquots and assaying for remaining activity, while the increase in absorbance at 340 nm was monitored by temporarily removing 1 mL and measuring its absorbance. The entire inactivation was completed in less than two hours (figure 3.10) and chromophore formation was seen to lag behind the inactivation. The inactivated, chromophoric HAL was then freed of excess cysteine by exchange into the tryptic digest buffer at pH 7.8. We opted to skip the carboxymethylation step that Hernandez et al. employed on the premise that the

![Graph](image_url)

*Figure 3.10 Time course of increase at 340 nm and of inactivation of HAL in the presence of L-cysteine*
peptides we were interested in contained no cysteine residues and therefore didn’t need to be capped.*

Trypsin (EC 3.4.21.4) from porcine pancreas was used in the first round of proteolysis of the inactivated enzyme. This protease cleaves peptide bonds on the carbonyl side of lysine and arginine residues. Therefore, according to the number and location of these residues in the predicted amino acid sequence of HAL, we expected to generate approximately 30 peptides, ranging in length from two to fifty-six residues (figure 3.11, lower right). We anticipated only one species absorbing at 340 nm based on the work of Hernandez et al. It soon became clear that we were generating two chromophoric peptides that were separated by several minutes in our analytical chromatogram of the tryptic digest (figure 3.11, lower left). After more than 48 hours of digestion the relative and absolute sizes of these peaks remained unchanged, indicating that they were minimal proteolytic fragments. We therefore chromatographed the entire digest on a preparative reverse phase column by multiple injections. Two peaks A and B were collected and concentrated by centrifuge filtration. These two fractions contained many non-chromophoric peptides, as re-chromatographing a sample of each and detecting at 220 nm revealed (not shown). N-terminal sequencing would have been ambiguous at this stage, so we simply assumed that we had isolated two compositional or structural variants of the 24 amino acid residue GSVGASGDLAPLATMSLVLLGEGK that Hernandez et al. had isolated. The bold symbols denote absolutely conserved residues in all known HAL and PAL, and include serine 143. We hoped that the next round of proteolysis would allow us to resolve the chromophoric octapeptides from non-chromophoric ones.

* This omission would prove to be a fortuitous decision in our efforts to determine the structure of the HAL chromophore. See section 3.2.3.2 for details.
Two major chromophoric peptides A and B.

Figure 3.11 HPLC profiles of tryptic digest of inactivated HAL showing appearance of two chromophoric peptides A and B. Chromatograms on the left taken with 340 nm detection and that on the right with 220 nm detection.
Figure 3.12 HPLC profiles of V8 protease digest of peptide B showing appearance of two chromophoric peptides B1 and B2 with 340 nm detection

The second protease employed was V8 protease (EC 3.4.21.19) from *Staphylococcus aureus*, which cleaves peptides on the carboxyl side of glutamate and aspartate residues. Peptides A and B were processed separately and each generated two daughter fragments that absorbed at 340 nm (data from B shown in figure 3.12). The masses of A1, A2, B1 and B2 were measured by 63
electrospray ionization mass spectrometry (ESIMS). A2 and B1 were concluded to be the same peptide, due to identical mass, retention time and UV absorbance maxima. Peptides A1 and B2 had different masses and retention times but identical UV absorbance maxima. Note that the true absorbance maxima for the Hernandez peptide, peptide A1 and peptide B2 are found at 333 nm, and not 340 nm. The latter wavelength was chosen by Hernandez et al. and us to follow the cysteine inactivations simply to facilitate comparisons with previous work. A framework model began to emerge from these findings.

<table>
<thead>
<tr>
<th>Name and Retention Time (min)</th>
<th>Structure</th>
<th>$\lambda_{\text{max}}$ (nm)</th>
<th>Low Resolution Mass (Da)</th>
<th>Difference (Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synthetic octapeptide</td>
<td>GSVGASGD</td>
<td>220</td>
<td>648.6</td>
<td>-</td>
</tr>
<tr>
<td>MIO octapeptide</td>
<td>GSVGASGD</td>
<td>315</td>
<td>612.6</td>
<td>-36</td>
</tr>
<tr>
<td>(in HAL)</td>
<td></td>
<td></td>
<td>(calc’d)</td>
<td></td>
</tr>
<tr>
<td>Hernandez octapeptide</td>
<td>GSVGASGD</td>
<td>333</td>
<td>789.4</td>
<td>+141</td>
</tr>
<tr>
<td>A1 (11)</td>
<td>GSVGASGD</td>
<td>333</td>
<td>850.5</td>
<td>+201.9</td>
</tr>
<tr>
<td>B2 (15.5)</td>
<td>GSVGASGD</td>
<td>333</td>
<td>807.5</td>
<td>+158.9</td>
</tr>
<tr>
<td>A2/B1 (13)</td>
<td>GSVGASGD</td>
<td>326</td>
<td>610.5</td>
<td>-38</td>
</tr>
</tbody>
</table>

*Table 3.3 Summary of octapeptides with basic model structures, absorbance maxima and neutral masses. The underlined residues are assumed to exist as a five-membered heterocycle. The symbols Q, X, Y and Z have been added to denote extra attachments to the MIO structure.*

Hernandez et al. purified a chromophoric octapeptide with a mass of 789.4 Da, which is 141 Da larger than the 648.6 Da predicted from the sequence. They further established that the extra mass was associated with the suspected dehydroalanine precursor serine 143. We on the other hand have isolated *three* chromophoric peptides of masses 610.5 Da (A2/B1), 807.5 Da (B2),
and 850.5 Da (A1) (table 3.3). If we assume that our GFP model and the published MIO structure for the electrophile are correct then the MIO octapeptide should have a mass of 648.6 Da less two molecules of water, or 612.6 Da. A2/B1 is only 2 Da smaller and is therefore likely to be quite similar in structure to the bare MIO. The difference in $\lambda_{\text{max}}$ between the MIO octapeptide, 315 nm, and A2/B1, 326 nm, must be reconciled with this 2 Da mass difference. Additionally, the fact that A2/B1 is smaller in mass than both A1 and B2 suggests that it is a common decomposition product. The mass difference between A1 and B2 is 43 Da, which may be explained by a decarboxylation. However, all these data fragments are insufficient to formulate a sound model for the chromophore.

At this juncture it was clear that we had isolated some interesting jigsaw pieces, and it was simply a matter of performing the right experiments to complete the picture. Peptide A2/B1 readily decomposed to smaller fragments that were difficult to interpret, therefore the following sections focus on the more stable peptides A1 and B2.

### 3.2.3 Characterization of the Octapeptides

#### 3.2.3.1 N-Terminal Analysis

Peptides A1 and B2 were submitted for N-terminal amino acid analyses at the Nucleic Acid and Protein Service (NAPS) unit at UBC. The procedure involves immobilizing the peptide on a membrane and using phenylisothiocyanate (PITC) and trifluoroacetic acid (TFA) to sequentially cleave derivatized amino acids at the amino terminus. To successfully generate the phenylthiocarbamyl (PTC) derivative it requires a free terminal $\alpha$-amino group to react with PITC and an alkaline environment (figure 3.13). Treatment with TFA liberates a cyclic derivative of the N-terminal residue, which rearranges in HCl to the phenylthiohydantoin (PTH) derivative. It is this amino acid-PTH species that is chromatographically identified by comparing its retention time to a
standard series. Since the octapeptide still retains an unmodified N-terminus, it is possible to repeat
the process for several residues. As in previous reports, the first four residues of the octapeptides
were confidently assigned as G, S, V and G with the fifth residue curiously blocked. This finding
is consistent with our model of a cyclic iminodehydroalanine in which the fifth residue, alanine, is
participating in a five-membered ring and is not expected to yield a PTH derivative.

\[
\text{Ph-}N\equiv C\equiv S \quad \xrightarrow{\text{PTC}} \quad \text{H}_2\text{N}\quad \text{Ph}-\text{N}-\text{C-}\text{O} \quad \xrightarrow{\text{PTC-amino acid derivative}} \quad \text{Ph}\quad \text{H}-\text{N}-\text{C-}\text{O} \quad \xrightarrow{\text{TFA}} \quad \text{PTC-amino acid derivative}
\]

\[
\text{MIO unable to participate in peptide cleavage reaction}
\]

Figure 3.13 Process of Edman degradation (top figure), rationalizing blockage at alanine due to
the presence of the MIO in the chromophoric octapeptides (bottom figure)
3.2.3.2 High Resolution Mass Spectrometry

In order to narrow down the number of possible compositional isomers for the chromophores, peptides A1 and B2 were submitted for high-resolution mass analysis (HRMS). This technique is sensitive enough to give masses to the ten-thousandth of a dalton. In order to properly interpret the data it is essential to use the isotopic (and not atomic) masses of atoms that correspond to the most abundant isotope for each atom in a given formula. Table 3.4 contains the calculated masses for the synthetic and MIO octapeptides and the observed masses for the peptides A1 and B2. We assume that the extra masses in A1 and B2 are attached through serine 143, as Hernandez et al. showed.

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence and Formula</th>
<th>Neutral Mass (Observed)</th>
<th>Mass Difference</th>
<th>Formula Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synthetic octapeptide</td>
<td>GSVGASGD</td>
<td>648.2715</td>
<td>+ 36.0211</td>
<td>H4O2</td>
</tr>
<tr>
<td></td>
<td>C24H40O13N8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MIO octapeptide</td>
<td>GSVGASGD</td>
<td>612.2504</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C24H36O11N8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1</td>
<td>GSVGASGD X</td>
<td>(850.2597)</td>
<td>+ 238.0093</td>
<td>see table 3.5</td>
</tr>
<tr>
<td></td>
<td>GSVGASGD Y</td>
<td>(807.2543)</td>
<td>+ 195.0039</td>
<td>see table 3.5</td>
</tr>
</tbody>
</table>

Table 3.5 High resolution neutral masses of octapeptides and differences from the MIO octapeptide. The underlined residues are assumed to exist as a five-membered heterocycle. The symbols X and Y have been added to denote extra attachments to the MIO structure.

There are hundreds of constitutional isomers that could give the masses shown for A1 and B2. In order to simplify the deductive process we initially assumed that A1 and B2 contained all of the atoms present in the MIO octapeptide GSVGASGD, C24H36O11N8. When the mass of this species was subtracted from the observed masses we had much smaller values to fit with empirical
formulae. The problem of building an isotopic model for a chromophore with a mass of 238.0093 Da is much simpler than one of 850.2597 Da.

<table>
<thead>
<tr>
<th>Observed Mass Difference</th>
<th>Number</th>
<th>Formula</th>
<th>Deviation (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1 238.0093</td>
<td>1</td>
<td>C6H2N6O5</td>
<td>-2.7</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>C6H10N2O4S2</td>
<td>-4.6</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>C6H4N5O6</td>
<td>+3.0</td>
</tr>
<tr>
<td>B2 195.0039</td>
<td>4</td>
<td>C5H1N5O4</td>
<td>-5.4</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>C5H9NO3S2</td>
<td>-7.8</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>C7H3N2O5</td>
<td>+1.5</td>
</tr>
</tbody>
</table>

Table 3.6 Summary of empirical formulae that match the observed mass differences between A1 and B2 and the MIO octapeptide. Search constraints ±10 ppm, #C 0-16, #H 0-16, #N 0-10, #O 0-10, #S 0-10

The three most reasonable formulae with masses close to 238.0093 Da are shown in table 3.6. Formula #2 contains two sulfur atoms and raised the possibility of cystine attachment. Cystine is simply two molecules of cysteine joined through a disulfide linkage. It seemed reasonable to us that inactivation with high concentrations of cysteine might lead to the subsequent formation of cystine, even though cystine itself is known not to cause permanent inactivation of HAL (table 3.2). The atomic formula of cystine is C6H12O4N2S2, which contains two more protons than formula #2. However, this species must somehow be covalently attached to the electrophile so this difference is justified. When fully assembled, the cystine model was found to be an almost exact match to the observed mass (figure 3.14). The conjugation of the cysteine amine to the MIO chromophore also helps explain why the UV absorbance of this peptide has shifted from 315 nm to 333 nm.
In examining the B2 half of table 3.5 it is apparent that #4 to 6 differ from #1 to 3 respectively by one atom each of C, H, N and O. If the cysteine model for A1 is correct then it is likely that the extra mass on B2 also contains two atoms of sulfur and therefore formula #5 is correct. An elegant solution to this puzzle is provided by replacing the pendant cysteine of cystine in A1 with a molecule of β-mercaptoethanol. The calculated mass from this model is an almost perfect match for the observed mass of B2 (figure 3.15). This explanation is reasonable because disulfide exchange could have easily occurred during proteolysis as the buffers contained β-mercaptoethanol. The replacement of an α-amino acid group by a primary alcohol also explains the larger retention time of B2 in the reverse-phase column, as it is more hydrophobic than A1.
One of the more interesting exercises performed after developing our model has been the re-interpretation of others’ results in this area. For instance, as stated in section 3.1, K. Weber and J. Retey\textsuperscript{57} proposed a structure for their chromophoric octapeptide that involved an arginine residue attached to an oxidized dehydroalanine moiety (figure 3.5, page 53). The mass of 789.4 Da that they (and Hernandez et al.) observe is easily and simply explained in the context of our model. Both groups employed reductive carboxymethylation with sodium borohydride and iodoacetate prior to proteolytic digestion that would likely give rise to the carboxymethylated cysteine product (figure 3.16).

Even though there is not as much physical data available for the A2/B1 peptide as there is for A1 and B2, it is possible to rationalize its mass and UV absorbance properties. From table 3.3 it is clear that A2/B1 is 2 Da smaller than the MIO octapeptide and it is a likely decomposition product of A1 and B2. A species of the same neutral mass of 610.8 Da is also observed in the fragmentation spectrum of the Hernandez octapeptide (figure 3.3). An internal trans-imination by the N-terminal glycine resulting in the loss of either cystine in A1, cysteine/\(\beta\)-mercaptoethanol in B2 or carboxymethyl cysteine in the Hernandez octapeptide can explain these observations (figure 3.17). The difference in UV maxima (333 to 326 nm) and peak shape is possibly due to a restriction on the number of available product isomers resulting from the tethered glycyl residue
The serine or aspartate residues in the octapeptide could also conceivably displace the cysteine, but it is unlikely that an oxygen linkage would produce a chromophore so similar to that of an amine-linked cysteine. A discussion about the existence of three isomeric forms of the chromophore in A1 and B2 will be elaborated in section 3.2.3.4.

Figure 3.17 Proposal for the formation of the A2/B1 decomposition product with a mass of 610 Da
A2

One isomer with single $\lambda_{\text{max}}$

326 nm

B2

Three isomers with slightly differing $\lambda_{\text{max}}$

333 nm

Figure 3.18 Comparison of UV absorbance maxima of A2 and B2. The slight asymmetry in the UV absorbance peak for B2 (and A1) could be due to the presence of three isomers of the chromophore, one of which has the same $\lambda_{\text{max}}$ as A2 (and B1), 326 nm

3.2.3.3 Amino Acid Analysis

In order to test for the presence of cysteine, both peptides were subjected to total amino acid analysis, in which they were hydrolyzed in concentrated acid, derivatized with PITC and chromatographed. Peaks for glycine, serine, valine, alanine, aspartate and cysteine were observed in the chromatograms of both peptides. Minor peaks that co-eluted with derivatized glutamate and isoleucine as well as a major unassigned peak were also observed (figure 3.19). It is difficult to quantitate the relative amounts of each amino acid liberated by hydrolysis, but glycine is clearly the most abundant constituent judging from peak height, consistent with three glycine residues in each octapeptide. The large, unassigned peak at twelve minutes or the apparent glutamate and isoleucine peaks could be decomposition products of the MIO function, such as an aldehydic serine residue. The unambiguous presence of cysteine in both A1 and B2 strengthened our model, but we desired a test that would both prove the presence of the disulfide bond and show a pendant cysteine for A1 but not B2.
Figure 3.19 Amino acid analysis of A1 and B2, indicating cysteine and all expected residues, plus peaks that correspond to glutamate and isoleucine and a large unidentified peak. Norleucine was added as an external standard.
Small amounts of A1 and B2 were treated with DTT in order to reduce the suspect disulfide bonds. The progress of the reaction was followed using HPLC and examining the shift in retention time of the peptides. In both cases the major peak shifted to an earlier time as the cleavage proceeded. After 24 hours of treatment each sample was analyzed for the presence of free cysteine by reaction with PITC followed by chromatography (figure 3.20). In the case of A1 but not B2 cysteine had clearly been liberated by disulfide reduction with DTT. Confident that the foundation of our model was solid, we turned our attention to confirming the connectivity of atoms through NMR spectroscopy.

Figure 3.20 Test for cysteine release from A1 and B2 via disulfide reduction. Norleucine was added as external standard
3.2.3.4 One and Two-Dimensional NMR Spectroscopy

Prior to running any NMR experiments on the peptides A1 and B2, a synthetic octapeptide GSVGASGD was prepared as a control to establish common chemical shifts and coupling constants. The standard spectra that were obtained enabled rapid interpretation of new peaks from cysteine in A1 and B2 and changes in chemical shifts due to the unsaturation of the MIO ring. In the following discussion the focus will be on chromophoric peptide B2, as it was isolated in greater quantity than A1, produced cleaner spectra and was generally simpler to interpret. To simplify the discussion of the signals arising from various $^1$H and $^{13}$C nuclei in B2, key atoms in the peptide have been labeled (figure 3.21) and are accompanied by table 3.6 to summarize all correlations observed in all the one and two-dimensional NMR experiments.
Figure 3.21 Structure of octapeptide B2 indicating amino acids and labeling conventions
<table>
<thead>
<tr>
<th>Residue</th>
<th>Atom#</th>
<th>δ&lt;sup&gt;13&lt;/sup&gt;C (ppm)</th>
<th>δ&lt;sup&gt;1&lt;/sup&gt;H (ppm)</th>
<th>COSY</th>
<th>HMBC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gly1</td>
<td>G1</td>
<td>NH&lt;sub&gt;2&lt;/sub&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>G2</td>
<td>CH&lt;sub&gt;2&lt;/sub&gt;</td>
<td>44.5</td>
<td>3.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>G3</td>
<td>CO</td>
<td>173</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>S1</td>
<td>NH</td>
<td></td>
<td>8.17</td>
<td>S2</td>
</tr>
<tr>
<td></td>
<td>S2</td>
<td>αCH</td>
<td>54.5</td>
<td>4.4</td>
<td>S1, S3a, S3b</td>
</tr>
<tr>
<td></td>
<td>S3a</td>
<td>βCH&lt;sub&gt;2&lt;/sub&gt;</td>
<td>41.8</td>
<td>3.51</td>
<td>S2, S3b</td>
</tr>
<tr>
<td></td>
<td>S3b</td>
<td></td>
<td>3.61</td>
<td></td>
<td>S2, S3a</td>
</tr>
<tr>
<td></td>
<td>S4</td>
<td>CO</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val</td>
<td>V1</td>
<td>NH</td>
<td></td>
<td>7.95</td>
<td>V2</td>
</tr>
<tr>
<td></td>
<td>V2</td>
<td>αCH</td>
<td>58</td>
<td>4.13</td>
<td>V1, V3</td>
</tr>
<tr>
<td></td>
<td>V3</td>
<td>βCH</td>
<td>30</td>
<td>2.05</td>
<td>V2, V4, V5</td>
</tr>
<tr>
<td></td>
<td>V4</td>
<td>γCH&lt;sub&gt;3&lt;/sub&gt;</td>
<td>18</td>
<td>0.85</td>
<td>V3</td>
</tr>
<tr>
<td></td>
<td>V5</td>
<td>γCH&lt;sub&gt;3&lt;/sub&gt;</td>
<td>18</td>
<td>0.85</td>
<td>V3, V4</td>
</tr>
<tr>
<td></td>
<td>V6</td>
<td>CO</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gly2</td>
<td>G'1</td>
<td>NH</td>
<td>8.35</td>
<td></td>
<td>G'2a, G'2b</td>
</tr>
<tr>
<td></td>
<td>G'2a</td>
<td>CH&lt;sub&gt;2&lt;/sub&gt;</td>
<td>62</td>
<td>3.55</td>
<td>G'1, G'2b</td>
</tr>
<tr>
<td></td>
<td>G'2b</td>
<td></td>
<td>3.60</td>
<td></td>
<td>G'1, G'2a</td>
</tr>
<tr>
<td></td>
<td>G'3</td>
<td>CO</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>A1</td>
<td>NH</td>
<td>8.25</td>
<td></td>
<td>A2</td>
</tr>
<tr>
<td></td>
<td>A2</td>
<td>αCH</td>
<td>41.5</td>
<td>4.75</td>
<td>A1, A3</td>
</tr>
<tr>
<td></td>
<td>A3</td>
<td>βCH&lt;sub&gt;3&lt;/sub&gt;</td>
<td>18.3</td>
<td>1.32</td>
<td>A3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A2, M2</td>
</tr>
<tr>
<td>MIO</td>
<td>M1</td>
<td>N1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M2</td>
<td>C2</td>
<td>151</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M3</td>
<td>N3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M4</td>
<td>CO</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M5</td>
<td>C5</td>
<td>114</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M6a</td>
<td>CH</td>
<td>7.1</td>
<td></td>
<td>C1a</td>
</tr>
<tr>
<td></td>
<td>M6b</td>
<td></td>
<td>139.5</td>
<td>7.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>M6c</td>
<td></td>
<td>145.5</td>
<td>7.4</td>
<td>C1b</td>
</tr>
<tr>
<td>Cys</td>
<td>Cl&lt;sub&gt;a&lt;/sub&gt;</td>
<td>NH</td>
<td>7.9</td>
<td></td>
<td>M6a</td>
</tr>
<tr>
<td></td>
<td>Cl&lt;sub&gt;b&lt;/sub&gt;</td>
<td></td>
<td>8.5</td>
<td></td>
<td>M6b</td>
</tr>
<tr>
<td></td>
<td>C2</td>
<td>αCH</td>
<td>59.7</td>
<td>3.85</td>
<td>C3a, C3b</td>
</tr>
<tr>
<td></td>
<td>C3a</td>
<td>βCH&lt;sub&gt;2&lt;/sub&gt;</td>
<td>44.5</td>
<td>3.0</td>
<td>C2, C3a</td>
</tr>
<tr>
<td></td>
<td>C3b</td>
<td></td>
<td>3.3</td>
<td></td>
<td>C2, C3b</td>
</tr>
<tr>
<td>βME</td>
<td>ME1</td>
<td>SCH&lt;sub&gt;2&lt;/sub&gt;</td>
<td>41.2</td>
<td>2.75</td>
<td>ME2</td>
</tr>
<tr>
<td></td>
<td>ME2</td>
<td>OCH&lt;sub&gt;2&lt;/sub&gt;</td>
<td>59.5</td>
<td>3.6</td>
<td>ME1</td>
</tr>
<tr>
<td>Gly3</td>
<td>G''1a</td>
<td>CH&lt;sub&gt;2&lt;/sub&gt;</td>
<td>42</td>
<td>4.20</td>
<td>G''1b</td>
</tr>
<tr>
<td></td>
<td>G''1b</td>
<td></td>
<td>4.27</td>
<td></td>
<td>G''1a</td>
</tr>
<tr>
<td></td>
<td>G''2</td>
<td>CO</td>
<td>167</td>
<td></td>
<td>G''1a, G''1b</td>
</tr>
<tr>
<td>Asp</td>
<td>D1</td>
<td>NH</td>
<td>8.05</td>
<td></td>
<td>D2</td>
</tr>
<tr>
<td></td>
<td>D2</td>
<td>αCH</td>
<td>49</td>
<td>4.1</td>
<td>D1, D3a, D3b</td>
</tr>
<tr>
<td></td>
<td>D3a</td>
<td>βCH&lt;sub&gt;2&lt;/sub&gt;</td>
<td>40.2</td>
<td>2.2</td>
<td>D2, D3b</td>
</tr>
<tr>
<td></td>
<td>D3b</td>
<td></td>
<td>2.4</td>
<td></td>
<td>D2, D3a</td>
</tr>
<tr>
<td></td>
<td>D4</td>
<td>αCO&lt;sub&gt;2&lt;/sub&gt;</td>
<td>173</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>D5</td>
<td>βCO&lt;sub&gt;2&lt;/sub&gt;</td>
<td>173.5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3.7 Summary of data from all NMR experiments on octapeptide B2 in DMSO
3.2.3.4.1 One-Dimensional Proton NMR

A one-dimensional $^1$H NMR spectrum of B2 in DMSO-$d_6$ was obtained and was found to differ from the synthetic standard in several areas (figure 3.22). The peaks for B2 are not as sharp and in several regions exist as broad multiplets, suggesting that more than one structural isomer may be present with slightly perturbed chemical shifts. The alanine $\alpha^1$H at 4.75 ppm (A2) is shifted downfield because of its proximity to the MIO ring, as are the protons from glycine 3 (G"1a and G"1b). New peaks are observed for both cysteine and $\beta$-mercaptoethanol and several resonances are obscured by the prominent water peak. Perhaps most informative are the three resonances between 7.1 and 7.4 ppm which are assigned to the vinylic proton on the exocyclic methylidene carbon. The variation in chemical shifts and the coupling behavior of these resonances suggest three isomers for the chromophore. In DMSO a signal at 7.1 ppm appears as a doublet, at 7.2 ppm as a sharp singlet and at 7.4 ppm as a broad peak (figure 3.23). When the solvent is changed to D$_2$O the doublet at 7.1 ppm collapses to a singlet and the broad peak at 7.4 ppm sharpens dramatically. This effect could result from the rapid exchange of the $\alpha$-amine proton of cysteine, which is possible in D$_2$O but not DMSO. The singlet at 7.2 ppm has no coupling partner and is unaffected by the solvent change.

It is not difficult to imagine that these three isomers would also show small differences in their UV spectra, as alluded to in figure 3.17. The gross features of each isomer involve one ring, three double bonds and several heteroatoms. It is beyond the scope of this work, however, to enter into a detailed discussion of the factors governing the positions of the UV maxima. The following sections on heteronuclear correlation spectroscopy also support a three-isomer model.
Figure 3.22 $^1$H spectrum of B2 in DMSO-$d_6$ indicating known resonances including attached cysteine and $\beta$-mercaptoethanol. The peak at 5.5 is an impurity of undetermined origin.
Figure 3.23 Vinylic resonances in octapeptide B2 in DMSO-\textit{d}_6 as rationalized by the existence of three isomers of the chromophore

The assignment of the absorbances between 7.1 and 7.4 ppm is in part based on other compounds with similar structural features.\textsuperscript{63} Figure 3.24 shows two compounds with almost identical vinylogous amide functional groups whose vinyl protons have been observed at 6.9 and 7.35 ppm in deuterated chloroform. No vinylic coupling or isomeric forms are observed due to the lack of a neighboring amine proton and (presumably) steric hindrance between the amine and MIO carbonyl groups. Further examination of the chemical shift assignments and role that the level of amine substitution plays on the distribution of isomers is left to section 3.2.3.5.
3.2.3.4.2 The COSY Experiment

An extremely useful experiment for graphically illustrating protons that are coupled to each other is known as COSY, which stands for Correlation Spectroscopy. The results of a COSY experiment are presented in two dimensions such that the off-diagonal signals, or cross peaks, correspond to coupled spin systems (i.e. protons) whose members can be inferred by examining the coordinates of the cross peak. A detailed analysis of the COSY spectrum allows us to decipher all the protons that a given proton is coupled to, usually through geminal or vicinal relationships.

The COSY spectrum of B2 in DMSO-d$_6$ shows cross peaks between two of the three vinylic resonances and two signals in the amide N-H region of 8-9 ppm (figure 3.25) consistent with a multi-isomer model. The majority of the remaining off-diagonal peaks are readily identified with neighboring protons and are illustrated in figure 3.26 by arrows. The full COSY spectrum is shown in figure 3.27 and expansions of important regions are shown in figures 3.28 and 3.29.
Figure 3.25 Expansion of COSY spectrum of B2 showing coupling between vinylic M6a, M6c and amide C1a, C1b protons. Note the much stronger cross peak for the doublet at 7.1 ppm.
Figure 3.26 Proton-proton couplings from a COSY experiment on B2 in DMSO-\textit{d}_6
Figure 3.27 COSY spectrum of B2 in DMSO-$d_6$. Boxed regions are shown expanded in figures 3.25, 3.28 and 3.29.
Figure 3.28 Expansion of COSY spectrum of B2 in DMSO-$d_6$
Figure 3.29 Expansion of COSY spectrum of B2 in DMSO-d6 showing coupling between $\alpha^1H$ and amide $^1H$
3.2.3.4.3 The HMQC Experiment

The most important proposed structural feature of peptide B2 is the MIO ring. Unfortunately it contains only one proton and it is impossible to prove the proposed structure from that resonance alone. Fortunately we can employ heteronuclear correlation experiments in two dimensions to determine the chemical shifts of carbon atoms that are associated with protons through one, two or three bonds. The \(^1\)H detected Heteronuclear Multiple Quantum Coherence (HMQC) experiment is used to determine geminal \(^1\)H/\(^{13}\)C connections via correlations between the \(^1\)H resonances on the x, or \(F_2\) axis and \(^{13}\)C on the y, or \(F_1\) axis. The major benefit of this experiment over other heteronuclear correlations is that it is much more sensitive since the more sensitive \(^1\)H nucleus is detected rather than the insensitive \(^{13}\)C nucleus. Unlike the \(^1\)H/\(^1\)H COSY, the HMQC data presented here does not contain a spectrum on the \(F_1\) axis; insufficient amounts of peptide B2 were available to permit acquisition and display of a \(^{13}\)C spectrum. The chemical shifts of all carbon nuclei with attached protons are readily observed from the cross-peaks however.

The data from the HMQC experiment confirmed a great deal of the assignments made from the COSY spectra, and caused some slight revision. The coupling between the vinylic proton of the major isomer and the geminal vinylic carbon M6b is readily observed (figure 3.30). A chemical shift value of 140 ppm for an exocyclic methylidene carbon is very reasonable. A smaller correlation is observed for one of the minor isomers M6c at 7.4 and 146 ppm. Correlations for the third isomer were not detected presumably because the concentration of this species was too low. Every other proton-bearing carbon can be assigned in this spectrum and are summarized in figures 3.31 and 3.32.
Figure 3.30 Expansion of HMQC spectrum of B2 indicating geminal $^1H/^13C$ coupling between the M6b $^1H$ and $^{13}C$, and between the M6c $^1H$ and $^{13}C$. 
Figure 3.31 Proton-carbon couplings from a HMQC experiment on B2 in DMSO-d$_6$. Values in ppm with protons first and carbons second.
Figure 3.32 HMQC spectrum of B2 indicating geminal $^1H/^13C$ couplings
3.2.3.4.4 The HMBC Experiment

The Heteronuclear Multiple Bond multiple quantum Coherence experiment (HMBC) determines long-range (two or three bond) $^1\text{H}-^{13}\text{C}$ connectivities and is very appropriate for our needs, namely to probe the carbon atoms of the MIO ring which don’t have any geminal protons. Unfortunately, significantly less data can be extracted from our HMBC experiment compared to the COSY and HMQC. The M6b proton from the major isomer of B2 is observed coupling to the M5 carbon of the MIO ring through two bonds (figure 3.33). More interesting are the two and three bond couplings of the A2 and A3 protons of alanine to the M2 carbon of the MIO ring. The alanine A3 protons are observed to be coupling to three closely grouped yet distinct carbon nuclei (figure 3.34). The electronic environment of these carbon atoms could be sufficiently distinct in the three isomers to allow their resolution in the correlation spectrum. The A3 protons are also observed as three closely grouped doublets, further supporting the three isomer model. The scant remainder of the heteronuclear shift correlations are summarized in figures 3.35 and 3.36.
Figure 3.33 Expansion of the HMBC spectrum of B2 showing long range coupling between the imine M6b $^1$H and the MIO M5 $^{13}$C.
Figure 3.34 Expansion of the HMBC spectrum of B2 showing three distinct cross-peaks for the alanine A3 $^1$H and the M2 $^{13}$C of the MIO ring.
Figure 3.35 Two and three-bond proton-carbon couplings from an HMBC experiment on B2 in DMSO-d_6. Values in ppm, protons first, carbons second.
Figure 3.36 HMBC spectrum of B2 showing several distant $^1H$/$^13C$ couplings. Boxed region is shown expanded in figure 3.34.
3.2.3.5 Further Evidence for Chromophore Structures

A structure-based search for compounds similar to our model for the MIO chromophore was performed for us by Prof. R. Andersen and yielded several papers containing corroborative information, two of which were mentioned in section 3.2.3.4.1. An additional compound was found that contains the essential features of the MIO ring with an amine attached through the methylidene carbon. The absorbance maximum of this compound occurs at 336 nm in ethanol, very close to the 333 nm that we observe for our peptides in water (figure 3.37).

\[
\text{UV } \lambda_{\text{max}} = 336 \text{ nm (ethanol)}
\]

*Figure 3.37 Structure of compound similar to the MIO chromophore in HAL*

The most useful information relating to our three-isomer model comes from a study on the synthesis of the sleep-inducing benzodiazepine Loprazolam. In 1977 J. B. Taylor et al. published a report on the synthesis of a novel class of annelated 1,4-benzodiazepines, one of which is now known as the potent hypnotic agent Loprazolam (figure 3.38). The authors discovered that several synthetic intermediates exist in isomeric forms with the amino group on the exocyclic methylidene double bond lying \textit{cis} or \textit{trans} to the carbonyl group (\textit{E} and \textit{Z} in I and II). They observed two \textsuperscript{1}H NMR signals for the vinylic proton of a freshly prepared solution of a monomethyl amine. The ratio of these signals changed over several weeks until the originally minor component predominated, consistent with production of an isomer containing a strong hydrogen bond, presumably the \textit{E} form I. In contrast, the spectrum of a dimethyl amine doesn’t change over time and suggests that, in the absence of any possibly hydrogen bonding, this compound and related
tertiary amines exist entirely in the $Z$ form III. Thus, a competition between the forces of destabilizing steric hindrance and stabilizing hydrogen bonding likely determines the ratio of the isomers of the chromophoric peptides that we end up observing as well. No mention was made in Taylor’s report of the hydroxy-imine tautomer that seems to predominate in our peptides.

![Figure 3.38](image)

*Figure 3.38 Structure of Loprazolam and isomeric forms of precursor compounds showing an important H-bond in monoalkyl amines. Cis/trans label defined by amine/carbonyl geometry*

From our spectra it is difficult to determine with absolute certainty which of the two coupled vinylic resonances corresponds to which isomer in B2. Our original tentative assignment\(^{\text{66}}\) was based on vinylic chemical shifts from a study of similar compounds\(^{\text{67}}\) in which the nitrogen at M1 has been replaced by a carbon. The authors found that the vinylic proton from the $Z$ isomer had a significantly higher chemical shift in DMSO than an analogous cis isomer. We therefore assigned the larger peak at 7.4 ppm to the $Z$ isomer and the much smaller doublet at 7.1 ppm to the
E. However, the broad form of the peak at 7.4 ppm and its larger size relative to the peak at 7.1 ppm would suggest that it actually represents the vinylic proton from the H-bonded cis isomer (figure 3.39). The amine-carbonyl hydrogen bond makes that isomer thermodynamically more stable and would tend to partially disengage the coupling to the vinylic proton. This weaker coupling is clearly evident in the COSY spectrum shown in figure 3.25, where the Z isomer shows very strong coupling between the amine and vinylic protons. In the long run though, we are less concerned with the assignment of isomer distributions than we are with how the observed structure can give us insight into the HAL mechanism.

![Diagram of cis (E) and trans (Z) isomers of the B2 chromophore.](image)

*Figure 3.39 Structures of cis (E) and trans (Z) isomers of the B2 chromophore*

3.3 Significance of Chromophore Structure to the HAL Mechanism

The original impetus for pursuing the chromophoric peptides generated from cysteine inactivated HAL was to determine the structure and chemical reactivity of the electrophile. This information could then give us valuable clues as to the mechanism of the lyase reaction. This hypothesis was based in part on the knowledge that cysteine is a good histidine analogue with a
low $K_1$ value, and that there is an extra mass associated with the electrophile progenitor serine 143 from earlier peptide studies. To our minds it seemed likely that some portion of cysteine had become attached to the enzyme, even though earlier reports strongly suggested that the sulfur of cysteine was readily removed from the chromophore under denaturing conditions without affecting its UV absorbance properties. After careful consideration of all the data collected during the characterization of peptides A1 and B2 a very reasonable model for the structure of each has emerged. In terms of the two competing mechanisms for HAL, our chromophore would at first glance seem to support the Hanson and Havir model. This is because cysteine is attached to the exocyclic methylidene carbon of the MIO group via its $\alpha$-amine, and the Hanson and Havir mechanism involves an initial attack by the $\alpha$-amine of histidine. A potential inactivation mechanism is therefore illustrated in figure 3.40 that incorporates many features of the Hanson and Havir mechanism and begins with the nucleophilic attack of the substrate amine at the electrophile.

In this mechanism cysteine adds to the MIO as histidine is modeled in figure 1.17 on page 19, but the 1,3 hydrogen shift does not occur. Instead, the thiolate I undergoes an electron transfer to generate superoxide and the thiol radical II, which abstracts a proton to give the allylic radical III. Combination with superoxide leads to the hydroperoxide IV that would readily be eliminated to form imine V. A subsequent deprotonation results in the penultimate inactivation product that becomes oxidatively coupled to another molecule of cysteine. The thiol promoted oxidation could explain why cysteine inactivates HAL while serine doesn’t. Alternatively, the role of the thiol may be to bind an active site metal ion. The oxidation leading to IV may occur via spontaneous autoxidation analogous to that observed with the GFP chromophore.

This proposal appears to justify the requirement for high pH and oxygen; thiolates have been shown to react with oxygen to generate a sulfur radical and a radical oxygen anion. High pH would also promote the elimination of hydrogen peroxide. Its major shortcoming is that it can’t explain the finding of Hernandez et al. that both N-acetyl cysteine and 3-mercaptopropionate can
permanently inactivate HAL but not form a 333 nm chromophore (figure 3.3, page 51). The lack of a free amine in each of these species precludes their participation in the initial addition step.

Figure 3.40 Potential inactivation mechanism incorporating the Hanson and Havir model
A more satisfying inactivation scheme mimics the Retey mechanism with initial attack by the thiolate (figure 3.41). According to the first step of the Retey mechanism the imidazole of histidine nucleophilically adds to the MIO group in an enzymatic Friedel-Crafts-type process. If the reason that cysteine is such a good competitive inhibitor of HAL at normal pH values is because the thiol can mimic the ‘electron-rich’ imidazole and add to the MIO, then perhaps the same is true at the high pH required for permanent inactivation. The thiolate is much more nucleophilic and its attack leads to the enolate VI which could readily undergo autoxidation to the hydroperoxide VII. A spontaneous elimination enables the regeneration of the MIO group with cysteine attached through sulfur in VIII. A subsequent S-to-N rearrangement takes place through an intermediate thiazoline IX to give the vinylogous amide that becomes oxidized to the final product (boxed).

This mechanism satisfies all of the requirements of the inactivation experiment: high pH to generate a nucleophilic thiolate, the presence of oxygen for the autoxidation of VI and, perhaps most significantly, a free amine to generate the chromophoric species after formation of the permanently inactivated VIII. This sequence of inactivation first and chromophore formation second matches Klee’s findings in figure 3.2 on page 49 and completely explains the Hernandez cysteine analogue data in figure 3.3 on page 51. These findings strongly support the Retey model for inactivation of HAL with cysteine and consequently also the Retey model for HAL with histidine.

It seems likely that VIII (less the cysteine amine) would accumulate as permanently inactivated, non-chromophoric enzyme if 3-mercaptopropionate were substituted for cysteine, as indicated in figure 3.3. We are left to conclude that a covalently attached amine, and not sulfur, is required to affect both a bathochromic shift of the MIO absorbance maximum from 315 to 333 nm and lead to a large increase in the molar absorptivity. This cannot happen until the thiolate has attacked the MIO and an oxidation has occurred.
Figure 3.41 Potential inactivation mechanism incorporating the Retey model
Additional support for the 'thiolate attack' mechanism comes from an experiment designed to test the oxygen requirement for inactivation of HAL with 3-mercaptopropionate. We began the experiment by mixing a solution of HAL at pH 10.5 with a solution of 3-mercaptopropionate. Oxygen had been removed from both solutions by repeatedly purging with helium, and was reintroduced at 70 minutes. Periodic removal and assay of aliquots indicated that no inactivation was occurring in the absence of oxygen, but that a rapid drop-off in activity occurred with its reintroduction. We therefore demonstrated that oxygen is required to permanently inactivate HAL with an amine-less cysteine analogue (figure 3.42), suggesting reversible formation of VI in figure 3.41 but irreversible elimination of hydroperoxide from VII.
In conclusion, although both of the proposed mechanisms of HAL with histidine may be extended to explain the inactivation products obtained from treatment of the enzyme with cysteine at high pH, only the Retey model fully addresses the results from amine-less cysteine analogues and is therefore a stronger candidate for the real mechanism of HAL with histidine.

3.4 The X-Ray Crystal Structure of HAL

One of the most powerful tools available in enzymology is X-ray crystallography, in which the three-dimensional structure of a protein is determined by examination of the diffraction patterns observed from the passage of X-rays through a crystalline sample of the protein. Unfortunately, the size and quality of the crystals required to obtain sufficient resolution for identification of key structural elements are often very difficult to achieve. For example, G. K. Farber et al. \textsuperscript{70} were the first to crystallize HAL from \textit{P. putida} but were unable to report much more than the dimensions of the asymmetric unit due to the poor quality of their crystals.

The following year Retey and Schulz\textsuperscript{71} reported similar problems during their attempts at wild type HAL crystallization. However, they discovered that mutating cysteine 273 to alanine (C273A) enabled them to eliminate thiols from their buffering solutions and thereby grow high quality crystals. Cysteine 273 was modeled to be on the surface of HAL\textsuperscript{72} and was the cause of the protein aggregation problem that required the deleterious buffer thiols. As mentioned in section 3.1 during the course of the work described in this thesis the crystal structure of HAL was published. A high level of parallel $\alpha$-helices and subunit organization explains the relatively high melting temperature of HAL (figure 3.8, page 58). Also revealed for the first time was the nature of the dehydroalanine moiety (figure 3.43).
The electrophilic prosthetic group 4-methylidine imidazole-5-one (MIO) contains iminodehydroalanine and forms autocatalytically by cyclization and dehydration of residues 142-144, ASG. Unfortunately, no substrate or substrate analogues were co-crystallized in the active site so that the roles of the residues in the vicinity remain speculative. No metals were observed either, casting further doubt on the hypothesis that imidazole-metal coordination helps acidify the β-protons of histidine. The authors claim that they are able to model in a molecule of histidine that shows the imidazole to be very close to the electrophilic MIO group.

The structure determination by Retey et al. allowed them to address a criticism of their original mechanism, namely the requirement for the loss of aromaticity in the imidazole of histidine. The presence of the MIO in the new mechanism allows the generation of aromaticity in the prosthetic group while it is lost in the substrate (figure 3.44). Most of the details of their previous proposal remain unchanged. The authors point out, however, that the presence of the
imino group instills greater electrophilicity at the exocyclic methylidene carbon because the lone pair on the sp² hybridized nitrogen cannot delocalize into the α,β unsaturated carbonyl. Up to this point they had steadfastly ignored the possibility of iminodehydroalanine in all of their models of the HAL electrophile. Perhaps seeing truly is believing.

![Diagram](image)

Figure 3.44 Retey mechanism incorporating the MIO prosthetic group

3.5 Materials and Methods

3.5.1 General Information

All buffers and synthetic reagents were purchased from the Aldrich or Sigma Chemical companies unless explicitly stated. Water for buffers and enzyme assays were either double distilled or filtered through a Barnstead NANOpure ultrapure water filtration system. Buffers for DE52 and MonoQ chromatography were filtered through a 0.22 μm sterilizing filter and degassed.
prior to use. *Escherichia coli* stocks with the pP$_L$-Lambda plasmid containing the *hutH* (HAL) gene from *Pseudomonas putida* were generously provided by A. T. Phillips. Protein concentrations were determined using the Coomassie dye-binding assay with BSA as a standard. Protein purity was assessed using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS PAGE) as described by Laemmeli. Molecular weight markers for SDS PAGE were carbonic anhydrase (29 kDa) and BSA (66 kDa). All enzyme kinetic assays and ultraviolet (UV) spectra were recorded on a Cary 3E UV-visible spectrophotometer. HPLC were performed on either a Waters 625 LC system with a tunable absorbance detector (active HAL with phosphate buffers) or a Waters 600E with a 484 tunable absorbance detector (chromophoric peptide resolution with organic solvents). Analytical reverse-phase chromatography were run on a Delta-Pak C18 300A 3.9 x 150 mm column. Preparative reverse-phase chromatography were run on a Delta-Pak C18 300A 19.0 x 300 mm column. A unit of HAL enzyme activity is defined as the amount of enzyme necessary to produce 1 µM urocanate min$^{-1}$ at 30°C. For all kinetic assay experiments an extinction coefficient of ε = 10 000 M$^{-1}$ at 277 nm was assumed for urocanate. ESIMS was performed by Mr. Shouming He on a Perkin-Elmer Sciex API300 electrospray mass spectrometer. High and low-resolution MS were performed on a Kratos Concept II HQ mass spectrometer. Amino acid analyses were performed at the UBC NAPS unit. One and two-dimensional $^1$H and $^{13}$C spectra were performed on a Bruker WH-400 MHz or Varian Aspect 500 MHz spectrometer by Liane Darge. Structure based searching was performed by Prof. Raymond Anderson.

### 3.5.2 Purification of HAL

HAL expressing *E. coli* cells$^{60}$ were grown in LB media at 30°C in a rotary incubator set to 200 rpm until the cells gave an absorbance of 0.7 at 600 nm. The temperature was then increased to 42°C C and the cells allowed to grow for a further seven hours. Cells were centrifuged
from growth media (5 000 rpm x 10 min) and resuspended in buffer A (20 mM potassium phosphate, pH 7.3, 2 mM β-mercaptoethanol, 2% glycerol (by volume)) at 5° C. The cells were lysed by passage through a French pressure cell at 20 000 psi and insoluble cell debris was pelleted by centrifugation (5 000 rpm x 30 min). The supernatant was then rapidly heated to 75° C with magnetic stirring in a water bath for 10 min. Insoluble coagulated protein was removed by ultracentrifugation (30 000 rpm x 15 min). The supernatant was concentrated by centrifuge filtration (Millipore Biomax 10 kDa NMWL membrane 15 mL volume) to a concentration of 10-15 mg mL\(^{-1}\). The protein was loaded onto a column of diethylaminoethyl cellulose (DE52) that had been preequilibrated with buffer A. The column was washed with 50 mL of buffer A and eluted with a linear gradient of sodium chloride in buffer A (0 to 500 mM). Fractions containing HAL activity were pooled and freed of salt by buffer exchange into buffer A by centrifuge filtration. Solutions of enzyme up to 20 mg mL\(^{-1}\) were rapidly frozen by emersion in liquid nitrogen and stored at -78° C. When greater purity was desired the thawed HAL stocks could be chromatographed on a Waters AP-1 MonoQ HPLC column (10 x 100 mm) that had been equilibrated with buffer A. The column was washed with 7 mL of buffer A and eluted with a linear gradient of sodium chloride in buffer A (0 to 200 mM over 40 min at 1 mL min\(^{-1}\)). The active fractions were pooled, frozen in liquid nitrogen and stored at -78° C. Analysis of the protein by SDS PAGE showed a single band with an approximate molecular weight of 53 kDa.

### 3.5.3 Activity Assay for HAL

HAL was assayed spectrophotometrically by following changes at 277 nm and 30°C due to increase in concentration of product urocanate. The pre-assay buffer consisted of 100 mM Tris-acetate, pH 9.0 containing 5 mM β-mercaptoethanol. The assay buffer contained an additional 63 mM L-histidine. The test for activity was run as follows: Pre-assay buffer (190 µL) was pipetted
into a quartz cuvette and HAL (10 μg/μL) was introduced, mixed and warmed to 30 °C. Assay buffer (800 μL) at 30 °C was then added, mixed and monitored at 277 nm. Slopes of the absorbance versus time curves were converted into initial rates using an extinction coefficient ε for urocanate of 10 000 L mol⁻¹ cm⁻¹.

3.5.4 Inactivation of HAL with L-Cysteine

DE52 purified HAL (3.5 g) was rapidly thawed and buffer exchanged by centrifuge filtration into CAPS (200 mL, 10 mM) at pH 10.5 and 20° C. A solution of L-cysteine (200 mL, 100 mM) in CAPS (10 mM) at pH 10.5 was added at once to give a final concentration of 50 mM L-cysteine. The solution was gently stirred and aliquots (10 μL) were removed periodically to test for remaining activity and absorbance at 340 nm. After 2 hrs no enzyme activity remained and no further increases in absorbance at 340 nm could be detected. The inactivated HAL was freed of excess cysteine by multiple passes through centrifuge filters and buffer exchange into the tryptic digest buffer (~180 mL buffer A, pH 7.8) using centrifuge filters. The approximate final concentration was approximately 20 mg mL⁻¹.

3.5.5. Digestion of HAL with Trypsin

Solid urea was gradually added to a solution of cysteine-inactivated HAL (10 mg/mL) in tryptic digest buffer (180 mL, 20 mg/mL) until a concentration of 5 M was reached. Portions of trypsin from porcine pancreas were periodically added over 48 hrs until a final concentration of 5% (w/w) relative to HAL was reached. The digestion was monitored by removing aliquots and examining the peptide profile by analytical reverse phase HPLC at 340 and 220 nm. After 48 hrs the digestion was complete and the entire solution was filtered through 0.22 μm syringe filters and
loaded on to a preparative reverse phase column equilibrated with deionized water containing 0.1% trifluoroacetic acid (TFA). Two chromophoric peptides A and B with retention times of 19 and 23 minutes respectively were isolated by elution with water/0.1% TFA and a linear gradient of acetonitrile containing 0.05% TFA, 0-50% at 45 min. Each of A and B were lyophilized and redissolved in V8 protease buffer (buffer A, pH 7.3) to an approximate concentration of 10 mg mL$^{-1}$.

**3.5.6 Digestion of Peptides A and B with V8 Protease**

Portions of V8 protease from *Staphylococcus aureus* were added to solutions of peptides A and B over 48 hrs to final concentrations of 5% (w/w) relative to total protein. The digestion was monitored by removing aliquots and examining the peptide profile by analytical reverse phase HPLC at 340 and 220 nm. Each of the chromophoric peptides A and B yielded two daughter fragments at shorter retention times. A gave rise to A1 (11 min) and A2 (13 min) while B gave rise to B1 (13 min) and B2 (15.5 min). After 48 hrs the digestion was complete and the solutions were filtered through a 0.22 µm syringe filter and loaded on to a preparative reverse phase column equilibrated with deionized water containing 0.1% trifluoroacetic acid (TFA). The four chromophoric peptides were isolated by elution with a linear gradient of acetonitrile containing 0.05% TFA, 0-30% at 30 min and were found to be free of non-chromophoric peptides by ESIMS. All of the peptides were lyophilized and stored at −78°C until analyses were performed.

**3.5.7 Cleavage of Disulfide Bonds in Peptides A1 and B2 with DTT**

Peptides A1 and B2 (approximately 100 µg each) were injected into potassium phosphate buffer (1 mL, 20 mM, pH 7.3). The concentration of each was equalized by monitoring absorbance levels at 333 nm and diluting with small volumes of buffer until A1 and B2 gave the same
absorbance value. Samples were removed from each (100 μL) and acted as controls. Solid DTT was then added to each to give a final concentration of 2 mM. The solutions were incubated at 37 °C for one hour and then submitted for amino acid analysis at the NAPS facility. The peptide A1 treated with DTT gave rise to a chromatogram indicating the presence of liberated cysteine; B2 and the two controls indicated that no cysteine had been liberated.
Chapter Four

Synthesis and Assay of Histidine Analogues
4.1 Introduction

The primary goal of this project was to establish the structure of the active site electrophile of the ammonia-lyases HAL and PAL. The preceding chapter described our successful efforts to isolate and rigorously characterize a chromophoric peptide fragment from HAL that contains a modified version of the electrophile following reaction with cysteine and oxygen. The point of attachment was established to be between the amine of cysteine and the exocyclic methylidene carbon of the MIO electrophile. As was discussed in Chapter Three, the mechanism of formation of this bond may be interpreted in two ways, each of which supports either the Hanson and Havir or Retey mechanisms for the HAL reaction with histidine. Thus, the structural study did not unambiguously differentiate between the two models and we turned our attention to synthesizing and testing substrate analogues of histidine. It was hoped that the degree of reactivity of these compounds would help us determine the true mechanism of HAL, the second major goal of this thesis. The core idea was to exploit the fundamental difference between the two competing mechanisms, that of the interaction between the enzyme electrophile and the substrate.

The basic tenet was that these compounds could only be substrates for HAL via one of the two possible mechanisms. For instance, if the Retey mechanism is operative then exchanging the amine of histidine for a much better leaving group, such as a halide, could still result in a viable substrate (figure 4.1). However, this compound would not be expected to be compatible with the Hanson and Havir mechanism since the first step involves nucleophilic attack by the α-amine. Equivalently, if the Hanson and Havir mechanism is correct then modification of the imidazole moiety might be expected to have only a minor effect on HAL activity. This same change would conceivably result in a compound that could not undergo an elimination via the Retey mechanism (figure 4.2). For either of these compounds to give useful information it is crucial that their structural deviations from histidine be minimal so as to not prevent their access to the active site.
As such, a lack of reactivity of either compound can’t be taken as evidence for a particular mechanism, as there may be other factors involved in the binding of substrates to HAL that our models fail to take into account.

The first part of Chapter Four describes the synthesis and testing of these two different types of histidine analogues. Subsequent sections outline other experiments that were designed to demonstrate HAL inhibition and substrate isotope exchange from the perspective of the Retey mechanism.

Figure 4.1 A comparison showing that the chlorine analogue of histidine could be a substrate for the Retey mechanism but not for the Hanson and Havir mechanism
4.1.1 Properties of Histidine

There are several features of histidine that make it unique among the amino acids. The imidazole has a $pK_a$ value near 7, making it one of the strongest bases that can exist at neutral pH. Its imidazole side chain contains a tertiary amine whose nucleophilicity is enhanced by the fact

Figure 4.2 A comparison showing that the pyrazole analogue of histidine could be a substrate for the Hanson and Havir mechanism but not for the Retey mechanism. Insert in middle left shows the normal C3 deprotonation event in the Retey mechanism.
that the five-membered ring structure holds back the atoms bonded to the nitrogen, making the lone pair more accessible. This nitrogen can therefore readily accept protons for hydrogen bonding. The other, hydrogen-bearing nitrogen can donate a proton for hydrogen bonding. Consequently, the imidazole side chain of histidine is very versatile – almost the chemical equivalent of being ambidextrous.  

The atoms of histidine may be labeled using a variety of descriptors. The standard used in this work is illustrated in figure 4.3. The non-ionized imidazole ring can exist as two tautomers, with the hydrogen atom on either N1 or N3. The actual position depends on the relative proton affinities of the two nitrogen atoms, and can vary with local environmental conditions. The imidazole is protonated below pH 7, destroying its nucleophilicity, and deprotonated above pH 14.4 to give an aromatic anion. Additionally, the C2 atom is observed to slowly exchange its hydrogen atom in protic solvents. This exchange reaction will be encountered in section 4.2.5. 

Figure 4.3 Numbering convention and pH behavior of the histidine imidazole
The Retey mechanism requires a non-protonated imidazole to nucleophilically attack the enzyme MIO. This is consistent with the fact that the pH optimum for the reaction of histidine with HAL is 9.0, although the microenvironment within the active site could perturb the imidazole pKₐ. A positively charged side-chain would seem to be of more benefit to the Hanson and Havir mechanism, since the imidazolium form could better stabilize partial negative-charge character at the β-carbon. Their mechanism also requires a neutral amine to attack the enzyme MIO. These factors could play important roles in the ability of our substrate analogues to interact with the enzyme.

4.2 Results and Discussion

4.2.1 Synthesis of Halogen Analogues of Histidine

The replacement of the histidine α-amine by either fluorine, chlorine or bromine was intended to produce substrates with much better leaving groups than ammonia but that are still small enough to fit into the enzyme active site. A simple one-step synthesis that avoids racemized product is halodeazotization where the amine is oxidized to a diazo function, which is an excellent leaving group (figure 4.3). It is believed that the participation of the carboxylate in the deazotization step results in a transient α-lactone which is subsequently attacked by a halide ion. This double displacement has been shown to result in 100% retention of configuration. For the bromo and chloro derivatives L-histidine was first suspended in concentrated hydrobromic or hydrochloric acid and cooled to 0 °C. A solution of sodium nitrite was then slowly dripped into the stirring mixture over approximately one hour. An audible hiss could be heard as the nitrogen gas rapidly bubbled out of the solution. After filtering off the precipitated salts and reducing the solvent, the product was recrystallized from hot water. The ¹H NMR, mass spectra and melting point data were consistent with published reports.
The preparation of the fluoro analogue became slightly problematic when it was discovered that water is a better nucleophile than fluoride ion in concentrated hydrofluoric acid, and the 3-hydroxy product resulted exclusively. This was rectified by using 70% hydrogen fluoride in pyridine, oven-dried sodium nitrite and maintaining strictly anhydrous conditions. The final product was isolated as a white solid and gave good $^1$H and $^{19}$F NMR spectra and showed the predicted mass value from LSIMS. There was no evidence for a rearrangement product that had been reported in the fluorodeazotization of phenylalanine. In that work the authors propose participation of the phenyl ring in a triple displacement reaction (figure 4.5). A fluoride ion ultimately attacks the cyclopropyl ring at the $\beta$-carbon resulting in 3-fluoro-2-phenylpropionic acid. We don’t observe this product, perhaps because the histidine imidazole is completely protonated in the highly acidic environment of the reaction and is not nucleophilic enough to participate in the lactone opening.
4.2.2 Assay of Halogen Analogues of Histidine with HAL

Each of the three halogenated compounds were tested for activity with HAL at pH 9.0. Unfortunately, no production of urocanate was detected for either the bromo, chloro or fluoro analogues after prolonged exposure to the enzyme. Fluorohistidine was also assayed for production of fluoride ions using a fluoride sensitive electrode (figure 4.6). No increase in fluoride ion concentration could be detected in the presence of a large amount (1 mg) of HAL. The chloro analogue was tested as an inhibitor of HAL at concentrations as high as 100 mM but no inhibition was observed. Our conclusion was that these compounds were not binding properly to the enzyme, possibly because it is necessary for the \( \alpha \)-amine of histidine to be protonated or able to hydrogen bond in the active site. Replacement by halogen negates this interaction and prevents us from interrogating the active site in the manner that the compounds were designed for. Therefore, this negative result doesn’t shed any light on the question of which of the two proposed mechanisms is correct.
4.2.3 Synthesis of Imidazole Analogues of Histidine

As illustrated in figure 4.2, the substitution of pyrazole for imidazole in histidine is intended to nullify a key step in the Retey mechanism. The presence of nitrogen adjacent to the abstractable proton eliminates an important resonance form that previously allowed the delocalization of the incipient negative charge into the imidazole ring. It was hoped that the change in position of this nitrogen atom in the imidazole ring wouldn’t perturb binding but would be sufficient to prevent a reaction via the Retey mechanism. However, there is no obvious reason why the Hanson and Havir mechanism could not be used to explain the elimination of ammonia from this substrate analogue. This analysis assumes that pyrazole can function as well as imidazole in coordinating to active site residues, or perhaps to a metal ion as suggested in section 1.6. Therefore, any evidence of ammonia elimination from this compound could be interpreted in favour of the Hanson and Havir mechanism.
The compound that we were interested in had previously been prepared by J. C. Vederas et al.\textsuperscript{80} Their method involves a nucleophilic attack on β-lactones prepared from protected L-serine using a modified Mitsunobu procedure (figure 4.7). The use of L-serine ensures that we only produce L-alanine derivatives. First, the amine of L-serine was protected with benzylchloroformate to give 10, which was subsequently recrystallized from hot ethyl acetate. The β-lactone required the preparation of dimethyl azodicarboxylate 9 (DMeAD) which was synthesized by treating hydrazine with two equivalents of methyl chloroformate and sodium carbonate in ethanol to give 8, followed by oxidation with N-bromosuccinimide and pyridine. Compound 9 was purified by careful vacuum distillation and was added drop-wise to a solution of dry triphenylphosphine in dry, freshly distilled THF at -78 °C. Pre-forming the adduct of 9 and triphenylphosphine in this way is essential for obtaining reasonable yields.\textsuperscript{81} To this slurry was slowly added a solution of 10 in dry THF. The protected lactone 11 was obtained as white crystals following flash chromatography on silica gel. Reaction with pyrazole in acetonitrile produced 12, which was followed by removal of the benzylxycarbonyl with hydrogen gas and 10\% palladium on carbon to give the final product, \( \beta \)-\((\text{pyrazol-1-yl})\)-L-alanine, 13.
Figure 4.7 Synthesis of β-(pyrazol-1-yl)-L-alanine (from reference 81)
Two other side-chain variants were attempted, with imidazole and triazole in place of pyrazole. In each case complex mixtures were obtained at the lactone opening stage. The $^1$H NMR spectra were very complicated and the mass spectra appeared to indicate that oligomeric species were forming with two or three lactones becoming attached to one imidazole or triazole (figure 4.8 showing imidazole). These compounds were not pursued further.

Figure 4.8 Mass spectrum of products from a reaction of imidazole and 11 showing masses rationalized by oligomeric species
4.2.4 Assay of Imidazole Analogues of Histidine with HAL

β-(Pyrazol-1-yl)-L-alanine was tested as a possible substrate for HAL and it was found that the compound is neither a substrate nor an inhibitor of the enzyme at concentrations up to 100 mM. The unsaturated product obtained from ammonia elimination of the pyrazole analogue would very likely have a different UV absorbance profile from urocanate. Therefore, two assays were performed, one examining overall UV absorbance changes and one looking at changes in the $^1$H NMR spectrum of the pyrazole analogue with HAL in deuterated buffers. In both cases no significant changes were observed, suggesting that the compound was not getting into the active site to react and/or that the thermodynamics of the elimination are not as favorable compared to the histidine/urocanate pair.

The fact that β-(pyrazol-1-yl)-L-alanine doesn't inhibit HAL indicates that the compound simply isn’t binding to the enzyme. The N1 nitrogen in the imidazole of histidine possibly plays a crucial role in properly orienting the substrate in HAL’s active site. Relocating it to the 4-position could destroy a key interaction. The results from our substrate analogue experiments indicated that there is a high degree of specificity for the natural substrate, as subtle structural changes can greatly impair binding. A new approach was therefore taken that looked at the potential for reversible proton exchange in substrate analogues.

4.2.5 Test for Exchange of Imidazole Protons

In 1995 J. Retey et al. published a study of substrate demonstrating that 3-hydroxyphenylalanine is a better substrate for PAL than phenylalanine. He rationalized that the intermediate formed after addition of the aromatic ring to the enzyme MIO was stabilized by resonance with the ring hydroxyl (figure 4.9). This work was expanded in 1998 when Gloge et al. showed that the two “halves” of 3-hydroxyphenylalanine, 3-hydroxytoluene and glycine, could...
synergistically inhibit PAL. By themselves however, these compounds exhibit little or no inhibition. The authors speculate that the aromatic ring of 3-hydroxytoluene is adding to the enzyme MIO in accordance with their proposed PAL mechanism (figure 1.20). No direct evidence of a covalent attachment is presented to back this claim however.

![Diagram of MIO and 3-hydroxyphenylalanine interaction](image)

**Figure 4.9** Addition of m-hydroxyphenylalanine to the MIO showing intermediate stabilization by resonance (top). Glycine and m-hydroxytoluene synergistically inhibiting HAL (from reference 84)

<table>
<thead>
<tr>
<th>Kinetic Constants</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K_f$(gly)</td>
<td>&gt; 20 mM</td>
</tr>
<tr>
<td>$K_f$(m-OH-tol)</td>
<td>= 17 mM</td>
</tr>
<tr>
<td>$K_f$(gly/m-OH-tol)</td>
<td>= 0.9 mM</td>
</tr>
</tbody>
</table>
We became interested in this result because of its possible extension to our experiments on HAL using modified substrates. In particular, we were intrigued with the idea of combining an isotope exchange experiment, similar to Furata's as outlined in section 1.10, with the synergistic inhibition experiment. Before going into the details of our proposal, some further interpretation of the significance of Furata's results is merited.

As illustrated in figure 1.21 on page 26 the observed isotope exchange at C5 of the histidine imidazole can reasonably be rationalized by the Retey mechanism. However, it can also be explained, with some qualification, in terms of the Hanson and Havir mechanism (figure 4.10, showing histidine already attached to the rearranged MIO).

If one accepts an E1cb mechanism, the transient allylic carbanion can rearrange to allow reversible isotope exchange at C5. This idea, as well as the Hanson and Havir mechanism in general, is more plausible if an imidazole-metal coordination is invoked to stabilize the negative charge. The major problems with this interpretation of the isotope exchange at C5 are the destruction of the aromaticity in the imidazole required to affect the exchange and the high pK_a of the α-proton, a problem inherent to the Hanson and Havir mechanism. Still, the apparent ambiguity of this isotope exchange further muddies the waters of distinction between the two proposed mechanisms. We therefore wanted to design an experiment that would clarify the isotope exchange result and give unambiguous support to one of the two models.
Figure 4.10 Isotope exchange at the 5-position of the histidine imidazole is possible to rationalize using the Hanson and Havir mechanism

Our experiment involves replacing histidine with two pairs of reagents, 4-methylimidazole and glycine, or imidazole and L-alanine (figure 4.11). It was anticipated that our pairs of reagents might synergistically bind to the active site and perhaps also show exchange of the C5 proton. The key concept behind using these disconnected pairs of reagents in this experiment is that while successful exchange of deuterium into 4-methylimidazole can be explained by both mechanisms,
only the Retey mechanism can reasonably be used to justify exchange of the C5 protons in imidazole. This concept is further illustrated in figures 4.12 and 4.13.

Figure 4.12 Justification for C5 proton exchange in 4-methylimidazole and imidazole using the Retey mechanism
Hanson and Havir mechanism

Allylic anion can pick up deuterium at C5 and the methyl carbon

Figure 4.13 Justification for C5 proton exchange in 4-methylimidazole only using the Hanson and Havir mechanism

There are some pitfalls within our plan, as illustrated using 4-methylimidazole and glycine in figure 4.14. The most significant is that if the imidazole does add to the MIO and is deprotonated at C5, aromaticity is restored to the substrate. This could lead to a dead-end complex, stabilized to such a degree that it effectively permanently inactivates the enzyme. No exchange at C5 would be observed. The same complex is reached by deprotonation at the methyl carbon followed by a 1,3-hydrogen shift. This in itself would not be such a disaster though; such a result could still be interpreted as a validation of the Retey mechanism. Assuming that a dead-end complex doesn’t occur, the reversibility of the steps allows for the possibility of deuterium wash-in at the methyl position or C5 if the assay is performed in deuterated buffers.
Figure 4.14 Scheme illustrating route for deuterium incorporation into 4-methylimidazole through a possible covalent intermediate
Solutions of L-alanine and imidazole or glycine and 4-methylimidazole in buffered D$_2$O were prepared in NMR tubes and incubated with HAL at 37 °C. Spectra were periodically obtained and examined for exchange of deuterium at all positions. After more than 72 hours incubation the only apparent exchange of protons for deuterium occurred at the C2 position of imidazole and 4-methylimidazole (figure 4.15, methyl protons and internal standard not shown). This exchange was anticipated owing to the slight acidity of these protons. Moreover, no synergistic inhibition of HAL was observed with either pair of histidine analogues. Perhaps the absence of an hydroxy moiety in the imidazole ring of our compounds precluded formation of a covalent species.

*Figure 4.15* $^1$H NMR spectra of 4-methylimidazole and HAL in buffered D$_2$O showing imidazole C2 and C5 protons with deuterium wash-in at C2 only (methyl protons not shown)
theorized in the Retey study with m-hydroxytoluene. The lack of any exchange at C5 meant that no conclusions about HAL’s mechanism could be drawn from this study.

4.2.6 Inhibition of HAL with 2-Methylimidazole

During the course of the previous study on isotope exchange it was discovered that 2-methylimidazole is a competitive inhibitor of HAL with a $K_i$ of 20 mM (figure 4.16). Imidazole is known to be an uncompetitive inhibitor,\textsuperscript{27} by definition not binding to the active site of HAL. The introduction of a methyl group at C2 could be making the imidazole more electron rich and more nucleophilic towards the electrophilic MIO group at the active site of HAL. Alternatively, the alkyl group could simply be making the molecule more hydrophobic and more suitable to the active site, which may contain a complementary hydrophobic pocket.

![Graph showing competitive inhibition of HAL with 2-methylimidazole](image)

*Figure 4.16 Graph showing competitive inhibition of HAL with 2-methylimidazole*
4.3 Overall Conclusions

The first major goal of this project was realized when the structure of the electrophile at the active site of HAL was unambiguously identified as a 4-methylidene imidazole-5-one, or MIO group. The electrophile was isolated as a chromophoric octapeptide and was found to contain a covalently attached cysteine residue, with three isomers proposed for the amine-methylidene carbon linkage. The formation of the chromophoric peptide can be rationalized by both the Hanson and Havir and Retey mechanisms, considering the requirements for cysteine, buffer pH above 10 and the presence of oxygen. The fact that cysteine is a good competitive inhibitor of HAL suggests that it may interact with the enzyme active site in a manner analogous to the natural HAL substrate histidine. This in turn raises the possibility that a link may be made between the mechanism of formation of the inactivation chromophore and the actual mechanism of HAL. Based on the inhibitory behavior of several amine-less cysteine analogues with HAL, in the presence and absence of oxygen, a much stronger case is made for a Retey mechanism-based model of inactivation of HAL with cysteine, which therefore strengthens the Retey mechanism for HAL with histidine. A major criticism of the original Retey mechanism is also alleviated with the participation of the MIO group, as aromaticity lost in the substrate is gained in the electrophile. Additionally, the Retey mechanism easily and simply explains the loss of deuterium at the C5 position in the imidazole ring of labeled histidine noted elsewhere, whereas the Hanson and Havir mechanism requires some unlikely intermediates. Considered in the context of the activity of dehydroalanine-less HAL mutants with 5-nitrohistidine and the real need to address the extremely non-acidic β-protons in histidine, our overall results very strongly support the Retey mechanism for HAL with histidine.
4.4 Materials and Methods

4.4.1 General Information

All synthetic reagents were purchased from the Aldrich or Sigma Chemical companies unless explicitly stated. Pyridine and acetonitrile were distilled from calcium hydride and THF was distilled from sodium immediately prior to use. 2-methyl and 4-methylimidazole were recrystallized from petroleum ether. Fluoride ion concentrations were determined using an ORION Model 94-09 fluoride electrode.

4.4.2 Synthesis of (S)-2-chloro/bromo-3-(imidazol-4-yl) Propionic Acid

The synthesis of (S)-2-chloro/bromo-3-(imidazol-4-yl) propionic acid has been described by J. C. Craig et al. After crystallizing the product from hot water it was found that injection on a Delta-Pak C18 300A 19.0 x 300 mm column with water as eluent allowed the removal of a minor, highly colored impurity which was absorbing strongly at 277 nm, the wavelength at which the activity assay is performed. The resulting product showed identical $^1$H NMR and MS as reported earlier.

4.4.3 Synthesis of (S)-2-fluoro-3-(imidazol-4-yl) Propionic Acid

The synthesis of many fluoro analogues of amino acids have been described by G. A. Olah et al. (S)-2-fluoro-3-(imidazol-4-yl) propionic acid was prepared by closely following the procedures established for other amino acids. A solution of hydrofluoric acid in pyridine (70%, 13 mL) in a plastic vial was fitted with a rubber septum and dry nitrogen was flushed through as it was cooled to 0 °C. Vacuum dried (<0.1 mmHg, 24 hrs) L-histidine (1.0 g, 6.4 mmol) was introduced in small portions by periodic removal of the septum. To this solution was slowly added
oven dried (140 °C, 24 hrs, cooled in desiccator) sodium nitrite (640 mg, 10 mmol) maintaining a temperature below 5 °C. As the reaction proceeded the solution turned a pale green color. After one hour the ice bath was removed and the solution allowed to warm to room temperature and stirred (1 hr). The pH was adjusted to 4.0 with sodium hydroxide (10 M) and insoluble salts filtered off. The remaining pale yellow solution was lyophilized to a fine white powder that was re-dissolved in a minimum of hot deionized water. The product crystallized upon slow cooling and was dried on a high vacuum line (300 mg, 30%). $^1$H NMR (400 MHz, D$_2$O) δ 3.3 (m, 2H), 5.05 (dt, $J_{HH}$=12 Hz, $J_{HF}$=100 Hz), 7.25 (s, 1H), 8.5 (s, 1H); $^{19}$F{$^1$H} (188 MHz, D$_2$O) δ -53.84 (s, 1F). Mass (LSIMS); 159 m/z (M+H)$^+$

As was the case with the chloro and bromo species, further purification on a reverse-phase column removed a minor, highly colored impurity. The resulting product gave an identical $^1$H NMR spectrum.

4.4.4 Spectrophotometric Assay of Fluoro, Chloro and Bromo Analogues of Histidine

Solutions of the three halogen analogues of histidine were prepared in concentrations varying between 5 and 100 mM using tris-acetate buffer (100 mM pH 9.0). Aliquots of the halogen analogues (990 μL) were added to cuvettes and equilibrated at 30 °C for several minutes. Aliquots of HAL (10 μg/10 μL) were added and the cuvettes rapidly mixed. Changes in absorbance (277 nm) were monitored to follow production of urocanate. In none of the cases were increases observed at this or any other wavelength.
4.4.5 Fluoride Ion Concentration Assay for (S)-2-fluoro-3-(imidazol-4-yl) Propionic Acid

A solution of (S)-2-fluoro-3-(imidazol-4-yl) propionic acid (100 mM) was prepared in potassium fluoride (5 mM, pH 9.0). A fluoride sensitive electrode was introduced, along with a miniature stirring bar. The electrode began recording a baseline fluoride ion concentration (200 seconds) before HAL (1 mg/200 μL) was injected into the stirring solution. The electrode immediately recorded the drop in fluoride ion concentration from the dilution but showed no further change over the next several hundred seconds. The enzyme was found to be fully functional under these conditions by the subsequent addition of histidine and production of urocanate.

4.4.6 Inhibition of HAL with (S)-2-chloro-3-(imidazol-4-yl) Propionic Acid

Solutions of L-histidine (5 mM) in tris-acetate (100 mM, pH 9.0) were prepared with and without the inclusion of (S)-2-chloro-3-(imidazol-4-yl) propionic acid (100 mM). When these two solutions were assayed for activity in the presence of HAL (10 μg) no difference in rates of production of urocanate were observed, indicating that no enzyme inhibition was occurring.

4.4.7 Synthesis of β-(pyrazol-1-yl)-L-Alanine

The synthesis of β-(pyrazol-1-yl)-L-alanine has been described by Vederas et al. The final product obtained after catalytic hydrogenation gave the expected mass from LSIMS and 1H NMR spectrum.
4.4.8 1H NMR Assay of β-(pyrazol-1-yl)-L-Alanine

10 mL of a solution of β-(pyrazol-1-yl)-L-alanine (50 mM) in tris-acetate (50 mM, pH 9.0) was prepared and lyophilized. D₂O (10 mL) was added and the solution lyophilized. D₂O (10 mL) was added and an aliquot (1 mL) was injected into an NMR tube. A 1H NMR spectrum was obtained as a standard. HAL (150 µg) was added via syringe and the tube incubated at 37 °C (1 hr). A second 1H NMR spectrum was obtained and showed no changes from the control other than a more pronounced water peak. The NMR tube was incubated for 24 hours but a third spectrum obtained after that time showed no additional change from the control.

4.4.9 Test for Deuterium Wash-in with 4-Methylimidazole with HAL

A solution of 4-methylimidazole (20 mM), glycine (20 mM) and tris-acetate (20 mM, pH 9.0) was prepared using 10 mL D₂O. The solution was lyophilized to remove traces of water and the residue was redissolved in D₂O (10 mL) and an aliquot (1 mL) was pipetted into an NMR tube. β-Mercaptoethanol was added to a final concentration of 1 mM. A 1H NMR spectrum was then obtained to establish a control spectrum. HAL (100 µg) was added by injection and the tube was suspended in a water bath at 37 °C. Spectra were periodically obtained over a period of 72 hours. The only significant spectral change over time was the disappearance of the singlet arising from the C2 proton, which was anticipated.

4.4.10 Inhibition of HAL with 2-Methylimidazole

A solution of 2-methylimidazole (200 mM) was prepared in tris-acetate buffer (100 mM, pH 9.0) and was combined with solutions of L-histidine in tris-acetate buffer (100 mM, pH 9.0). These were diluted to create a series of solutions which varied in both imidazole and histidine
concentration. An aliquot of each (1 mL) was pipetted into a cuvette and equilibrated at 37 °C for a few minutes before HAL (10 μg) was introduced. Increases in absorbance at 277 nm were converted into initial rates of production of urocanate. Plots of inverse initial rate versus the inverse of histidine concentration resulted in a series of lines, one for each different concentration of imidazole inhibitor. The pattern of these lines revealed that 2-methylimidazole is competitively inhibiting HAL, with a $K_i$ value of 20 mM.


4 Hofmann, A. W., *Ber.* **14** 659 (1881)


36 Friedel, C., Crafts, J. M., *Compt. Rend.* **84** 1392 (1877)


72 Rost, B., Methods Enzymol. 266 525 (1996)


75 Edlbacher, S., von Bidder, H., Hoppe-Seylers Z. Physiol. Chem. 276 126 (1942)


